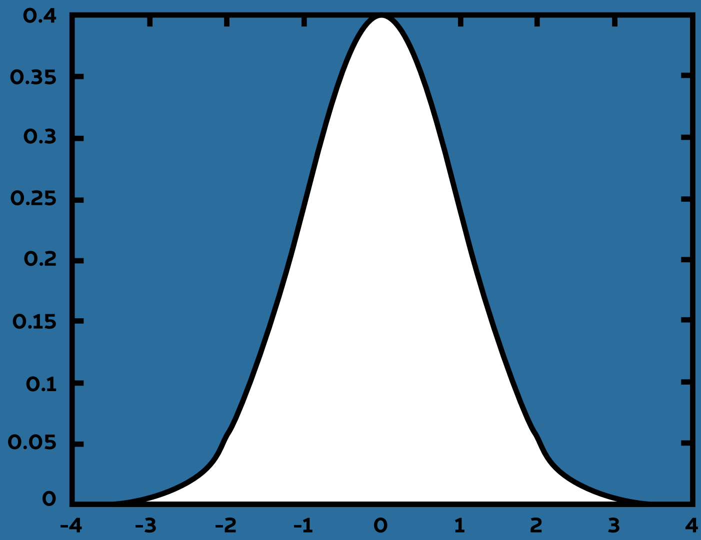
Olvass és tanulj !

#05

Bevezetés a valószínűségszámításba

A „véletlen” az, amikor Isten névtelen akar maradni.



Szerző: Medvegyev Péter

BUDAPESTI CORVINUS EGYETEM
MAGYAR NEMZETI BANK

Medvegyev Péter

Bevezetés a valószínűségszámításba

Közgazdaságtudomanyi Kar
Matematika Tanszék

Budapesti Corvinus Egyetem
Közgazdaságtudomanyi Kar
Matematika Tanszék

„A Budapesti Corvinus Egyetem és a Magyar Nemzeti Bank
együttműködési megállapodása keretében támogatott mű.”

ISBN 978-963-503-647-9
Kiadás 2017

Nyomdai kivitelezés: CC Printing Kft.

TARTALOM

I.

ALAPFOGALMAK

A valószínűségszámítás tárgyalásakor a szokásos kiindulási pont a valószínűségi mező, az eseménytér és az események alapvető tulajdonságainak bevezetése. A feladatok és a konkrét valószínűségszámítási problémák meghatározásakor a valószínűségeket gyakran közvetlenül nem ismerjük, és az alkalmazások során csak a feltételes valószínűségekre tudunk következtetni és a feltételes valószínűségekből a teljes valószínűség tétele segítségével számoljuk ki a valószínűségeket. Ennek megfelelően a teljes valószínűség tételét mint a valószínűségszámítás egyik alapeszközét az alapdefiníciókkal és alapfogalmakkal együtt tárgyaljuk.

1.1. Eseménytér

Minden valószínűségszámítási modell megadásakor meg kell adni a lehetséges kimenetelek Ω halmazát, az Ω halmaz részhalmazaiból álló megfigyelhető események $\mathscr{A}$ családját és az $A \in \mathscr{A}$ megfigyelhető események $\mathbf{P}(A)$ valószínűségét. Az Ω alaphalmazt szokás biztos eseménynek is nevezni. Az $(\Omega, \mathscr{A}, \mathbf{P})$ hármast együtt valószínűségi mezőnek mondjuk. Az $\mathscr{A}$ halmaz elemei tehát maguk is halmazok, mégpedig az Ω biztos esemény részhalmazai. Vagyis ha $A \in \mathscr{A}$, akkor $A \subseteq \Omega$. Bár nyilvánvaló, azért érdemes hangsúlyozni, hogy a $\mathbf{P}$ valószínűség egy az $\mathscr{A}$ halmazrendszeren értelmezett halmazfüggvény, vagyis egyrészt halmazokhoz rendel számokat, másrészt csak az $\mathscr{A}$ halmazrendszerben szereplő halmazokhoz rendel számot. Ha az Ω valamely részhalmaza nem eleme a lehetséges események $\mathscr{A}$ családjának, akkor ezt a halmazt értelemszerűen nem tekintjük eseménynek, és ezért nem is tulajdonítunk neki valószínűséget. Szemléletesen a lehetséges események $\mathscr{A}$ halmaza azokból az eseményekből áll, amelyeket, legalábbis elvileg, meg tudunk figyelni, vagyis amelyekről el tudjuk dönteni, hogy bekövetkeztek vagy sem. Elvileg tehát különbséget kell tenni események és (rész)halmazok között, vagyis minden esemény részhalmaz, de nem megfordítva. Ennek ellenére, némiképpen pontatlanul, hacsak nem okoz zavart, a halmaz és az esemény elnevezést egymás szinonimájaként fogjuk kezelni.

1.1.1. Definíció. A valószínűségszámítás axiómái szerint az $(\Omega, \mathscr{A}, \mathbf{P})$ valószínűségi mező a következő tulajdonságokkal rendelkezik:

1. Minden $A \in \mathscr{A}$ esetén a $\mathbf{P}(A)$ értelmes és $\mathbf{P}(A) \geq 0$.

2. $\Omega \in \mathscr{A}$ és $\mathbf{P}(\Omega) = 1$. Az Ω biztos esemény egyúttal lehetséges esemény is, és a biztos esemény valószínűsége 1.

3. Ha $A \in \mathscr{A}$, akkor az $A^c \in \mathscr{A}$ szintén teljesül. Vagyis egy megfigyelhető esemény komplementere is megfigyelhető, így lehetséges esemény.

4. Ha $A_1, A_2, \ldots$ legfeljebb megszámlálható sok esemény, vagyis $A_k \in \mathscr{A}$, akkor $\cup_{k=1}^{\infty} A_k \in \mathscr{A}$. Tehát az $\mathscr{A}$ zárt a legfeljebb megszámlálható egyesítés műveletére.

5. Ha $A_1, A_2, \ldots$ legfeljebb megszámlálható sok páronként diszjunkt esemény, akkor

$$\mathbf{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbf{P}(A_k),$$

vagyis legfeljebb megszámlálható sok páronként diszjunkt esemény egyesítésének valószínűsége éppen az egyesítésben szereplő események valószínűségének összege. Érdemes felidézni, hogy páronként való diszjunktságon értelemszerűen azt értjük, hogy ha $n \neq m$ két különböző index, akkor $A_n \cap A_m = \emptyset$.

Mivel $\cap_k A_k = \cap_k \left(A_k^c\right)^c = \left(\cup_k A_k^c\right)^c$, ezért a lehetséges események $\mathscr{A}$ halmaza nemcsak a legfeljebb megszámlálható egyesítésre, hanem a legfeljebb megszámlálható metszetre nézve is zárt, vagyis megszámlálható sok lehetséges esemény metszete is lehetséges esemény.

1.1.2. Példa. Két kockával való dobáshoz tartozó valószínűségi mező.

Tegyük fel, hogy két kockát egyszerre feldobunk. Ilyenkor a lehetséges kimenetelek halmaza az (i, j) számpárokból áll, ahol i és j az $1,2,3,4,5,6$ számok valamelyikét vehetik fel. Vagyis az Ω elemeinek száma 36. Az $\mathscr{A}$ elemei azok a halmazok, amelyek megfigyelhetőek. Hogy mely számpárok figyelhetőek meg, az attól függ, hogy a kockák különbözőek vagy sem. Ha megkülönböztethetőek, akkor az Ω minden részhalmaza eleme az $\mathscr{A}$ eseményrendszernek. Ha azonban nem, akkor például az $(1,2)$ kimenetelből álló halmaz, vagyis az $\{(1,2)\}$ egy elemű halmaz nem esemény, csak az $\{(1,2),(2,1)\}$ két elemű halmaz lesz eleme az $\mathscr{A}$ megfigyelhető eseményrendszernek. Ennek megfelelően a $\mathbf{P}$ valószínűség nem lesz feltétlenül értelmezve az összes részhalmazra, és lesznek olyan kimenetelek is, amelyekhez tartozó egy elemű halmazoknak nem lesz valószínűsége. Például a $\mathbf{P}(\{(1,1)\}) = 1/36$ egyenlőség teljesül, függetlenül attól, hogy a kockák megkülönböztethetőek vagy sem, de a $\mathbf{P}(\{(1,2)\}) = 1/36$ egyenlőség csak akkor értelmes, ha a kockák megkülönböztethetőek. Ennek kapcsán érdemes nyomatékosan hangsúlyozni, hogy a $\mathbf{P}$ a részhalmazokon és nem a kimeneteleken van értelmezve, ezért semmi sem biztosítja azt hogy az egy elemű halmazok rendelkeznek valószínűséggel.

□

1.1.3. Példa. Sztochasztikus folyamatok, filtrációk.

Az előző példában szereplő kockákat a legegyszerűbben úgy tudjuk megkülönböztetni, ha egymás után dobjuk fel őket. Ha az Ω kimenetelei függnek az időtől, akkor sztochasztikus folyamatról beszélünk. Ilyenkor az Ω elemeit azonosíthatjuk a sztochasztikus folyamat trajektóriáival, ahol trajektórián a folyamat lehetséges lefutásait megadó (időváltozós) függvényeket értjük. Sztochasztikus folyamatok esetén kézenfekvő beszélni valamely időpontig megfigyelhető eseményekről. Ha $\mathscr{F}_t$ jelöli a t időpontig bezárólag megfigyelhető események összességét, akkor nyilván $\mathscr{F}_t \subseteq \mathscr{F}_s$ valahányszor $t \leq s$. Az $(\Omega, \mathscr{F}_t, \mathbf{P})$ hármas minden t-re egy önálló valószínűségi mezőt definiál. A

t időparaméterrel indexelt $(\mathscr{F}_t)$ eseményrendszert szokás filtrációnak mondani. Például ha egy kockát egymás után kétszer dobunk fel, akkor az $\mathscr{A}$ eseményrendszer a 36 elemből álló

$$\Omega = \{(i,j) \mid i,j = 1,2,\ldots,6\} = R_6 \times R_6,$$

Descartes-szorzat összes részhalmazából áll, ahol $R_6 \triangleq \{1,2,3,4,5,6\}$. Értelemszerűen $\mathscr{F}_2 = \mathscr{A}$. Ugyanakkor az $\mathscr{F}_1$ elemei az olyan $A \subseteq \Omega$ halmazok, amely első koordinátája az R_6 egy részhalmaza, a második koordinátája azonban mindig a teljes R_6 halmaz, ugyanis az első kocka feldobása után csak a számpár első elemét tudjuk meghatározni, a második elem ekkor még az R_6 tetszőleges eleme lehet.

□

1.1.4. Példa. Egy kockával addig dobunk, amíg 6-os nem jön ki. Adjuk meg a valószínűségi mezőt!

A lehetséges kimenetelek Ω halmaza az $(i_1, i_2, \ldots, i_{n-1}, 6)$ alakú véges sorozatok halmaza, ahol az i_k az $1,2,3,4,5$ számok valamelyike. Mivel előfordulhat, hogy végtelen sok dobás esetén sem lesz a dobott szám hatos, ezért az Ω alaptérhez hozzá kell még csapni az $(i_1, i_2, \ldots)$ alakú végtelen sorozatokat is. A lehetséges események $\mathscr{A}$ halmazának vehetjük az Ω összes részhalmazának halmazát. A végtelen sorozatok halmazának valószínűsége nulla, az $\omega = (i_1, i_2, \ldots, i_{n-1}, 6)$ alakú kimenetelek valószínűsége $1/6^n$. Alternatív módon, kényelmi megfontolásokból, az Ω alaptérnek tekinthetjük az $R_6 \triangleq \{1,2,3,4,5,6\}$ értékű végtelen sorozatok halmazát is, vagyis feltehetjük, hogy a kockát mindig végtelen sokszor feldobjuk. De a $\mathbf{P}$ valószínűség meghatározásakor a már elmondottak szerint kell eljárni. Ha N az olyan sorozatok halmaza, amelyek nem tartalmazzák a hatos számjegyet, akkor $\mathbf{P}(N) = 0$, a hatos számjegyet tartalmazó sorozatok valószínűsége $1/6^n$, ahol n az első hatos indexe. Bármennyire is ártatlannak tűnik a definíció, potenciálisan tartalmaz egy matematikai ellentmondást. A példában $\mathscr{A}$ az összes részhalmazok halmaza. Az axiómák szerint az $\mathscr{A}$ minden elemének kell valószínűséget tulajdonítanunk. Az Ω összes részhalmaza között vannak olyanok is, amelyek elemeinek száma nem megszámlálható. A $\mathbf{P}$ valószínűség megadásakor azonban látszólag csak az egyes kimenetelek valószínűségeit adtuk meg. Ha ismerjük az egyes kimenetelek, az Ω biztos esemény $\omega \in \Omega$ elemeinek valószínűségét, akkor a valószínűség definíciója alapján csak az olyan $A \subset \Omega$ események valószínűségét tudjuk megadni, amelyek elemeinek számossága legfeljebb megszámlálható. A hatosra végződő véges sorozatok számossága megszámlálható, így az Ω minden részhalmazának tudunk valószínűséget definiálni[1].

□

[1]Ha az Ω téren az összes sorozatok halmazát értjük, akkor azokat a sorozatokat, amelyek az első hatosig megegyeznek ekvivalensnek tekintjük. Így az Ω elemei valójában ekvivalenciaosztályok. Vegyük észre, hogy mivel az axiómák nem teszik lehetővé a valószínűség minden további megfontolás nélküli kiterjesztését, elvileg matematikai ellentmondások származhatnak az $(\Omega, \mathscr{A}, \mathbf{P})$ definiálásakor. Ezen problémák tárgyalásától azonban eltekintünk.

A valószínűségi mező definíciójából egy sor kézenfekvő tulajdonság vezethető le. Ezek közül tekintsük néhány igen egyszerűt:

1. Tetszőleges $A \in \mathscr{A}$ esetén $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$. Valóban: mivel $A \cup A^c = \Omega$ és $A \cap A^c = \emptyset$, ezért

$$1 = \mathbf{P}(\Omega) = \mathbf{P}(A \cup A^c) = \mathbf{P}(A) + \mathbf{P}(A^c),$$

amiből az egyenlőség már triviális.

2. Ha $A, B \in \mathscr{A}$ és $A \subseteq B$, akkor $\mathbf{P}(B \backslash A) = \mathbf{P}(B) - \mathbf{P}(A)$, speciálisan a $\mathbf{P}(B \backslash A) \geq 0$ miatt $\mathbf{P}(B) \geq \mathbf{P}(A)$. A bizonyításhoz emlékeztetünk arra, hogy $B \backslash A \stackrel{\circ}{=} B \cap A^c$, ahol az $\stackrel{\circ}{=}$ egyenlőség a definíció szerint való egyenlőséget jelenti. A $B \backslash A$ és az A halmazok diszjunktak, és mivel a feltétel szerint $A \subseteq B$, ezért

$$(B \backslash A) \cup A \stackrel{\circ}{=} (B \cap A^c) \cup A = (B \cap A^c) \cup (B \cap A) = B \cap (A \cup A^c) = B \cap \Omega = B,$$

ebből

$$\mathbf{P}(B \backslash A) + \mathbf{P}(A) = \mathbf{P}(B),$$

amiből az egyenlőség már evidens.

3. Tetszőleges $A, B \in \mathscr{A}$ esetén $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$. A bizonyításhoz vegyük észre, hogy

$$A \cup B = (A \backslash B) \cup (A \cap B) \cup (B \backslash A),$$

ahol az egyesítés páronként diszjunkt, ugyanis például

$$\begin{aligned}(A \backslash B) \cap (A \cap B) &= (A \cap B^c) \cap (A \cap B) = \\ &= (A \cap A) \cap (B \cap B^c) = \emptyset.\end{aligned}$$

Ebből, kihasználva, hogy $B \cap A \subseteq A$ és $B \cap A \subseteq B$

$$\begin{aligned}\mathbf{P}(A \cup B) &= \mathbf{P}(A \backslash B) + \mathbf{P}(A \cap B) + \mathbf{P}(B \backslash A) = \\ &= \mathbf{P}(A) - \mathbf{P}(A \cap B) + \mathbf{P}(A \cap B) + \mathbf{P}(B) - \mathbf{P}(A \cap B),\end{aligned}$$

amiből az egyenlőség már ismét evidens.

4. Ha $A_n \nearrow A$, vagyis ha $A_n \subseteq A_{n+1}$ és $A = \cup_{n=1}^{\infty} A_n$, akkor $\mathbf{P}(A_n) \nearrow \mathbf{P}(A)$. A bizonyításhoz vezessük be az $A_0 \stackrel{\circ}{=} \emptyset$ és $B_k \stackrel{\circ}{=} A_k \backslash A_{k-1}$ halmazokat. Az $A_n \subseteq A_{n+1}$ tartalmazás miatt $A_n = \cup_{k=1}^{n} B_k$. Nyilván a B_k halmazok páronként diszjunktak, következésképpen

$$\begin{aligned}\mathbf{P}(A) &= \mathbf{P}(\cup_n A_n) = \mathbf{P}(\cup_k B_k) = \sum_{k=1}^{\infty} \mathbf{P}(B_k) = \lim_{n \to \infty} \sum_{k=1}^{n} \mathbf{P}(B_k) = \\ &= \lim_{n \to \infty} \mathbf{P}(A_n)\end{aligned}$$

és a $(\mathbf{P}(A_n))$ sorozat nyilván monoton nő.

5. Ha $A_n \searrow A$, vagyis ha $A_{n+1} \subseteq A_n$ és $A = \cap_{n=1}^{\infty} A_n$, akkor $\mathbf{P}(A_n) \searrow \mathbf{P}(A)$. Az összefüggés igazolásához elég a komplementer halmazokra áttérni. Ha $B_n \stackrel{\circ}{=} A_n^c$ és $B \stackrel{\circ}{=} A^c$, akkor $B_n \nearrow B$. Ebből

$$1 - \mathbf{P}(A_n) = \mathbf{P}(B_n) \nearrow \mathbf{P}(B) = 1 - \mathbf{P}(A),$$

amiből a $\mathbf{P}(A_n) \searrow \mathbf{P}(A)$ konvergencia már evidens. A két utóbbi tulajdonságot szokás a valószínűség folytonossági tulajdonságainak is mondani.

6. A valószínűség megszámlálhatóan szubadditív, vagyis tetszőleges (A_n) legfeljebb megszámlálható halmazból álló eseményrendszer esetén $\mathbf{P}(\cup_n A_n) \leq \sum_n \mathbf{P}(A_n)$. Speciálisan ha az (N_n) eseményekre $\mathbf{P}(N_n) = 0$, akkor $\mathbf{P}(\cup_n N_n) = 0$. Valóban: véges eseményrendszer esetén az egyenlőtlenség a korábbi tulajdonságokból indukcióval kapható:

$$\begin{aligned} \mathbf{P}\left(\cup_{k=1}^{n+1} A_k\right) &= \mathbf{P}(\cup_{k=1}^{n} A_k) + \mathbf{P}(A_{n+1}) - \mathbf{P}((\cup_{k=1}^{n} A_k) \cap A_{n+1}) \leq \\ &\leq \sum_{k=1}^{n} \mathbf{P}(A_k) + \mathbf{P}(A_{n+1}) = \sum_{k=1}^{n+1} \mathbf{P}(A_k). \end{aligned}$$

A végtelen eset igazolásához vegyük észre, hogy $B_n \stackrel{\circ}{=} \cup_{k=1}^{n} A_k \nearrow B = \cup_{k=1}^{\infty} A_k$, így az egyenlőtlenség mind a két oldalán határértéket véve és felhasználva a valószínűség már belátott folytonossági tulajdonságát

$$\mathbf{P}(\cup_{k=1}^{\infty} A_k) \stackrel{\circ}{=} \mathbf{P}(B) = \lim_{n\to\infty} \mathbf{P}(B_n) \leq \lim_{n\to\infty} \sum_{k=1}^{n} \mathbf{P}(A_k) = \sum_{k=1}^{\infty} \mathbf{P}(A_k).$$

□

1.1.5. Példa. Mi a valószínűsége, hogy három kockával történő dobás esetén a legnagyobb dobott érték 5-ös?

A_n jelentse azt, hogy a legnagyobb dobott érték nem haladja meg n-et, vagyis, hogy egyik kockával sem dobunk n-nél nagyobbat. Nyilván $\mathbf{P}(A_5) = (5/6)^3$ és $\mathbf{P}(A_4) = (4/6)^3$. Annak a valószínűsége, hogy a legnagyobb dobott érték pontosan 5 egyenlő $\mathbf{P}(A_5 \backslash A_4) = \mathbf{P}(A_5) - \mathbf{P}(A_4) = (5/6)^3 - (4/6)^3$ ugyanis $A_4 \subset A_5$. Speciálisan $(5/6)^3 - (4/6)^3 = 61/216$.

□

1.1.6. Példa. A társasági bridzs játéknál szokás szerint húzás dönti el, hogy ki kivel játszik: a két legnagyobb, illetve a két legkisebb lapot húzó játszik együtt. Péter a 12. legnagyobbat, Kati a 28. legnagyobbat húzza az 52 lapból. Mi a valószínűsége, hogy együtt fognak játszani?

A jelentse azt az eseményt, hogy ők húzták a két legnagyobbat, B pedig azt, hogy a két legkisebbet húzták. A és B diszjunkt események, tehát $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$. Az

A esemény akkor következik be, ha a további két húzás a Katiénál kisebb 24 lapból történik, ennek valószínűsége:

$$\mathbf{P}(A) = \frac{\binom{24}{2}}{\binom{50}{2}} = \frac{24 \cdot 23}{50 \cdot 49} = \frac{276}{1225}.$$

Hasonlóan

$$\mathbf{P}(B) = \frac{\binom{11}{2}}{\binom{50}{2}} = \frac{11 \cdot 10}{50 \cdot 49} = \frac{11}{245}.$$

Így a keresett valószínűség a két szám összege, vagyis $0,2702$.

□

1.2. Feltételes valószínűség

A valószínűségszámítás konkrét alkalmazásaiban a $\mathbf{P}$ valószínűséget általában közvetett módon, a feltételes valószínűségek megadásával határozzuk meg.

1.2.1. Definíció. Ha A és B lehetséges események és $\mathbf{P}(B) > 0$, akkor definíció szerint

$$\mathbf{P}(A \mid B) \mathrel{\hat{=}} \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

A feltételes valószínűség segítségével megfogalmazhatjuk a teljes valószínűség tételét. Ehhez szükségünk van egy további definícióra.

1.2.2. Definíció. Egy legfeljebb megszámlálható darab eseményből álló (B_k) eseményrendszert teljes eseményrendszernek mondunk, ha

1. $\mathbf{P}(\cup_k B_k) = 1$,
2. $\mathbf{P}\left(B_k \cap B_j\right) = 0$ valahányszor $j \neq k$.

1.2.3. Tétel (Teljes valószínűség tétele)**.** *Ha (B_k) egy teljes eseményrendszer és minden k-ra $\mathbf{P}(B_k) > 0$, akkor minden $A \in \mathscr{A}$ halmazra*

$$\mathbf{P}(A) = \sum_k \mathbf{P}(A \mid B_k) \cdot \mathbf{P}(B_k).$$

Bizonyítás: A tétel bizonyítása majdnem triviális. Az egyetlen „nehézség” abból ered, hogy a (B_k) teljes eseményrendszer definíciójában nem tettük fel, hogy $B_k \cap B_j = \emptyset$, csak az ennél gyengébb $\mathbf{P}\left(B_k \cap B_j\right) = 0$ összefüggést követeltük meg. Vezessük be az $N_{kj} \mathrel{\hat{=}} B_k \cap B_j$ halmazokat. Ha $k \neq j$, akkor a feltételek szerint $\mathbf{P}\left(N_{kj}\right) = 0$. Legyen továbbá $N_0 \mathrel{\hat{=}} \Omega \setminus (\cup_k B_k)$. A teljes eseményrendszer definíciójából evidens, hogy $\mathbf{P}(N_0) = 0$. Jelölje N az imént bevezetett nulla valószínűségű halmazok egyesítését.

Mivel az N legfeljebb megszámlálható nulla valószínűségű halmaz egyesítése, ezért a valószínűség már belátott szubadditivitása miatt az N valószínűsége szintén nulla. Nyilvánvaló módon tetszőleges C eseményre

$$\begin{aligned}\mathbf{P}(C\cap N^c) &= \mathbf{P}(C\cap N^c)+0=\mathbf{P}(C\cap N^c)+\mathbf{P}(C\cap N)= \\ &= \mathbf{P}((C\cap N^c)\cup(C\cap N))=\mathbf{P}(C\cap(N\cup N^c))=\mathbf{P}(C).\end{aligned}$$

Következésképpen a $\mathbf{P}$ valószínűséget leszűkítve az N^c halmazra újra valószínűségi mezőt kapunk, így feltehetjük, hogy a teljes eseményrendszer definíciójában a különböző B_k halmazok diszjunktak és az egyesítésük az Ω biztos esemény. Felhasználva, hogy a B_k halmazok valószínűsége pozitív,

$$\begin{aligned}\mathbf{P}(A) &= \mathbf{P}(A\cap\Omega)=\mathbf{P}(A\cap(\cup_k B_k))=\sum_k \mathbf{P}(A\cap B_k)= \\ &= \sum_k \frac{\mathbf{P}(A\cap B_k)}{\mathbf{P}(B_k)}\cdot\mathbf{P}(B_k)\stackrel{\circ}{=}\sum_k \mathbf{P}(A\mid B_k)\cdot\mathbf{P}(B_k).\end{aligned}$$

A bizonyítás elemzéséből világos, hogy az egyetlen lépés, ahol a $\mathbf{P}(B_k)>0$ feltételt felhasználtuk az a

$$\mathbf{P}(A\cap B_k)=\frac{\mathbf{P}(A\cap B_k)}{\mathbf{P}(B_k)}\cdot\mathbf{P}(B_k)$$

egyenlőség, amely a $\mathbf{P}(B_k)=0$ esetben értelmetlen. A valószínűségszámításban gyakran szokás élni avval a konvencióval, hogy ha egy szorzat valamelyik tényezője nulla, akkor az egész szorzat értéke is nulla, függetlenül attól, hogy a másik tényező véges, értelmes, vagy esetleg végtelen. Ha a teljes valószínűség tételében szereplő szorzatokra alkalmazzuk ezt a megállapodást, akkor a $\mathbf{P}(B_k)>0$ feltételek elhagyhatóak. Ilyenkor célszerű a teljes eseményrendszer definícióját is módosítani. Gyakran teljes eseményrendszeren az Ω tér eseményekből álló, legfeljebb megszámlálható számosságú partícióját szokás érteni, megengedve a nulla valószínűséggel rendelkező halmazokat is. Vagyis mivel legfeljebb megszámlálható darab nulla valószínűségű halmazt szabadon hozzácsatolhatunk az eseményrendszerhez, ezért ilyenkor egy $B_k\in\mathscr{A}$ halmazokból álló legfeljebb megszámlálható halmazrendszert akkor mondunk teljes eseményrendszernek, ha a (B_k) család egy partíció, vagyis ha $\cup_k B_k=\Omega$ és $B_k\cap B_j=\emptyset$ valahányszor $k\neq j$.

□

Miként a bizonyításból látszik, a teljes valószínűség tétele a feltételes valószínűség definíciójából következő triviális azonosság. Ugyanakkor konkrét feladatokban általában a $\mathbf{P}(A)$ valószínűség nem ismert, csak a feltételes valószínűségekre, illetve a teljes eseményrendszer valószínűségeire tudunk következtetni. Ilyenkor az azonosságot a $\mathbf{P}$ meghatározására használjuk.

1.2.4. Példa. Másodikra kihúzott golyó színének valószínűsége az urnamodellben.

Az elemi valószínűségszámítás klasszikus modellje az urnamodell. Az urnamodellben megfogalmazható legegyszerűbb példában egy urnában két különböző színű golyó van, mondjuk öt darab piros és három darab fekete. A legegyszerűbb kérdés a következő: Ha egymás után kétszer húzunk, akkor mi lesz annak a valószínűsége, hogy a másodszorra kihúzott golyó piros lesz. A szokásos megoldás a következő: Az elsőre kihúzott golyó színe egy teljes eseményrendszert alkot. Ha B_1 a piros és B_2 a fekete golyóhoz tartozó esemény[2], akkor $\mathbf{P}(B_1) = 5/8$ és $\mathbf{P}(B_2) = 3/8$. Ha A jelöli azt az eseményt, hogy a második golyó piros lesz, akkor $\mathbf{P}(A \mid B_1) = 4/7$ és $\mathbf{P}(A \mid B_2) = 5/7$, ugyanis a második húzás előtt az urnában már csak hét golyó maradt, és attól függően, hogy pirosat vagy feketét húztunk az első húzásra, a hét golyóból négy vagy öt lesz a piros. Így a teljes valószínűség tétele miatt

$$\mathbf{P}(A) = \frac{4}{7}\frac{5}{8} + \frac{5}{7}\frac{3}{8}.$$

Ugyanakkor vegyük észre, hogy hallgatólagosan több feltételezéssel is éltünk. Egyrészt nyilván feltettük, hogy egy adott szín kihúzásának valószínűsége az urnában levő golyók arányával azonos, de azt is feltettük, hogy az első húzás után valamilyen módon nem változik meg a golyók színe. Természetesen a feladat megfogalmazásából ez kézenfekvő, de ugyanakkor ez a hallgatólagos feltétel tipikus példáját adja annak, hogy konkrét példákban a feltételes valószínűség definiálja a valószínűséget és nem fordítva, miként azt az alfejezet elején tettük, vagy miként azt a feltételes valószínűség definíciója alapján gondolnánk.

□

1.2.5. Példa. Átmenetvalószínűségek a Markov-láncokban.

Az egyik legegyszerűbb sztochasztikus folyamatok a Markov-láncok. Tegyük fel, hogy egy rendszer véges számú állapot valamelyikében lehet. A legegyszerűbb esetben a folyamat a $t = 1, 2, \ldots$ időpontokban az egyik állapotból átugrik egy másik állapotba. Az alapfeltétel, vagyis a Markov-láncot definiáló tulajdonság, hogy annak a valószínűsége, hogy hova ugrik a következő időpontban a rendszer csak az aktuális állapottól függ, és nem függ például attól, hogy miként jutott a rendszer az aktuális állapotba. A modell megadásához meg kell mondani, hogy milyen valószínűséggel ugrik a rendszer az egyik állapotból a másikba. Jelölje p_{ij} annak a valószínűségét, hogy a rendszer az i állapotból a j állapotba ugrik. Vegyük észre, hogy hallgatólagosan azt is feltettük, hogy a p_{ij} átmenetvalószínűségek nem függnek az időtől. A p_{ij} egy feltételes valószínűség: annak a valószínűsége, hogy a rendszer a $t+1$ időpontban a j állapotban lesz, feltéve, hogy a t időpontban az i állapotban volt. Ha a rendszer lehetséges állapotainak száma N, akkor a p_{ij} átmenetvalószínűségekből alkothatunk egy $N \times N$-es P mátrixot. A rendszer $t = 0$ időpontban felvett állapotai egy teljes eseményrendszert alkotnak. Ha a $t = 0$ időpontban a rendszer r_i valószínűséggel van az i állapotban, akkor a teljes valószínűség tétele alapján annak a valószínűsége, hogy a rendszer a $t = 1$ időpontban a j

[2] Természetesen az első húzáshoz tartozó színekről van szó.

állapotban lesz $\sum_{i=1}^{N} r_i p_{ij}$. Ha az r_i feltétel nélküli valószínűségeket egy q_0 sorvektorba rendezzük, akkor a $\sum_{i=1}^{N} r_i p_{ij}$ éppen a $q_0 P$ szorzat j-edik eleme. Ha $q_1 \stackrel{\circ}{=} q_0 P$, akkor a gondolatmenetet a $t = 1$ és a $t = 2$ időpontok között alkalmazva a $q_2 \stackrel{\circ}{=} q_1 P = q_0 P^2$ sorvektor a rendszer lehetséges állapotainak valószínűségét adja meg a $t = 2$ időpontban. Az eljárást nyilván tetszőleges t időpontra folytathatjuk. A példában ismét azt látjuk, hogy közvetlenül a valószínűségek nem adottak és a modell alapadatai (részben) feltételes valószínűségek.

$\square$

1.2.6. Példa. Péter és Pál pingpongoznak. Mindkettő $1/2$ valószínűséggel nyer minden játszmát. A játék tétje egy tábla csokoládé, és ezt az nyeri el, aki három játszmát tud egymás után megnyerni. Az első játszmát Péter nyeri, mi a valószínűsége, hogy övé lesz a csokoládé?

Jelöljük egyessel, ha Péter nyer egy játszmát, és nullával, ha veszít. A keresett valószínűséget pedig jelöljük $\mathbf{P}(A) = p$-vel. Ha a következő két játszma közül Péter bármelyiket elveszti, vagyis alább a feltétel valamelyik indexe nulla, onnantól a nyerésének valószínűsége $1 - p$ lesz, hiszen Pál kerül ugyanolyan helyzetbe, mint Péter volt korábban. A következő két játszma alakulása szerint bontsuk fel az eseményteret, és alkalmazzuk a teljes valószínűség tételét a B_{11}, B_{10} és B_0 teljes eseményrendszerre. Eszerint

$$\begin{aligned}\mathbf{P}(A) &= \mathbf{P}(A \mid B_{11})\mathbf{P}(B_{11}) + \mathbf{P}(A \mid B_{10})\mathbf{P}(B_{10}) + \mathbf{P}(A \mid B_0)\mathbf{P}(B_0) = \\ &= 1\frac{1}{4} + (1-p)\frac{1}{4} + (1-p)\frac{1}{2} = 1 - \frac{3}{4}p.\end{aligned}$$

Ebből $p = 4/7$. A figyelmes olvasó felvetheti, hogy előfordulhat-e, hogy pozitív valószínűséggel nem lesz a játéknak nyertese. A játéknak akkor van vége, ha valamelyik játékos háromszor egymás után nyer. Az Ω tekinthető a nulla és egy számokból álló végtelen sorozatok halmazának. Jelölje A azokat a kimeneteleket, azokat a sorozatokat, amely során valaki nyer. Az A esemény olyan sorozatokból áll, amelyben van legalább három egymás után következő nulla vagy egyes. Ha B jelöli azokat a sorozatokat, amelyekben legalább három darab egymás utáni egyes van, mégpedig úgy, hogy az egyesek az $1, 2, 3$ vagy az $4, 5, 6$ stb. helyeken vannak, akkor $B \subset A$. A B halmaz által leírt kísérlet tekinthető egy geometriai eloszlású változónak, ahol a siker valószínűsége $p = 1/8$. A geometriai eloszlással később részletesen foglalkozni fogunk. Most csak azt jegyezzük meg, hogy egy független kísérletsorozatban annak a valószínűsége, hogy az n-edik lépésben következik be valamilyen p valószínűségű kívánt esemény először, $p(1-p)^{n-1} \stackrel{\circ}{=} pq^{n-1}$. Az, hogy az esemény valamikor bekövetkezik

$$\sum_{k=1}^{\infty} p(1-p)^{k-1} \stackrel{\circ}{=} \sum_{k=1}^{\infty} pq^{k-1} = \frac{p}{1-q} = \frac{p}{1-(1-p)} = 1.$$

$\square$

1.2.7. Példa. 52 lapos kártyából 3 piros lap elveszett. Mi a valószínűsége, hogy a csomagból ászt húzunk?

Jelölje A azt az eseményt, hogy ászt húzunk, és B_0, B_1, illetve B_2 azt, hogy nulla, egy vagy két piros ász veszett el. Vegyük észre, hogy $B_3 = \emptyset$, ugyanis mivel piros lapok vesztek el, ezért maximum két ász veszhetett el. Mivel 26 darab piros lap van, amelyek közül 24 nem ász és kettő pedig ász, ezért

$$\mathbf{P}(B_k) = \frac{\binom{24}{3-k}\binom{2}{k}}{\binom{26}{3}}, \quad k = 0, 1, 2.$$

Ugyanakkor mivel összesen négy ász van, és a megmaradt lapok száma 49, ezért

$$\mathbf{P}(A \mid B_k) = \frac{4-k}{49}.$$

Ebből a teljes valószínűség tétele alapján

$$\mathbf{P}(A) = \frac{\binom{24}{3}\binom{2}{0}}{\binom{26}{3}} \frac{4}{49} + \frac{\binom{24}{2}\binom{2}{1}}{\binom{26}{3}} \frac{3}{49} + \frac{\binom{24}{1}\binom{2}{2}}{\binom{26}{3}} \frac{2}{49} = \frac{1}{13}.$$

□

A teljes valószínűség tételének egy gyakran használt következménye a következő:

1.2.8. Következmény (Bayes-tétel)**.** *Legyen* (B_k) *egy teljes eseményrendszer és tegyük fel, hogy az A esemény valószínűsége pozitív. Ekkor*

$$\mathbf{P}(B_k \mid A) = \frac{\mathbf{P}(A \mid B_k) \cdot \mathbf{P}(B_k)}{\sum_n \mathbf{P}(A \mid B_n) \cdot \mathbf{P}(B_n)}.$$

Bizonyítás: Bayes-tétele elemi következménye a feltételes valószínűség definíciójának és a teljes valószínűség tételének. Mivel a feltétel szerint $\mathbf{P}(A) > 0$, ezért a $\mathbf{P}(B_k \mid A)$ feltételes valószínűség értelmezhető. A feltételes valószínűség definíciója szerint

$$\mathbf{P}(B_k \mid A) \circeq \frac{\mathbf{P}(A \cap B_k)}{\mathbf{P}(A)} \circeq \frac{\mathbf{P}(A \mid B_k) \cdot \mathbf{P}(B_k)}{\mathbf{P}(A)}.$$

A teljes valószínűség tétele alapján

$$\mathbf{P}(A) = \sum_n \mathbf{P}(A \mid B_n) \cdot \mathbf{P}(B_n),$$

amit a nevezőbe beírva éppen a bizonyítandó összefüggést kapjuk. Érdemes megjegyezni, hogy a $\mathbf{P}(A) > 0$ feltételre valójában nincs szükség, ugyanis ha $\mathbf{P}(A) = 0$, akkor mind a két oldal értelmetlen, így valójában az egyenlőség akkor is azonos értelmű kifejezést eredményez, vagyis a két oldal egyszerre értelmes vagy értelmetlen.

□

1.2.9. Példa. Tesztvizsgán minden kérdésre a megadott három válaszból egy a helyes. A vizsgázó p valószínűséggel tudja a helyes választ, továbbá ha nem tudja, akkor tippel, és $1/3$ valószínűséggel találja el a helyes választ.

1. Milyen valószínűséggel ad a vizsgázó helyes választ?

2. Ha helyes választ adott a vizsgázó egy kérdésre, akkor mi a valószínűsége, hogy tudta a választ?

A jelentse azt, hogy tudja a választ, B azt, hogy helyes választ ad. A teljes valószínűség illetve Bayes tétele alapján

$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(B \mid A)\,\mathbf{P}(A) + \mathbf{P}(B \mid A^c)\,\mathbf{P}(A^c) = \\ &= 1 \cdot p + \frac{1}{3} \cdot (1-p) = \frac{1}{3} + \frac{2}{3}p\end{aligned}$$

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(B \mid A)\,\mathbf{P}(A)}{\mathbf{P}(B \mid A)\,\mathbf{P}(A) + \mathbf{P}(B \mid A^c)\,\mathbf{P}(A^c)} = \frac{p}{\frac{1}{3} + \frac{2}{3}p} = \frac{3p}{2p+1}.$$

□

1.3. Események függetlensége

A feltételes valószínűség fogalma szoros kapcsolatban van a függetlenség fogalmával.

1.3.1. Definíció. Az A és B eseményeket függetlennek mondjuk, ha $\mathbf{P}(A \cap B) = \mathbf{P}(A)\,\mathbf{P}(B)$, illetve általában az $A_1, A_2, \ldots, A_n$ eseményeket függetlenek mondjuk, ha tetszőleges

$$1 \le i_1 < i_2 < \ldots < i_k \le n$$

esetén

$$\mathbf{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = \mathbf{P}(A_{i_1})\,\mathbf{P}(A_{i_2}) \cdots \mathbf{P}(A_{i_k}).$$

Speciálisan

$$\mathbf{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbf{P}(A_1)\,\mathbf{P}(A_2) \cdots \mathbf{P}(A_n).$$

Az elnevezés logikája alapján az A és a B eseményeket függetlennek mondjuk, ha $\mathbf{P}(A \mid B) = \mathbf{P}(A)$. De ilyenkor természetesen fel kell tenni, hogy $\mathbf{P}(B) > 0$. A feltételes valószínűség definíciója szerint

$$\frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \mathbf{P}(A),$$

amiből a $\mathbf{P}(A \cap B) = \mathbf{P}(A)\,\mathbf{P}(B)$ egyenlőség már evidens. Ugyanakkor ha $\mathbf{P}(B) = 0$, akkor a feltételes valószínűségről nem beszélhetünk, de a függetlenség fogalma akkor is értelmes marad. A függetlenség a valószínűségszámítás leggyakrabban használt fogalma, amelyet már a korábbi példákban is implicite többször használtuk.

1.3.2. Példa. K városból L városba és L városból M városba két-két út vezet. Hófúvás alkalmával minden út egymástól függetlenül p valószínűséggel járhatatlan. Az Útinform szerint K-ból M-be nem lehet eljutni. Mi a valószínűsége, hogy K-ból L-be azért el lehet jutni?

A jelölje azt az eseményt, hogy K-ból L-be el lehet jutni, B azt, hogy L-ből M-be át lehet menni. Nyilván $\mathbf{P}(A) = \mathbf{P}(B) = 1 - p^2 \triangleq P$. Világos, hogy az $1 - p^2$ felírásakor kihasználtuk, hogy az egyes utakon a hófúvás egymástól független, és akkor lehet az egyik városból a vele szomszédos másikba átmenni ha nem járhatatlan mind a két út. A feladat szerint a $\mathbf{P}(A \mid (A \cap B)^c)$ feltételes valószínűséget kell kiszámítani, ugyanis az Útinform jelentése szerint az $A \cap B$ esemény nem teljesül.

$$\begin{aligned}
\mathbf{P}(A \mid (A \cap B)^c) &= \frac{\mathbf{P}(A \cap (A \cap B)^c)}{\mathbf{P}((A \cap B)^c)} = \frac{\mathbf{P}(A \cap (A^c \cup B^c))}{\mathbf{P}((A \cap B)^c)} = \\
&= \frac{\mathbf{P}(A \cap B^c)}{1 - \mathbf{P}(A \cap B)} = \frac{\mathbf{P}(A)\,\mathbf{P}(B^c)}{1 - \mathbf{P}(A \cap B)} = \frac{P(1-P)}{1-P^2} \\
&= \frac{P}{1+P}.
\end{aligned}$$

A számolás során felhasználtuk, hogy ha A és B függetlenek, akkor az A és a B^c is függetlenek. Valóban,

$$\begin{aligned}
\mathbf{P}(A) &= \mathbf{P}(A \cap (B \cup B^c)) = \mathbf{P}((A \cap B^c) \cup (A \cap B)) = \\
&= \mathbf{P}(A \cap B^c) + \mathbf{P}(A \cap B) = \\
&= \mathbf{P}(A \cap B^c) + \mathbf{P}(A)\,\mathbf{P}(B).
\end{aligned}$$

Ezt rendezve

$$\begin{aligned}
\mathbf{P}(A \cap B^c) &= \mathbf{P}(A) - \mathbf{P}(A)\,\mathbf{P}(B) = \mathbf{P}(A)(1 - \mathbf{P}(B)) = \\
&= \mathbf{P}(A)\,\mathbf{P}(B^c),
\end{aligned}$$

amely éppen a függetlenséget definiáló egyenlőség.

□

Végezetül érdemes a függetlenség és a feltételes valószínűség kapcsán egy általános megjegyzést tenni. A három halmazelméleti művelet közül a komplementer valószínűségének kiszámolása igen egyszerű. Az egyesítés valószínűségének kiszámolása csak annyiban bonyolult, hogy az általános esetben vissza kell vezetni a metszet valószínűségének meghatározására. Az igazi problémát a metszet valószínűségének kiszámolása jelenti. Ezt lényegében egyetlen módon tehetjük meg, a feltételes valószínűség kiszámolásán keresztül, feltéve, ha azt ismerjük, vagy az alkalmazás szempontjából értelmes módon definiálni tudjuk. A feltételes valószínűség kiszámolásának legegyszerűbb módja a függetlenség feltételezése. Nem meglepő tehát, hogy a függetlenség feltétele szinte minden valószínűségszámítási feladatban megjelenik. Más kérdés az, hogy az alkalmazás szempontjából mikor jogos ez a feltételezés. A valószínűségszámítás gyakorlati alkalmazása során a függetlenség indokolatlan használata okozza a legnagyobb problémát.

II.

ELEMI VALÓSZÍNŰSÉGSZÁMÍTÁS

Ebben a fejezetben röviden felidézzük az elemi valószínűségszámítás néhány feladatát. A valószínűségszámítás népszerűségének egyik oka, hogy szinte minden matematikai probléma átfogalmazható valószínűségszámítási feladattá. Az elemi valószínűségszámítás lényegében a kombinatorika és a geometria eszköztárára támaszkodik. A kombinatorika feladait az úgynevezett klasszikus valószínűségi mező segítségével alakíthatjuk valószínűségszámítási problémákká, a geometria feladatait pedig a geometriai valószínűség fogalmának bevezetésével transzformálhatjuk valószínűségszámítási feladattá. Annak ellenére, hogy az ilyen típusú példák az elemi valószínűségszámítás témakörébe tartoznak, távolról sem egyszerű feladatok és igen gyakran komoly matematikai ismereteket igényelnek. Az elemi jelző nagyrészt arra utal, hogy a feladatok megoldásához nincs szükség a matematikai analízis ismeretére.

2.1. A klasszikus valószínűségi mező

A klasszikus valószínűségi mező esetén az Ω biztos esemény n darab kimenetelből áll. Az egyes kimenetelek „esélye” azonos, így egy $A \in \mathscr{A}$ esemény valószínűsége k/n, ahol k az A elemeinek száma. A megfogalmazás során óvatosan kell eljárni, ugyanis nem mondhatjuk azt, hogy az Ω minden kimenetele azonos valószínűségű, ugyanis nem tudjuk azt, hogy az egyes kimenetelekből álló egyelemű halmazok elemei az $\mathscr{A}$ halmazrendszernek. Vegyük észre, hogy az előző fejezetben szereplő példák legtöbbje a klasszikus valószínűségi mező használatára épült. Egy klasszikus valószínűségszámítási feladat lényegében két kombinatorikai feladatból áll: Ki kell számolni a k és az n értékét. Ennek megfelelően röviden áttekintjük a kombinatorika néhány elemi fogalmát.

2.1.1. Példa. Mintavételezési eljárások.

A legtöbb klasszikus feladat központi elve a szorzatelv: Ha két „kritérium” mentén kell egy halmaz elemeit megszámolni, és az első kritérium szerint k_1 fajta elem létezik, a másik kritérium szerint k_2 fajta elem létezik, akkor az összes lehetséges elemek száma $k_1 \cdot k_2$. Vagyis, ha a lehetséges párokat egy táblázatba rendezzük, akkor a táblázat elemeinek száma megegyezik a sorok és az oszlopok darabszámának szorzatával. Nézzünk néhány példát:

1. *Visszatevéses mintavétel a sorrend figyelembevételével:* A legtöbb kombinatorikai feladat megfogalmazható az urnamodell segítségével. Tegyük fel, hogy egy urnában m darab golyó van, és mindegyik golyóra egy-egy különböző szám van írva. Ezek után kiveszünk n darab golyót. Minden egyes kivétel után felírjuk a számot, majd visszatesszük a golyót. Hány darab számsor lehetséges? A szorzatelv alapján a válasz nyilván m^n, ugyanis minden egyes kivételnél a lehetséges számok száma m és ezt kell n-szer megismételni.

2. *Visszatevés nélküli mintavétel a sorrend figyelembevételével:* A feladat azonos az előzővel, de a golyókat nem tesszük vissza. Ilyenkor nyilván $m \geq n$ és a válasz ismételten a szorzatelv triviális alkalmazásával

$$(m)_n \mathrel{\hat{=}} m\,(m-1)\,(m-2)\cdots(m-n+1)\,.$$

Az $(m)_n$ jelölés bevezetését az indokolja, hogy az ilyen típusú sorozatok igen sok feladatban előfordulnak. Ha $m = n$, akkor $(n)_n$ helyet szokás az $n!$ jelölést használni. Ilyenkor szokás permutációról beszélni. Vagyis $n!$ a lehetséges permutációk számát jelöli. Kényelmi okokból érdemes bevezetni a $0! \mathrel{\hat{=}} 1$ jelölést.

3. *Visszatevés nélküli mintavétel a sorrend figyelembevétele nélkül:* A feladat ismételten azonos az előzővel, de most nem vesszük figyelembe a sorrendet. Világos, hogy ilyenkor azokat a számsorozatokat, amelyek azonos számokból állnak ekvivalensnek tekintjük. Egy ekvivalenciaosztályban $n!$ elem van, így a lehetséges elvivalenciaosztályok száma $(m)_n/n!$. Mivel ez a kifejezés is gyakran előfordul érdemes egy külön jelölés bevezetni rá:

$$\binom{m}{n} \mathrel{\hat{=}} \frac{(m)_m}{n!} = \frac{m!}{n!\,(m-n)!}$$

Az $\binom{m}{n}$ kifejezést szokás binomiális együtthatónak mondani, illetve kombinációkról beszélni. A $0! = 1$ jelölés értelemszerű alkalmazásával $\binom{m}{0} = 1$. Ennek a feladatnak egy kézenfekvő általánosítása a következő: Tegyük fel, hogy az m darab golyó k darab csoportba van beosztva, mondjuk színek szerint. Ugyanakkor az azonos színű golyókat már nem tudjuk egymástól megkülönböztetni. Ilyenkor a lehetséges színsorozatok száma

$$\frac{m!}{n_1!n_2!\cdots n_k!}$$

ahol n_i az i-edik színű golyók száma. Nyilvánvalóan $\Sigma_{i=1}^{k} n_i = m$.

4. *Visszatevéses mintavétel a sorrend figyelembevétele nélkül:* A golyókat visszatesszük, a golyókon levő számok különbözők, de nem vesszük figyelembe, hogy a golyókat mikor húztuk ki, csak azt, hogy hányszor vettük ki az egyes golyókat. A korábbi esetek mindegyike lényegében nyilvánvaló volt. Az egyetlen eset, amikor némiképpen gondolkodni kell az éppen a jelen eset. Ahhoz, hogy a lehetséges konfigurációk számát meg tudjuk adni meg kell adni, hogy hogyan kell rögzíteni az egyes kísérleteket. Összesen m fajta golyónk van, ezért vegyünk m darab rubrikát egy papíron és minden esetben amikor egy golyót kiveszünk, húzzunk egy vonalat a megfelelő m rubrika valamelyikébe. Így n darab golyó kivétele után n darab vonalunk van. Az egyes vonalak között ott vannak a rubrikákat elhatároló jelek, legyenek ezek csillagok. A lényeges észrevétel az, hogy az m darab rubrika felírásához $m-1$ darab csillagra van szükség. Ennek megfelelően a papíron $m+n-1$ jel van, amelyek közül $m-1$ darab csillag és n

darab vonal van. A feladat tehát az, hogy miként lehet az $m+n-1$ jelből vagy az $m-1$ darab csillagot, vagy az n darab vonalat kivenni. Ezek száma az előző eset alapján

$$\binom{m+n-1}{n} = \binom{m+n-1}{m-1}.$$

□

2.1.2. Példa. A lottótalálatok valószínűsége.

A klasszikus valószínűségszámítás legnépszerűbb példája az ötöslottó találatok valószínűségének kiszámolása. Mivel 90 számból kell 5 számot eltalálni és a számok sorrendje nem számít, ezért a lehetséges húzások száma

$$\binom{90}{5} = 43\,949\,268.$$

Mivel csak egy számsor lesz nyerő, ezért az öttalálatos valószínűsége

$$P_5 = \frac{1}{\binom{90}{5}} = 2,2754 \times 10^{-8}.$$

Bárki felvetheti, hogy a lottó kihúzásakor nem az $1/\binom{m}{n}$ szabályt használják, hanem egymás után húzzák ki az öt számot. Ha m elemből n darabot kiveszünk egymás után, úgy hogy minden egyes kivételkor a még az urnában levők azonos eséllyel kerülnek kihúzásra, akkor annak a valószínűsége, hogy a megadott n egyed kerül kihúzásra

$$\frac{n}{m}\frac{n-1}{m-1}\cdots\frac{1}{m-n+1} = 1/\binom{m}{n}.$$

Amire érdemes felfigyelni az a 10^{-8} nagyságrend[1]. A négytalálatos szelvények lehetséges száma a szorzatelv alapján

$$\binom{5}{4}\binom{85}{1} = 5 \cdot 85 = 425,$$

ugyanis az öt nyerő szám közül ki kell venni az eltalált négyet, vagy el kell hagyni egyet, amit nem találunk el, majd ki kell venni a rossz 85 szám közül egyet, és ezt a két számot össze kell szorozni. Ennek megfelelően

$$P_4 = \frac{\binom{5}{4}\binom{85}{1}}{\binom{90}{5}} = 9,6702 \times 10^{-6},$$

[1] Magyarországon körülbelül 10^7 ember él, így ha mindenki kitölt egy szelvényt, akkor $10^7/\binom{90}{5} = 0,22754$ annak a valószínűsége, hogy egy héten lesz öttalálatos. Ugyanakkor ennél valójában sokkal kevesebb szelvényt szoktak vásárolni, ugyanis általában minden család játszik egy szelvénnyel. Általában egy lottóláz csúcsán, amikor a lehetséges nyeremény már 2 milliárd forint felett van körülbelül 6 milló szelvényt szoktak vásárolni. Ilyenkor a nyerési valószínűség $6 \cdot 10^6/\binom{90}{5} = 0,13652$. Egy átlagos héten körülbelül 3 millió szelvény kerül megvásárlásra, így az öttalálat valószínűsége is a fele, durván 7%.

amely még mindig 10^{-5} nagyságrendű. Hasonlóan

$$\begin{aligned} P_3 &= \frac{\binom{5}{3}\binom{85}{2}}{\binom{90}{5}} = 8,123 \times 10^{-4} \\ P_2 &= \frac{\binom{5}{2}\binom{85}{3}}{\binom{90}{5}} = 2,2474 \times 10^{-2} \end{aligned}$$

Annak a valószínűsége, hogy egy adott héten semmit nem nyerünk a lottón,

$$P_1 + P_0 = \frac{\binom{5}{1}\binom{85}{4}}{\binom{90}{5}} + \frac{\binom{5}{0}\binom{85}{5}}{\binom{90}{5}} = 0,9767.$$

Ha valaki hatvan év alatt minden héten vesz egy lottószelvényt, akkor összesen körülbelül $60 \cdot 52 = 3120$ szelvényt vesz. Ebből annak a valószínűsége, hogy élete során soha sem nyer

$$(P_1 + P_0)^{3120} = \left(\frac{\binom{5}{1}\binom{85}{4}}{\binom{90}{5}} + \frac{\binom{5}{0}\binom{85}{5}}{\binom{90}{5}} \right)^{3120} = 1,1507 \times 10^{-32}.$$

Annak a valószínűsége, hogy lesz legalább egy hármasa

$$\begin{aligned} 1 - (P_2 + P_1 + P_0)^{3120} &= 1 - \left(\frac{\binom{5}{1}\binom{85}{4}}{\binom{90}{5}} + \frac{\binom{5}{0}\binom{85}{5}}{\binom{90}{5}} + \frac{\binom{5}{2}\binom{85}{3}}{\binom{90}{5}} \right)^{3120} = \\ &= 0,92313, \end{aligned}$$

ami viszonylag nagy. Ugyanakkor annak a valószínűsége, hogy lesz legalább egy öttalálata

$$1 - (1 - P_5)^{3120} = 1 - \left(1 - \frac{1}{\binom{90}{5}} \right)^{3120} = 7,0988 \times 10^{-5}.$$

Hogy lesz legalább egy négyese

$$1 - (1 - P_5 - P_4)^{3120} = 1 - \left(1 - \frac{1}{\binom{90}{5}} - \frac{\binom{5}{4}\binom{85}{1}}{\binom{90}{5}} \right)^{3120} = 2,9790 \times 10^{-2},$$

ami már nem is annyira reménytelen[2].

□

2.1.3. Példa. Hatoslottó és az ötöslottó összehasonlítása.

[2] A konkrét valószínűségeket a Scientific Workplace segítségével számoltam ki. Mivel nagyon kicsi számokról van szó a kerekítési hibákért nem kezeskedem.

A hatoslottó során 45 számból kell hatot eltalálni. A főnyeremény valószínűsége

$$P_6 = \frac{1}{\binom{45}{6}} = 1,2277 \times 10^{-7}.$$

Az ötöslottó négytalálatát összehasonlítva a hatoslottó hattalálatával

$$\frac{1}{\binom{45}{6}} / \frac{\binom{5}{4}\binom{85}{1}}{\binom{90}{5}} = 1,2696 \times 10^{-2},$$

vagyis körülbelül százszor nagyobb az ötöslottóban a négytalálat valószínűsége mint a hattoslottóban a hattalálat valószínűsége.

□

2.1.4. Példa. Eurojackpot nyerési valószínűsége.

A játék során az A oszlop 50 számából kell ötöt eltalálni, és a B oszlop tíz számjegyéből kell kettőt eltalálni. Az összes lehetséges szelvények száma

$$\binom{50}{5}\binom{10}{2} = 95\,344\,200.$$

Az ötöslottóhoz képest ez

$$\frac{\binom{50}{5}\binom{10}{2}}{\binom{90}{5}} = 2,1694$$

arány, vagyis az ötöslottóban a főnyeremény valószínűsége körülbelül kétszer akkora mint a eurojackpotban. Viszonylag jó nyereménynek számít még az $5+1$ találat is. Ilyenkor

$$\binom{50}{5}\binom{2}{1}\binom{8}{1} = 33\,900\,160$$

darab szelvény közül kell az egyiket eltalálni. Ezt összevetve az ötöslottó összes szelvényével

$$\frac{\binom{50}{5}\binom{2}{1}\binom{8}{1}}{\binom{90}{5}} = 0,77135$$

Az $5+1$ találat valószínűsége

$$\frac{\binom{2}{1}\binom{8}{1}}{\binom{50}{5}\binom{10}{2}} = 1,6781 \times 10^{-7},$$

ami jobb az ötöslottó nyerési valószínűségénél.

□

2.1.5. Példa. Többszörös találat valószínűsége.

Mi annak a valószínűsége, hogy egy öthetes lottószelvénnyel valakinek háromszor lesz két találata? A két találat valószínűsége

$$P_2 = \frac{\binom{5}{2}\binom{85}{3}}{\binom{90}{5}} = 2,2474 \times 10^{-2} \approx \frac{1}{45}.$$

Ahhoz, hogy pontosan három két találatunk legyen ki kell választani az öt hétből azt a hármat, amikor nyerünk, így a keresett valószínűség

$$\binom{5}{3} P_2^3 \left(1 - P_2^2\right) = 1,1345 \times 10^{-4},$$

ami körülbelül hetede egy hármas és tizenegyszerese egy négyes valószínűségének. Annak a valószínűsége, hogy négyszer lesz kettese

$$\binom{5}{4} P_2^4 \left(1 - P_2\right) = 1,2468 \times 10^{-6}.$$

□

2.1.6. Példa. Azonos születésnapok valószínűsége.

Legyen adva m személy, akik a 365 nap valamelyikén született. Mi annak a valószínűsége, hogy legalább két embernek azonos lesz a születésnapja? Az összes esetek száma $n = 365^m$. Azon esetek száma, amikor nincsen két azonos születésnap $k = (365)_m$, Így annak a valószínűsége, hogy lesz legalább két azonos születésnap

$$P_m = 1 - \frac{k}{n} = 1 - \frac{(365)_m}{365^m}.$$

Talán némiképpen meglepő, de ha $m \geq 23$, akkor $P_m \geq 1/2$ és $P_{55} \geq 0,99$.

□

2.1.7. Példa. Egy konkrét szám kihúzásának valószínűsége az ötöslottóban.

Mi annak a valószínűsége, hogy a kihúzott öt szám között egy fix szám, mondjuk a 13, szerepelni fog? Az összes kimenetelek száma $\binom{90}{5}$. Ha mondjuk a 13-as számot már eleve kivettük, akkor a megmaradt 89 számból kell kivenni a maradék négyet, így a keresett valószínűség

$$\frac{\binom{89}{4}}{\binom{90}{5}} = \frac{\frac{89\cdot 88\cdot 87\cdot 86}{1\cdot 2\cdot 3\cdot 4}}{\frac{90\cdot 89\cdot 88\cdot 87\cdot 86}{1\cdot 2\cdot 3\cdot 4\cdot 5}} = \frac{5}{90}.$$

Természetesen egy feladatnak nem csak egy lehetséges megoldási módja van. Érdemes néhány alternatív megközelítést is megmutatni, ugyanis mindegyik rávilágít a valószínűségszámítás szabályaira. A komplementer szabály segítségével a következőképpen számolhatunk: A komplementer esemény, hogy a 13-at egyetlen egyszer sem húzzuk ki. Vagyis az öt szám mindegyike a megmaradt 89 számból kerül ki. Így annak a valószínűsége, hogy a 13-as kihúzásra kerül

$$1-\frac{\binom{89}{5}}{\binom{90}{5}}=\frac{\binom{90}{5}-\binom{89}{5}}{\binom{90}{5}}=$$
$$=\frac{90\cdot 89\cdot 88\cdot 87\cdot 86-89\cdot 89\cdot 88\cdot 87\cdot 86\cdot 85}{90\cdot 89\cdot 88\cdot 87\cdot 86}=$$
$$=\frac{(90-85)\cdot 89\cdot 89\cdot 88\cdot 87\cdot 86}{90\cdot 89\cdot 88\cdot 87\cdot 86}=\frac{5}{90}.$$

Érdemes átgondolni a feladatot a teljes valószínűség tétele szempontjából is. A 13-as kihúzását öt diszjunk részre oszthatjuk aszerint, hogy melyik húzásra jön ki a 13-as szám. Értelemszerű jelöléssel ha $\mathbf{P}(A\cap B_1)$ jelöli annak a valószínűségét, hogy az első húzásra már megkapjuk a 13-ast, akkor nyilván $\mathbf{P}(A\cap B_1)=1/90$. A második húzás esetén már csak 89 számból húzhatunk, így ilyenkor a 13 valószínűsége már csak $1/89$. De a

$$\mathbf{P}(A\cap B_2)=\mathbf{P}(A\mid B_2)\mathbf{P}(B_2)$$

szabály miatt ezt meg kell szorozni annak a valószínűségével, hogy a 13-as még bent van a kihúzandó számok között. Ennek a valószínűsége $89/90$, így $\mathbf{P}(A\cap B_2)=1/90$. Általában, annak a valószínűsége, hogy éppen a k-adik húzásra jön ki a 13-as szorzatént írható fel. Egyrészt a még bent levő $90-k+1$ számból ki kell húzni a 13-ast, aminek a valószínűsége $1/(90-k+1)$. De ezt meg kell szorozni avval, hogy a korábbi húzások során a 13-as nem lett kihúzva. Ennek a valószínűsége

$$\frac{89}{90}\frac{88}{89}\cdots\frac{90-k+1}{90-k+2}=\frac{90-k+1}{90},$$

ugyanis az utolsó előtti lépésben még $90-k+2$ számnak kell lenni, ahhoz, hogy az utolsó lépésben éppen $90-k+1$ szám maradjon. A teljes valószínűség tétele alapján tehát

$$\mathbf{P}(A)=\sum_{k=1}^{5}\mathbf{P}(A\mid B_k)\mathbf{P}(B_k)=\frac{5}{90}.$$

□

2.1.8. Példa. Newton-féle binomiális tétel

$$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}$$

A szorzás szabályai alapján az $(x+y)^n$ elvégzésekor pontosan 2^n darab $x^k y^{n-k}$ alakú tag összege lesz a kifejezés. Mivel az szorzat értékének meghatározásakor a sorrend nem számít, adott k-ra éppen $\binom{n}{k}$ darab $x^k y^{n-k}$ tag lesz a 2^n tagú összegben. Speciálisan $2^n=\Sigma_{k=0}^{n}\binom{n}{k}$. A 2^n egy n elemű halmaz összes részhalmazainak számát adja meg. A jobboldal a részhalmazok elemszám szerinti csoportositásban szereplő halmazok számát adja meg.

□

2.2. Geometriai valószínűség

A geometriai valószínűség kiszámolásakor két terület arányát kell megadni. A területeket az elemi geometria segítségével számoljuk ki.

2.2.1. Példa. Óvatosság városában a párbaj ritkán végződik tragikusan. A helyi szokások szerint ugyanis a párbajozó feleknek reggel 6 és 7 óra között meg kell jelenniük a városszéli tisztáson és 7 percet kell várakozniuk az ellenfélre. Ha találkoznak, akkor a párbaj létrejön, ha nem, akkor az ügy el van intézve. Mi a valószínűsége a párbaj létrejöttének, ha mindkét fél véletlenszerűen választott időpontban érkezik a tisztásra?

Ábrázoljuk koordináta-rendszerben (óra törtrészében) a kiérkezési időpontokat. Az egységnégyzetben rajzoljuk meg a találkozás létrejöttét reprezentáló pontok halmazát. Nyilván azokban a pontokban jöhet létre a párbaj, amelyekre $|y-x| \leq 7/60$. Vagyis $-7/60 \leq y-x \leq 7/60$. Egyszerűbb a komplementer esemény valószínűségével számolni. Az $y = x+7/60$ egyenes feletti háromszög területe $(1-7/60)^2/2$. Mivel ezt kétszer kell venni, a geometriai valószínűséggel számolva:

$$P = 1 - \left(1 - \frac{7}{60}\right)^2 = 0,21972.$$

□

2.2.2. Példa. Egy háromszög α szöge egyenletes eloszlású a $(0, \pi/2)$ intervallumon, a β szöge az α rögzített értéke mellett egyenletes eloszlású a $(0, \pi-\alpha)$ intervallumon. Mi a valószínűsége annak, hogy a háromszög legnagyobb szöge a harmadik, a γ szög lesz?

Koordinátarendszerben ábrázolva az α és a β értékeit egyenletes eloszlású pontot kapunk az $x = 0$, $y = 0$, $x+y = \pi$, és az $x = \pi/2$ egyenesek által határolt trapézon[3]. A trapéz területe

$$T = \frac{\pi+\pi/2}{2}\frac{\pi}{2} = \pi^2\frac{3}{8}.$$

Ha most $\alpha > \beta$, akkor az $\alpha+\beta+\gamma = \pi$ egyenlőségbe $\alpha = \gamma$ esetet téve a $2\alpha+\beta = \pi$ egyenletet kapjuk. Ha $\beta > \alpha$ akkor a $\gamma = \beta$ helyettesítéssel az $\alpha+2\beta = \pi$ egyenest kapjuk. A jó pontok területe a két egyenes alatti terület. Az egyenesek az $(\pi/3, \pi/3)$ pontban metszik egymást, így a keresett terület a $(0,0)$, $(\pi/2,0)$, $(\pi/3, \pi/3)$, $(0, \pi/2)$ négyszög területe. Ebbe egy $\pi/3$ oldalhosszal rendelkező négyzetet beírva további két

[3] Vegyük észre, hogy a feladat megfogalmazása feltételes valószínűségeket tartalmazott, és ebből konstruáltuk meg a valószínűségi mezőt, ahol most Ω a trapéz, $\mathscr{A}$ a trapéz részhalmazai a $\mathbf{P}$ pedig az ezekhez rendelt területek, osztva természetesen a trapéz területével. Az $\mathscr{A}$ nem áll feltétlenül a trapéz összes részhalmazából, csak azokból, amelyekre tudunk területet definiálni.

háromszög területét kell kiszámolni. Ezek együttes területe $\pi/3 \cdot (\pi/2 - \pi/3)$. Ebből a keresett valószínűség

$$P = \frac{\pi^2 (1/9 + 1/18)}{3\pi^2/8} = \frac{\pi^2/6}{3\pi^2/8} = \frac{4}{9}.$$

□

2.2.3. Példa. Két pontot a $(0,1)$ intervallumra dobva mi lesz annak a valószínűsége, hogy a középső szakasz lesz a legnagyobb?

A $(0,1)$ intervallumba eső pontokat egy koordinátarendszerben ábrázolva az egységnégyzet egy pontját kapjuk. Feltehetjük, hogy az x tengelyre mért érték a kisebb, ugyanis a fordított esetben analóg módon kell eljárni. Így a lehetséges események halmaza az egységnégyzetben az $y = x$ egyenes feletti rész, amelynek területe $T = 1/2$. Jelölje $\xi < \eta$ a $(0,1)$ intervallumba eső két pontot. A két szélső szakasz hossza, ξ és $1-\eta$, így a harmadik szakasz hossza $\zeta \stackrel{\circ}{=} 1 - (\xi + 1 - \eta) = \eta - \xi$. A kedvező esetekben $\zeta \geq \xi$ és $\zeta \geq 1 - \eta$, vagyis $\eta - \xi \geq \xi$ és $\eta - \xi \geq 1 - \eta$. Ezeket a pontokat határoló egyeneseket beírva az $y \geq 2x$ és $y \geq x/2 + 1/2$ összefüggéssel felírható tartományhoz jutunk. Az egyenesek metszéspontja az $(1/2, 2/3)$ pont. Az $y = 2x$ egyenes az $y = 1$ egyenest az $x = 1/2$ pontban metszi. Így a kedvező kimenetek halmaza az $(0, 1/2)$, $(1/3, 2/3)$, $(1/2, 1)$ és $(0,1)$ pontok által meghatározott négyszög. Ebbe beírva egy $1/3$ oldalhosszú négyzetet a kimaradó két háromszög együttes területe $1/3 \cdot (1/2 - 1/3) = 1/3 \cdot 1/6$. Így a kedvező kimenetek területe $(1/3)^2 + 1/18$. Így a „jó" pontok területe $1/6$. Ezt osztva a $T = 1/2$ értékkel a nem túl meglepő $P = 1/3$ valószínűséget kapjuk.

□

III.

ELOSZLÁS- ÉS SŰRŰSÉGFÜGGVÉNYEK

A valószínűségszámítás központi fogalmai a valószínűségi változók és a hozzájuk tartozó eloszlások. A valószínűségszámításban előre haladva a valószínűségi változók egyre absztraktabb definíciójával találkozhatunk. Ezekre az igen általános megközelítésekre azonban a továbbiakban nem lesz szükségünk és csak a legegyszerűbb definícióval fogunk élni. Első lépésben az egydimenziós valószínűségi változókat tárgyaljuk, majd röviden foglalkozunk a többdimenziós eloszlásokkal.

3.1. Valószínűségi változók és eloszlásaik

Első lépésként a valós értékű valószínűségi változó definícióját tárgyaljuk.

3.1.1. Definíció. Legyen adva egy $(\Omega, \mathscr{A}, \mathbf{P})$ valószínűségi mező. Az Ω téren értelmezett, a valós számok halmazába képező ξ függvényeket valós értékű valószínűségi változóknak mondjuk. A továbbiakban az egyszerűbb és főleg a rövidebb szóhasználat miatt igen gyakran csak változókról fogunk beszélni.

A valószínűségi változó fenti definíciója nem teljesen pontos, ugyanis valójában nem minden függvény tekinthető valószínűségi változónak. Például, ha $\Omega \stackrel{\circ}{=} [0,1]$ intervallum és $\mathscr{A}$ a triviális eseményrendszer, vagyis az $\mathscr{A}$ csak az Ω és az $\emptyset$ halmazokból áll, akkor az $(\Omega, \mathscr{A}) = ([0,1], \{\emptyset, [0,1]\})$ eseménytéren valójában egyedül a konstans függvények lesznek valószínűségi változók. Például a

$$\xi(\omega) \stackrel{\circ}{=} \begin{cases} 1 & \text{ha} \quad \omega < 1/2 \\ 0 & \text{ha} \quad \omega \geq 1/2 \end{cases}$$

függvényt nem tekintjük valószínűségi változónak. Ennek oka az, hogy például a

$$\{\xi = 0\} \stackrel{\circ}{=} \{\omega \mid \xi(\omega) = 0\} = \{\omega \mid \xi(\omega) < 1\} = [1/2, 1]$$

az Ω egy részhalmaza, de nem megfigyelhető esemény, ugyanis nem eleme az $\mathscr{A}$-nak. További talán szemléletesebb példa, hogy amikor megkülönböztethetetlen két kockával dobunk, akkor az egyik, előre rögzített kockán dobott szám értéke nem lesz valószínűségi változó, ugyanakkor valószínűségi változó lesz, ha a kockák különbözőek. Miként említettük, sztochasztikus folyamatok esetén a megfigyelhető események halmaza időben változhat, így előfordulhat, hogy egy ξ függvény egy t időpontban nem valószínűségi változó, de egy későbbi s időpontban már az. Például az imént említett kockadobás példában egy t időpontban még nem tudjuk a két kockát megkülönböztetni, mert például sötét van, de egy későbbi időpontban, miután felkelt a nap, már meg tudjuk őket különböztetni. Másképpen a valószínűségi változók családja függ a mezőn értelmezett lehetséges, megfigyelhető események körétől. A pontos részletek tisztázása a tárgyalás általunk megcélzott elemi szintjén nagyon messze vezetne. Némi heurisztikával azt mondhatjuk, hogy egy ξ függvényt egy $(\Omega, \mathscr{A})$ eseménytéren csak akkor tekintünk valószínűségi változónak, ha az $\mathscr{A}$ elég gazdag ahhoz, hogy a ξ „leírható” legyen az $\mathscr{A}$-ban szereplő eseményekkel.

Következő fontos fogalmunk az eloszlásfüggvény fogalma.

3.1.2. Definíció. Egy a valós számokon értelmezett $F(x)$ függvényt a ξ változó eloszlásfüggvényének mondjuk, ha minden x valós szám esetén

$$F(x) = \mathbf{P}(\xi < x).$$

A definíció kapcsán érdemes hangsúlyozni, hogy az eloszlásfüggvények minden valós szám esetén értelmezve vannak. A figyelmes olvasó azonnal megkérdezheti, hogy honnan tudható, hogy a

$$\{\xi < x\} \stackrel{\circ}{=} \{\omega \mid \xi(\omega) < x\} \subseteq \Omega$$

alakú halmazoknak van valószínűsége, vagyis hogy ezek a halmazok nem pusztán az Ω részhalmazai, hanem egyúttal elemei a megfigyelhető események $\mathscr{A}$ családjának is. Ezt természetesen általában, tetszőleges függvényre, miként a fenti példában is, nem tudjuk. Ezért is hangsúlyoztuk, hogy valójában nem minden függvény tekinthető valószínűségi változónak. Ismételten csak azt tudjuk mondani, hogy a pontos részletek szükségtelenül és talán egyenlőre zavarólag is túl messze vezetnek. Elégedjünk meg avval, hogy a második, pontosított definíció szerint ha egy $\xi : \Omega \to \mathbb{R}$ függvény valószínűségi változó, akkor a $\{\xi < x\} \subseteq \Omega$ halmaz egy esemény, aminek van valószínűsége, így a ξ-nek értelmezhető az eloszlásfüggvénye.

3.1.3. Definíció. Azt mondjuk, hogy két valószínűségi változónak azonos az eloszlása, ha azonos az eloszlásfüggvénye.

Elvileg egy valószínűségi változó nem keverendő össze az ő eloszlásfüggvényével, ugyanis teljesen más matematikai objektumokról van szó. Ennek ellenére igen gyakran fogunk élni, például a $\xi = N(0,1)$ vagy $\xi = B(n,p)$ típusú, elvileg hibás, jelölésekkel, ahol a ξ valószínűségi változó, a másik oldalon pedig egy eloszlás jele szerepel. Természetesen hasonlóan hibás, de használatos a $\xi \in N(0,1)$ jelölés is, ugyanis továbbra is a ξ egy az Ω halmazon értelmezett függvény az $N(0,1)$ pedig egy eloszlás, amely az $\mathbb{R}$ bizonyos részhalmazaihoz rendel számokat.

3.1.4. Állítás (Az eloszlásfüggvények tulajdonságai)**.** *Ha F egy $\xi : \Omega \to \mathbb{R}$ valószínűségi változó eloszlásfüggvénye, akkor*

1. *az F monoton nő,*
2. *az F balról folytonos, vagyis minden x valós számra*

$$F(x-0) \stackrel{\circ}{=} \lim_{z \nearrow x} F(z) = F(x),$$

3. $\lim_{x\to\infty} F(x) = 1,$
4. $\lim_{x\to-\infty} F(x) = 0.$

Bizonyítás: Az összefüggések a valószínűség elemi tulajdonságai alapján evidensek.

1. Ha $x < y$, akkor $\{\xi < x\} \subseteq \{\xi < y\}$, így $F(x) \circeq \mathbf{P}(\xi < x) \leq \mathbf{P}(\xi < y) \circeq F(y)$. Az F monotonitása miatt az alábbi határértékek létezését elegendő monoton növekedő, illetve csökkenő sorozatok mentén ellenőrizni.

2. Ha $x_n \nearrow x$, akkor $A_n \circeq \{\xi < x_n\} \nearrow A \circeq \{\xi < x\}$, ugyanis ha $x_n \leq x_{n+1}$, akkor $\{\xi < x_n\} \subseteq \{\xi < x_{n+1}\}$ és mivel $x_n \nearrow x$, ezért valamely ω kimenetelre $\xi(\omega) < x$ pontosan akkor, ha van olyan n, hogy $\xi(\omega) < x_n$. Ebből következően a valószínűség folytonossága miatt $F(x_n) = \mathbf{P}(A_n) \nearrow \mathbf{P}(A) = F(x)$.

3. Teljesen analóg módon, ha $x_n \nearrow \infty$, akkor mivel a valószínűségi változók, definíció szerint, nem vehetnek fel végtelen értéket, ezért

$$A_n \circeq \{\xi < x_n\} \nearrow \Omega,$$

így a valószínűség imént már említett elemi tulajdonsága miatt

$$F(x_n) \circeq \mathbf{P}(\xi < x_n) \circeq \mathbf{P}(A_n) \nearrow \mathbf{P}(\Omega) = 1.$$

4. Hasonlóan, ha $x_n \searrow -\infty$, akkor mivel a valószínűségi változók értékei valós számok $A_n \circeq \{\xi < x_n\} \searrow \emptyset$, ugyanis nincs olyan ω, amelyre minden n-re $\xi(\omega) < x_n$ lenne. Ebből a valószínűség folytonossági tulajdonsága miatt

$$F(x_n) \circeq \mathbf{P}(\xi < x_n) = \mathbf{P}(A_n) \searrow \mathbf{P}(\emptyset) = 0.$$

□

3.1.5. Példa. Az egyenletes eloszlás eloszlásfüggvénye.

A legegyszerűbb eloszlás az úgynevezett egyenletes eloszlás. Egy egyenletes eloszlású változó esetén egy $[a,b]$ szakasz pontjai közül választunk ki egyet. Az egyenletes eloszlás definíciója szerint annak a valószínűsége, hogy egy $[c,d] \subseteq [a,b]$ szakaszból választjuk a pontot $(d-c)/(b-a)$. Az egyenletes eloszlás eloszlásfüggvénye

$$F(x) = \mathbf{P}(\xi < x) = \begin{cases} 0 & \text{ha} \quad x \leq a \\ \frac{x-a}{b-a} & \text{ha} \quad a < x \leq b \\ 1 & \text{ha} \quad b < x \end{cases}.$$

Könnyen látható, hogy az egyenletes eloszlás definíciójában az a és b végpontoknak nincs szerepe, ugyanis ezek a pontok felvételének valószínűsége nulla. Így beszélhetünk például az (a,b) nyílt intervallumon értelmezett egyenletes eloszlásról is, amely eloszlásfüggvénye szintén a fenti $F(x)$.

□

3.1.6. Példa. A Cauchy-eloszlás eloszlásfüggvénye.

Definíció szerint a Cauchy-eloszlás, pontosabban a standard Cauchy-eloszlás, eloszlásfüggvénye

$$F(x) = \frac{1}{\pi}\arctan x + \frac{1}{2}.$$

Az $\arctan x$ függvény alakjából azonnal látható, az $F(x)$ kielégíti az eloszlásfüggvénytől megkövetelt tulajdonságokat. A figyelmes olvasó azonnal megkérdezheti, hogy miért létezik olyan ξ valószínűségi változó, amely eloszlásfüggvénye éppen $F(x)$. Legyen η a $(-\pi/2, \pi/2)$ szakaszon egyenletes eloszlású valószínűségi változó[1] és legyen $\xi \overset{\circ}{=} \tan\eta$, vagyis a ξ értéke legyen az η tangense. Ekkor kihasználva, hogy az $\arctan x$ függvény szigorúan monoton nő,

$$\begin{aligned} F(x) &\overset{\circ}{=} \mathbf{P}(\xi < x) \overset{\circ}{=} \mathbf{P}(\tan\eta < x) = \\ &= \mathbf{P}(\eta < \arctan x) = G(\arctan x), \end{aligned}$$

ahol $G(x)$ az η eloszlásfüggvénye. Minden x valós számra $\arctan x \in (-\pi/2, \pi/2)$ és

$$G(x) = \frac{x + \pi/2}{\pi}, \quad -\pi/2 < x < \pi/2,$$

ezért

$$F(x) = G(\arctan x) = \frac{1}{\pi}\left(\arctan x + \frac{\pi}{2}\right),$$

amiből a képlet már evidens. Miként látni fogjuk, a Cauchy-eloszlásnak nincsen várható értéke, ezért a szokásos valószínűségszámítási paraméterekkel nem lehet a Cauchy-eloszlásokat jellemezni. Cauchy-eloszláson időnként nem egy eloszlást, vagyis nem a standard Cauchy-eloszlást, hanem egy eloszláscsaládot szokás érteni. Vagyis ha ξ standard Cauchy-eloszlású, akkor a $\zeta = a\xi + b$ alakú változókat is Cauchy-eloszlású, pontosabban $Cauchy(b, a)$ változóknak mondjuk. A b paramétert lokációs paraméternek szokás mondani és a később tárgyalt várható értékkel analóg fogalom, az a paramétert skálaparaméternek szokás mondani és a szintén később tárgyalt szóráshoz hasonló szerepet játszik. Általában azonban, ha nem hangsúlyozzuk, Cauchy-eloszláson a standard Cauchy-eloszlást értjük.

□

3.1.7. Példa. Exponenciális eloszlás eloszlásfüggvénye.

Legyen $\lambda > 0$. A λ paraméterrel rendelkező exponenciális eloszlás eloszlásfüggvénye

$$F(x) = \begin{cases} 0 & \text{ha} \quad x \le 0 \\ 1 - \exp(-\lambda x) & \text{ha} \quad x > 0 \end{cases}.$$

Ismételten világos, hogy az $F(x)$ eleget tesz az állításban szereplő tulajdonságoknak. Legyen most η egyenletes eloszlású a $(0,1)$ szakaszon és legyen $\xi \overset{\circ}{=} -\lambda^{-1}\ln\eta$. A

[1] A még figyelmesebb olvasóban felmerülhet, hogy akkor miért is létezik olyan változó, amely a $[0,1]$ szakaszon egyenletes eloszlású? Ennek tárgyalása azonban valóban messze vezetne és az elemi valószínűségszámításban mint intuitíve világos tényt elfogadjuk, hogy a $[0,1]$ szakaszon létezik egyenletes eloszlású változó.

logaritmusfüggvény a $(0,1)$ szakaszon negatív, így az η mindig pozitív, következésképpen ha $x \leq 0$, akkor $\mathbf{P}(\xi < x) = 0$. Ha $x > 0$, akkor

$$\begin{aligned} \mathbf{P}(\xi < x) &\stackrel{\circ}{=} \mathbf{P}\left(-\lambda^{-1} \ln \eta < x\right) = \mathbf{P}(\ln \eta > -\lambda x) = \\ &= \mathbf{P}(\eta > \exp(-\lambda x)) = 1 - \exp(-\lambda x). \end{aligned}$$

□

3.1.8. Példa. Diszkrét eloszlások eloszlásfüggvénye.

A példákból evidens, hogy az eloszlásfüggvények képlete még a legegyszerűbb esetekben sem egyetlen „formula”, hanem sokkal inkább egy „kapcsoszárójeles” definícióval adható meg. Ez még inkább így van a diszkrét eloszlások esetén. Példaként tekintsük a kockadobáshoz tartozó eloszlás eloszlásfüggvényét.

$$F(x) = \begin{cases} 0 & \text{ha} & x \leq 1 \\ 1/6 & \text{ha} & 1 < x \leq 2 \\ 2/6 & \text{ha} & 2 < x \leq 3 \\ 3/6 & \text{ha} & 3 < x \leq 4 \\ 4/6 & \text{ha} & 4 < x \leq 5 \\ 5/6 & \text{ha} & 5 < x \leq 6 \\ 1 & \text{ha} & x > 6 \end{cases}.$$

Vegyük észre, hogy az eloszlásfüggvény definíciójának megfelelően az ugrások nagysága éppen a megfelelő $1,2,3,4,5,6$ értékhez tartozó $1/6$ valószínűség, de a függvény értéke, a balról való folytonosságnak megfelelően az ugrás helyén még a korábbi érték.

□

3.1.9. Példa. Intervallumokba esés valószínűségének meghatározása eloszlásfüggvény segítségével.

Az eloszlásfüggvényekhez kapcsolódó legegyszerűbb példa a különböző intervallumokba esés valószínűségének meghatározása. Legyen ξ egy valószínűségi változó és legyen F a ξ eloszlásfüggvénye. Az eloszlásfüggvény közvetlenül az $I \stackrel{\circ}{=} (-\infty, x)$ nyílt félegyenesbe esés valószínűségét adja meg. A $\mathbf{P}(A \setminus B) = \mathbf{P}(A) - \mathbf{P}(B)$ szabály alapján ha $a < b$, akkor $F(b) - F(a) = \mathbf{P}(a \leq \xi < b)$. Tetszőleges x pont esetén $\{x\} = \cap_n [x, x+1/n)$, vagyis $[x, x+1/n) \searrow \{x\}$. Ezért a valószínűség folytonossági tulajdonsága alapján

$$\mathbf{P}(\xi = x) = \lim_{n \nearrow \infty} F\left(x + \frac{1}{n}\right) - F(x) \stackrel{\circ}{=} F(x+0) - F(x),$$

vagyis miként már az előző példában is megjegyeztük, az x pont felvételének valószínűsége éppen az x pontban való szakadás nagysága. Érdemes megjegyezni, hogy mivel az eloszlásfüggvény monoton nő, ezért az $F(x+0)$ határértéket elegendő volt egyetlen sorozat mentén meghatározni. Ebből ha $a<b$, akkor

$$\begin{aligned}\mathbf{P}(\xi\in(a,b)) &= \mathbf{P}((\xi\in[a,b))\setminus\{\xi=a\})= \\ &= (F(b)-F(a))-(F(a+0)-F(a))= \\ &= F(b)-F(a+0).\end{aligned}$$

□

3.1.10. Példa. Számoljuk ki a ξ^+ és a ξ^- eloszlásfüggvényét.

Emlékeztetünk, hogy $x^+ \circeq \max(0,x)$ és $x^- = \max(0,-x) \circeq -\min(0,x)$. Legyen $F(x)$ egy ξ valószínűségi változó eloszlásfüggvénye. Jelölje $G(x)$ a ξ^+ eloszlásfüggvényét. A balról való folytonosság miatt könnyen látható, hogy

$$G(x)\circeq\mathbf{P}\left(\xi^+<x\right)=\begin{cases}0 & \text{ha} \quad x\le 0\\ F(x) & \text{ha} \quad x>0\end{cases}.$$

Ugyanakkor ismételten a balról való folytonosság miatt

$$U(x)\circeq\mathbf{P}\left(-\xi^-<x\right)=\begin{cases}F(x) & \text{ha} \quad x\le 0\\ 1 & \text{ha} \quad x>0\end{cases}.$$

Ha $H(x)$ a ξ^- eloszlásfüggvénye, és $x>0$, akkor

$$\begin{aligned}H(x) &= \mathbf{P}\left(\xi^-<x\right)=\mathbf{P}\left(-\xi^->-x\right)= \\ &= 1-\mathbf{P}\left(-\xi^-\le -x\right)=1-U(-x+0)= \\ &= 1-F(-x+0).\end{aligned}$$

Így

$$H(x)=\mathbf{P}\left(\xi^-<x\right)=\begin{cases}0 & \text{ha} \quad x\le 0\\ 1-F(-x+0) & \text{ha} \quad x>0\end{cases}.$$

□

3.1.11. Példa. Eloszlásfüggvény folytonossági pontjai.

Egy eloszlásfüggvénynek legfeljebb n darab olyan ugrása lehet, amely nagysága nagyobb vagy egyenlő mint $1/n$. Ebből következően egy eloszlásfüggvénynek legfeljebb megszámlálható ugrása lehet. Így az összes szakadási ponthoz található olyan balról konvergens sorozat, amelynek pontjaiban az eloszlásfüggvény folytonos. Így a balról való folytonosság miatt elég egy eloszlásfüggvényt a folytonossági pontokban ismerni. Hasonlóan egy eloszlásfüggvényt elegendő a racionális pontokban ismerni.

□

Az eloszlásfüggvény mellett szükségünk lesz a sűrűségfüggvények fogalmára.

3.1.12. Definíció. Egy f integrálható függvényt az F eloszlásfüggvény sűrűségfüggvényének mondjuk, ha tetszőleges $a < b$ esetén

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

A definíció kapcsán érdemes felidézni, hogy ahhoz, hogy egy függvény integrálható legyen, nem szükséges, hogy minden pontban értelmezve legyen. Ugyancsak közismert, hogy egy függvényt véges számú pontban megváltoztatva az integrál értéke nem változik. Ezért, szemben az eloszlásfüggvénnyel, a sűrűségfüggvény egyrészt nem feltétlenül értelmes a számegyenes minden pontjában, másrészt a sűrűségfüggvény nem egyértelmű. A sűrűségfüggvények meghatározása igen gyakran az analízisből ismert Newton–Leibniz-tételre épül:

3.1.13. Állítás. *Ha az $F(x)$ eloszlásfüggvény folytonosan deriválható, akkor az $f(x) \circeq F'(x)$ derivált az F egy sűrűségfüggvénye.*

3.1.14. Példa. A Cauchy-eloszlás sűrűségfüggvénye.

A (standard) Cauchy-eloszlás egy sűrűségfüggvénye

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$

□

Sajnos az állítás nem mindig alkalmazható, ugyanis a legtöbb eloszlásfüggvény „kapcsos zárójellel“ adott, vagyis nem minden pontban deriválható:

3.1.15. Példa. Egyenletes eloszlás sűrűségfüggvénye.

Az (a,b) intervallumon való egyenletes eloszlás egy sűrűségfüggvénye

$$f(x) \circeq \begin{cases} \frac{1}{b-a} & \text{ha} \quad x \in (a,b) \\ 0 & \text{ha} \quad x \notin (a,b) \end{cases}.$$

Az a és b pontokban a $f(x)$ nincs definiálva, de a sűrűségfüggvény definíciója szerint erre nincsen is szükség. Szemben az eloszlásfüggvényekkel a sűrűségfüggvények értelmezési tartománya nem feltétlenül a teljes számegyenes.

□

3.1.16. Példa. Az exponenciális eloszlás sűrűségfüggvénye.

Az exponenciális eloszlás egy sűrűségfüggvénye

$$f(x) \circeq \begin{cases} 0 & \text{ha} \quad x \le 0 \\ \lambda \exp(-\lambda x) & \text{ha} \quad x > 0 \end{cases}.$$

□

3.1.17. Példa. Diszkrét eloszlás sűrűségfüggvényének kérdése.

Mivel a sűrűségfüggvény definíciója alapján amennyiben létezik sűrűségfüggvény, akkor az eloszlásfüggvény a sűrűségfüggvény integrálfüggvénye. Minden integrálfüggvény folytonos, így egy diszkrét eloszlásnak nincs sűrűségfüggvénye, annak ellenére, hogy az eloszlásfüggvény deriváltja esetleg véges számú ponttól eltekintve létezik és folytonos.

□

A példák alapján nem érdektelen a következő észrevétel:

3.1.18. Állítás. *Ha az $F(x)$ eloszlásfüggvény szakaszonként folytonosan deriválható és az $F(x)$ függvény folytonos, akkor az $f(x) = F'(x)$ deriváltfüggvény az F egy sűrűségfüggvénye.*

Bizonyítás: Emlékeztetünk, hogy a szakaszonként folytonosan deriválhatóság definíciója szerint létezik véges sok $x_1 < x_2 < \ldots < x_n$ pont, hogy ezeken kívül az F folytonosan deriválható. Jelölje f a deriváltat. Ha $x_{k-1} < a < b < x_k$, akkor a Newton–Leibniz-formula miatt

$$F(b) - F(a) = \int_a^b f(t)\,dt. \tag{3.1.1}$$

Tekintsük most az $x_{k-1} < a < b = x_k$ esetet. Mivel az integrál a felső határ folytonos függvénye és mivel, az F balról folytonos, ezért az egyenlőség ebben az esetben is érvényben marad. Ugyanakkor az $x_{k-1} = a < b < x_k$ esetben általában csak az

$$F(b) - F(a+0) = \int_a^b f(t)\,dt$$

egyenlőség teljesül. Ha azonban az F folytonos, akkor $F(a) = F(a+0)$, így a fenti (3.1.1) egyenlőség ilyenkor érvényben marad. Ebből az állítás igazolása már evidens.

□

3.2. Többdimenziós valószínűségi változók

Ezidáig csak az egydimenziós valószínűségi változókat definiáltuk. Teljesen analóg módon definiálhatjuk a többdimenziós valószínűségi változókat, amelyeket mint az Ω alaptér $\mathbb{R}^n$-be való leképezéseit értelmezzük. A szokásos jelölés szerint többdimenziós eloszlások esetén kiírjuk a koordinátákat, vagyis n-dimenziós vektor értékű függvények helyett $\xi_1, \xi_2, \ldots, \xi_n$ véges sorozatokról beszélünk. Az egydimenziós esethez hasonlóan értelmezhetjük az eloszlás és a sűrűségfüggvényeket. Az egyszerűbb jelölés céljából csak a kétdimenziós eloszlásokkal foglalkozunk, az általános eset tárgyalása teljesen analóg.

3.2.1. Definíció. A (ξ,η) kétdimenziós valószínűségi változó $F(x,y)$ eloszlásfüggvényén a minden (x,y) számpárra értelmezett

$$F(x,y) \mathrel{\hat{=}} \mathbf{P}(\xi < x, \eta < y)$$

módon definiált kétváltozós függvényt értjük. Az $f(x,y)$ kétváltozós integrálható függvényt az F eloszlásfüggvény sűrűségfüggvényének mondjuk, ha minden I és J intervallum esetén

$$\mathbf{P}(\xi \in I, \eta \in J) = \int_I \int_J f(x,y)\, dydx.$$

A sűrűségfüggvény definíciójából nyilvánvalóan, tetszőleges x,y esetén

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\, dvdu. \tag{3.2.1}$$

Ha a sűrűségfüggvény folytonos, akkor

$$\frac{\partial^2 F}{\partial x \partial y}(x,y) = f(x,y),$$

vagyis ha az F rendelkezik folytonos sűrűségfüggvénnyel, akkor az F vegyes parciális deriváltjai megadják a sűrűségfüggvényt. Másoldalról, ha f egy olyan nem negatív, integrálható függvény, amelyre $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dydx = 1$, akkor a (3.2.1) sor alapján definiált $F(x,y)$ függvény tekinthető eloszlásfüggvénynek.

A többdimenziós eloszlásokkal kapcsolatos legfontosabb fogalom a peremeloszlás fogalma. Peremeloszláson értelemszerűen a ξ és az η mint egyváltozós valószínűségi változók eloszlását értjük. A (ξ,η) pár eloszlására mint együttes eloszlásra szokás hivatkozni. Az együttes eloszlásból igen egyszerű előállítani a peremeloszlásokat, ugyanis nyilvánvalóan igazak a következő állítások:

3.2.2. Tétel. *Legyen $F(x,y)$ egy (ξ,η) pár eloszlásfüggvénye, illetve $f(x,y)$ jelölje a sűrűségfüggvényt.*

1. *Tetszőleges x esetén*

$$G(x) \mathrel{\hat{=}} \mathbf{P}(\xi < x) = \lim_{y\to\infty} F(x,y) \mathrel{\hat{=}} F(x,\infty),$$

vagyis ha az y változóval a végtelenbe tartunk, akkor az így kapott függvény, amely természetesen egyedül az x függvénye, éppen a ξ eloszlásfüggvénye. Hasonlóan a

$$H(y) \mathrel{\hat{=}} \mathbf{P}(\eta < y) = \lim_{x\to\infty} F(x,y) \mathrel{\hat{=}} F(\infty,y)$$

éppen az η eloszlásfüggvénye.

2. *Ha az $f(x,y)$ sűrűségfüggvény létezik, akkor a ξ, illetve az η változók eloszlásának is van sűrűségfüggvénye és a*

$$g(x) \mathrel{\stackrel{\circ}{=}} \int_{-\infty}^{\infty} f(x,y)\,dy$$

paraméteres integrál a ξ és a

$$h(y) \mathrel{\stackrel{\circ}{=}} \int_{-\infty}^{\infty} f(x,y)\,dx$$

paraméteres integrál az η eloszlásának sűrűségfüggvénye.

Bizonyítás: A tétel első állításának bizonyítása egyszerű következménye az alábbi észrevételnek, amely a valószínűség elemi folytonossági tulajdonságaiból evidens:

$$\begin{aligned} G(x) &\mathrel{\stackrel{\circ}{=}} \mathbf{P}(\xi < x) = \mathbf{P}(\xi < x, \eta \text{ tetszőleges}) = \\ &= \mathbf{P}(\xi < x, \cup_n \{\eta < n\}) = \mathbf{P}(\cup_n \{\xi < x, \eta < n\}) = \\ &= \lim_{n \nearrow \infty} F(x,n) \mathrel{\stackrel{\circ}{=}} F(x,\infty). \end{aligned}$$

Vegyük észre, hogy minden x-re az F az y változó monoton növekedő függvénye, így a határértéket elegendő egyetlen sorozatra ellenőrizni. A sűrűségfüggvényre vonatkozó észrevétel igazolása analóg módon történik: Vezessük be az állításban szereplő g függvényt. Ekkor tetszőleges $a < b$ esetén, ismételten a valószínűség folytonossági tulajdonságai miatt

$$\begin{aligned} \int_a^b g(x)\,dx &\mathrel{\stackrel{\circ}{=}} \int_a^b \int_{-\infty}^{\infty} f(x,y)\,dydx = \int_{-\infty}^{\infty} \int_a^b f(x,y)\,dxdy = \\ &= \lim_{n \nearrow \infty} \int_{-n}^{n} \int_a^b f(x,y)\,dxdy = \\ &= \lim_{n \nearrow \infty} \mathbf{P}(a \le \xi < b, -n \le \eta < n) = \\ &= \mathbf{P}(a \le \xi < b, -\infty < \eta < \infty) = \\ &= \mathbf{P}(a \le \xi < b, \eta \text{ tetszőleges}) = \mathbf{P}(a \le \xi < b). \end{aligned}$$

vagyis a g valóban sűrűségfüggvénye a ξ változónak.

A figyelmes olvasó észreveheti, hogy a kettős integrálok néhány fontos tulajdonságát a számolás során kihasználtuk. Egyrészt kihasználtuk, hogy a g függvény értelmezhető, illetve kihasználtuk azt is, hogy az integrál értéke független attól, hogy milyen sorrendben vettük az integrálokat. Ezen tulajdonságok igazolása ha nem is nehéz, de nem is annyira nyilvánvaló, miként azt esetleg az első ránézésre feltételeznénk. Ha az $f(x,y)$ függvényre feltételezzük a folytonosságot, akkor a bizonyítás többé-kevésbé elemi eszközökkel elvégezhető, de nagyon sok esetben az együttes sűrűségfüggvény folytonossága, illetve mindenhol való definiáltsága nem feltétlenül teljesül. Például, ha a ξ változó nem negatív, miközben az η tetszőleges, akkor az $x = 0$ egyenes mentén az együttes sűrűségfüggvény várhatóan nem lesz feltétlenül folytonos. A nem folytonos eset kezelése azonban nem teljességgel triviális, ugyanis már két változóban is igen bonyolult esetek fordulhatnak elő. Emellett a folytonosság feltételezése azért is problémás, mert a kulcstulajdonság, amely az integrálok felcserélhetőségét, illetve a paraméteres integrálás értelmességét biztosítja, az az együttes sűrűségfüggvény nem negativitása. A pontos állítás, amely a bizonyítás korrektségét biztosítja az úgynevezett Fubini-tétel, amely szintén azok közé a tételek közé tartozik, amelyek indoklása nagyon messze vezetne. Ugyanakkor azt is érdemes észrevenni, hogy az integrálok felcserélhetősége vagyis a tetszőleges I és J intervallumokra teljesülő

$$\int_I \int_J f(x,y)\,dxdy = \int_J \int_I f(x,y)\,dydx$$

feltétel már implicite a sűrűségfüggvény definíciójából is következik, ugyanis mind a két integrál megegyezik a

$$\begin{aligned}
\mathbf{P}(\xi \in I, \eta \in J) &\mathrel{\overset{\circ}{=}} \mathbf{P}(\{\xi \in I\} \cap \{\eta \in J\}) = \\
&= \mathbf{P}(\{\eta \in J\} \cap \{\xi \in I\}) = \\
&= \mathbf{P}(\eta \in J, \xi \in I)
\end{aligned}$$

valószínűséggel. Mivel a definícióban az I és a J tetszőleges intervallumok lehetnek, ezért a két integrál felcserélhetősége, illetve a paraméteres integrálok értelmessége implicit módon része a többváltozós sűrűségfüggvény definíciójának.

□

Kézenfekvően felmerül az előző állítás megfordításának kérdése: Miként lehet a peremeloszlásokból rekonstruálni az együttes eloszlást? A válasz nem túl meglepően az, hogy további információk nélkül sehogy. Az egydimenziós eloszlások tárgyalása során bemutatott példákból világos, hogy az eloszlások körében az egyenletes eloszlás kitüntetett szereppel bír. A $(0,1)$ intervallumon egyenletes eloszlás többdimenziós megfelelője a kopula.

3.2.3. Definíció. Egy $(\xi_1, \xi_2, \ldots, \xi_n)$ többdimenziós valószínűségi változót kopulának mondunk, ha a peremeloszlások, vagyis a ξ_k változók, eloszlása egyenletes a $(0,1)$ intervallumon.

Tegyük fel, hogy ismerjük az $\eta_1, \eta_2, \ldots, \eta_n$ többdimenziós valószínűségi változó peremeloszlásait. A megfelelő eloszlásfüggvények legyenek $F_1, F_2, \ldots, F_n$. Az egyszerűség kedvéért tegyük fel, hogy ezek az eloszlásfüggvények folytonosak és szigorúan monoton nőnek, következésképpen léteznek az inverzfüggvényeik. Vegyük észre, hogy a

$$(\xi_1, \xi_2, \ldots, \xi_n) \mathrel{\overset{\circ}{=}} (F_1(\eta_1), F_2(\eta_2), \ldots, F_n(\eta_n))$$

egy kopula, ugyanis a minden koordinátája egyenletes eloszlású. Valóban, mivel tetszőleges y esetén $0 \le F_k(y) \le 1$, ezért az $F_k(\eta_k)$ eloszlásfüggvényét elegendő a $[0,1]$

szakaszon kiszámolni. Ilyenkor, felhasználva az inverzfüggvény létezését

$$\mathbf{P}(\xi_k < x) \stackrel{\circ}{=} \mathbf{P}(F_k(\eta_k) < x) = \mathbf{P}\left(\eta_k < F_k^{-1}(x)\right) = F_k\left(F_k^{-1}(x)\right) = x,$$

következésképpen a ξ_k egyenletes eloszlású, így a $(\xi_1, \xi_2, \ldots, \xi_n)$ egy kopula. Megfordítva, ha adott egy $(\xi_1, \xi_2, \ldots, \xi_n)$ kopula és az F_k eloszlásfüggvények szigorúan monoton nőnek és folytonosok, akkor az

$$\left(F_1^{-1}(\xi_1), F_2^{-1}(\xi_2), \ldots, F_n^{-1}(\xi_n)\right)$$

egy olyan többdimenziós eloszlás, amely peremeloszlásai éppen F_k-val reprezentálhatóak. Valóban, mivel a ξ_k egyenletes eloszlású a $[0,1]$ szakaszon, illetve mivel az F_k szigorúan monoton nő, ezért

$$\mathbf{P}\left(F_k^{-1}(\xi_k) < x\right) = \mathbf{P}\left(F_k\left(F_k^{-1}(\xi_1)\right) < F_k(x)\right) = \mathbf{P}(\xi_k < F_k(x)) = F_k(x),$$

ugyanis ha $0 \leq y \leq 1$, akkor $\mathbf{P}(\xi_k < y) = y$. Mindezek alapján a többdimenziós eloszlások megadásához a peremeloszlásokon kívül az együttmozgást megadó kopulát is meg kell adni. Szemben az egydimenziós esettel, ahol a $[0,1]$ szakaszon való egyenletes eloszlás fogalma egyértelmű, a kopula nem egyértelműen meghatározott fogalom, ezért adott peremeloszlásokhoz számos egymástól eltérő többdimenziós eloszlás adható meg. A legegyszerűbb kopula a függetlenségből származó kopula, amikor a kopulát megadó változók együttes sűrűségfüggvénye azonosan 1 a $[0,1]^n$ egységnégyzeten és nyilván azon kívül pedig mindenhol 0. Könnyen látható, hogy ilyenkor a kopula együttes sűrűségfüggvénye a peremsűrűségfüggvények szorzata, következésképpen az együttes eloszlásfüggvény is a peremeloszlás-függvények szorzata. Ilyenkor például kétváltozóban az F_1 és F_2 peremeloszlásokhoz tartozó többdimenziós eloszlás

$$\begin{aligned} F(x,y) &= \mathbf{P}(\xi_1 < x, \xi_2 < y) = \mathbf{P}\left(F_1^{-1}(\eta_1) < x, F_2^{-1}(\eta_2) < y\right) = \\ &= \mathbf{P}\left(F_1^{-1}(\eta_1) < x\right)\mathbf{P}\left(F_2^{-1}(\eta_2) < y\right) = F_1(x) F_2(y). \end{aligned}$$

Ez indokolja a következő definíciót:

3.2.4. Definíció. A $\xi_1, \xi_2, \ldots \xi_n$ változókat függetlennek mondjuk, ha az együttes eloszlásfüggvényük a peremeloszlásfüggvények szorzata.

3.2.5. Állítás. *Ha a $\xi_1, \xi_2, \ldots, \xi_n$ változók függetlenek, és $I_1, I_2, \ldots, I_n$ tetszőleges intervallumok, akkor*

$$\mathbf{P}(\xi_1 \in I_1, \xi_2 \in I_2, \ldots, \xi_n \in I_n) = \mathbf{P}(\xi_1 \in I_1) \cdots \mathbf{P}(\xi_n \in I_n).$$

Bizonyítás: Az állítás igen természetesnek tűnik, valójában a számbajöhető esetek nagy száma miatt közvetlen úton, pusztán a definíciókra támaszkodva az igazolása nehézkes, ugyanakkor kézenfekvő, ezért az olvasóra bízzuk[2].

[2] Absztrakt nyelvezetet és eszköztárat használva az indoklás igen egyszerű és természetes és jó példát szolgáltat arra, hogy az elemi valószínűségszámítás legegyszerűbb tételeinek igazolása is az eszközök elemi jellege miatt sokszor szükségtelenül nehéz tud lenni.

□

Miként már jeleztük a függetlenség a valószínűségszámítás legfontosabb fogalma. Fontosságát az adja, hogy igen általános körülmények között, ha a $\xi_1, \xi_2 \ldots, \xi_n$ változók függetlenek, és $h_1, h_2, \ldots, h_n$ igen általános függvények, akkor a

$$h_1(\xi_1), h_2(\xi_2,), \ldots, h_n(\xi_n)$$

változók is függetlenek maradnak. Vagyis a függetlenség tulajdonsága nem vész el, ha a változókon nem lineáris transzformációkat végzünk.

3.2.6. Példa. Ha a ξ_1 és a ξ_2 változók függetlenek, akkor függetlenek az $\exp(\xi_1)$ és az $\exp(\xi_2)$ változók is.

Tekintsük az együttes eloszlásfüggvényt. Ha $x, y > 0$, akkor

$$\begin{aligned}\mathbf{P}(\exp(\xi_1) < x, \exp(\xi_2) < y) = \mathbf{P}(\xi_1 < \ln x, \xi_2 < \ln y) = \\ = \mathbf{P}(\xi_1 < \ln x)\mathbf{P}(\xi_2 < \ln y) = \\ = \mathbf{P}(\exp(\xi_1) < x)\mathbf{P}(\exp(\xi_2) < y).\end{aligned}$$

□

3.2.7. Példa. Ha a ξ_1 és a ξ_2 változók függetlenek, akkor függetlenek a $|\xi_1|$ és a $|\xi_2|$ változók is.

Tekintsük az együttes eloszlásfüggvényt. Ha $x, y > 0$, akkor

$$\begin{aligned}\mathbf{P}(|\xi_1| < x, |\xi_2| < y) = \mathbf{P}(-x < \xi_1 < x, -y < \xi_2 < y) = \\ = \mathbf{P}(-x < \xi_1 < x)\mathbf{P}(-y < \xi_2 < y) = \\ = \mathbf{P}(|\xi_1| < x)\mathbf{P}(|\xi_2| < y).\end{aligned}$$

□

3.2.8. Példa. Ha a ξ_1 és a ξ_2 változók függetlenek, akkor függetlenek a ξ_1^2 és a ξ_2^2 változók is.

Tekintsük az együttes eloszlásfüggvényt. Ha $x, y > 0$, akkor

$$\begin{aligned}\mathbf{P}\left(\xi_1^2 < x, \xi_2^2 < y\right) = \mathbf{P}\left(|\xi_1| < \sqrt{x}, |\xi_2| < \sqrt{y}\right) = \\ = \mathbf{P}\left(|\xi_1| < \sqrt{x}\right)\mathbf{P}\left(|\xi_2| < \sqrt{y}\right) = \\ = \mathbf{P}\left(\xi_1^2 < x\right)\mathbf{P}\left(\xi_2^2 < y\right).\end{aligned}$$

□

3.2.9. Példa. Számoljuk ki a $(0,0), (1,0)$ és $(0,1)$ csúcspontokkal rendelkező háromszögben egyenletes eloszlású valószínűségi változó peremeloszlásait.

Jelölje Δ a háromszöget. A háromszög területe $1/2$ és mivel az eloszlás egyenletes, ezért definíció szerint a sűrűségfüggvény konstans a Δ háromszögön, vagyis

$$f(x,y)=\begin{cases} c & \text{ha} \quad (x,y)\in\Delta \\ 0 & \text{ha} \quad (x,y)\notin\Delta \end{cases}.$$

A kérdés csak az, hogy mennyi a c értéke? Mivel sűrűségfüggvényről van szó, ezért teljesülni kell az

$$1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = c\cdot\lambda(\Delta)$$

összefüggésnek, ahol $\lambda(\Delta)$ a háromszög területe. Ebből $c=2$. A szimmetria miatt elegendő az egyik változó peremeloszlását meghatározni. A peremsűrűségfüggvény

$$g(x)=\int_{-\infty}^{\infty} f(x,y)\,dy.$$

Világos, hogy ha $x\notin[0,1]$, akkor az $f(x,y)=0$, így elegendő az $x\in[0,1]$ szakaszra koncentrálni. Ilyenkor a sűrűségfüggvény az $y<0$ és az $y>x$ tartományokon nulla, így

$$g(x)=\int_{-\infty}^{\infty} f(x,y)\,dy=\int_{0}^{1-x} 2dy = 2(1-x), \quad 0\le x\le 1.$$

Tehát

$$g(x)=\begin{cases} 2(1-x) & \text{ha} \quad x\in[0,1] \\ 0 & \text{ha} \quad x\notin[0,1] \end{cases}.$$

Az így kapott függvény valóban sűrűségfüggvény, ugyanis $g(x)\ge 0$ és

$$\int_{-\infty}^{\infty} g(x)\,dx=\int_0^1 g(x)\,dx=\int_0^1 2(1-x)\,dx = 2\left[x-\frac{x^2}{2}\right]_0^1 = 1.$$

□

3.2.10. Példa. Egy félkörben egyenletes eloszlású valószínűségi változó sűrűségfüggvénye.

Legyen K most az $y\ge 0$ félsíkba eső egységnyi sugarú félkör. K területe $\lambda(K)=\pi/2$. A már az előző példában látott módon eljárva

$$f(x,y)=\begin{cases} 2/\pi & \text{ha} \quad (x,y)\in K \\ 0 & \text{ha} \quad (x,y)\notin K \end{cases}.$$

Az x változó szerinti sűrűségfüggvény ha $x\in(-1,1)$

$$g(x)=\int_{-\infty}^{\infty} f(x,y)\,dy=\int_0^{\sqrt{1-x^2}}\frac{2}{\pi}dy=\frac{2}{\pi}\sqrt{1-x^2},$$

így

$$g(x)=\begin{cases}\frac{2}{\pi}\sqrt{1-x^2} & \text{ha} \quad -1<x<1\\ 0 & \text{ha} \quad x\notin(-1,1)\end{cases}.$$

Ellenőrzésképpen talán nem teljesen érdektelen ellenőrizni, hogy a g valóban sűrűségfüggvény. Egyrészt $g(x)\geq 0$, másrészt

$$\int_{-\infty}^{\infty} g(x)\,dx=\frac{2}{\pi}\int_{-1}^{1}\sqrt{1-x^2}dx=\frac{4}{\pi}\int_{0}^{1}\sqrt{1-x^2}dx=1,$$

ugyanis az elemi analízisből ismert, hogy az integrál értéke $\pi/4$. Ugyanakkor ha $y\in(0,1)$, akkor

$$h(y)=\int_{-\infty}^{\infty} f(x,y)\,dx=\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\frac{2}{\pi}dx=\frac{4}{\pi}\sqrt{1-y^2}.$$

Ismételten sűrűségfüggvényről van szó, ugyanis a $(0,1)$ intervallumon kívül a $h(y)$ nulla, így

$$\int_{-\infty}^{\infty} g(y)\,dy=\frac{4}{\pi}\int_{0}^{1}\sqrt{1-y^2}dy=1.$$

□

3.2.11. Példa. Az $F(x,y)=\left(x^{-1}+y^{-1}-1\right)^{-1}$, ha $x,y\in(0,1)$, tekinthető egy kopula eloszlásfüggvényének.

A peremeloszlásokat az $x=1$ és az $y=1$ helyettesítéssel kaphatjuk, amiből evidens, hogy a peremeloszlások egyenletesek. Többdimenziósban nincs egyszerű, könnyen használható kritérium annak eldöntésére, hogy eloszlásfüggvényről van-e szó. A legegyszerűbben használható kritérium, ha megmutatjuk, hogy a másodrendű vegyes parciális derivált folytonos és nem negatív, vagyis használható sűrűségfüggvénynek[3].

$$\begin{aligned}
\frac{\partial^2 F}{\partial x\partial y}(x,y) &= \frac{\partial^2}{\partial x\partial y}\frac{xy}{x+y-xy}=\frac{\partial}{\partial y}\frac{y(x+y-xy)-(1-y)xy}{(x+y-xy)^2}=\\
&= \frac{\partial}{\partial y}\frac{y^2}{(x+y-xy)^2}=\frac{\partial}{\partial y}\left(\frac{y}{x+y-xy}\right)^2=\\
&= 2\frac{y}{x+y-xy}\frac{x+y-xy-y(1-x)}{(x+y-xy)^2}=\\
&= 2\frac{yx}{(x+y-xy)^3}\geq 0
\end{aligned}$$

valahányszor $x,y\in(0,1)$.

□

[3] Többdimenzióban gyakran előfordul, hogy az eloszlás folytonos, de nincs sűrűségfüggvény. Érdemes megjegyezni, hogy ez egydimenzióban is előfordulhat, de ilyen ellenpélda konstruálása nem tartozik az elemi valószínűségszámítás tárgykörébe.

IV.

STIELTJES-INTEGRÁLÁS

Ebben a fejezetben egy rövid kitérőt teszünk. A kitérő során áttekintjük az integrálelmélet néhány kérdését. Ismételten nem a legáltalánosabb megközelítést követjük és csak néhány elemi állítást foglalunk össze. Az itt leírtak szervesen kapcsolódnak az egyetemi bevezető matematika tananyaghoz, amelynek ismeretét feltételezzük. Mivel az olvasó már ismeri az analízis legfontosabb fogalmait, így a logikai szerkezetben a lineáris kifejtési sorrendtől eltekintünk. A fejezet legfontosabb mondanivalója, hogy a Riemann-integrál, illetve annak közvetlen és igen természetes általánosítása a Riemann–Stieltjes-integrál, lényegében csak a folytonos függvények körében definiálható. Már a legegyszerűbb szakadásos függvényekre is előfordulhat, hogy a Riemann–Stieltjes-integrál az itt kifejtett klasszikus megközelítéssel nem definiálható. A Riemann–Stieltjes-integrál ezen hiányossága ellenére a fogalom nagyban egyszerűsíti a várható érték tárgyalását, amely a következő fejezet tárgya. Mivel az integrálás eredendően a várható értékhez kapcsolódik, ebben a fejezetben már érintőlegesen használjuk a várható érték fogalmát. Első olvasásra, amennyiben az idevágó megjegyzések nem érthetőek, figyelmen kívül hagyhatóak.

4.1. Newton–Leibniz-szabály

A legtöbb ember az integrálelmélettel való ismerkedés során általában nem érti, hogy miért van szükség több fajta, Riemann, Lebesgue, Stieltjes stb. integrálokra, vagy ami ugyanaz, miközben, egyre több integrálfogalmat tanul meg, mégsem tud több konkrét integrált kiszámolni. Hogy egy kicsit pontosabban fogalmazzunk, nem világos az integrálás és a Newton–Leibniz-szabály viszonya. Kezdjük tehát ezzel a problémával.

A Riemann-integrál definíciója és tulajdonságai

Indulásképpen megjegyezzük, hogy első közelítésben az f integrálandó függvényt folytonosnak tekinthetjük, illetve legfeljebb olyannak, amelynek véges sok szakadási pontja van. Az integrál Riemann-féle definíciója közismert. Rögzítsünk egy $[a,b]$ véges, zárt intervallumot, és osszuk fel[1]

$$a = x_0 < x_1 < \cdots < x_n = b$$

osztópontokkal n részre. Jelölje m_k, illetve M_k, a k-dik részintervallumon az integrálandó függvény infimumát, illetve szuprémumát, és tekintsük az

$$s \mathrel{\overset{\circ}{=}} \sum_{k=1}^{n} m_k \left(x_k - x_{k-1}\right), \quad S \mathrel{\overset{\circ}{=}} \sum_{k=1}^{n} M_k \left(x_k - x_{k-1}\right)$$

[1] A következők megértése szempontjából, bár nyilvánvaló, érdemes hangsúlyozni, hogy a közelítő összegek mindig véges összegek, így mindig értelmezettek, feltéve, hogy a függvény korlátos, így a szuprémum és az infimum véges.

alsó, illetve felső, közelítő összegeket. Világos, hogy $s \leq S$, de az is könnyen belátható, hogy ez akkor is érvényben marad, ha a két közelítő összeg más-más felosztáshoz tartozik. Másképpen, egyetlen alsó közelítő összeg sem lehet nagyobb egyetlen felső közelítő összegnél.

4.1.1. Definíció. Az f függvényt Riemann-integrálhatónak nevezzük, ha az alsó közelítő összegek szuprémuma megegyezik a felső közelítő összegek infimumával. Az így kapott közös

$$-\infty < \sup s = \inf S < \infty$$

értéket az f függvény Riemann-integráljának mondjuk, és

$$\int_a^b f(x)\,dx, \quad \text{vagy} \quad \int_{[a,b]} f(x)\,dx$$

módon jelöljük.

A definíció alapján evidens, hogy az f függvénynek korlátosnak kell lenni, a nem korlátos függvények automatikusan nem Riemann-integrálhatóak. Minden bevezető analízis könyvben szerepel a következő állítás, amely az integrál definíciója alapján az oszcillációs összeg felhasználásával igazolható.

4.1.2. Állítás. *A Riemann-integrál a következő tulajdonságokkal rendelkezik:*

1. *Minden az $[a,b]$ szakaszon folytonos függvény az $[a,b]$ szakaszon integrálható.*
2. *Ha az f az $[a,b]$ szakaszon integrálható, akkor ott az $|f|$ is integrálható és*

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx,$$

vagyis az integrálhatóságból következik az abszolút integrálhatóság, de nem azonos vele.

3. *Ha az f és a g integrálható, akkor az $f+g$ is integrálható, és*

$$\int_a^b f(x)\,dx + \int_a^b g(x)\,dx = \int_a^b (f(x)+g(x))\,dx,$$

vagyis az integrál additív funkcionál. Ha az f integrálható és a λ egy tetszőleges konstans, akkor a λf is integrálható és

$$\int_a^b \lambda f(x)\,dx = \lambda \int_a^b f(x)\,dx,$$

vagyis az additivitást is figyelembe véve az integrálható függvények körében az integrál az integrandus lineáris funkcionálja[2].

[2] Fontos hangsúlyozni, hogy abból, hogy az $f+g$ integrálható, nem következik, hogy az f és a g is integrálható. Ha az f nem integrálható, és g éppen a $-f$, akkor az $f+g=0$ függvény integrálható.

4. *Ha az $f \geq 0$ függvény integrálható, akkor $\int_a^b f(x)\,dx \geq 0$, vagyis az integrál nem negatív funkcionál.*

5. *Ha az f integrálható az $[a,b]$ intervallumon, és $c \in [a,b]$, akkor az f integrálható az $[a,c]$ és $[c,b]$ intervallumokon, és*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx, \tag{4.1.1}$$

valamint megfordítva, ha az f integrálható az $[a,c]$ és a $[c,b]$ intervallumokon, akkor integrálható az $[a,b]$ intervallumon is, és teljesül a (4.1.1). Ez alapján az integrál az integrációs tartomány additív intervallumfüggvénye. Ugyanakkor[3]

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,\chi_{[a,c]}(x)\,dx.$$

Vegyük észre, hogy az $f\chi_{[a,c]}$ függvénynek általában az a és c pontokban szakadása van, így a formula teljesüléséhez szükség van az integrál értelmezésében nem folytonos integrandusokat is megengedni.

6. *Az integrál előállítható tetszőlegesen választott közbülső pontokhoz tartozó felosztássorozathoz tartozó közelítő összegek határértékeként, vagyis ha tekintjük az*

$$a = x_0^{(n)} < x_1^{(n)} \ldots < x_{m_n}^{(n)} = b$$

felosztáshoz tartozó $\sum_{k=1}^n f\left(\tau_k^{(n)}\right)\left(x_k^{(n)} - x_{k-1}^{(n)}\right)$ közelítő összeget és

$$\lim_{n\to\infty} \max_k \left(x_k^{(n)} - x_{k-1}^{(n)}\right) = 0,$$

akkor a $\tau_k^{(n)} \in \left[x_{k-1}^{(n)}, x_k^{(n)}\right]$ pontok választásától függetlenül

$$\lim_{n\to\infty} \sum_{k=1}^{m_n} f\left(\tau_k^{(n)}\right)\left(x_k^{(n)} - x_{k-1}^{(n)}\right) = \int_a^b f(x)\,dx.$$

A Riemann-integrál improprius kiterjesztése

Az előző 4.1.1. definíció alapján nem korlátos függvényeknek, vagy nem zárt, vagy nem korlátos intervallumon értelmezett függvényeknek nem tudunk integrált tulajdonítani. Mit jelentenek akkor az analízisből ismert $\int_0^1 1/\sqrt{x}dx = 2$, vagy az

$$\int_{-\infty}^{\infty} \exp\left(-x^2\right) dx = \sqrt{\pi}$$

összefüggések? Hogyan terjesszük ki az integrált nem korlátos tartományokra, vagy nem korlátos függvényekre? Erre szolgál az improprius integrál bevezetése.

[3] A továbbiakban χ_A alatt az A halmaz karakterisztikus vagy indikátorfüggvényét értjük. Definíció szerint χ_A az A pontjaiban egy és az A^c halmazon pedig nulla.

4.1.3. Definíció. Legyen I olyan véges vagy végtelen intervallum, amelyen az f függvény integrálja nincsen definiálva, vagy azért mert az intervallum nem korlátos, vagy azért mert az intervallumon az f függvény nem korlátos.

1. Tegyük fel, hogy az I előáll olyan $I_n \subseteq I_{n+1}$ zárt, véges intervallumok monoton növekedő sorozatának egyesítéseként amelyek mindegyikén az f integrálható. Ha a
$$\lim_{n\to\infty} \int_{I_n} f(x)\,dx$$
határérték létezik, véges és független az $I_n \nearrow I$ sorozat megválasztásától, akkor a határértéket az f függvény I intervallumra vonatkozó improprius integráljának nevezzük, és
$$\int_I f(x)\,dx$$
módon jelöljük[4]. Értelemszerűen, ha például az I végpontjai a és b, ahol az a és a b lehetnek végtelenek is, akkor az improprius integrált a Riemann-integrálhoz hasonlóan
$$\int_a^b f(x)\,dx$$
módon is szokás jelölni.

2. Természetesen az f függvény nem csak az I integrációs tartomány végpontjai körül lehet korlátlan, hanem az intervallum belső pontjai körül is. Ilyenkor az integrációs tartományt a szakadási pontok szerint több részre bontjuk, az egyes részintervallumokon külön-külön integrálunk, majd az így kapott részintegrálokat összegezzük.

Vegyük észre, hogy az integrál ilyen irányú kiterjesztése esetén nem világos, hogy mit jelent az
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$
egyenlőség. Ez most állítás, vagy definíció? Érdemes megjegyezni, hogy az egész konstrukció némiképpen ad hoc. Ha véges számú szakadási pont van az integrációs szakaszon belül, akkor az eljárás kézenfekvő. Ha végtelen sok szakadási pont van, és a szakadási pontokból álló, mondjuk monoton növekedő, sorozatnak létezik határértéke, az eljárás szintén végrehajtható. Elképzelhető továbbá, hogy az integrációs tartomány több olyan részintervallumra bontható, amelyek mindegyikén a szakadási pontok sorozata konvergens, így az integrál definiálható, és az integrál a részintervallumokon vett integrálok összegeként írható fel. És így tovább. Az egyszerűség kedvéért azonban általában csak olyan improprius integrálokat szokás tekinteni, ahol a szakadási pontok száma véges.

[4]Megjegyezzük, hogy a definíció alapján könnyen belátható, hogy az integrálás formális szabályai érvényben maradnak.

4.1.4. Példa. Vizsgáljuk meg a Cauchy-eloszlás várható értékét megadó

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x}{1+x^2}dx$$

integrált.

Az improprius integrál definíciójában lényeges, hogy a határérték független az I_n intervallumok megválasztásától. Az $x/\left(1+x^2\right)$ függvény páratlan, ezért tetszőleges n esetén $\int_{-n}^{n} x/\left(1+x^2\right)dx = 0$, ami alapján az integrált nullának várhatnánk. Ugyanakkor a Newton–Leibniz-szabály alapján, ha például $I_n \stackrel{\circ}{=} [-n, 2n]$

$$\int_{-n}^{2n}\frac{x}{1+x^2}dx = \frac{1}{2}\left[\ln\left(1+x^2\right)\right]_{-n}^{2n} = \ln\sqrt{\frac{1+(2n)^2}{1+n^2}} \to \ln 2,$$

tehát a függvény improprius integrálja nem létezik, ami a valószínűségszámítás terminológiájával azt jelenti, hogy a Cauchy-eloszlásnak nincs várható értéke.

□

Nem csekély probléma forrása, ami végig kísérteni fog, hogy ha erre külön nem ügyelünk, akkor az integrál jelölésén nem látszik, hogy milyen integrálról van szó. Általában az elemi analízis tárgyalásakor integrál alatt mindig improprius integrált szokás érteni, és az improprius jelzőt általában el szokás hagyni. Természetesen ez az általánosan követett gyakorlat pontatlan, és ezért nem helyes. A jelölés pontatlansága bizonyos részletkérdéseket elfed. Ilyen például, amikor helyettesítéssel Riemann-integrált impropriussá alakítunk, vagy fordítva. Ugyanakkor az integrálok típusának jelölése általában inkább precízkedés, mint precízség, és a típus pontos jelölése valószínűleg inkább gátolná, mint elősegítené a megértést. Ha azonban véges intervallum esetén hangsúlyozni kell, hogy improprius integrálról van szó, akkor ezt külön szokás jelölni. Ha például hangsúlyozni kell, hogy az integrál a b pontban improprius, akkor az integrált az $\int_a^{b-} f(x)\,dx$, $\int_a^{b-0} f(x)\,dx$, esetleg az $\int_a^{\to b} f(x)\,dx$ vagy $\int_{[a,b)} f(x)\,dx$ módon szokás jelölni. Az integrálfogalom kiterjesztése csak akkor helyes, ha ebből nem eredhet félreértés. Az improprius Riemann-integrál esetén ez teljesül[5]:

4.1.5. Állítás. *Ha f Riemann-integrálható az $[a,b]$ intervallumon, akkor ott impropriusan is integrálható, és a két integrál értéke megegyezik. Másképpen fogalmazva, ha az f függvény Riemann-integrálható az $[a,b]$ intervallumon, akkor az*

$$x \mapsto F(x) \stackrel{\circ}{=} \int_a^x f(t)\,dt$$

integrálfüggvény az $[a,b]$ szakaszon folytonos.

[5] Bár az állítást triviálisnak szokás tekinteni, de például a Stieltjes-integrálokra már nem teljesül.

Bizonyítás: Az f Riemann-integrálható, ezért korlátos. Ha $[a_n, b_n] = I_n \subseteq [a,b]$, és K az f korlátja, akkor az integrál elemi tulajdonságai alapján

$$\begin{aligned}
\left|\int_a^b f(x)\,dx - \int_{a_n}^{b_n} f(x)\,dx\right| &= \left|\int_a^{a_n} f(x)\,dx + \int_{b_n}^{b} f(x)\,dx\right| \leq \\
&\leq \int_a^{a_n} |f(x)|\,dx + \int_{b_n}^{b} |f(x)|\,dx \leq \\
&\leq K(a_n - a + b - b_n) \to 0.
\end{aligned}$$

□

Az integrál kiterjesztése nem áldozat nélküli. Az integrál továbbra is nem negatív, lineáris funkcionál marad, de ha az f integrálható, akkor ebből már nem következik, hogy az $|f|$ függvény is integrálható lesz.

4.1.6. Példa. Az $\int_{-\infty}^{\infty} \frac{\sin x}{x} dx$ véges, de az $\int_{-\infty}^{\infty} \left|\frac{\sin x}{x}\right| dx$ végtelen, vagyis az

$$f(x) = \frac{\sin x}{x}$$

függvény impropriusan integrálható, de az abszolút értéke nem integrálható.

Mivel az integrandus páros, és a $[0,1]$ szakaszon folytonosnak tekinthető, elegendő az

$$\int_1^{\infty} \frac{\sin x}{x} dx \tag{4.1.2}$$

integrállal foglalkozni. Parciálisan integrálva

$$\int_1^{\infty} \frac{\sin x}{x} dx = \left[-\frac{\cos x}{x}\right]_1^{\infty} - \int_1^{\infty} \frac{\cos x}{x^2} dx = \cos 1 - \int_1^{\infty} \frac{\cos x}{x^2} dx.$$

Mivel $\left|x^{-2} \cos x\right| \leq 1/x^2$ ezért a majoráns kritérium alapján az utolsó integrál konvergens. Másrészt viszont

$$\begin{aligned}
\int_1^{\infty} \left|\frac{\sin x}{x}\right| dx &\geq \int_1^{\infty} \frac{\sin^2 x}{x} dx = \int_1^{\infty} \frac{1-\cos 2x}{2x} dx = \\
&= \left[\frac{1}{2} \ln x\right]_1^{\infty} - \int_1^{\infty} \frac{\cos 2x}{2x} dx.
\end{aligned}$$

Mivel az első kifejezés értéke végtelen, az integrálról pedig a (4.1.2) integrálhoz hasonlóan megmutatható, hogy véges, ezért az utolsó kifejezés végtelen.

□

A felesleges általánosítás elkerülése céljából foglalkozzunk egyedül az $\int_0^{\infty} g(x)\,dx$ típusú kifejezésekkel. Az integrál improprius kiterjesztése kapcsán kézenfekvően merül fel a kérdés, hogy miért nem használjuk a következő konstrukciót? Tegyük fel, hogy a g integrandus nem negatív, és tekintsük egy

$$x_0 = 0 < x_1 < x_2 < \ldots$$

monoton növekedő sorozatot. Ez követően tekintsük az

$$s \mathrel{\mathring{=}} \sum_k m_k (x_k - x_{k-1}), \quad S \mathrel{\mathring{=}} \sum_k M_k (x_k - x_{k-1}),$$

alsó és felső közelítő összegeket. Mivel $g \geq 0$ és az (x_k) sorozat monoton nő, ezért az összegek nem negatív tagokból állnak, így az összegek értéke mindig értelmes, de esetlegesen végtelen is lehet. Vegyük észre, hogy kihasználtuk, hogy $g \geq 0$, ugyanis ellenkező esetben a s és S sorok konvergenciáját nem tudnánk feltétlenül garantálni. Ezt követően tekintsük az alsó közelítő összegek s^* felső határát és a felső közelítő összegek S^* alsó határát. Igen kézenfekvő módon ha $s^* = S^* < \infty$, akkor azt mondjuk, hogy a g integrálható és az integrált $\int_0^\infty g(x)\,dx$ módon jelöljük.

4.1.7. Állítás. *Ha $g \geq 0$ és folytonos, akkor az imént bevezetett integrálás és az improprius integrálás megegyezik.*

Bizonyítás: Tegyük fel, hogy az I improprius integrál létezik. Definíció szerint az I ilyenkor véges. Kézenfekvő módon minden s alsó, illetve S felső közelítő összegre $s \leq I \leq S$. Ebből következően $s^* \leq I \leq S^*$. Ugyanakkor a g folytonossága miatt tetszőleges zárt szakaszon a felosztás finomításával az alsó és a felső közelítő összegek közötti eltérés tetszőlegesen kicsivé tehető. Ha a számegyenest felbontjuk megszámlálható szakaszra és az egyes szakaszokon az eltérést $\varepsilon/2^n$ alá visszük, akkor a teljes egyenesen az s és S eltérése kisebb lesz mint ε, vagyis mivel az ε tetszőleges, ezért $s^* = I = S^*$ lesz. Tegyük fel, hogy $s^* = S^* < \infty$. Világos módon tetszőleges x esetén $I(x) \mathrel{\mathring{=}} \int_0^x g(t)\,dt \leq S^*$. Mivel $g \geq 0$, ezért az $I(x)$ integrálfüggvény monoton növekedő, így az $I = \sup_x I(x)$ improprius integrál létezik és véges. De ekkor a már elmondottak alapján $I = S^* = s^*$.

□

Integrálok kiszámolása

Az improprius kiterjesztés legfontosabb indoka, hogy megőrzi az elemi analízis központi tételét, a Newton–Leibniz-szabályt.

4.1.8. Tétel (Newton–Leibniz-szabály). *Legyen f integrálható az $[a,b]$ intervallumon, ahol az a és b határok végtelenek is lehetnek.*

1. *Ha az f antideriváltja az (a,b) intervallumon F, és az F folytonos az $[a,b]$ szakaszon*[6]*, akkor*
$$\int_a^b f(x)\,dx = F(b) - F(a).$$

2. *Ha az f folytonos az (a,b)-n, akkor az*
$$F(x) \stackrel{\circ}{=} \int_a^x f(t)\,dt$$
integrálfüggvény az f antideriváltja az (a,b) nyílt intervallumon.

Bizonyítás: Hangsúlyozzuk, hogy az f integrálhatósága impropius értelemben értendő.

1. Legyen f először Riemann-integrálható. Ilyenkor az $[a,b]$ véges. Tetszőleges $a = x_0 < x_1 \ldots < x_n = b$ felosztására a középértéktétel alapján
$$F(b) - F(a) = \sum_k [F(x_k) - F(x_{k-1})] = \sum_k f(\xi_k)(x_k - x_{k-1}),$$
következésképpen
$$s_n \leq F(b) - F(a) \leq S_n, \tag{4.1.3}$$
így az f integrálhatóságának definíciója szerint
$$\int_a^b f(x)\,dx \leq F(b) - F(a) \leq \int_a^b f(x)\,dx.$$
Ha most az f impropius értelemben integrálható, és $I_n \stackrel{\circ}{=} [a_n, b_n]$ a megfelelő közelítő intervallum sorozat, akkor az F feltételezett folytonossága miatt
$$\int_I f(x)\,dx \stackrel{\circ}{=} \lim_{n\to\infty} \int_{a_n}^{b_n} f(x)\,dx = \lim_{n\to\infty} [F(b_n) - F(a_n)] = F(b) - F(a).$$
Mi történik akkor, ha az f az integrációs tartományon belül nem korlátos? Tegyük mondjuk fel, hogy egy c pontban az f végtelenhez tart. Ha az $[a,c]$ és $[c,b]$ intervallumokon az f (impropriusan) integrálható, akkor az integrált a két intervallumon vett integrál összegével definiáljuk, vagyis
$$\int_a^b f(x)\,dx \stackrel{\circ}{=} \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$
Formailag a Newton–Leibniz-szabály érvényben marad, hiszen ha az F függvény folytonos, és a c ponttól eltekintve az f antideriváltja, akkor
$$\begin{aligned}\int_a^b f(x)\,dx &\stackrel{\circ}{=} \int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \\ &= F(b) - F(c) + F(c) - F(b) = F(b) - F(a).\end{aligned}$$

[6] Értelemszerűen végtelen pontokban folytonosságon a határérték létezését értjük.

2. Térjünk rá a fordított irányra. Az integrál esetleg improprius, és az a pont lehet véges, de lehet $-\infty$ is. Rögzítsük az x pontot. Ha $\varepsilon > 0$, akkor az f folytonossága miatt létezik olyan $\delta > 0$, hogy ha $|x-t| < \delta$ akkor $|f(x) - f(t)| < \varepsilon$. A bizonyítás kulcsa, hogy ha $\delta > h > 0$, akkor[7]

$$\begin{aligned} \left|\frac{F(x+h) - F(x)}{h} - f(x)\right| &= \left|\frac{1}{h}\int_x^{x+h} f(t) - f(x)\,dt\right| \leq \\ &\leq \frac{1}{h}\int_x^{x+h} |f(t) - f(x)|\,dt \leq \frac{\varepsilon}{h}\int_x^{x+h} dx = \varepsilon, \end{aligned}$$

amelyből

$$\lim_{h\searrow 0} \frac{F(x+h) - F(x)}{h} = f(x).$$

Hasonló gondolatmenet érvényes, ha $h < 0$, ilyenkor az $\int_{x+h}^{x} f(t)\,dt$ integrált kell vizsgálni. Egy valós függvény pontosan akkor deriválható, ha jobbról és balról deriválható, és a két derivált megegyezik, ezért $F'(x) = f(x)$.

□

Nem lehet a tétel jelentőségét eléggé hangsúlyozni. Ha ismerjük az f antideriváltját, akkor ki tudjuk számolni az integrálját! Vegyük észre, hogy az F-nek, mindaddig amíg folytonos, nem kell minden pontban az f antideriváltjának lenni. Ha az F folytonos, és véges sok ponttól eltekintve az f antideriváltja, akkor is érvényes a képlet. Ha például a c kivételes pont, és az $[a,c]$ és a $[c,b]$ intervallumokon érvényes a formula, akkor

$$\begin{aligned} \int_a^b f(x)\,dx &= \int_a^c f(x)\,dx + \int_c^b f(x)\,dx = F(c) - F(a) + F(b) - F(c) = \\ &= F(b) - F(a). \end{aligned}$$

Tipikus példa az

$$F(x) \circeq |x|, \quad f(x) \circeq \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases}, \quad \int_{-1}^{1} f(x)\,dx = F(1) - F(-1) = 0.$$

Ha az F nem folytonos az $[a,b]$ intervallumon, mondjuk a c pontban szakadása van, de a c pontban van jobb és bal oldali határértéke, és az F egyébként az f antideriváltja, akkor

$$\begin{aligned} \int_a^b f(x)\,dx &= \int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \\ &= F(c-0) - F(a) + F(b) - F(c+0) = \\ &= F(b) - F(a) + F(c-0) - F(c+0). \end{aligned}$$

[7] Az egyetlen pont, ahol óvatosan kell eljárni, hogy vajon az abszolút értéket be lehet-e vinni az integrál mögé. Ha Riemann-integrálról van szó, akkor biztosan, de improprius integrál esetében is, legfeljebb az integrál végtelen lesz. Ez utóbbi most azonban nem állhat elő, mivel a folytonosság miatt az $[x, x+h]$ halmazon az f korlátos, az integrációs tartomány pedig véges, vagyis egyszerű Riemann-integrálról van szó.

4.1.9. Példa. Számoljuk ki az $\int_{-1}^{1} 1/\sqrt{|x|}dx$ és az $\int_{-1}^{1} 1/x^2 dx$ integrálokat!

Az

$$F(x) \stackrel{\circ}{=} \begin{cases} -2\sqrt{|x|} & \text{ha} \quad x < 0 \\ 2\sqrt{x} & \text{ha} \quad x \geq 0 \end{cases}$$

függvény a 0 ponttól eltekintve az f antideriváltja, és mivel az F folytonos, ezért

$$\int_{-1}^{1} \frac{1}{\sqrt{|x|}} dx = F(1) - F(-1) = 4.$$

Az $1/x^2$ antideriváltja $-1/x$, amely nincs definiálva a 0 pontban, és nem is terjeszthető ki oda folytonos módon, ezért a Newton–Leibniz-formula nem használható:

$$\int_{-1}^{1} \frac{1}{x^2} dx = 2\int_{0}^{1} \frac{1}{x^2} dx = \infty,$$

bár $[-1/x]_{-1}^{1} = 0$. Vegyük észre, hogy mivel a végtelen értékű integrálokat definíció szerint kizártuk, ezért egyidejüleg az integrál sem létezik.

□

4.2. Stieltjes-integrál

Miként már említettük, az elemi valószínűségszámításban gyakran felteszik, hogy két fajta eloszlás van: folytonos és diszkrét. Ennek megfelelően a várható érték két különböző definícióval rendelkezik. A várható értéket folytonos esetben integrállal, diszkrét esetben pedig összeggel szokás értelmezni. Ez a megközelítés eltakarja a várható érték fogalma mögött meghúzódó egységes koncepciót.

4.2.1. Definíció. Az I intervallumon értelmezett F függvényt súlyfüggvénynek mondjuk, ha az F az I intervallumon monoton nő.

Súlyfüggvényre legfontosabb példa valamely valószínűségi változó F eloszlásfüggvénye. A valószínűségszámításban időnként előforduló másik súlyfüggvény a változó

$$V(x) \stackrel{\circ}{=} 1 - F(x) + F(-x) = \mathbf{P}(\xi \geq x) + \mathbf{P}(\xi < -x)$$

úgynevezett farokeloszlása. A V monoton csökken, ezért ilyenkor integrátornak a $-V$ függvényt szokás választani. A Stieltjes-integrál tárgyalásakor a súlyfüggvényről elegendő csak annyit feltenni, hogy monoton nő, az eloszlásfüggvények egyéb tulajdonságaira, mint például a balról való folytonosságra, vagy a korlátosságra a Stieltjes-integrálás bevezetése során nem fogunk hivatkozni. Valójában még a monotonitás megkövetelése is szükségtelenül erős. A Stieltjes-integrálás elméletében elegendő feltenni, hogy az F súlyfüggvény két monoton függvény különbsége. Ha egy függvény két monoton függvény különbsége, akkor korlátos változásúnak szokás mondani. Ha

$F = F_1 - F_2$, akkor az F-re vonatkozó Stieltjes-integrál könnyen értelmezhető mint az F_1 és az F_2 szerinti integrálok különbsége. Mivel nem törekszünk a Stieltjes-integrál minél teljesebb bemutatására, így megelégszünk a monoton növekedő súlyfüggvényekre vonatkozó integrálás bevezetésével, és a korlátos változású súlyfüggvényekre csak érintőlegesen utalunk.

Stieltjes-integrálás véges szakaszokon

Legyenek $a < b$ valós számok és tekintsük az $[a,b]$ véges, zárt intervallum valamilyen

$$a = x_0 < x_1 < \ldots < x_n = b \tag{4.2.1}$$

felosztását. Legyen g az $[a,b]$ intervallumon értelmezett korlátos függvény, az F pedig legyen monoton növekedő. Tekintsük az

$$\begin{aligned} s &\overset{\circ}{=} \sum_{i=1}^{n} \inf_{x\in[x_{i-1},x_i]} g(x)\left[F(x_i) - F(x_{i-1})\right], \\ S &\overset{\circ}{=} \sum_{i=1}^{n} \sup_{x\in[x_{i-1},x_i]} g(x)\left[F(x_i) - F(x_{i-1})\right], \\ \sigma &\overset{\circ}{=} \sum_{i=1}^{n} g(\tau_i)\left[F(x_i) - F(x_{i-1})\right], \quad \tau_i \in [x_{i-1},x_i] \end{aligned}$$

alsó, felső és közbülső közelítő összegeket. Az F monoton nő, ezért

$$F(x_i) - F(x_{i-1}) \geq 0,$$

így $s \leq \sigma \leq S$. A korlátos és zárt $[a,b]$ szakaszon az F súlyfüggvény biztosan korlátos, a közelítő összegek végesek. Miként a Riemann-integrál esetén, most is könnyen igazolható, hogy ha s és S két különböző felosztáshoz tartozik, akkor az alapul vett felosztásoktól függetlenül $s \leq S$.

4.2.2. Definíció. Ha a felső közelítő összegek infimuma és az alsó közelítő összegek szuprémuma véges és megegyezik, akkor a g függvényt az F súlyfüggvényre nézve Riemann–Stieltjes-integrálhatónak mondjuk. A szuprémum és az infimum közös értékét a g F szerinti Riemann–Stieltjes-integráljának, vagy egyszerűen Stieltjes-integráljának mondjuk. Az integrált az

$$\int_a^b g(x)\, dF(x), \quad \int_a^b g dF \tag{4.2.2}$$

módon jelöljük[8]. Azt mondjuk, hogy az $\int_a^b g dF$ integrál előáll a közelítő összegek határértékeként, ha minden $\varepsilon > 0$ esetén létezik $\delta > 0$, hogy ha a fenti (4.2.1) felosztásban

[8] Az imént definiált (4.2.2) integrálban a g függvényre szokás az integrandus, az F súlyfüggvényre pedig az integrátor elnevezést is használni.

$\max_k \Delta x_k < \delta$, akkor a τ_k közbülső pontok választásától függetlenül

$$\left| \int_a^b g dF - \sum_{k=1}^{n} g(\tau_k) \Delta F(x_k) \right| < \varepsilon,$$

ahol értelemszerűen

$$\Delta F(x_k) \overset{\circ}{=} F(x_k) - F(x_{k-1}).$$

4.2.3. Állítás. *Ha g az $[a,b]$ véges intervallumon folytonos, akkor tetszőleges F monoton növekedő súlyfüggvény esetén az $\int_a^b g dF$ integrál létezik, és előáll a közelítő összegek határértékeként.*

Bizonyítás: A feltétel szerint g az $[a,b]$ halmazon folytonos, ezért ott egyenletesen is folytonos, és így tetszőleges $\varepsilon > 0$-hoz létezik olyan $\delta > 0$, hogy ha $|x' - x''| < \delta$, akkor[9] $|g(x') - g(x'')| < \varepsilon / [F(b) - F(a)]$. Ha (4.2.1) az $[a,b]$ szakasz olyan felosztása, ahol az osztópontok távolsága kisebb mint δ, és s az alsó és S a felső közelítő összeg, akkor az egyenletes folytonosság miatt

$$\begin{aligned} 0 &\leq S - s = \\ &= \sum_{k=1}^{n} \left(\sup_{t\in[x_{i-1},x_i]} g(t) - \inf_{t\in[x_{i-1},x_i]} g(t) \right) [F(x_k) - F(x_{k-1})] \leq \varepsilon, \end{aligned}$$

vagyis a g integrálható, és $\int_a^b g dF \in [s,S]$. Ha σ egy δ-nál finomabb felosztáshoz tartozó tetszőleges közelítő összeg, akkor

$$\int_a^b g dF - \varepsilon \leq s \leq \sigma \leq S \leq \int_a^b g dF + \varepsilon,$$

vagyis az integrál előáll mint a közelítő összegek határértéke.

□

4.2.4. Következmény. *Ha g az $[a,b]$ véges intervallumon folytonos, akkor az $\int_a^b g dF$ integrál értéke független az F súlyfüggvény (a,b) intervallumban levő szakadási pontjaiban felvett értékeitől.*

Bizonyítás: Mivel az F monoton nő, ezért legfeljebb megszámlálható sok szakadási pontja lehet. Így a folytonossági pontok sűrűen vannak a számegyenesen. Ennek következtében a közelítő összegek választhatók úgy, hogy minden osztópont folytonossági pont. Mivel az integrál a közelítő összegek határértéke, ezért az integrál csak a súlyfüggvény folytonossági pontokban felvett értékétől függ.

□

[9] Feltehetjük, hogy $F(b) - F(a) > 0$, hiszen ellenkező esetben az F monotonitása miatt az F konstans és ezért a g nyilván integrálható az F-re nézve, ugyanis minden alsó és felső közelítő összeg nulla.

4.2.5. Következmény. *Ha f és g az $[a,b]$ intervallumon folytonos függvények, és F tetszőleges monoton növekedő súlyfüggvény, akkor fennállnak a következők:*

1. $\left|\int_a^b g dF\right| \leq \int_a^b |g| dF.$

2. *Az integrál lineáris funkcionál*[10]*:*

$$\begin{aligned}\int_a^b f dF + \int_a^b g dF &= \int_a^b (f+g) dF, \\ \int_a^b \lambda g dF &= \lambda \int_a^b g dF, \quad \lambda \in \mathbb{R}.\end{aligned}$$

3. *Ha* $g \geq 0$, *akkor* $\int_a^b g dF \geq 0$, *tehát az integrál nem negatív funkcionál.*

4. *Az integrál az integrációs tartomány additív függvénye, vagyis*

$$\int_a^b g dF = \int_a^c g dF + \int_c^b g dF.$$

5. *Ha* $u(x)$ *egy szigorúan monoton növekedő, folytonos függvény, akkor*

$$\int_a^b g(u(x)) dF(x) = \int_{u(a)}^{u(b)} g(y) dG(y), \tag{4.2.3}$$

ahol $G(y) \triangleq F\left(u^{-1}(y)\right)$.

6. *Érvényes a parciális integrálás következő formulája: Ha f és g a folytonosságon kívül még monoton növekedőek*[11] *is, akkor*

$$[fg]_a^b = f(b)g(b) - f(a)g(a) = \int_a^b f dg + \int_a^b g df.$$

Bizonyítás: A g folytonossága miatt a $|g|$ is folytonos, és ezért integrálható. Mivel az F súlyfüggvény monoton nő, ezért a közelítő összegekre

$$\left|\sum_{k=1}^n g(\tau_k)\left[F(x_k) - F(x_{k-1})\right]\right| \leq \sum_{k=1}^n |g(\tau_k)| \left|F(x_k) - F(x_{k-1})\right|,$$

így mivel folytonos függvényekre az integrál a közelítő összegek határértéke, ezért $\left|\int_a^b g dF\right| \leq \int_a^b |g| dF$. Ha a g folytonos, akkor a λg is folytonos, tehát integrálható.

[10] Az integrál a súlyfüggvények szerint is additív, de erre nem lesz szükségünk.

[11] Illetve általában f és g folytonos és korlátos változású.

Mivel folytonos függvények integrálja előállítható a közelítő összegek határértékeként, ezért

$$\begin{aligned}\int_a^b \lambda g dF &= \lim_{n\to\infty}\sum_k \left(\lambda g\left(\tau_k^{(n)}\right)\right)\Delta F\left(x_k^{(n)}\right) = \\ &= \lambda \lim_{n\to\infty}\sum_k g\left(\tau_k^{(n)}\right)\Delta F\left(x_k^{(n)}\right) = \lambda\int_a^b g dF.\end{aligned}$$

Az összeg integrálására vonatkozó állítás bizonyítás analóg módon végezhető el. Legyen c egy közbülső pont. Mivel folytonos függvény leszűkítése is folytonos, ezért az $\int_a^c g dF$, $\int_c^b g dF$ integrálok léteznek. Tekintsük az $[a,b]$ intervallumhoz tartozó $\sum_{k=1}^n g(\tau_k)[F(x_k)-F(x_{k-1})]$ közelítő összeget. Mivel a közelítő összegek a közbülső pontok választásától függetlenül az integrálhoz tartanak, ezért feltehetjük, hogy csak olyan közelítő összegeket tekintünk, ahol a c pont része a felosztásnak, mondjuk $c = x_m$. Ilyenkor

$$\sum_{k=1}^n g(\tau_k)\Delta F(x_k) = \sum_{k=1}^m g(\tau_k)\Delta F(x_k) + \sum_{k=m+1}^n g(\tau_k)\Delta F(x_k),$$

amely összegek az $\int_a^c g dF + \int_c^b g dF$ közelítő összegei. A helyettesítéses integrálás formulájának igazolásához tekintsük a közelítő összegeket: Ha (x_i) az $[a,b]$ szakasz egy felosztása, akkor az $y_i \triangleq u(x_i)$ sorozat a $[u(a), u(b)]$ egy felbontása. Továbbá

$$\begin{aligned}&\sum_i g(u(x_{i-1}))(F(x_i)-F(x_{i-1})) = \\ &= \sum_i g(y_{i-1})\left(F\left(u^{-1}(y_i)\right) - F\left(u^{-1}(y_{i-1})\right)\right) = \\ &= \sum_i g(y_{i-1})(G(y_i)-G(y_{i-1})).\end{aligned}$$

Az u függvény egyenletesen folytonos, így ha $\max_i \Delta x_i \to 0$, akkor $\max \Delta y_i \to 0$, amiből, kihasználva, hogy a két integrál létezik és tetszőleges közelítő összegek határértéke, a fenti (4.2.3) formula már evidens. A többi állítás hasonlóan könnyen igazolható.

□

4.2.6. Következmény. *Ha az F függvény két monoton függvény különbsége*[12]*, vagyis ha $F = F_1 - F_2$, ahol F_1 és F_2 monotonon növekedő, és a g függvény folytonos, akkor az*

$$\int_a^b g dF \triangleq \int_a^b g dF_1 - \int_a^b g dF_2$$

[12] Vagyis, ha az F korlátos változású.

módon definiált integrál független az F két monoton növekedő függvény különbségére való felbontásától, és előáll a

$$\sigma_n \stackrel{\circ}{=} \sum_{k=1}^{n} g(\tau_k)\Delta F(x_k) \stackrel{\circ}{=} \sum_{k=1}^{n} g(\tau_k)(F(x_k) - F(x_{k-1})) \tag{4.2.4}$$

közelítő összegek határértékeként.

Bizonyítás: A g függvény folytonos, a (4.2.4) közelítő összeg konvergens, ugyanis

$$\sigma_n = \sum_{k=1}^{n} g(\tau_k)\Delta F_1(x_k) - \sum_{k=1}^{n} g(\tau_k)\Delta F_2(x_k),$$

és az összeg két tagja konvergens. Világos továbbá, hogy (σ_n) határértéke független az F súlyfüggvény F_1 és F_2 függvényekre való felbontásától.

□

4.2.7. Példa. Ha az integrátor súlyfüggvénynek és az integrandusnak közös szakadási pontja van, akkor az integrál nem feltétlenül létezik. Előfordulhat az is, hogy bár az integrál létezik, de nem lesz a közbülső közelítő összegek határértéke.

A legegyszerűbb példa, ha a g a 0 pontból álló halmaz karakterisztikus függvénye, és F a 0 pontban 1-et ugró súlyfüggvény. Ekkor az $\int_{-1}^{1} gdF$ integrál minden alsó közelítő összege 0, hiszen minden részintervallum tartalmaz olyan pontot, ahol a g nulla, ugyanakkor az összes felső közelítő összeg 1, hiszen minden felosztáshoz tartozik olyan a 0 pontot tartalmazó részintervallum, ahol a súlyfüggvény növekménye 1 és a g maximuma is 1. Másik példaként tekintsük a

$$g(x) \stackrel{\circ}{=} \begin{cases} 0 & \text{ha} \quad x < 0 \\ 1 & \text{ha} \quad x \geq 0 \end{cases}, \quad F(x) \stackrel{\circ}{=} \begin{cases} 0 & \text{ha} \quad x \leq 0 \\ 1 & \text{ha} \quad x > 0 \end{cases}$$

függvényeket, és számoljuk ki az $\int_{-1}^{1} gdF$ integrált! Világos, hogy az $x_0 = -1$, $x_1 = 0$, $x_2 = 1$ felosztásra $s = 0 \cdot 0 + 1 \cdot 1 = 1$ és $S = 1 \cdot 0 + 1 \cdot 1 = 1$, vagyis az integrál 1. De ugyanakkor minden olyan felosztásra, amelynek a 0 nem osztópontja, a közbülső pontok megfelelő választásával az integrál közelítő összeg vagy 0, vagy 1, és ezért az integrál nem tekinthető a tetszőlegesen választott közbülső közelítő összegek határértékének!

□

Improprius Stieltjes-integrál

Miként definiáljuk a Stieltjes-integrálokat ha az integrációs tartomány nem zárt, illetve ha korlátlan? Lényegében ugyanúgy, ahogy ezt már bemutattuk: az integrált improprius módon terjesztjük ki. Ha az $I_n \nearrow I$ és az I_n intervallumok véges zárt szakaszok, akkor

$$\int_I gdF \stackrel{\circ}{=} \lim_{n\to\infty} \int_{I_n} gdF,$$

ahol értelemszerűen megköveteljük, hogy a határérték nem függhet az $I_n \nearrow I$ sorozat megválasztásától. Értelemszerűen használni fogjuk az $\int_{-\infty}^{\infty} gdF$ és az $\int_{a+}^{b-} gdF$ és hasonló szimbólumokat. Az improprius kiterjesztés tárgyalásakor egyetlen dologra kell odafigyelni: Mivel az F súlyfüggvény nem biztos, hogy folytonos, a nyílt és a zárt intervallumokon vett integrálok értéke nem feltétlenül azonos. A Riemann-integrálra bemutatott bizonyítás értelemszerű módosításával belátható, hogy ha az a és a b véges pontok az F súlyfüggvény folytonossági pontjai és g folytonos az $[a,b]$ szakaszon, akkor $\int_a^b gdF = \int_{a+}^{b-} gdF$, vagyis ilyenkor az improprius és a valódi integrálok egybeesnek. Hasonlóan miként a Riemann-integrál esetén, ha az integrandus nem negatív és folytonos, akkor az improprius integrálok előállíthatók végtelen sok osztópontból álló alsó és felső közelítőösszegek közös szuprémumaként, illetve infimumaként, miközben az osztópontok választhatók az F folytonossági pontjainak[13].

4.2.8. Állítás. *Ha g a teljes számegyenesen korlátos és folytonos függvény, akkor tetszőleges F korlátos súlyfüggvényre létezik az $\int_{-\infty}^{\infty} gdF$ integrál.*

Bizonyítás: Mivel véges intervallumra létezik az integrál, ezért szimmetria okok miatt elegendő belátni, hogy a $\lim_{b\to\infty} \int_a^b gdF$ határérték létezik és véges. Mivel F monoton nő és korlátos, ezért létezik a $\lim_{x\to\infty} F(x) < \infty$ határérték. Ha K a g integrandus egy korlátja, akkor tetszőleges $\varepsilon > 0$ számhoz létezik olyan N, hogy ha $b'' > b' \geq N$, akkor

$$\begin{aligned} \left| \int_a^{b'} gdF - \int_a^{b''} gdF \right| &= \left| \int_{b'}^{b''} gdF \right| \leq \int_{b'}^{b''} |g| \, dF \leq \\ &\leq K \int_{b'}^{b''} dF = K \left[F\left(b''\right) - F\left(b'\right) \right] < \varepsilon, \end{aligned}$$

amiből a Cauchy-kritérium alapján a határérték létezik és véges, így az improprius integrál létezik.

□

Stieltjes-integrálok kiszámolása

Hogyan lehet kiszámolni egy Stieltjes-integrált? Erre vonatkozólag tekintsük a bevezetőben említett két esetet, amikor a ξ valószínűségi változónak van folytonos sűrűségfüggvénye illetve amikor ξ eloszlása diszkrét.

4.2.9. Állítás. *Ha a g folytonos függvényre létezik az $\int_I gdF$ integrál, továbbá az F folytonos és véges számú ponttól eltekintve folytonosan deriválható, és $f \stackrel{\circ}{=} F'$, akkor*

$$\int_I gdF = \int_I g(x) f(x) \, dx.$$

[13] Ez utóbbi kitétel nem tűnik fontosnak, de alapvetően azért lényeges, mert a várható érték definíciójában szereplő $\int_{-\infty}^{\infty} xdF(x)$ integrál ilymódon független attól, hogy az eloszlásfüggvényt a $\mathbf{P}(\xi < x)$ vagy $\mathbf{P}(\xi \leq x)$ módon definiáljuk.

Bizonyítás: Legyen először $I \circeq [a,b]$ korlátos, zárt intervallum, és tekintsük az $\int_a^b gdF$ integrál tetszőleges

$$\sum_{i=1}^{n} g(\varepsilon_i)[F(x_i) - F(x_{i-1})] \tag{4.2.5}$$

közelítő összegét. Mivel a Stieltjes-integrál az integrációs intervallum additív függvénye, ezért feltehető, hogy az (a,b) nyílt intervallumban F deriválható, így alkalmazható a középértéktétel. Ez alapján

$$F(x_i) - F(x_{i-1}) = f(\tau_i)(x_i - x_{i-1}), \quad x_{i-1} < \tau_i < x_i.$$

A (4.2.5) sorban az ε_i közbülső pontot τ_i-nek választva a fenti (4.2.5) közelítő összeg

$$\sigma_n = \sum_{i=1}^{n} g(\tau_i) f(\tau_i)(x_i - x_{i-1}).$$

Ha $\max_i \left(x_i^{(n)} - x_{i-1}^{(n)}\right) \to 0$, akkor f és g folytonossága alapján az alábbi számolás során az integrálok előállnak a közelítő összegek határértékeként, vagyis

$$\begin{aligned} \int_a^b gdF &= \lim_{n\to\infty} \sum_{i=1}^{n} g\left(\tau_i^{(n)}\right) \Delta F\left(x_i^{(n)}\right) = \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} g\left(\tau_i^{(n)}\right) f\left(\tau_i^{(n)}\right) \Delta x_i^{(n)} = \int_a^b gfdx. \end{aligned}$$

Ha például $b \to \infty$, akkor a feltétel alapján a $\lim_{b\to\infty} \int_a^b gdF \circeq \int_a^\infty gdF$ határérték létezik, tehát létezik a $\lim_{b\to\infty} \int_a^b gfdx$ határérték is, ami definíció szerint éppen $\int_a^\infty gfdx$.

□

4.2.10. Példa. Számoljuk ki az $\int_0^1 x^2 dx^2$ integrált!

Az elmondottak szerint

$$\int_0^1 x^2 dx^2 = \int_0^1 x^2 2x dx = 2\int_0^1 x^3 dx = \frac{2}{4} = \frac{1}{2}.$$

Nem érdektelen kiszámolni az integrált $y = x^2$ helyettesítéssel. Az $\left(\sqrt{y}\right)^2 = y$ felhasználásával $\int_0^1 x^2 dx^2 = \int_0^1 ydy = 1/2$. De használhatjuk a parciális integrálás formuláját is. Ez alapján

$$2\int_0^1 x^2 dx^2 = \left[x^2 x^2\right]_0^1 = 1.$$

□

Mielőtt folytatnánk, érdemes tisztázni, hogy mit értünk diszkrét eloszláson. Általános esetben egy ξ változót akkor mondunk diszkrét eloszlásúnak, ha van olyan (x_k) legfeljebb megszámlálható halmaz, hogy a $\{\xi = x_k\}$ halmazok egy teljes eseményrendszer.

Az egyszerűség kedvéért azonban egy legfeljebb megszámlálható halmazra koncentrálódó eloszlást csak akkor fogunk diszkrét eloszlásnak mondani, ha tetszőleges $[a,b]$ szakasz esetén csak véges számú olyan $t_k \in [a,b]$ pont van, amely valószínűsége pozitív[14]. Ha a t_k ponthoz tartozó $\mathbf{P}(\xi = t_k)$ valószínűség p_k, akkor nyilván $\Sigma_k p_k = 1$. A (t_k) pontokra támaszkodó diszkrét eloszlás eloszlásfüggvénye $F(x) = \sum_{t_k<x} p_k$.

4.2.11. Állítás. *Ha F egy diszkrét eloszlás eloszlásfüggvénye, akkor minden g folytonos függvény esetén tetszőleges I intervallumra*

$$\int_I gdF = \sum_{t_k \in I} g(t_k)(F(t_k+0) - F(t_k)) = \sum_{t_k \in I} g(t_k) p_k.$$

Bizonyítás: Először tegyük fel, hogy $I = [a,b]$. Ilyenkor az ugrások száma véges. Tekintsük az $[a,b]$ egy tetszőleges $a = x_0, \ldots, x_n = b$ felosztását. Ha most valamilyen részintervallumba nem esik t_i ugrópont, akkor a megfelelő tag nulla, hiszen $F(x_i) - F(x_{i-1}) = 0$. Feltehető, hogy a felosztás már olyan finom, hogy minden t_i egyetlen intervallum belsejében szerepel. A $\tau_i = t_i$ pontot közbülső pontnak választva a közelítő összeg a felbontástól függetlenül $\sigma \mathrel{\hat{=}} \sum_k g(t_k) p_k$, vagyis ilyenkor az állítás érvényes. Az általános eset, az a,b végpontok szerinti határértékkel kapható.

□

4.2.12. Példa. A sorok összege mint Stieltjes-integrál.

Ha a $\sum_{k=0}^{\infty} a_k$ sor konvergens, és

$$F(x) \mathrel{\hat{=}} \begin{cases} 0 & \text{ha} \quad x \leq 0 \\ n & \text{ha} \quad x \in (n-1, n] \end{cases},$$

akkor analóg módon érvelve $\sum_{k=0}^{\infty} a_k = \int_0^{\infty} gdF$, ahol g tetszőleges olyan folytonos „interpolációs" függvény, amelyre $a_k = g(k)$.

□

[14] Legyen (r_k) a racionális számok egy felsorolása. Ha az r_k súlya 2^{-k}, akkor ez egy megszámlálható halmazra koncentrálódó eloszlás, amit most nem tekintünk diszkrét eloszlásnak.

V.

A VÁRHATÓ ÉRTÉK

A várható érték, az eloszlás mellett, tulajdonképpen a valószínűségszámítás legfontosabb fogalma. Segítségével definiálhatjuk a szórást, illetve a korrelációs együtthatót. A tényleges gyakorlatban legtöbbször az eloszlás pontos ismerete illúzió és az alkalmazásokban megelégszünk a várható érték és a szórás ismeretével. A várható érték mellett időnként szükségünk van az eloszlások momentumaira. A momentumok segítségével definiálhatjuk a ferdeséget és a laposságot. Ugyancsak a várható érték segítségével definiáljuk a különböző függvénytranszformáltakat mint például a generátorfüggvényt, a momentumgeneráló függvényt, illetve a karakterisztikus függvényt.

5.1. A várható érték definíciója

Ha az eloszlás diszkrét, akkor a várható érték definíciója viszonylag egyszerű és igen természetes.

5.1.1. Definíció. Ha a ξ valószínűségi változó eloszlása diszkrét és a t_k lehetséges értékek felvételének valószínűsége p_k, akkor a ξ változó $\mathbf{E}(\xi)$ módon jelölt várható értékét az

$$\mathbf{E}(\xi) \circeq \sum_k t_k \mathbf{P}(\xi = t_k) = \sum_k t_k p_k$$

véges, vagy végtelen összeggel definiáljuk. A várható értéket csak akkor tekintjük definiáltnak, ha az összeg abszolút konvergens, vagyis az összeg értéke független attól, hogy miként, milyen sorrendben indexeltük a t_k pontokat.

Az abszolút konvergencia megkövetelése igen természetes, ugyanis amennyiben egy konvergens sor nem abszolút konvergens, akkor feltételesen konvergens, így az összeg értéke függ az összegzés sorrendjétől, vagyis ilyenkor az $\mathbf{E}(\xi)$ érték nem lenne az eloszlás által egyértelműen adott. Az abszolút konvergencia miatt az összeg értéke úgy is meghatározható, hogy előbb kiszámoljuk a pozitív tagok összegét, majd ehhez hozzáadjuk a negatív tagok összegét. Emlékeztetünk, hogy $x^+ \circeq \max(0,x)$ és $x^- = \max(0,-x) \circeq -\min(0,x)$. Ezzel a jelöléssel[1]

$$\mathbf{E}(\xi) = \sum_k t_k^+ p_k - \sum_k t_k^- p_k = \mathbf{E}\left(\xi^+\right) - \mathbf{E}\left(\xi^-\right).$$

Nem negatív tagú sorok összege a sorrendtől függetlenül definiálható, feltételesen konvergens sorok esetén azonban a negatív és a pozitív tagok összege külön-külön végtelen, így ilyenkor az $\mathbf{E}(\xi)$ értékét egy $\infty - \infty$ típusú kifejezéssel kellene definiálni. Időnként meg szokás engedni, hogy a várható érték nagysága végtelen legyen. Ezen értelemszerűen azt értjük, hogy a pozitív vagy a negatív rész összege végtelen, miközben a másik rész értéke véges. Mi azonban ezt nem engedjük meg és a várható értéket csak

[1] A figyelmes olvasó észreveheti, hogy a $t = 0$ pont mind a két összegben szerepel, de ennek nincs jelentősége, ugyanis mind a két alkalommal nulla értékkel.

akkor tekintjük értelmesnek, ha az értéke véges. A Stieltjes-integrálokról elmondottak alapján evidens, hogy diszkrét eloszlásokra

$$\mathbf{E}(\xi) = \int_{-\infty}^{\infty} x dF(x).$$

Hogyan definiáljuk nem diszkrét eloszlások várható értékét? Kézenfekvő módon az előző sort mint definíciót tekintjük.

5.1.2. Definíció. Tetszőleges $F(x)$ eloszlásfüggvénnyel rendelkező ξ valószínűségi változó esetén

$$\mathbf{E}(\xi) \stackrel{\circ}{=} \int_{-\infty}^{\infty} x dF(x)$$

feltéve, hogy az integrál abszolút konvergens, vagyis

$$\int_{-\infty}^{\infty} |x| \, dF(x) < \infty.$$

Érdemes egy rövid kitérőt tenni, és röviden áttekinteni a definíció tartalmát. Először is vegyük észre, hogy a várható érték, mint minden Stieltjes-integrál értéke, nem függ az F súlyfüggvény szakadási pontjaiban felvett értékétől. Így a várható érték független attól, hogy az eloszlásfüggvényt a $\mathbf{P}(\xi < x)$, vagy például a $\mathbf{P}(\xi \leq x)$ módon definiáltuk. A Stieltjes-integrál additivitása miatt

$$\int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{0} x dF(x) + \int_{0}^{\infty} x dF(x).$$

Emlékeztetünk, hogy a ξ^+ eloszlásfüggvénye

$$G(x) = \mathbf{P}\left(\xi^+ < x\right) = \begin{cases} 0 & \text{ha} \quad x \leq 0 \\ F(x) & \text{ha} \quad x > 0 \end{cases}$$

és a ξ^- eloszlásfüggvénye

$$H(x) = \mathbf{P}\left(\xi^- < x\right) = \begin{cases} 0 & \text{ha} \quad x \leq 0 \\ 1 - F(-x+0) & \text{ha} \quad x > 0 \end{cases}.$$

A közelítő összegekbe az $x = 0$ nulla értéket belevéve, illetve kihasználva, hogy az integrál értéke nem függ a közelítő pontok választásától, valamint azt, hogy folytonos integrandus esetén a Stieltjes-integrál létezik, egyszerűen belátható, hogy

$$\mathbf{E}\left(\xi^+\right) = \int_{-\infty}^{\infty} x dG(x) = \int_{0}^{\infty} x dF(x).$$

Ismételten mivel az integrál értéke nem függ a szakadási pontokban felvett értéktől, ezért

$$\begin{aligned}\int_{-\infty}^{0} xdF(x) &= \int_{-\infty}^{0} xdF(x+0) = -\int_{-\infty}^{0} -xdF(x+0) = \\ &= -\int_{0}^{\infty} xd(-F(-x+0)) = -\int_{0}^{\infty} xd(1-F(-x+0)) = \\ &= -\int_{0}^{\infty} xdH(x) = -\mathbf{E}\left(\xi^{-}\right).\end{aligned}$$

Vegyük észre, hogy némiképpen pontatlanul jártunk el, ugyanis csak az integrációs tartomány belső pontjaiban változtathatjuk meg a szakadási pontokban az integrátor értékét, de ez most sem okoz gondot, ugyanis miként a ξ^{+} esetén az $x=0$ pontot a közelítő pontnak választva, a számolásban szereplő integrálok értéke megegyezik.

5.1.3. Állítás. *Ezzel beláttuk a következő egyenlőségeket:*

$$\int_{0}^{\infty} xdF(x) = \mathbf{E}\left(\xi^{+}\right), \int_{-\infty}^{0} xdF(x) = -\mathbf{E}\left(\xi^{-}\right),$$
$$\mathbf{E}(\xi) = \mathbf{E}\left(\xi^{+}\right) - \mathbf{E}\left(\xi^{-}\right).$$

Mivel $\mathbf{E}(\xi) = \mathbf{E}\left(\xi^{+}\right) - \mathbf{E}\left(\xi^{-}\right)$, ezért elég a $\xi^{+}, \xi^{-} \geq 0$ nem negatív változók várható értékének interpretálására koncentrálni. Legyen tehát $\xi \geq 0$. Tekintsük a $[0,\infty)$ félegyenes

$$0 = x_0 < x_1 < x_2 < \ldots < x_k < \ldots$$

sorozattal megadott felbontását, és tekintsük az Ω téren definiált

$$A_k \circeq \left\{\xi^{-1}\left([x_k, x_{k+1})\right)\right\} \circeq \left\{\omega \mid \xi(\omega) \in [x_k, x_{k+1})\right\}$$

eseményeket. Az $A_0, A_1, \ldots$ események egy teljes eseményrendszert alkotnak, ugyanis diszjunktak, és az egyesítésük kiadja az Ω alaphalmazt. Nyilván a $\xi^{(a)} \circeq \sum_k x_k \chi_{A_k}$ és $\xi^{(f)} \circeq \sum_k x_{k+1} \chi_{A_k}$ függvényekre $\xi^{(a)} \leq \xi \leq \xi^{(f)}$. Az is nyilvánvaló, hogy ha az

$$0 = y_0 < y_1 < y_2 < \ldots$$

az előző (x_k) sorozat egy finomítása, vagyis további osztópontok hozzáadásával képződik, akkor az alsó közelítő összeg nem csökken, illetve a felső közelítő összeg nem nő. Így felmerül, hogy a ξ várható értékét megegyezik-e az

$$\mathbf{E}\left(\xi^{(a)}\right) = \sum_k x_k^{(n)} \mathbf{P}\left(A_k^{(n)}\right) = \sum_k x_k^{(n)} \mathbf{P}\left(\xi \in \left[x_k^{(n)}, x_{k+1}^{(n)}\right)\right)$$

lehetséges alsó közelítő összegek szuprémumával és a

$$\mathbf{E}\left(\xi^{(f)}\right) = \sum_k x_{k+1}^{(n)} \mathbf{P}\left(A_k^{(n)}\right) = \sum_k x_{k+1}^{(n)} \mathbf{P}\left(\xi \in \left[x_k^{(n)}, x_{k+1}^{(n)}\right)\right)$$

lehetséges felső közelítő összegek infimumával. Ez a nem negatív függvények Stieltjes-integráljára az előző fejezetben bemutatott tulajdonság miatt evidens, ugyanis az eloszlásfüggvény definíciója szerint

$$\mathbf{P}(\xi \in [x_k, x_{k+1})) = \mathbf{P}(x_k \leq \xi < x_{k+1}) = F(x_{k+1}) - F(x_k),$$

tehát

$$\begin{aligned} \mathbf{E}\left(\xi^{(a)}\right) &= \sum_k x_k \left(F(x_{k+1}) - F(x_k)\right). \\ \mathbf{E}\left(\xi^{(f)}\right) &= \sum_k x_{k+1} \left(F(x_{k+1}) - F(x_k)\right). \end{aligned}$$

és így

$$\sup \mathbf{E}\left(\xi^{(a)}\right) = \inf \mathbf{E}\left(\xi^{(f)}\right) = \int_0^\infty x dF(x) = \int_{-\infty}^\infty x dF(x),$$

ahol az utolsó egyenlőség azért teljesül, mert $\xi \geq 0$.

5.1.4. Tétel. *A várható értékre teljesülnek a következőek*[2]*:*

1. *A várható érték lineáris funkcionál*[3]*, vagyis tetszőleges a és b konstansok esetén $\mathbf{E}(a\xi + b\eta) = a\mathbf{E}(\xi) + b\mathbf{E}(\eta)$, feltéve, hogy a ξ és az η változóknak létezik véges várható értéke.*

2. *Ha $\xi \geq 0$, akkor $\mathbf{E}(\xi) \geq 0$. Vagyis a várható érték nem negatív funkcionál. Ebből ha a ξ és az η változóknak létezik véges várható értéke és $\xi \geq \eta$, akkor $\mathbf{E}(\xi) \geq \mathbf{E}(\eta)$. Ennek megfelelően szokás azt is mondani, hogy a várható érték monoton funkcionál.*

3. *Tetszőleges $A \in \mathscr{A}$ esetén $\mathbf{E}(\chi_A) = \mathbf{P}(A)$, speciálisan $\mathbf{E}(\chi_\Omega) = \mathbf{E}(1) = 1$, ahol a várható érték mögött az 1 szimbólum az azonosan 1 konstans függvényt jelöli és*

$$\chi_A(\omega) \stackrel{\circ}{=} \begin{cases} 1 & ha \quad \omega \in A \\ 0 & ha \quad \omega \notin A \end{cases}$$

az A esemény karakterisztikus, vagy más néven indikátorfüggvénye[4].

---

[2] Az alábbi tulajdonságok kimondásakor kihangsúlyoztuk, hogy a várható értéknek végesnek kell lenni. A továbbiakban azonban a várható értékről mindig automatikusan feltesszük, hogy véges. Például, ha a $\xi \geq 0$ várható értéke végtelen, akkor $0 = \mathbf{E}(\xi - \xi) = \infty - \infty$ egyenlőség értelmetlen.

[3] Funkcionálon általában függvényekhez számokat rendelő leképezéseket szokás érteni.

[4] A χ_A elnevezése körüli bonyodalom forrása, hogy karakterisztikus függvényen szokás a később tárgyalt $\varphi(t) = \mathbf{E}(\exp(it\xi))$ függvényt is érteni. Ugyanakkor a φ nem egy eseményhez, hanem egy valószínűségi változóhoz, pontosabban annak eloszlásához tartozik, így nem jelent félreértést, ha egy esemény karakterisztikus függvényéről beszélünk.

Érdemes hangsúlyozni, hogy a tulajdonságok igazolása távolról sem egyszerű. Például a linearitás igazolásához ismerni kellene a $\xi+\eta$ eloszlásfüggvényét, amelyet, legalábbis a tárgyalás jelen szintjén, nem ismerünk. Az elemi valószínűségszámításban[5] a várható érték kezelése némiképpen elnagyolt és ezen a Stieltjes-integrál bevezetése nem segít.

5.1.5. Példa. Diszkrét valószínűségi változók összegének várható értéke.

Legyenek ξ és η diszkrét valószínűségi változók. A ξ eloszlását adja meg az (x_k, p_k) sorozat, az η eloszlását pedig adja meg az (y_k, q_k) sorozat. Ha az $\mathbf{E}(\xi)$ és $\mathbf{E}(\eta)$ várható értékek léteznek, akkor a $\Sigma_k x_k p_k$ és a $\Sigma_k y_k q_k$ sorozatok abszolút konvergensek. Tekintsük a $\zeta \stackrel{\circ}{=} \xi+\eta$ összeget. A ζ lehetséges értékei $u_{ij} \stackrel{\circ}{=} x_i + y_j$ alakú számok, amelyek között természetesen ismétlődések is előfordulhatnak, a lehetséges legfeljebb megszámlálható különböző értékeket jelölje $z_k, k \in \mathbb{N}$. Ahhoz, hogy a ζ várható értékét meg tudjuk adni ismerni kell a ζ eloszlását, ami a „legrosszabb” esetben azt jelenti, hogy tudni kell a u_{ij} számokhoz tartozó

$$r_{ij} \stackrel{\circ}{=} \mathbf{P}\left(\xi = x_i, \eta = y_j\right)$$

együttes eloszlást, illetve a

$$\mathbf{P}(\zeta = z_k) = \sum_{z_k = x_i + y_j} r_{ij}$$

valószínűségeket, ahol értelemszerűen arról van szó, hogy azon (i, j) indexekre kell az összegzést elvégezni, amelyekre $z_k = x_i + y_j$. A peremeloszlásokra nyilván érvényesek a $p_i = \Sigma_j r_{ij}$ és a $q_j = \Sigma_i r_{ij}$ egyenlőségek.

$$\begin{aligned}
\mathbf{E}(\xi)+\mathbf{E}(\eta) &= \sum_i x_i p_i + \sum_j y_j q_j = \\
&= \sum_i x_i \left(\sum_j r_{ij}\right) + \sum_j y_j \left(\sum_i r_{ij}\right) = \\
&= \sum_i \sum_j x_i r_{ij} + \sum_j \sum_i y_j r_{ij} = \\
&= \sum_i \sum_j x_i r_{ij} + \sum_i \sum_j y_j r_{ij} = \\
&= \sum_i \sum_j \left(x_i + y_j\right) r_{ij} = \\
&= \sum_k z_k \sum_{z_k = x_i + y_j} r_{ij} = \\
&= \sum_k z_k \mathbf{P}(\xi+\eta = z_k) = \mathbf{E}(\xi+\eta).
\end{aligned}$$

[5] Elemi valószínűségszámításról általában akkor beszélünk, ha nem jelenik meg benne a Kolmogorov-féle, azaz a mértékelméletre alapozott valószínűségszámítási modell.

A bizonyítás kulcsa, hogy egyrészt a második kettős összegben az összegzés sorrendjét felcseréltük, illetve az utolsó előtti sorban az összegzés sorrendjét megváltoztattuk, és az összeget a $\xi+\eta$ által felvehető z_k értékek szerint csoportosítottuk. Ezt csak akkor tehetjük meg, ha az összegzés abszolút konvergens, ami teljesül, ugyanis a várható érték definíciója szerint például

$$\sum_j \sum_i |y_j| r_{ij} = \sum_j |y_j| \sum_i r_{ij} = \sum_j |y_j| q_j < \infty.$$

Hasonlóan látható, hogy a $\sum_i \sum_j \left(x_i + y_j\right) r_{ij}$ összeg is abszolút konvergens, így az összegzésben a tagok tetszőleges módon csoportosíthatóak és zárójelezhetőek, így az azonos értékű $z_k = x_i + y_j$ tagokhoz tartozó valószínűségek összevonhatók.

□

5.1.6. Példa. Diszkrét eloszlások konvolúciója.

Az előző példa speciális eseteként tegyük fel, hogy az x_i és y_j egész számok. Ilyenkor az $x_i + y_j$ értékek is egészek. Valamely k egészhez tartozó valószínűségek összege

$$u_k \stackrel{\circ}{=} \mathbf{P}(\xi + \eta = k) = \sum_j r_{jk-j}.$$

Ha a ξ és az η függetlenek, akkor $r_{ij} = p_i q_j$, és ilyenkor

$$u_k = \sum_j p_j q_{k-j},$$

amely kifejezésből álló sorozatot a (p_i) és $\left(q_j\right)$ sorozatok konvolúciójának mondjuk. Ha i és j tetszőleges egész lehet, akkor

$$u_k = \sum_{j=-\infty}^{\infty} p_j q_{k-j},$$

ha pedig az i és a j nem negatív egészek, akkor

$$u_k = \sum_{j=0}^{k} p_j q_{k-j}.$$

□

5.2. Szórás és momentumok

A várható érték segítségével számos, a valószínűségszámításban fontos fogalom definiálható. A hosszú sorban az első a momentumok fogalma.

5.2.1. Definíció. Ha ξ valószínűségi változó, és $n > 0$ tetszőleges egész szám, akkor az $\mathbf{E}(\xi^n)$ várható értékét a ξ változóhoz tartozó eloszlás n-edik momentumának mondjuk. Természetesen egy változó, illetve eloszlás momentumáról csak akkor beszélhetünk, ha az őt definiáló várható érték véges.

A második momentum segítségével definiáljuk a szórást:

5.2.2. Definíció. Ha ξ valószínűségi változó, akkor a

$$\mathbf{D}(\xi) \stackrel{\circ}{=} \sqrt{\mathbf{E}(\xi - \mathbf{E}(\xi))^2}$$

számot a ξ szórásának mondjuk. Feltéve persze, hogy a kifejezésben minden értelmes. A $\mathbf{D}^2(\xi)$ kifejezést a ξ varianciájának mondjuk.

Mivel a várható érték egy lineáris funkcionál, ezért érvényes a szórás kiszámolását nagyban segítő következő összefüggés:

$$\begin{aligned}
\mathbf{D}^2(\xi) &\stackrel{\circ}{=} \mathbf{E}(\xi - \mathbf{E}(\xi))^2 = \mathbf{E}\left(\xi^2 - 2\mathbf{E}(\xi)\xi + \mathbf{E}^2(\xi)\right) = \\
&= \mathbf{E}\left(\xi^2\right) - 2\mathbf{E}(\mathbf{E}(\xi)\xi) + \mathbf{E}\left(\mathbf{E}^2(\xi)\right) = \\
&= \mathbf{E}\left(\xi^2\right) - 2\mathbf{E}(\xi)\mathbf{E}(\xi) + \mathbf{E}^2(\xi)\mathbf{E}(1) = \\
&= \mathbf{E}\left(\xi^2\right) - 2\mathbf{E}^2(\xi) + \mathbf{E}^2(\xi) = \mathbf{E}\left(\xi^2\right) - \mathbf{E}^2(\xi).
\end{aligned}$$

Diszkrét eloszlások várható értéke és szórása

A következőkben néhány diszkrét eloszlás várható értékét és szórását számoljuk ki. Ezen a ponton az eloszlások hátterét vagy tulajdonságait még nem vizsgáljuk.

5.2.3. Példa. Számoljuk ki a Poisson-eloszlás várható értékét és szórását.

A Poisson-eloszlás a $k = 0, 1, 2, \ldots$ pontokra koncentrálódik. Az eloszlás definíciója szerint

$$p_k \stackrel{\circ}{=} \mathbf{P}(\xi = k) \stackrel{\circ}{=} \frac{\lambda^k}{k!}\exp(-\lambda),$$

ahol $\lambda > 0$ az eloszlás paramétere.

1. Először vizsgáljuk meg, hogy valóban eloszlásról van-e szó. Mivel $\lambda > 0$, ezért minden k indexre $p_k > 0$. Ugyanakkor az exponenciális függvény hatványsora miatt

$$\begin{aligned}
\sum_{k=0}^{\infty} p_k &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\exp(-\lambda) = \exp(-\lambda)\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = \\
&= \exp(-\lambda)\exp(\lambda) = 1,
\end{aligned}$$

vagyis valóban valószínűségeloszlásról van szó.

2. Számoljuk ki a várható értéket. Definíció szerint, felhasználva, hogy $x_k = k$,

$$\begin{aligned}
\mathbf{E}(\xi) &= \sum_{k=0}^{\infty} x_k p_k = \exp(-\lambda) \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} = \\
&= \exp(-\lambda) \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = \exp(-\lambda) \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \\
&= \frac{\lambda}{\exp(\lambda)} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \frac{\lambda}{\exp(\lambda)} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \\
&= \frac{\lambda}{\exp(\lambda)} \exp(\lambda) = \lambda.
\end{aligned}$$

3. A szórás kiszámolásához számoljuk ki a második momentumot.

$$\begin{aligned}
\mathbf{E}\left(\xi^2\right) &= \sum_{k=0}^{\infty} x_k^2 p_k = \exp(-\lambda) \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} = \\
&= \exp(-\lambda) \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} = \exp(-\lambda) \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} = \\
&= \frac{\lambda}{\exp(\lambda)} \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!} = \frac{\lambda}{\exp(\lambda)} \sum_{k=0}^{\infty} (k+1) \frac{\lambda^k}{k!} = \\
&= \frac{\lambda}{\exp(\lambda)} \left(\sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right) = \lambda^2 + \lambda,
\end{aligned}$$

ahol az utolsó sorban kihasználtuk a várható érték már ismert értékét.

4. Végül a szórás értéke

$$\mathbf{D}(\xi) = \sqrt{\mathbf{E}\left(\xi^2\right) - \mathbf{E}^2(\xi)} = \sqrt{\lambda^2 + \lambda - \lambda^2} = \sqrt{\lambda}.$$

5. Legyenek ξ és η független λ és μ paraméterű Poisson-eloszlású valószínűségi változók. Ekkor a $\xi + \eta$ eloszlása szintén Poisson-eloszlású lesz $\lambda + \mu$ paraméterrel:

$$\begin{aligned}
\mathbf{P}(\xi + \eta = n) &= \sum_{k=0}^{n} \frac{\lambda^k}{k!} \exp(-\lambda) \frac{\mu^{n-k}}{(n-k)!} \exp(-\mu) = \\
&= \frac{\exp(-(\lambda+\mu))}{n!} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} \lambda^k \mu^{n-k} = \\
&= \frac{\exp(-(\lambda+\mu))}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda^k \mu^{n-k} = \\
&= \frac{\exp(-(\lambda+\mu))}{n!} (\lambda+\mu)^n.
\end{aligned}$$

Az egyenlőségnek van egy fontos következménye, nevezetesen, hogy a Poisson-eloszlás úgymond korlátlanul osztható. Ezen azt értjük, hogy ha ξ egy λ paraméterű Poisson-eloszlású változó, akkor tetszőleges n esetén megadhatóak $\xi_1, \xi_2, \ldots, \xi_n$ azonos eloszlású és független változók, hogy $\xi = \Sigma_{k=1}^{n} \xi_k$. A Poisson-eloszlás esetén a ξ_k változók választhatók szintén Poisson-eloszlásúnak λ/n paraméterrel.

$\square$

5.2.4. Példa. Számoljuk ki a geometriai eloszlás várható értékét és szórását.

A geometriai eloszlás a $k = 1, 2, 3, \ldots$ pontokra koncentrált eloszlás. Definíció szerint $p_k = pq^{k-1}$, ahol $0 < p < 1$ és $q \circeq 1 - p$.

1. Első lépésként belátjuk, hogy valóban eloszlásról van szó. Mivel nyilvánvalóan $p_k > 0$, ezért elég belátni, hogy $\sum_k p_k = 1$. A geometriai sor összegképlete alapján[6]

$$\begin{aligned} \sum_k p_k &= \sum_{k=1}^{\infty} pq^{k-1} = p \sum_{k=1}^{\infty} q^{k-1} = p \sum_{k=0}^{\infty} q^k = \\ &= \frac{p}{1-q} = 1. \end{aligned}$$

2. Következő lépésként számoljuk ki a várható értéket. Kihasználva, hogy a $\Sigma_{k=0}^{\infty} x^k$ hatványsor a konvergencia tartományon belül tagonként deriválható, illetve, hogy az $x^0 = 1$ konstans tag deriváltja nulla

$$\begin{aligned} \mathbf{E}(\xi) &= \sum_k x_k p_k = \sum_{k=1}^{\infty} kpq^{k-1} = p \sum_{k=1}^{\infty} kq^{k-1} = \\ &= p \sum_{k=1}^{\infty} \left(q^k\right)' = p \sum_{k=0}^{\infty} \left(q^k\right)' = \\ &= p \left(\sum_{k=0}^{\infty} q^k\right)' = p \left(\frac{1}{1-q}\right)' = \\ &= \frac{p}{(1-q)^2} = \frac{1}{p}. \end{aligned}$$

[6] Az eloszlás elnevezését éppen a geometriai sorral való szoros kapcsolata indokolja.

3. Számoljuk ki a második momentumot. A várható érték már kiszámolt értékét használva, illetve hogy a hatványsor első két tagjának második deriváltja nulla

$$\begin{aligned}
\mathbf{E}\left(\xi^2\right) &= \sum_k x_k^2 p_k = \sum_{k=1}^{\infty} k^2 pq^{k-1} = \sum_{k=1}^{\infty} (k+1)kpq^{k-1} - \sum_{k=1}^{\infty} kpq^{k-1} = \\
&= p\sum_{k=1}^{\infty}\left(q^{k+1}\right)'' - \frac{1}{p} = p\sum_{k=2}^{\infty}\left(q^{k}\right)'' - \frac{1}{p} = \\
&= p\sum_{k=0}^{\infty}\left(q^{k}\right)'' - \frac{1}{p} = p\left(\sum_{k=0}^{\infty} q^{k}\right)'' - \frac{1}{p} = \\
&= p\left(\frac{1}{1-q}\right)'' - \frac{1}{p} = p\left(\frac{1}{(1-q)^2}\right)' - \frac{1}{p} = \\
&= 2p\frac{1}{(1-q)^3} - \frac{1}{p} = \frac{2p-p^2}{p^3} = \frac{2-p}{p^2}.
\end{aligned}$$

4. Végül számoljuk ki a szórást.

$$\mathbf{D}(\xi) = \sqrt{\mathbf{E}\left(\xi^2\right) - \mathbf{E}^2(\xi)} = \sqrt{\frac{2-p}{p^2} - \frac{1}{p^2}} = \sqrt{\frac{1-p}{p^2}} = \frac{\sqrt{q}}{p}.$$

□

5.2.5. Példa. A geometriai eloszlás és a negatív binomiális vagy Pascal-eloszlás.

Legyen A egy esemény, amely valószínűsége legyen p. Ha egymás után az A eseményre nézve független kísérleteket végzünk, akkor mi annak a valószínűsége, hogy az A esemény éppen az n-edik kísérletre következik be? Ha ξ jelöli azt, hogy hányadik kísérletre következik be az A esemény, akkor a ξ változó geometriai eloszlású: $\mathbf{P}(\xi = n) = pq^{n-1}$, ahol $n = 1, 2, \ldots$ és $q \stackrel{\circ}{=} 1 - p$. Ha η_r jelöli azt, hogy r-en felül hány kísérletet kell ahhoz elvégezi ahhoz, hogy r-szer következzen be az A, akkor

$$\mathbf{P}(\eta_r = n) = \binom{r+n-1}{n} p^r q^n, \quad n = 0, 1, \ldots$$

ugyanis az r-edik bekövetkezés előtt $r+n-1$ kísérlet volt és ebből kell az n darab „helytelen” kísérlet helyét kiválasztani. Az így kapott eloszlást negatív binomiális vagy Pascal-eloszlásnak mondjuk[7]. Vegyük észre, hogy a geometriai eloszlás esetén $n = 1, 2, \ldots$. A geometriai eloszlás definíciója az irodalomban nem teljesen egyértelmű. Mi geometriai eloszláson az imént definiált eloszlást értjük, vagyis amikor $n = 1, 2, 3, \ldots$. Ha a geometriai eloszlás definíciójában csupán a helytelen kísérleteket

[7] Miként később látni fogjuk, a negatív binomiális eloszlás tágabb fogalom, mint a Pascal-eloszlás. A negatív binomiális eloszlás képletében r helyébe α-t szokás írni, ahol $\alpha > 0$ tetszőleges.

számoljuk, akkor $\mathbf{P}(\xi = n) = pq^n$, ahol $n = 0,1,2\ldots$. Ilyenkor a félreértés elkerülése céljából inkább elsőrendű Pascal-eloszlásról fogunk beszélni[8]. Vegyük észre, hogy az $r = 2$ paraméterű Pascal-eloszlás felírható két darab független, azonos paraméterű elsőrendű Pascal-eloszlás összegeként. Az első Pascal-eloszlás során az első sikerre, a második esetén a második sikerre várunk. Valóban a konvolúciós képlet alapján, felhasználva hogy az elsőrendű Pascal-eloszlás pq^k, $k = 0,1,2,\ldots$ ha az összeg eloszlása (s_n)

$$\begin{aligned} s_n &= \sum_{j=0}^{n} pq^j pq^{n-j} = (n+1)\,p^2q^n = \\ &= \binom{2+n-1}{n} p^2q^n = \mathbf{P}(\eta_2 = n). \end{aligned}$$

Hasonlóan a három elsőrendű Pascal-eloszlás összege, ismételten a konvolúciós képlet alapján, felhasználva, hogy a konvolúció asszociatív, vagyis három tag összegét úgy számolhatjuk, hogy két tag összegéhez hozzáadjuk a harmadik tagot. Ha az összeg eloszlása (s_n), akkor

$$s_n = \sum_{j=0}^{n} (j+1)\,p^2q^j pq^{n-j} =$$

$$= p^3q^n \sum_{j=0}^{n} (j+1) = \frac{(n+1)(n+2)}{2} p^3q^n =$$

$$= \binom{3+n-1}{2} p^3q^n = \mathbf{P}(\eta_3 = n).$$

Általában egy r-ed rendű és egy elsőrendű Pascal-eloszlás konvolúciója $(r+1)$-ed rendű Pascal-eloszlás, következésképpen minden Pascal-eloszlás felbontható független elsőrendű Pascal eloszlások összegére[9]. Valóban, ismételten a konvolúciós képlet alapján az összeg eloszlása

$$\sum_{j=0}^{n} \binom{r+j-1}{j} p^r q^j pq^{n-j} = p^{r+1}q^n \sum_{j=0}^{n} \binom{r+j-1}{j}.$$

Meg kell mutatni, hogy

$$\sum_{j=0}^{n} \binom{r+j-1}{j} = \binom{r+1+n-1}{n} = \binom{r+n}{n}.$$

[8] Evvel persze a probléma nincsen megoldva. A Pascal-eloszlás esetén nem egyértelmű, hogy mi lesz az eloszlás tartója. Mi az eloszlást az r értékétől függetlenül a $0,1,\ldots$ számokon definiáljuk. De az irodalomban az is előfordul, hogy az $r, r+1,\ldots$ számokat értik az eloszlás tartóján.

[9] Ez igen szemléletes, ugyanis először az első bekövetkezésre várunk, aztán a másodikra stb. Bármennyire is szemléletes ez az összefüggés, formálisan is igazolni kell, ugyanis például nem teljesen nyilvánvaló, hogy az egymást követő bekövetkezések időtartama független és azonos eloszlású. Ez lényegében a később tárgyalt erős Markov-tulajdonság alkalmazása.

Ez teljes indukcióval igazolható: $n=0$ esetén az egyenlőség triviálisan teljesül. Hasonlóan triviális az $n=1$ eset is. Ha az egyenlőséget már n-re igazoltuk, akkor

$$\begin{aligned}
\sum_{j=0}^{n+1}\binom{r+j-1}{j} &= \binom{r+n}{n}+\binom{r+n}{n+1}= \\
&= \frac{(r+n)!}{n!r!}+\frac{(r+n)!}{(n+1)!\,(r-1)!}= \\
&= \frac{(n+1)\,(r+n)!+r\,(r+n)!}{r!\,(n+1)!}= \\
&= \frac{(r+n+1)!}{r!\,(n+1)!}=\binom{r+n+1}{n+1},
\end{aligned}$$

ami éppen az egyenlőség $(n+1)$-re. Az elsőrendű Pascal-eloszlás várható értéke[10]

$$\sum_{n=0}^{\infty} npq^n = q\sum_{n=1}^{\infty} npq^{n-1} = \frac{q}{p},$$

ahol kihasználtuk, hogy a geometriai eloszlás várható értéke $1/p$. Ebből az r-edrendű Pascal-eloszlás várható értéke, felhasználva, hogy az összeg várható értéke a várható értékek összege $\mathbf{E}(\xi)=rq/p$. Végezetül számoljuk ki a szórást. Elég az elsőrendű esettel foglalkozni. Felhasználva, hogy a geometriai eloszlás második momentuma $(2-p)/p^2$

$$\mathbf{E}\left(\xi^2\right)=\sum_{n=0}^{\infty} n^2pq^n = q\sum_{n=1}^{\infty} n^2pq^{n-1} = q\frac{2-p}{p^2}.$$

Ebből

$$\begin{aligned}
\mathbf{D}^2(\xi) &= q\frac{2-p}{p^2}-\frac{q^2}{p^2}=\frac{q\,(2-p)-q^2}{p^2}=\frac{q\,(2-p-q)}{p^2}= \\
&= \frac{q\,(2-p-(1-p))}{p^2}=\frac{q}{p^2},
\end{aligned}$$

ami azonos a geometriai eloszlás varianciájával[11]. Miként később látni fogjuk, független eloszlások összegének varianciája megegyezik a varianciák összegével. Ebből az r-edrendű Pascal-eloszlás szórása

$$\mathbf{D}(\xi)=\frac{\sqrt{rq}}{p}.$$

□

[10] Ha ξ geometriai eloszlású, akkor a $\xi-1$ elsőrendű Pascal-eloszlást követ. Ebből a várható érték $1/p-1=q/p$.

[11] A szórás nem változik, ha egy változót egy konstanssal eltolunk.

Folytonos eloszlások várható értéke és szórása

Ebben az alpontban néhány folytonos eloszlás várható értékét és szórását számoljuk ki.

5.2.6. Példa. Számoljuk ki az exponenciális eloszlás várható értékét és szórását.

A $\lambda > 0$ paraméterű exponenciális eloszlás sűrűségfüggvénye

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{ha} \quad x \geq 0 \\ 0 & \text{ha} \quad x < 0 \end{cases}.$$

1. Mivel $f(x) \geq 0$, ezért ahhoz, hogy az $f(x)$ valóban sűrűségfüggvényt elegendő belátni, hogy a teljes számegyenesen vett integrálja 1 :

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= \lambda \int_{0}^{\infty} \exp(-\lambda x)\,dx = \lambda \left[\frac{\exp(-\lambda x)}{-\lambda}\right]_0^{\infty} = \\ &= 0 + \frac{\lambda}{\lambda} = 1. \end{aligned}$$

2. A várható érték kiszámolása a következő: Parciálisan integrálva

$$\begin{aligned} \int_{-\infty}^{\infty} x dF(x) &= \int_{0}^{\infty} x f(x)\,dx = \lambda \int_{0}^{\infty} x \exp(-\lambda x)\,dx = \\ &= \lambda \left[\frac{x \exp(-\lambda x)}{-\lambda}\right]_0^{\infty} + \lambda \int_0^{\infty} \frac{\exp(-\lambda x)}{\lambda} dx = \\ &= 0 + \left[\frac{\exp(-\lambda x)}{-\lambda}\right]_0^{\infty} = \frac{1}{\lambda}. \end{aligned}$$

3. A szórás kiszámolásához ki kell számolni a ξ második momentumát, vagyis a ξ^2 várható értékét. A tárgyalás jelen szintjén még nem ismerjük a transzformált valószínűségi változók várható értékére vonatkozó formulát, mitöbb, éppen ennek bevezetése a célunk, így némiképpen körülményesen kell eljárnunk. Ahhoz, hogy a ξ^2 várható értékét meg tudjuk határozni, ki kell számolni a ξ^2 eloszlását. Ehhez szükségünk lesz a ξ változó eloszlásfüggvényére. Korábban már láttuk, hogy

$$F(x) = 1 - \exp(-\lambda x), \quad x \geq 0.$$

Mivel $\mathbf{P}(\xi \geq 0) = 1$, ezért az alábbi számolásban az abszolútérték elhagyható. Ha $x \geq 0$, akkor

$$\begin{aligned} G(x) &\circeq \mathbf{P}\left(\xi^2 < x\right) = \mathbf{P}\left(\sqrt{\xi^2} < \sqrt{x}\right) = \mathbf{P}\left(|\xi| < \sqrt{x}\right) = \\ &= \mathbf{P}\left(\xi < \sqrt{x}\right) = F\left(\sqrt{x}\right) = 1 - \exp\left(-\lambda \sqrt{x}\right). \end{aligned}$$

Ebből az összetett függvény deriválási szabálya miatt ha $x > 0$, akkor a ξ^2 sűrűségfüggvénye

$$g(x) = G'(x) = \left(1 - \exp\left(-\lambda\sqrt{x}\right)\right)' = \lambda \exp\left(-\lambda\sqrt{x}\right)\frac{1}{2\sqrt{x}}.$$

$\mathbf{P}\left(\xi^2 \geq 0\right) = 1$, így

$$\begin{aligned}\mathbf{E}\left(\xi^2\right) &= \int_{-\infty}^{\infty} x dG(x) = \int_{-\infty}^{\infty} x g(x)\, dx = \\ &= \int_0^{\infty} x\lambda \exp\left(-\lambda\sqrt{x}\right)\frac{1}{2\sqrt{x}} dx = \frac{\lambda}{2}\int_0^{\infty}\sqrt{x}\exp\left(-\lambda\sqrt{x}\right) dx.\end{aligned}$$

Az integrálban az $u = \sqrt{x}$ kézenfekvő helyettesítést végrehajtva, mivel $x = u^2$

$$\begin{aligned}\mathbf{E}\left(\xi^2\right) &= \frac{\lambda}{2}\int_0^{\infty} u\exp(-\lambda u)\frac{dx}{du}du = \\ &= \frac{\lambda}{2}\int_0^{\infty} u\exp(-\lambda u)\, 2u du = \\ &= \lambda\int_0^{\infty} u^2\exp(-\lambda u)\, du = \int_0^{\infty} u^2 f(u)\, du.\end{aligned}$$

Parciálisan integrálva

$$\begin{aligned}\mathbf{E}\left(\xi^2\right) &= \left[\lambda u^2\frac{\exp(-\lambda u)}{-\lambda}\right]_0^{\infty} + 2\int_0^{\infty} u\exp(-\lambda u)\, du = \\ &= \frac{2}{\lambda}\int_0^{\infty} u\lambda\exp(-\lambda u)\, du = \frac{2}{\lambda}\mathbf{E}(\xi) = \frac{2}{\lambda^2}.\end{aligned}$$

4. Ebből már a szórás egyszerűen kiszámolható:

$$\mathbf{D}(\xi) = \sqrt{\mathbf{E}\left(\xi^2\right) - \mathbf{E}^2(\xi)} = \sqrt{\frac{2}{\lambda^2} - \frac{1}{\lambda^2}} = \sqrt{\frac{1}{\lambda^2}} = \frac{1}{\lambda},$$

ami a jelen esetben ugyanaz, mint a várható érték.

□

5.2.7. Példa. Számoljuk ki az egyenletes eloszlás várható értékét és szórását.

Először a standard egyenletes eloszlás várható értékét és szórását számoljuk ki, az általános esetet a várható érték és a szórás általános tulajdonságai segítségével határozzuk meg.

1. Standard egyenletes eloszláson a $[0,1]$ szakaszra koncentrált egyenletes eloszlást érjük. Ilyenkor a sűrűségfüggvény

$$f(x)=\begin{cases} 0 & \text{ha} \quad x<0 \\ 1 & \text{ha} \quad 0<x<1 \\ 0 & \text{ha} \quad x>1 \end{cases}.$$

Mivel $\int_0^1 dx=1$, ezért az $f(x)$ valóban egy eloszlás sűrűségfüggvénye.

2. Tekintsük a várható értéket.

$$\begin{aligned} \mathbf{E}(\xi) &= \int_{-\infty}^{\infty} xdF(x)=\int_{-\infty}^{\infty} xf(x)\,dx= \\ &= \int_0^1 x\cdot 1dx=\left[\frac{x^2}{2}\right]_0^1=\frac{1}{2}. \end{aligned}$$

3. A szórás kiszámolásához számoljuk ki a ξ^2 eloszlását. Mivel a ξ a $[0,1]$ szakaszra koncentrálódik, ezért ugyanez igaz a ξ^2-re is, így elég a sűrűségfüggvényt ott meghatározni. A már látott módon az eloszlásfüggvény segítségével érvelve: ha $0<x<1$, akkor

$$G(x)\stackrel{\circ}{=}\mathbf{P}\left(\xi^2<x\right)=\mathbf{P}\left(\xi<\sqrt{x}\right)=F\left(\sqrt{x}\right)=\sqrt{x},$$

ugyanis a $[0,1]$ szakaszon az egyenletes eloszlás $F(x)$ eloszlásfüggvénye éppen x. Ebből a ξ^2 sűrűségfüggvénye a $(0,1)$ intervallumban $G'(x)\stackrel{\circ}{=}g(x)=1/\left(2\sqrt{x}\right)$. Így a második momentum

$$\mathbf{E}\left(\xi^2\right)=\int_{-\infty}^{\infty} xg(x)\,dx=\int_0^1 x\frac{1}{2\sqrt{x}}dx=\frac{1}{2}\int_0^1\sqrt{x}dx.$$

Bár az integrált most közvetlenül ki tudjuk számolni, mégis ismét hajtsuk végre az $u=\sqrt{x}$ helyettesítést. Ekkor, mivel $x=u^2$, ezért

$$\mathbf{E}\left(\xi^2\right)=\frac{1}{2}\int_0^1 u\frac{dx}{du}du=\int_0^1 u^2du=\int_0^1 u^2 1du=\int_{-\infty}^{\infty} u^2 f(u)\,du=\frac{1}{3}$$

4. Ebből a szórás értéke

$$\mathbf{D}(\xi)=\sqrt{\mathbf{E}\left(\xi^2\right)-\mathbf{E}^2(\xi)}=\sqrt{\frac{1}{3}-\left(\frac{1}{2}\right)^2}=\sqrt{\frac{1}{12}}.$$

5. Végül legyen a ξ egyenletes eloszlású valamely $[a,b]$ szakaszon. Azonnal látszik, hogy az $\eta\stackrel{\circ}{=}(\xi-a)/(b-a)$ egyenletes eloszlású a $[0,1]$ szakaszon, vagyis $\xi=(b-a)\eta+a$, ahol az η egyenletes eloszlású a $[0,1]$ szakaszon. Ebből a várható érték linearitása miatt

$$\mathbf{E}(\xi)=(b-a)\frac{1}{2}+a=\frac{a+b}{2}.$$

A szórás kiszámolásához egyrészt vegyük figyelembe, hogy ha egy változóhoz hozzáadunk egy konstanst, akkor a szórás értéke nem változik, ugyanis a várható érték is eltolódik a hozzáadott konstanssal, és így mivel a szórás kiszámolásakor a várható értéket le kell vonni, ezért a konstans végső soron kiesik. Ugyancsak a szórás képletéből, felhasználva, hogy a várható értékből a konstans szorzó kiemelhető,

$$\mathbf{D}(a\xi) = \sqrt{\mathbf{E}\left(a\xi^2\right) - \mathbf{E}^2(a\xi)} = \sqrt{a^2\left(\mathbf{E}\left(\xi^2\right) - \mathbf{E}^2(\xi)\right)} = |a|\,\mathbf{D}(\xi).$$

Ebből a szórás

$$\mathbf{D}(\xi) = \frac{b-a}{\sqrt{12}}.$$

□

5.3. Transzformált változók várható értéke

A bemutatott példák alapján nem túl meglepő a következő tétel:

5.3.1. Tétel. *Ha a ξ változó eloszlásának van $f(x)$ sűrűségfüggvénye, akkor*

$$\mathbf{E}\left(\xi^2\right) = \int_{-\infty}^{\infty} x^2 f(x)\,dx,$$

ahol a két oldal egyszerre véges vagy végtelen.

Bizonyítás: A már látott gondolatmenetet követve, ha $x > 0$, akkor

$$\mathbf{P}\left(\xi^2 < x\right) = \mathbf{P}\left(|\xi| < \sqrt{x}\right) = \mathbf{P}\left(-\sqrt{x} < \xi < \sqrt{x}\right) = F\left(\sqrt{x}\right) - F\left(-\sqrt{x}\right),$$

ahol felhasználtuk, hogy az F folytonos, ugyanis létezik sűrűségfüggvénye. Ezt x szerint deriválva, a ξ^2 sűrűségfüggvénye

$$g(x) \overset{\circ}{=} \frac{f(\sqrt{x}) + f(-\sqrt{x})}{2}\frac{1}{\sqrt{x}}, \quad x > 0.$$

A várható érték képlete alapján

$$\begin{aligned}\mathbf{E}\left(\xi^2\right) &= \int_0^{\infty} x g(x)\,dx = \int_0^{\infty} x\frac{f(\sqrt{x}) + f(-\sqrt{x})}{2}\frac{1}{\sqrt{x}}dx = \\ &= \int_0^{\infty} \sqrt{x}\frac{f(\sqrt{x}) + f(-\sqrt{x})}{2}dx.\end{aligned}$$

Miként a példákban ismét $u = \sqrt{x}$ helyettesítéssel

$$\begin{aligned}\mathbf{E}\left(\xi^2\right) &= \int_0^{\infty} u\frac{f(u) + f(-u)}{2}\frac{dx}{du}du = \int_0^{\infty} u^2\left(f(u) + f(-u)\right)du = \\ &= \int_{-\infty}^{\infty} u^2 f(u)\,du.\end{aligned}$$

□

Teljesen analóg módon látható be a következő:

5.3.2. Tétel. *Ha a ξ változó eloszlásának van $f(x)$ sűrűségfüggvénye, akkor a ξ n-edik momentuma*

$$\mathbf{E}(\xi^n) = \int_{-\infty}^{\infty} x^n f(x)\,dx$$

módon számolható, ahol a két oldal egyszerre véges vagy végtelen, illetve létezik vagy nem létezik.

A gondolatmenet tovább általánosítható. Legyen g egy szigorúan monoton, differenciálható leképezés. A g lehet szigorúan monoton növekedő, de lehet csökkenő is. A szigorú növekedés miatt a g-nek létezik g^{-1} inverze, és a differenciálhatóságból következő folytonosság miatt ez az inverz szintén egy intervallumon van értelmezve. Tekintsük a $g(\xi)$ változó eloszlását. Ha a ξ eloszlásfüggvénye $F(x)$ és a g nő, akkor

$$\mathbf{P}(g(\xi) < x) = \mathbf{P}\left(\xi < g^{-1}(x)\right) = F\left(g^{-1}(x)\right), \tag{5.3.1}$$

ha pedig a g csökken, akkor

$$\mathbf{P}(g(\xi) < x) = \mathbf{P}\left(\xi > g^{-1}(x)\right) = 1 - F\left(g^{-1}(x)\right).$$

Az utolsó lépésben kihasználtuk, hogy mivel a ξ-nek van sűrűségfüggvénye ezért az F folytonos. Ebből az összetett függvény deriválásával a $g(\xi)$ transzformált valószínűségi változó sűrűségfüggvénye[12]

$$f\left(g^{-1}(x)\right)\left|\frac{d}{dx}g^{-1}(x)\right|.$$

Az abszolútérték azért kerül a képletbe, mert amikor a g szigorúan monoton csökken, akkor a g^{-1} is szigorúan csökken és így a deriváltja nem pozitív. Ebből adódik a következő:

5.3.3. Tétel. *Ha a ξ változó eloszlásának van $f(x)$ sűrűségfüggvénye és g egy szigorúan monoton, folytonosan differenciálható függvény, akkor a $g(\xi)$ transzformált változónak is van $h(x)$ sűrűségfüggvénye és*

$$h(x) = f\left(g^{-1}(x)\right)\left|\frac{d}{dx}g^{-1}(x)\right|.$$

5.3.4. Példa. Számoljuk ki az $a\xi + b$ sűrűségfüggvényét.

[12] Vegyük észre, hogy hallgatólagosan feltettük, hogy a g deriválhatóságából következik a g^{-1} deriválhatósága. Ez azonban általában nem teljesül. Például a $g(x) = x^3$ függvény szigorúan monoton nő, deriválható, de az inverze az $x = 0$ pontban nem deriválható. Hallgatólagosan feltesszük, hogy ez a probléma csak véges sok pontban jelentkezik, így a sűrűségfüggvények értelmezését a kivételes pontok nem zavarják, ugyanis a sűrűségfüggvények amúgy is csak véges számú ponttól való eltekinthetőség erejéig vannak definiálva.

Értelemszerűen feltehetjük, hogy $a \neq 0$. Ha a ξ változó sűrűségfüggvénye $f(x)$, akkor az $a\xi + b$ transzformált valószínűségi változó sűrűségfüggvénye

$$g(x) = f\left(\frac{x-b}{a}\right)\frac{1}{|a|} \tag{5.3.2}$$

□

5.3.5. Példa. Számoljuk ki az általános Cauchy-eloszlás sűrűségfüggvényét.

A standard Cauchy-eloszlás sűrűségfüggvénye

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$

Ha ξ standard Cauchy-eloszlású, akkor az $a\xi + b$ változó eloszlást (a,b) paraméterű általános Cauchy-eloszlásnak hívjuk. Az előző példa alapján az $a\xi + b$ alakú változó eloszlásának sűrűségfüggvénye

$$\begin{aligned} g(x) &= \frac{1}{\pi|a|}\frac{1}{1+\left(\frac{x-b}{a}\right)^2} = \frac{1}{\pi|a|}\frac{a^2}{a^2+(x-b)^2} = \\ &= \frac{1}{\pi}\frac{|a|}{a^2+(x-b)^2}. \end{aligned}$$

□

5.3.6. Példa. A μ várható értékű és σ szórású normális eloszlás sűrűségfüggvénye

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Miként később látni fogjuk, az

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$$

sűrűségfüggvénnyel rendelkező eloszlás, amit szokás standard normális eloszlásnak mondani várható értéke 0, szórása pedig 1. Ha a ξ standard normális eloszlású, akkor a $\sigma\xi + \mu$ várható értéke μ, és ha $\sigma > 0$, akkor a szórása pedig σ. Ebből a már látott (5.3.2) képlet alapján az $f(x)$ fenti képlete evidens.

□

Végezetül megjegyezzük, hogy a transzformált valószínűségi változók várható értékére vonatkozó képlet sokkal általánosabb körülmények között is érvényes. Igen általános feltételek mellett teljesül a

$$\mathbf{E}(g(\xi)) = \int_{-\infty}^{\infty} g(x)\,dF(x) \tag{5.3.3}$$

képlet, amiből a sűrűségfüggvény létezése esetén már következik az

$$\mathbf{E}(g(\xi)) = \int_{-\infty}^{\infty} g(x)\,dF(x) = \int_{-\infty}^{\infty} g(x)\,f(x)\,dx$$

szabály. A képlet például érvényes minden folytonos g transzformációs függvényre, de a g lehet nem folytonos is. Az elemi valószínűségszámítás tárgyalása során erre az általános esetre azonban nincs szükségünk, és indoklására sincsen módunk.

5.3.7. Példa. Vizsgáljuk meg az

$$\mathbf{E}\left(\xi^2\right) = \int_{-\infty}^{\infty} x^2 dF(x) \tag{5.3.4}$$

összefüggést.

Érdemes némiképpen általánosabban megközelíteni a problémát. Legyen g egy szigorúan monoton növekedő, folytonos függvény. A fenti (5.3.1) sor alapján a $g(\xi)$ változó eloszlásfüggvénye

$$G(x) = \mathbf{P}(g(\xi) < x) = \begin{cases} 0 & \text{ha} \quad x \leq g(-\infty) \\ F\left(g^{-1}(x)\right) & \text{ha} \quad g(-\infty) < x \leq g(\infty) \\ 1 & \text{ha} \quad g(\infty) < x \end{cases}.$$

Ekkor tetszőleges $g(-\infty) < a < b < g(\infty)$ esetén a helyettesítéses integrálás formulája alapján[13]

$$\int_a^b g(x)\,dF(x) = \int_{g(a)}^{g(b)} x\,dG(x).$$

Ha most $a \to -\infty$ és $b \to \infty$, akkor, feltéve, hogy az $\int_{-\infty}^{\infty} g(x)\,dF(x)$ improprius Stieltjes-integrál konvergens

$$\int_{-\infty}^{\infty} g(x)\,dF(x) = \int_{g(-\infty)+0}^{g(\infty)-0} x\,dG(x). \tag{5.3.5}$$

Felhasználva, hogy $\mathbf{P}(g(\xi) = g(\pm\infty)) = 0$, vagyis hogy a $g(\pm\infty)$ a G eloszlásfüggvény folytonossági pontjai

$$\mathbf{E}(g(\xi)) \stackrel{\circ}{=} \int_{-\infty}^{\infty} x\,dG(x) = \int_{-\infty}^{\infty} g(x)\,dF(x).$$

[13] V. ö. (4.2.3) sor, 54. oldal.

Ha tehát g szigorúan monoton nő és folytonos, akkor a fenti (5.3.3) formula teljesül. Példaként gondoljuk az $\mathbf{E}(\exp(\xi))$ várható értékre. Ha $G(x)$ az $\exp(\xi)$ eloszlásfüggvénye, akkor a fenti (5.3.5) alapján

$$\mathbf{E}(\exp(\xi)) = \int_0^\infty x dG(x) = \int_{-\infty}^\infty \exp(x)\, dF(x).$$

Ha g csak monoton nő és folytonos, viszont a h szigorúan monoton nő és korlátos és folytonos[14], akkor az integrál és a várható érték additivitása miatt

$$\begin{aligned}
&\int_{-\infty}^\infty g(x)\, dF(x) + \int_{-\infty}^\infty h(x)\, dF(x) = \\
&= \int_{-\infty}^\infty g(x) + h(x)\, dF(x) = \\
&= \mathbf{E}(g(\xi) + h(\xi)) = \\
&= \mathbf{E}(g(\xi)) + \mathbf{E}(h(\xi)) = \\
&= \mathbf{E}(g(\xi)) + \int_{-\infty}^\infty h(x)\, dF(x).
\end{aligned}$$

Ebből, felhasználva, hogy az $\int_{-\infty}^\infty h(x)\, dF(x)$ integrál véges, kapjuk a kívánt egyenlőséget monoton növekedő függvényekre. Ha most $g(x) = g_1(x) - g_2(x)$, ahol a $g_1(x)$ és $g_2(x)$ monoton nő, akkor

$$\begin{aligned}
\mathbf{E}(g(\xi)) &= \mathbf{E}(g_1(\xi)) - \mathbf{E}(g_2(\xi)) = \\
&= \int_{-\infty}^\infty g_1(x)\, dF(x) - \int_{-\infty}^\infty g_2(x)\, dF(x) = \\
&= \int_{-\infty}^\infty g(x)\, dF(x),
\end{aligned}$$

feltéve, hogy a két várható érték véges és így az integrál és a várható érték additivitása használható. Ez utóbbi teljesül a $g(x) = x^2$ függvény

$$\begin{aligned}
g_1(x) &= \begin{cases} 0 & \text{ha} \quad x \le 0 \\ x^2 & \text{ha} \quad x > 0 \end{cases} \\
g_2(x) &= \begin{cases} -x^2 & \text{ha} \quad x \le 0 \\ 0 & \text{ha} \quad x > 0 \end{cases}
\end{aligned}$$

felbontása esetén, ugyanis könnyen látható, hogy ha az x^2 Stieltjes-integrálja végtelen, akkor a megfelelő várható érték is végtelen és fordítva.

□

[14] Például $h(x) = \arctan x$.

5.3.8. Példa. Tetszőleges ξ változóra

$$\begin{aligned}\mathbf{E}(\xi^n) &= \int_{-\infty}^{\infty} x^n dF(x), \\ \mathbf{E}(|\xi|) &= \int_{-\infty}^{\infty} |x| dF(x), \\ \mathbf{E}(\exp(a\xi)) &= \int_{-\infty}^{\infty} \exp(ax) dF(x).\end{aligned}$$

ahol $F(x)$ a ξ eloszlásfüggvénye.

5.3.9. Példa. Ha a ξ változó eloszlásának van $f(x)$ sűrűségfüggvénye és g egy folytonos és monoton függvény, akkor

$$\mathbf{E}(g(\xi)) = \int_{-\infty}^{\infty} g(x) f(x) dx,$$

ahol a két oldal egyszerre létezik vagy nem létezik, illetve véges vagy végtelen.

Az előzőek alapján

$$\mathbf{E}(g(\xi)) = \int_{-\infty}^{\infty} g(x) dF(x) = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

$\square$

5.3.10. Példa. Vizsgáljuk meg az

$$\mathbf{E}(\sin(\xi)) = \int_{-\infty}^{\infty} \sin x dF(x)$$

formulát.

Ismételten érdemes egy kicsit általánosabban tekinteni a problémát. Legyen g egy korlátos, folytonosan deriválható függvény. Az egyszerűbb jelölés céljából tegyük fel, hogy $|g| \leq 1$. Legyen $G(x)$ a $g(\xi)$ eloszlásfüggvénye. Mivel g korlátos, ezért a $g(\xi)$ eloszlása egy korlátos tartományra koncentrálódik, így az $\mathbf{E}(g(\xi)) = \int_{-\infty}^{\infty} x dG(x)$ integrál, miként az $\int_{-\infty}^{\infty} g(x) dF(x)$ is létezik. A kérdés csak az, hogy a két integrál értéke azonos-e vagy sem. Mivel a g folytonosan deriválható, ezért

$$\begin{aligned}g(x) &= g(0) + \int_0^x g'(t) dt = \\ &\quad g(0) + \int_0^x \left(g'(t)\right)^+ dt - \int_0^x \left(g'(t)\right)^- dt,\end{aligned}$$

vagyis a g felírható a monoton növekedő g_1 és g_2 függvények különbségeként. A problémát az jelenti, hogy miközben a $g(\xi)$ várható értéke véges, előfordulhat, hogy az $\mathbf{E}(g_1(\xi))$ és az $\mathbf{E}(g_2(\xi))$ várható értékek végtelenek, így a korábban tárgyalt gondolatmenet némi módosításra szorul. Vegyük észre, hogy ez a probléma nem léphet fel, ha a g tartója, korlátos, vagyis ha van olyan N szám, hogy ha $|x| > N$, akkor a $g(x) = 0$. Ha $A_n \overset{\circ}{=} \{|\xi| > n\}$, akkor $A_n \searrow \emptyset$, így a valószínűség folytonossága miatt tetszőleges ε esetén ha N elég nagy, akkor $\mathbf{P}(|\xi| > N) < \varepsilon$. Ebből, kihasználva, hogy $|g| \leq 1$

$$\left|\int_{-\infty}^{\infty} g(x)\,dF(x) - \int_{-N}^{N} g(x)\,dF(x)\right| < \varepsilon,$$

és a várható érték monotonitása miatt

$$\begin{aligned}
&|\mathbf{E}(g(\xi)) - \mathbf{E}(g(\xi)\chi(|\xi| \leq N))| \leq \\
&\leq \mathbf{E}(|g(\xi)|(1-\chi(|\xi| \leq N))) \leq \mathbf{E}(1-\chi(|\xi| \leq N)) = \\
&= \mathbf{P}(|\xi| > N) < \varepsilon.
\end{aligned}$$

Legyen h egy olyan korlátos tartójú, folytonosan deriválható függvény, amely a $[-N.N]$ szakaszon megegyezik a g-vel. A már elmondottak miatt $\mathbf{E}(h(\xi)) = \int_{-\infty}^{\infty} h(x)\,dF(x)$. Ha $-N$ és N az F eloszlásfüggvény folytonossági pontja, akkor az

$$\int_{-\infty}^{\infty} g(x)\chi(|x| \leq N)\,dF(x)$$

integrál létezik és ha a h függvényt elég közel választjuk a $g(x)\chi(|x| \leq N)$ függvényhez, akkor

$$\left|\int_{-\infty}^{\infty} g(x)\chi(|x| \leq N)\,dF(x) - \int_{-\infty}^{\infty} h(x)\,dF(x)\right| < \varepsilon.$$

A fenti két becslés alapján így könnyen igazolható, hogy ha $-N$ és N az F eloszlásfüggvény folytonossági pontjai, akkor

$$\begin{aligned}
|\mathbf{E}(h(\xi)) - \mathbf{E}(g(\xi))| &< 2\varepsilon, \\
\left|\int_{-\infty}^{\infty} h(x)\,dF(x) - \int_{-\infty}^{\infty} g(x)\,dF(x)\right| &< 2\varepsilon,
\end{aligned}$$

amiből az

$$\mathbf{E}(g(\xi)) = \int_{-\infty}^{\infty} g(x)\,dF(x)$$

egyenlőség már tetszőleges korlátos és folytonosan deriválható g esetére igazolható.

□

5.4. A szórásnégyzet additivitása

A várható érték linearitása mellett felmerül a szórás linearitása. Miként már láttuk, és az általános esetben is igen egyszerűen igazolható, hogy ha a egy tetszőleges konstans, akkor $\mathbf{D}(a\xi) = |a|\,\mathbf{D}(\xi)$. A szórás definíciójában szereplő négyzetgyök miatt a szórás additivitása érdemben nem teljesül. Vizsgáljuk meg a szórásnégyzet additivitását: A várható érték linearitása miatt

$$\mathbf{D}^2(\xi+\eta) = \mathbf{E}\left((\xi+\eta)^2\right) - \mathbf{E}^2(\xi+\eta) =$$

$$= \mathbf{E}\left(\xi^2\right) + \mathbf{E}\left(\eta^2\right) + 2\mathbf{E}(\xi\eta) - \mathbf{E}\left(\xi^2\right) - \mathbf{E}\left(\eta^2\right) - 2\mathbf{E}(\xi)\mathbf{E}(\eta) =$$

$$\mathbf{D}^2(\xi) + \mathbf{D}^2(\eta) + 2(\mathbf{E}(\xi\eta) - \mathbf{E}(\xi)\mathbf{E}(\eta)).$$

Ebből a szórásnégyzet pontosan akkor additív, ha $\mathbf{E}(\xi\eta) - \mathbf{E}(\xi)\mathbf{E}(\eta) = 0$. Mivel ez a kifejezés sokszor előfordul érdemes elnevezni.

5.4.1. Definíció. Az $\mathbf{E}(\xi\eta) - \mathbf{E}(\xi)\mathbf{E}(\eta)$ kifejezést a ξ és az η változók kovarianciájának nevezzük és $\mathrm{cov}(\xi,\eta)$ módon jelöljük. Ebből következően ha ξ és η tetszőleges szórással rendelkező valószínűségi változók, akkor

$$\mathbf{D}^2(\xi+\eta) = \mathbf{D}^2(\xi) + \mathbf{D}^2(\eta) + 2\mathrm{cov}(\xi,\eta).$$

A ξ és az η változókat korrelálatlannak mondjuk, ha $\mathrm{cov}(\xi,\eta) = 0$.

A korrelálatlansághoz kapcsolódó fontos fogalom a függetlenség fogalma, amelyet már röviden érintettünk, illetve amelyre később még vissza fogunk térni. Most csak események függetlenségére emlékeztetünk.

5.4.2. Definíció. Az A és B eseményeket függetlennek mondjuk, ha $\mathbf{P}(A\cap B) = \mathbf{P}(A)\mathbf{P}(B)$, illetve általában az $A_1, A_2, \ldots, A_n$ eseményeket függetlenek mondjuk, ha az $1,\ldots,n$ indexek tetszőleges $n_1, n_2, \ldots, n_k$ részhalmaza esetén

$$\mathbf{P}(A_{n_1} \cap A_{n_2} \cap \ldots \cap A_{n_k}) = \mathbf{P}(A_{n_1})\mathbf{P}(A_{n_2})\cdots\mathbf{P}(A_{n_k}).$$

Az A és a B események pontosan akkor függetlenek, ha a hozzájuk tartozó χ_A és χ_B valószínűségi változók korrelálatlanok, ugyanis

$$\mathbf{E}(\chi_A\chi_B) = \mathbf{E}(\chi_{A\cap B}) = \mathbf{P}(A\cap B), \text{ és } \mathbf{P}(A)\mathbf{P}(B) = \mathbf{E}(\chi_A)\mathbf{E}(\chi_B).$$

Ez fontos szerepet játszik az alábbi példában:

5.4.3. Példa. Számoljuk ki a binomiális eloszlás várható értékét és szórását.

$B(n,p)$ binomiális eloszláson egy olyan a $k = 0,1,2,\ldots,n$ számokra koncentrálódó diszkrét eloszlást értünk, amelyre

$$p_k = \binom{n}{k} p^k q^{n-k}, \tag{5.4.1}$$

ahol értelemszerűen $q \stackrel{\circ}{=} 1-p$. A binomiális eloszlás szokásos származtatási módja, hogy egy p valószínűségű sikerrel rendelkező kísérletet n-szer egymástól függetlenül végrehajtunk, és azt kérdezzük, hogy mi lesz annak a valószínűsége, hogy pontosan

k-szor következik be a sikeres kimenet. Mivel a kísérletek függetlenek, ezért annak a valószínűsége, hogy a k darab sikeres kimenet egy rögzített sorrendben fordul elő $p^k q^{n-k}$, ugyanis az együttes bekövetkezés valószínűsége a valószínűségek szorzata és éppen k-szor volt a kísérlet sikeres és $n-k$ esetben sikertelen. De mivel n elemből k elem $\binom{n}{k}$ módon vehető ki, ezért a keresett valószínűséget éppen a fenti (5.4.1) képlet adja meg. A Newton-féle binomiális formulából evidens, hogy a p_k számok összege 1, így valóban eloszlásról van szó. Számoljuk ki a várható értéket és a szórást.

1. A várható érték képletét közvetlenül alkalmazva a számolás némiképpen bonyolult. Ugyanakkor vegyük észre[15], hogy ha $\xi \stackrel{\circ}{=} \sum_{k=1}^{n} \chi_{A_k}$, ahol az (A_k) események függetlenek és mindegyik bekövetkezési valószínűsége éppen p, akkor a ξ eloszlása éppen $B(n,p)$. A várható érték additivitása miatt a keresett várható érték éppen az $\mathbf{E}\left(\chi_{A_k}\right)$ várható értékek összege. De $\mathbf{E}\left(\chi_{A_k}\right) = 1 \cdot p + 0 \cdot q = p$, ezért a $B(n,p)$ binomiális eloszlás várható értéke np.

2. A függetlenség miatt a szórásnégyzet additív, így tetszőleges k-ra $\mathbf{D}^2(\xi) = n\mathbf{D}^2\left(\chi_{A_k}\right)$. Ugyanakkor $\mathbf{E}\left(\chi_{A_k}^2\right) = \mathbf{E}\left(\chi_{A_k}\right) = p$, amiből

$$\mathbf{D}^2(\xi) = n\left(p - p^2\right) = np(1-p) = npq,$$

vagyis a $B(n,p)$ eloszlás szórása $\sqrt{npq}$.

3. Mivel a binomiális eloszlás független χ_{A_k} változók összege, ezért nyilvánvaló módon ha a ξ_1 eloszlása $B(n_1,p)$ és ξ_2 eloszlása $B(n_2,p)$, és a ξ_1 és a ξ_2 függetlenek, akkor $\xi_1+\xi_2$ eloszlása $B(n_1+n_2,p)$. Figyeljünk rá, hogy a p paraméter a két eloszlás esetén azonos. Ha a ξ_1 eloszlása $B(1,1/2)$ és a ξ_2 eloszlása $B(1,1/3)$, akkor a $\xi_1+\xi_2$ várható értéke $1/2+1/3=5/6$. A varianciájuk a függetlenség miatt $1/2 \cdot 1/2 + 1/3 \cdot 2/3 = 17/36$. Ebből ha az eloszlás binomiális lenne, akkor

$$\frac{5}{6} = np, \quad \frac{17}{36} = np(1-p).$$

Vagyis

$$\frac{np(1-p)}{np} = 1 - p = \frac{17/36}{5/6} = \frac{17}{30},$$

amiből

$$n = \frac{5/6}{13/30} = \frac{25}{13},$$

ami nem egész szám. Ha a triviális konstanstól eltekintve egy eloszlás korlátlanul osztható, akkor tartóhalmazának korlátlannak kell lenni. Ez a binomiális eloszlás esetén nem teljesül. A $B(n,p)$ az n-ben additív, de a p-ben nem. Az n változó nem osztható úgymond korlátlanul, a p igen, de p szerint nincs additivitás. A binomiális eloszlás tehát nem korlátlanul osztható.

□

[15] Tulajdonképpen így is lehetne a binomiális eloszlást definiálni.

5.4.4. Példa. A konstans esettől eltekintve korlátos változók nem lehetnek korlátlanul oszthatóak.

A binomiális eloszlás korlátos, így nem is lehet korlátlanul osztható. De hasonló okból nem korlátlanul osztható az egyenletes, vagy a később bevezetett béta eloszlás sem. Ezen igen hasznos tulajdonság igazolása viszonylag egyszerű. Legyen ξ korlátlanul osztható és tegyük fel, hogy korlátos és legyen K a ξ változó egy korlátja. Ha most $\left(\xi_i^{(n)}\right)_{i=1}^n$ a ξ változó egy felbontása, akkor például minden n-re $\left(\xi_i^{(n)}\right) \leq K/n$, ugyanis ha egy A_i halmazon pozitív valószínűséggel esne a $\xi_i^{(n)}$ változó a K/n fölé, akkor ξ a

$$\mathbf{P}\left(\cap_{i=1}^n A_i\right) = \prod_{i=1}^n \mathbf{P}\left(A_i\right) > 0$$

szabály miatt pozitív valószínűséggel esne a $\cap_{i=1}^n A_i$ halmazon a K korlát fölé, ami lehetetlen. Ebből, a függetlenség felhasználásával,

$$\mathbf{D}^2\left(\xi\right) = n\mathbf{D}^2\left(\xi_i^{(n)}\right) \leq n\mathbf{E}\left(\left(\xi_i^{(n)}\right)^2\right) \leq n\left(\frac{K}{n}\right)^2 \to 0,$$

vagyis $\mathbf{D}^2\left(\xi\right) = 0$, következésképpen ξ konstans.

$\square$

VI.

TÖBBDIMENZIÓS ELOSZLÁSOK

Ebben a fejezetben a többdimenziós eloszlások néhány jellemzőjét tárgyaljuk. A valószínűségszámítás kerek, egyszerű és hatékony matematikai elmélet, ha egyetlen valószínűségi változó jellemzőit kell tárgyalni. Az igen gyakran használt függetlenség feltétele azt biztosítja, hogy a többváltozós eloszlások vizsgálatát hatékonyan visszavezessük egyváltozós eloszlások vizsgálatára. Érdemes arra gondolni, hogy egy egyváltozós függvényről egyszerűen eldönthető, hogy eloszlásfüggvény vagy sem, de többváltozóban ez már távolról sem egyszerű kérdés. A függetlenség feltétele azonban a legtöbb vizsgálat során nem tartható, így valamiképpen kezelni kell a változók kapcsolatát. Ebben a fejezetben néhány ehhez kapcsolódó elemi fogalmat tárgyalunk.

6.1. Feltételes valószínűség és feltételes várható érték

Valószínűségi változók kapcsolatának kulcsa a feltételes valószínűség. A feltételes valószínűség egyértelműen definiált és egyszerűen megérthető és interpretálható, ha a feltétel valószínűsége pozitív. A nehézségek abból erednek, hogy nem világos, miként kell definiálni a feltételes valószínűséget akkor, ha a feltétel valószínűsége nulla. Tegyük fel, hogy definiálni akarjuk a $\mathbf{P}(A \mid N)$ valószínűséget, ahol az N esemény valószínűsége nulla. Ilyenkor a $\mathbf{P}(A \cap N) / \mathbf{P}(N)$ kifejezés értelmetlen, mert $0/0$ alakú kifejezésről van szó. Természetesen jön a gondolat, hogy miért nem definiáljuk a kifejezést határértékkel? Ha az N helyébe egy olyan $\widetilde{N}$ eseményt teszünk, amelyre $\mathbf{P}\left(\widetilde{N}\right)$ pozitív, akkor a $\mathbf{P}\left(A \mid \widetilde{N}\right)$ feltételes valószínűség már jól definiálható, és semmi más dolgunk nincs, mint tartani az $\widetilde{N}$ eseménnyel az N eseményhez. Ez valóban jó ötletnek tűnik, de sajnálatos módon az események körében nem definiálható a konvergencia fogalma[1]. Így tehát rögzített N esetén, ha $\mathbf{P}(N) = 0$, akkor a $\mathbf{P}(A \mid N)$ feltételes valószínűség még határértékkel sem definiálható, így értelmetlen. Vegyük észre, hogy ez a jelenség rokon a következő szituációval: Adott f függvény esetén keresendő az a g függvény, amelyre minden x-re $f(x) = \int_0^x g(t)\,dt$. Ilyen egyértelműen meghatározott g függvény nincsen, ugyanis egy függvényt csak akkor tekintünk meghatározottnak, ha az értelmezési tartományának minden pontjában egyértelműen megmondható, hogy mi az értéke. Ha tehát van egy g függvény, amely integráljaként az f előáll, akkor minden olyan függvény, amely csak néhány[2] pontban különbözik tőle szintén választható g függvényként. Vagyis a g mint egyedi függvény nem létezik, de mint függvények egy osztálya esetleg jól definiálható. Az már egy másik kérdés, hogy a lehetséges g függvények osztályában van-e olyan függvény, amely mondjuk folytonos, és ha igen, akkor érdemes-e ezt a reprezentánst azonosítani a számbajöhető g függvények osztályával.

A feltételes valószínűség legfőbb haszna, hogy a teljes valószínűség tétele segítségével

[1] Pontosabban az események valójában nem halmazok, hanem ekvivalenciaosztályok, így a konvergencia csak nulla valószínűségű halmazok erejéig definiálható.

[2] A kérdés erősen függ attól, hogy milyen integrálkonstrukcióról beszélünk. Az elemi valószínűségszámításban Riemann-integrálok szerepelnek, így a pontos karakterizáció nehezen adható meg.

a feltételes valószínűségekből kiszámolhatjuk a feltétel nélküli valószínűségeket. Ha az η változó diszkrét, akkor minden x esetén

$$\mathbf{P}(\xi < x) = \sum_k \mathbf{P}(\xi < x \mid \eta = y_k)\, \mathbf{P}(\eta = y_k),$$

ahol (y_k) az η lehetséges értékei. Ha $H(x)$ a ξ eloszlásfüggvénye és $G(y)$ az η eloszlásfüggvénye, akkor ezt a Stieltjes-integrálás jelölésével

$$H(x) = \int_{\mathbb{R}} \mathbf{P}(\xi < x \mid \eta = y)\, dG(y)$$

módon írhatjuk fel. Ebből látszik, hogy a feltételes valószínűséget bár nem tudjuk deriváltként értelmezni, de alkalmas integrálegyenlet megoldásaként esetleg meg tudjuk ragadni. A feltételes valószínűség szempontjából a fenti egyenlőség legfőbb problémája, hogy csak a H és a G peremeloszlásokat tartalmazza, így nem tükrözi az együttes eloszlást. Ezen könnyen segíthetünk, ha az általánosabb

$$F(x,y) \mathrel{\mathring{=}} \mathbf{P}(\xi < x, \eta < y) = \int_{-\infty}^{y} \mathbf{P}(\xi < x \mid \eta = u)\, dG(u)$$

egyenlőséget követeljük meg[3]. Ebből tetszőleges I intervallum esetén

$$\mathbf{P}(\xi < x, \eta \in I) = \int_I \mathbf{P}(\xi < x \mid \eta = y)\, dG(y).$$

Ha az η diszkrét és $I = [y_k, y_k]$, akkor

$$\mathbf{P}(\xi < x, \eta = y_k) = \mathbf{P}(\xi < x \mid \eta = y_k)\, \mathbf{P}(\eta = y_k),$$

amiből átosztva éppen a feltételes valószínűség már ismert definícióját kapjuk.

6.1.1. Definíció. Legyen $F(x,y)$ a (ξ,η) pár együttes eloszlásfüggvénye és jelölje $G(y)$ az η peremeloszlását. Ha egy alkalmas $F(x \mid y)$ módon jelölt függvényre tetszőleges x és y pontokra

$$F(x,y) = \int_{-\infty}^{y} F(x \mid u)\, dG(u), \tag{6.1.1}$$

akkor az $F(x \mid y)$ függvényt a ξ változó η-ra vonatkozó feltételes eloszlásfüggvényének mondjuk[4].

---

[3] A figyelmes olvasó észreveheti, hogy nem világos, hogy milyen értelemben létezik az integrál. Ez valóban probléma és ismét csak azt mondhatjuk, hogy a rendelkezésünkre álló eszközökkel nem tudjuk a kérdést megválaszolni.

[4] Vegyük észre, hogy sem az $F(x \mid y)$ létezésről, sem egyértelműségéről, sem tulajdonságairól nem beszélünk. Ezek tárgyalása elemi keretek között nem lehetséges.

6.1.2. Állítás. *Tegyük fel, hogy a (ξ,η) együttes eloszlásának van $f(x,y)$ sűrűségfüggvénye. Ha $g(y)$ az η sűrűségfüggvénye, akkor*

$$f(x \mid y) \circeq \frac{f(x,y)}{g(y)}$$

az $F(x \mid y)$ egy sűrűségfüggvénye, vagyis

$$F(x \mid y) = \int_{-\infty}^{x} f(v \mid y)\, dv,$$

ahol az $f(x \mid y)$ definíciója és a fenti egyenlőség az y változóban egy az η eloszlása szerint egy valószínűséggel rendelkező halmaz erejéig érvényes. Ha valamely y-ra $g(y) = 0$, akkor az egyszerűség kedvéért a hányados értékét is nullának definiáljuk és ilyenkor feltesszük, hogy az $f(x,y)$ is nulla, amely feltétel a sűrűségfüggvény definíciója alapján megengedett.

Bizonyítás: Fontos hangsúlyozni, hogy az egyenlőségek nem minden y esetén teljesülnek. Ha valamely y-ra a $g(y) = 0$, akkor az $f(x \mid y)$ képlete nincs is értelmezve. Nem beszélve arról, hogy már maguk a $g(y)$ és az $f(x,y)$ kifejezések sem függvények, hanem ekvivalenciaosztályok. Ez indokolja az állításban szereplő körülményes megfogalmazást. Az állítás indoklása viszonylag egyszerű: Meg kell mutatni, hogy az itt definiált $F(x \mid y)$ kielégíti a fenti (6.1.1) integrálegyenletet. Mivel az együttes eloszlásnak van sűrűségfüggvénye, így a peremeloszlásoknak is van, ezért

$$\begin{aligned}
\int_{-\infty}^{y} F(x \mid u)\, dG(u) &= \int_{-\infty}^{y}\int_{-\infty}^{x} f(v \mid u)\, dv dG(u) = \\
&= \int_{-\infty}^{y}\int_{-\infty}^{x} f(v \mid u)\, dv g(u)\, du = \\
&= \int_{-\infty}^{y}\int_{-\infty}^{x} f(v \mid u)\, g(u)\, dv du = \\
&= \int_{-\infty}^{y}\int_{-\infty}^{x} \frac{f(v,u)}{g(u)} g(u)\, dv du = \\
&= \int_{-\infty}^{y}\int_{-\infty}^{x} f(v,u)\, dv du = F(x,y).
\end{aligned}$$

$\square$

6.1.3. Definíció. Az elemi valószínűségszámítás keretében az $\mathbf{E}(\xi \mid \eta = y)$ feltételes várható értéket két esetben definiáljuk:

1. Ha az η változó diszkrét és a lehetséges értékei közül az egyik y_k, akkor az

$$F(x \mid y = y_k) \stackrel{\circ}{=} \frac{\mathbf{P}(\xi < x, \eta = y_k)}{\mathbf{P}(\eta = y_k)}$$

az x változóban egy közönséges eloszlásfüggvény, így az

$$\mathbf{E}(\xi \mid \eta = y_k) \stackrel{\circ}{=} \int_{\mathbb{R}} x dF(x \mid y = y_k)$$

definíció értelmes.

2. Ha a (ξ, η) párnak létezik sűrűségfüggvénye, akkor

$$\mathbf{E}(\xi \mid \eta = y) \stackrel{\circ}{=} \int_{\mathbb{R}} x \cdot f(x \mid y) dx.$$

Az olvasóban felmerülhet a kérdés, hogy miért nem alkalmaztuk az

$$\mathbf{E}(\xi \mid \eta = y) \stackrel{\circ}{=} \int_{\mathbb{R}} x dF(x \mid y)$$

jóval általánosabb és látszólag kézenfekvőbb definíciót. Bár nem teljesen nyilvánvaló, de erre jó okunk van, ugyanis az $F(x \mid y)$ kifejezés rögzített y-ra nem az x egy monoton függvénye és az sem világos, hogy van-e ilyen reprezentánsa a megfelelő függvényosztálynak[5]. Természetesen felmerülhet az a kérdés is, hogy a folytonos, a sűrűségfüggvénnyel rendelkező esetben, ez miért nem okoz gondot. Természetesen ilyenkor is gondot jelent a feltételes várható érték definíciója, és a konstrukcióból világos, hogy minden y esetén több érték is lehetséges, vagyis a feltételes várható érték ilyenkor sem egy függvény, sokkal inkább egy ekvivalenciaosztály. Ugyanakkor a feltételes várható értékként kapott ekvivalenciaosztály elemei nem lehetnek tetszőlegesek, ugyanis érvényes a következő állítás:

6.1.4. Állítás (Teljes várható érték tétele)**.** *Tegyük fel, hogy a ξ változónak van várható értéke. Ha a (ξ, η) eloszlásának van sűrűségfüggvénye és $G(y)$ az η eloszlásfüggvénye, akkor tetszőleges I intervallum esetén*

$$\int_I \mathbf{E}(\xi \mid \eta = y) dG(y) = \mathbf{E}(\xi \cdot \chi(\eta \in I)).$$

Speciálisan

$$\mathbf{E}(\xi) = \int_{-\infty}^{\infty} \mathbf{E}(\xi \mid \eta = y) dG(y).$$

[5] Ismét csak azt tudjuk mondani, hogy az ilyen reprezentáns létezésének igazolása nagyon messzire vezetne.

Bizonyítás: A transzformált valószínűségi változók várható értékére vonatkozó képlet szerint[6]

$$\int_I \mathbf{E}(\xi \mid \eta = y)\, dG(y) = \int_I \mathbf{E}(\xi \mid \eta = y)\, g(y)\, dy \overset{\circ}{=}$$

$$\overset{\circ}{=} \int_I \int_{\mathbb{R}} x \cdot f(x \mid y)\, dx g(y)\, dy \overset{\circ}{=} \int_I \int_{\mathbb{R}} x \frac{f(x,y)}{g(y)} g(y)\, dx dy =$$

$$= \int_I \int_{\mathbb{R}} x f(x,y)\, dx dy = \int_{\mathbb{R}} \int_I x f(x,y)\, dy dx = \int_{\mathbb{R}} \int_{\mathbb{R}} \chi(y \in I)\, x f(x,y)\, dy dx =$$

$$= \mathbf{E}(\chi(\eta \in I) \cdot \xi).$$

Érdemes megjegyezni, hogy az általános esetben az állításban szereplő első egyenlőség tekinthető a feltételes várható érték definíciójának is. Ugyanakkor a feltételes várható értéknek általában nincsen folytonos verziója, így az általános esetben az elemi valószínűségszámítás fegyverzetének ismeretében nem világos, hogy miként kell az integrált érteni.

□

6.1.5. Példa. Legyen (ξ, η) egyenletes eloszlású a $(0,0)$, $(1,0)$ és $(0,1)$ pontok által kifeszített háromszögön. Határozzuk meg az $\mathbf{E}(\xi \mid \eta = y)$ és az $\mathbf{E}(\eta \mid \xi = x)$ feltételes várható értékeket.

Szimmetriaokokból elég az egyiket kiszámolni. Jelölje Δ a háromszöget. Az együttes eloszlás sűrűségfüggvénye[7]

$$f(x,y) = \begin{cases} 2 & \text{ha} \quad (x,y) \in \Delta \\ 0 & \text{ha} \quad (x,y) \notin \Delta \end{cases}.$$

Az η peremeloszlása

$$g(y) = \int_{-\infty}^{\infty} f(x.y)\, dx = \begin{cases} 2(1-y) & \text{ha} \quad y \in (0,1) \\ 0 & \text{ha} \quad y \notin (0,1) \end{cases}.$$

Ebből

$$f(x \mid y) = \begin{cases} 1/(1-y) & \text{ha} \quad (x,y) \in \Delta \\ 0 & \text{ha} \quad (x,y) \notin \Delta \end{cases}.$$

Ha $0 < y < 1$, akkor a feltételes várható érték

$$\begin{aligned} \mathbf{E}(\xi \mid \eta = y) &= \int_{-\infty}^{\infty} x f(x \mid y)\, dx = \int_0^{1-y} \frac{x}{1-y} dx = \\ &= \frac{1}{1-y} \left[\frac{x^2}{2}\right]_0^{1-y} = \frac{1}{2}(1-y). \end{aligned}$$

[6] A két integrál felcserélhetőségét az biztosítja, hogy a ξ változónak létezik várható értéke és emiatt a kettős integrál abszolút konvergens.

[7] Pontosabban annak egy verziója.

Vegyük észre, hogy a $0 < y < 1$ tartományon kívül az $\mathbf{E}(\xi \mid \eta = y)$ kifejezés értelmetlen, vagy inkább bármi lehet. Így akár azt is mondhatjuk, hogy

$$\mathbf{E}(\xi \mid \eta = y) = (1-y)/2,$$

de például az

$$\mathbf{E}(\xi \mid \eta = y) = \chi_{(0,1)}(1-y)/2$$

egyenlőség is helyes.

□

6.1.6. Példa. Legyen a (ξ, η) pár együttes sűrűségfüggvénye

$$f(x,y) = \exp(-y), \quad y > x \geq 0.$$

Számoljuk ki a feltételes várható értékeket.

Először igazoljuk, hogy valóban sűrűségfüggvényről van szó. Mivel $f(x,y) \geq 0$ elég megmutatni, hogy a függvény integrálja 1 :

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = \int_0^{\infty}\int_0^{y} f(x,y)\,dxdy =$$
$$= \int_0^{\infty} e^{-y}\int_0^{y} 1dxdy = \int_0^{\infty} e^{-y}ydy = \Gamma(2) = 1.$$

Számoljuk ki a peremeloszlásokat és a feltételes sűrűségfüggvényeket:

$$\begin{aligned}
g(y) &\mathrel{\hat{=}} \int_{\mathbb{R}} f(x,y)\,dx = \int_0^{y} e^{-y}dx = ye^{-y}, \quad y > 0\\
f(x) &\mathrel{\hat{=}} \int_{\mathbb{R}} f(x,y)\,dy = \int_x^{\infty} e^{-y}dy = \left[-e^{-y}\right]_x^{\infty} = e^{-x}, \quad x > 0,\\
f(x \mid y) &= \frac{f(x,y)}{g(y)} = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y}, \quad 0 < x < y,\\
f(y \mid x) &= \frac{f(x,y)}{h(x)} = \frac{e^{-y}}{e^{-x}} = e^{x-y}, \quad 0 < x < y.
\end{aligned}$$

Így ha $y > 0$, illetve ha $x > 0$, akkor

$$\begin{aligned}
\mathbf{E}(\xi \mid \eta = y) &= \int_{\mathbb{R}} xf(x \mid y)\,dx = \int_0^{y} x\frac{1}{y}dx =\\
&= \frac{1}{y}\int_0^{y} xdx = \frac{y}{2},\\
\mathbf{E}(\eta \mid \xi = x) &= \int_{\mathbb{R}} yf(y \mid x)\,dy = \int_x^{\infty} ye^{x-y}dy =\\
&= e^x\int_x^{\infty} ye^{-y}dy = e^x\left(\left[-ye^{-y}\right]_x^{\infty} + \int_x^{\infty} e^{-y}dy\right) =\\
&= e^x\left(xe^{-x} + e^{-x}\right) = x + 1.
\end{aligned}$$

□

6.2. Korrelációs együttható

A többdimenziós eloszlások vizsgálatának leggyakrabban használt eszköze a korrelációs együttható[8].

6.2.1. Definíció. Legyenek ξ és η valószínűségi változók. A

$$\operatorname{cov}(\xi,\eta) \stackrel{\circ}{=} \mathbf{E}((\xi-\mathbf{E}(\xi))(\eta-\mathbf{E}(\eta)))$$

várható értéket a ξ és az η kovarianciájának, a

$$\operatorname{corr}(\xi,\eta) \stackrel{\circ}{=} \frac{\operatorname{cov}(\xi,\eta)}{\mathbf{D}(\xi)\mathbf{D}(\eta)}$$

hányadost pedig a ξ és az η korrelációs együtthatójának mondjuk.[9]

Emlékeztetünk, hogy a kovariancia elnevezést a következő állítás indokolja: A szórásnégyzetet szokás kovarianciának is mondani és tetszőleges ξ és η valószínűségi változókra

$$\mathbf{D}^2(\xi+\eta) = \mathbf{D}^2(\xi)+\mathbf{D}^2(\eta)+2\cdot\operatorname{cov}(\xi,\eta).$$

Ennek bizonyítása során beláttuk a következő fontos összefüggést: Ha ξ és η valószínűségi változók, akkor

$$\operatorname{cov}(\xi,\eta) = \mathbf{E}(\xi\eta)-\mathbf{E}(\xi)\mathbf{E}(\eta).$$

Speciálisan, ha ξ és η függetlenek, akkor a kovarianciájuk nulla.

6.2.2. Példa. Valószínűségi változók korrelációs együtthatója nulla, de a változók nem függetlenek.

A számos lehetséges példa közül talán a legegyszerűbb a következő: Legyen ξ standard normális eloszlású változó és $\eta \stackrel{\circ}{=} \xi^2$. Ilyenkor $\xi\eta = \xi^3$, amely várható értéke nulla. De ugyancsak nulla az $\mathbf{E}(\xi)\mathbf{E}(\eta)$ szorzat is, így a korrelációs együttható értéke nulla. Egy

[8] Általában, némiképpen pontatlanul, szokás pusztán korrelációról is beszélni, amely alatt az együttmozgás „szorosságát" szokás érteni. A korrelációs együttható a korrelációt próbálja, jól-rosszul, mérni. Szokás a korreláció erősségéről is beszélni. Ilyenkor az együttmozgás „szorosságát" akarjuk megragadni, előjeltől függetlenül. A korreláció mérésére vannak egyéb mutatók is nem csak a korrelációs együttható.

[9] A korrelációs együttható létezéséhez a két szórásnak létezni kell és pozitívnak kell lenni. Vagyis egy konstans változó és egy másik változónak nincsen korrelációs együtthatója ámbár a két változónak van szórása és függetlenek. Ez analóg avval, hogy a nulla vektornak nincs közrezárt szöge egy másik nem nulla vektorral.

másik tanulságos példa a következő: Ha a (ξ,η) eloszlása az y tengelyre szimmetrikus, akkor a (ξ,η) eloszlása megegyezik a $(-\xi,\eta)$ eloszlásával. Ebből következően

$$\operatorname{cov}(\xi,\eta)=\operatorname{cov}(-\xi,\eta)=-\operatorname{cov}(\xi,\eta),$$

vagyis $\operatorname{cov}(\xi,\eta)=0$. Így például a $(-1,0),(1,0),(0,1)$ háromszögön egyenletes eloszlású valószínűségi változó korrelációs együtthatója nulla, de a ξ és az η nem függetlenek. De példaként vehetjük az egységkörlapon egyenletes eloszlású (ξ,η) valószínűségi változókat is, amelyek korrelációs együtthatója nulla, de a ξ és az η nem független.

□

6.2.3. Állítás (Cauchy-egyenlőtlenség)**.** *Tetszőleges ξ és η változók esetén*

$$\mathbf{E}^2(\xi\eta)\le\mathbf{E}\left(\xi^2\right)\mathbf{E}\left(\eta^2\right)$$

Speciálisan $|\operatorname{corr}(\xi,\eta)|\le 1$, *és az egyenlőség szükséges és elegendő feltétele, hogy a ξ lineáris függvénye legyen az η-nak.*

Bizonyítás: Tetszőleges λ esetén

$$0\le\mathbf{E}\left((\xi+\lambda\eta)^2\right)=\lambda^2\mathbf{E}\left(\eta^2\right)+2\lambda\mathbf{E}(\xi\eta)+\mathbf{E}\left(\xi^2\right).$$

Ez minden λ esetén csak akkor lehetséges, ha a $D\stackrel{\circ}{=}4\mathbf{E}^2(\xi\eta)-4\mathbf{E}\left(\xi^2\right)\mathbf{E}\left(\xi^2\right)$ diszkrimináns nem pozitív, vagyis

$$D=4\mathbf{E}^2(\xi\eta)-4\mathbf{E}\left(\xi^2\right)\mathbf{E}\left(\xi^2\right)\le 0,$$

amiből az egyenlőtlenség nyilvánvaló. Az egyenlőtlenség egy fontos következménye, hogy ha ξ és η második momentumai végesek, akkor a kovarianciájuk is véges. Éppen ezért a korrelációs együtthatót és a kovarianciát csak olyan valószínűségi változókra definiáljuk, amelyek második momentuma véges. A ξ és az η helyébe a $(\xi-\mathbf{E}(\xi))/\mathbf{D}(\xi)$ és az $(\eta-\mathbf{E}(\eta))/\mathbf{D}(\eta)$ centralizált és normalizált változókat téve az egyenlőség vizsgálatához elegendő feltenni, hogy $\mathbf{E}(\xi)=\mathbf{E}(\eta)=0$ és $\mathbf{D}(\xi)=\mathbf{D}(\eta)=1$. Ha $\operatorname{corr}(\xi,\eta)=1$, akkor

$$\begin{aligned}\mathbf{E}\left((\xi-\eta)^2\right)&=\mathbf{E}\left(\xi^2\right)+\mathbf{E}\left(\eta^2\right)-2\mathbf{E}(\xi\eta)=\\&=\mathbf{D}^2(\xi)+\mathbf{D}^2(\eta)-2\operatorname{corr}(\xi,\eta)=\\&=1+1-2=0,\end{aligned}$$

vagyis $\xi=\eta$. A $\operatorname{corr}(\xi,\eta)=-1$ tárgyalása analóg, illetve a fordított irány evidens.

□

6.2.4. Példa. Legyen a (ξ,η) pár egyenletes eloszlású a $(0,0)$, $(1,0)$, $(0,1)$ háromszögön. Számoljuk ki a ξ és az η korrelációs együtthatóját.

Emlékeztetünk, hogy a ξ és az η sűrűségfüggvénye

$$g(x) = h(x) = 2(1-x), \quad 0 < x < 1.$$

Ebből

$$\begin{aligned}
\mathbf{E}(\xi) &= \mathbf{E}(\eta) = 2\int_0^1 x(1-x)\,dx = \frac{1}{3}. \\
\mathbf{E}\left(\xi^2\right) &= \mathbf{E}\left(\eta^2\right) = 2\int_0^1 x^2(1-x)\,dx = \frac{1}{6} \\
\mathbf{D}^2(\xi) &= \mathbf{D}^2(\eta) = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}.
\end{aligned}$$

Ahhoz, hogy a korrelációs együtthatót ki tudjuk számolni elegendő meghatározni az $\mathbf{E}(\xi\eta)$ várható értéket. A (ξ,η) együttes sűrűségfüggvénye a háromszögön 2, így

$$\begin{aligned}
\mathbf{E}(\xi\eta) &= \int_{\mathbf{R}}\int_{\mathbf{R}} xyf(x,y)\,dydx = \int_0^1\int_0^{1-x} 2xydydx = \\
&= 2\int_0^1 x\left[\frac{y^2}{2}\right]_0^{1-x} dydx = \int_0^1 x(1-x)^2\,dx = \frac{1}{12}.
\end{aligned}$$

Ebből a korrelációs együttható

$$\operatorname{corr}(\xi,\eta) = \frac{1/12 - 1/9}{1/18} = -\frac{1}{2}.$$

□

6.2.5. Példa. Legyen (ξ,η) egyenletes eloszlású a

$$K \mathrel{\hat{=}} \left\{(x,y) \mid x^2+y^2 \leq 1, x,y \geq 0\right\}$$

negyedkörlapon. Számoljuk ki a korrelációs együtthatóját.

Az együttes eloszlás sűrűségfüggvénye

$$f(x,y) = \begin{cases} 4/\pi & \text{ha} \quad (x,y)\in K \\ 0 & \text{ha} \quad (x,y)\notin K \end{cases}.$$

Szimmetriaokokból elegendő az egyik peremeloszlással foglalkozni. Ha $0 < x < 1$, akkor

$$h(x) = \int_{\mathbb{R}} f(x,y)\,dy = \int_0^{\sqrt{1-x^2}} \frac{4}{\pi}dy = \frac{4}{\pi}\sqrt{1-x^2}.$$

A várható érték

$$\begin{aligned}
\mathbf{E}(\xi) &= \frac{4}{\pi}\int_0^1 x\sqrt{1-x^2}dx = \frac{4}{\pi}\frac{1}{3}.\\
\mathbf{E}\left(\xi^2\right) &= \frac{4}{\pi}\int_0^1 x^2\sqrt{1-x^2}dx = \frac{4}{\pi}\frac{1}{16}\pi = \frac{1}{4}.\\
\mathbf{D}^2(\xi) &= \frac{1}{4}-\frac{16}{9\pi^2}.
\end{aligned}$$

A korrelációs együttható meghatározásához ki kell még számolni az $\mathbf{E}(\xi\eta)$ várható értéket.

$$\begin{aligned}
\mathbf{E}(\xi\eta) &= \frac{4}{\pi}\int_0^1\int_0^{\sqrt{1-x^2}} xydydx =\\
&= \frac{4}{\pi}\int_0^1 x\int_0^{\sqrt{1-x^2}} ydydx =\\
&= \frac{4}{\pi}\int_0^1 x\left[\frac{y^2}{2}\right]_0^{\sqrt{1-x^2}} dx =\\
&= \frac{2}{\pi}\int_0^1 x\left(1-x^2\right)dx = \frac{2}{\pi}\frac{1}{4} = \frac{1}{2\pi}.
\end{aligned}$$

A korrelációs együttható tehát

$$\operatorname{corr}(\xi,\eta) = \frac{1/(2\pi)-16/\left(9\pi^2\right)}{1/4-16/\left(9\pi^2\right)} = -0,30014.$$

Vegyük észre, hogy hasonlóan az előző példához a korrelációs együttható negatív, de értéke közelebb van a nullához, ugyanis a K negyedkör tartalmazza az előző példában szereplő háromszöget, vagyis a ξ és az η közötti kapcsolat a negyedkör esetén gyengébb.

□

6.2.6. Példa. Legyen a (ξ,η) egyenletes eloszlású a

$$P \doteq \left\{(x,y)\in[0,1]^2 \mid y\le x^2\right\}$$

halmazon. Számoljuk ki a $\operatorname{corr}(\xi,\eta)$ korrelációs együtthatót.

A parabola alatti terület $1/3$, így

$$f(x,y) = \begin{cases} 3 & \text{ha} \quad (x,y)\in P\\ 0 & \text{ha} \quad (x,y)\notin P\end{cases}.$$

A peremeloszlások $0 \leq x \leq 1$, illetve $0 \leq y \leq 1$ esetén

$$\begin{aligned} h(x) &= \int_0^{x^2} 3dy = 3x^2, \\ g(y) &= \int_{\sqrt{y}}^1 3dx = 3(1-\sqrt{y}). \end{aligned}$$

Ebből

$$\begin{aligned} \mathbf{E}(\xi) &= \int_0^1 3x^3 dx = \frac{3}{4}, \quad \mathbf{E}\left(\xi^2\right) = \int_0^1 3x^4 dx = \frac{3}{5}, \\ \mathbf{D}(\xi) &= \sqrt{\frac{3}{5} - \frac{9}{16}} = \sqrt{\frac{3}{80}}, \end{aligned}$$

illetve

$$\begin{aligned} \mathbf{E}(\eta) &= 3\int_0^1 y(1-\sqrt{y})\,dy = \frac{3}{10}, \\ \mathbf{E}\left(\eta^2\right) &= 3\int_0^1 y^2(1-\sqrt{y})\,dy = \frac{1}{7}, \\ \mathbf{D}(\eta) &= \sqrt{\frac{1}{7} - \frac{9}{100}} = \sqrt{\frac{37}{700}}. \end{aligned}$$

A szorzat várható értéke

$$\begin{aligned} \mathbf{E}(\xi\eta) &= \int_0^1 \int_0^{x^2} 3xydydx = 3\int_0^1 x\left[\frac{y^2}{2}\right]_0^{x^2} dx = \\ &= \frac{3}{2}\int_0^1 x^5 dx = \frac{1}{4}. \end{aligned}$$

Így a korrelációs együttható

$$\operatorname{corr}(\xi,\eta) = \frac{1/4 - 3/4 \cdot 3/10}{\sqrt{3/80}\sqrt{37/700}} = 0,56153.$$

Vegyük észre, hogy az y-tengelyre való tükrözés során a korrelációs együttható előjelet vált, illetve az x-tengely mentén való eltolás nem módosítja a korrelációs együtthatót, így a $(0,0)$, $(1,0)$, $(1,1)$ háromszögön egyenletes eloszlású változópár korrelációs együtthatója $1/2$. Mivel a parabola az $y = x$ egyenes alatt halad a P része ennek a háromszögnek, így a korrelációs együtthatója nagyobb.

$\square$

VII.

GAMMA-BÉTA ARITMETIKA

Az előző fejezettel befejeztük a valószínűségszámítás alapfogalmainak ismertetését. Ebben a rövid technikai fejezetben a gamma és a béta függvényekkel kapcsolatos legfontosabb ismereteket elevenítjük fel. A gamma és a béta függvények a valószínűségszámítás legfontosabb eszközei közé tartoznak és korábban is már érintőlegesen használtuk őket.

7.1. A gamma és béta függvények

Először tekintsük a definíciókat[1]:

7.1.1. Definíció. A

$$\Gamma(x) \circeq \int_0^\infty t^{x-1}\exp(-t)\,dt, \quad x>0$$

függvényt gamma függvénynek, a

$$B(x,y) \circeq \int_0^1 t^{x-1}(1-t)^{y-1}\,dt, \quad x,y>0$$

függvényt béta függvénynek mondjuk. A gamma függvény mellett vezessük be a

$$\Gamma(x,\lambda) \circeq \int_0^\infty t^{x-1}\exp(-\lambda t)\,dt, \quad x,\lambda>0.$$

függvényt. Egyszerű $u=t\lambda$ helyettesítéssel

$$\Gamma(x,\lambda) = \int_0^\infty \left(\frac{u}{\lambda}\right)^{x-1}\exp(-u)\,\frac{du}{\lambda} = \frac{\Gamma(x)}{\lambda^x}.$$

7.1.2. Példa. Ha $x>0$, akkor $\Gamma(x+1)=x\Gamma(x)$. Speciálisan ha $n>0$ egész, akkor $\Gamma(n)=(n-1)!$

Parciálisan integrálva, és felhasználva, hogy $x>0$, így a t^x a $t=0$ pontban nulla,

$$\begin{aligned}\Gamma(x+1) &= \int_0^\infty t^x\exp(-t)\,dt = \left[t^x\frac{\exp(-t)}{-1}\right]_0^\infty + \int_0^\infty xt^{x-1}\exp(-t)\,dt = \\ &= 0+x\int_0^\infty t^{x-1}\exp(-t)\,dt = x\Gamma(x).\end{aligned}$$

□

7.1.3. Példa. Igazoljuk a

$$B(x,y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

formulát!

[1] A gamma és béta függvényeket szokás szélesebb értelmezési tartományok felett is definiálni. Ilyenkor az integrálelőállítás a szélesebb értelmezési tartományokon már nem érvényes.

A formula indoklásának számos módja ismert. Mindegyik bizonyítás során valamilyen analízisből ismert tételt fel kell használni. Az itt bemutatott bizonyítás arra alapszik, hogy nem negatív függvények esetén az integrálás sorrendje felcserélhető. Ezt a tételt az analízisben Fubini-tételként szokás emlegetni. $s = t/(1-t)$ helyettesítéssel

$$\begin{aligned}
I &\overset{\circ}{=} \int_0^\infty \Gamma(x+y, 1+s)\, s^{x-1} ds = \int_0^\infty \Gamma(x+y)\, \frac{s^{x-1}}{(1+s)^{x+y}} ds = \\
&= \Gamma(x+y) \int_0^\infty (1+s)^{-(x+y)}\, s^{x-1} ds = \\
&= \Gamma(x+y) \int_0^1 \left(1 + \frac{t}{1-t}\right)^{-(x+y)} \left(\frac{t}{1-t}\right)^{x-1} \frac{1}{(1-t)^2} dt = \\
&= \Gamma(x+y) \int_0^1 \left(\frac{1}{1-t}\right)^{-(x+y)} \left(\frac{t}{1-t}\right)^{x-1} \frac{1}{(1-t)^2} dt = \\
&= \Gamma(x+y) \int_0^1 t^{x-1} (1-t)^{y-1} dt \overset{\circ}{=} \Gamma(x+y)\, B(x,y).
\end{aligned}$$

Az integrandus folytonos és nem negatív, ezért a Fubini-tételnek megfelelően az alább a két integrál felcserélhető:

$$\begin{aligned}
I &\overset{\circ}{=} \int_0^\infty \Gamma(x+y, 1+s)\, s^{x-1} ds \overset{\circ}{=} \\
&\overset{\circ}{=} \int_0^\infty \int_0^\infty \exp(-t(1+s))\, t^{x+y-1} s^{x-1} dt ds = \\
&= \int_0^\infty \int_0^\infty \exp(-t(1+s))\, t^{x+y-1} s^{x-1} ds dt = \\
&= \int_0^\infty t^{x+y-1} \exp(-t) \int_0^\infty \exp(-ts)\, s^{x-1} ds dt \overset{\circ}{=} \\
&\overset{\circ}{=} \int_0^\infty t^{x+y-1} \exp(-t)\, \Gamma(x,t)\, dt = \\
&= \int_0^\infty t^{x+y-1} \exp(-t)\, \frac{\Gamma(x)}{t^x} dt = \\
&= \Gamma(x) \int_0^\infty t^{y-1} \exp(-t)\, dt = \Gamma(x)\, \Gamma(y).
\end{aligned}$$

Ebből

$$\Gamma(x+y)\, B(x,y) = \Gamma(x)\, \Gamma(y),$$

amiből az azonosság már evidens módon adódik.

□

7.1.4. Példa. Ha $x > 0$, akkor

$$B(x+1, y) = \frac{x}{x+y} B(x,y).$$

A fenti összefüggés szerint

$$B(x+1,y)=\frac{\Gamma(x+1)\Gamma(y)}{\Gamma(x+y+1)}=\frac{x\Gamma(x)\Gamma(y)}{(x+y)\Gamma(x+y)}=\frac{x}{x+y}B(x,y).$$

□

Az integrálok felcserélése egy igen hatékony eszköz. Erre további példaként tekintsük a következőt:

7.1.5. Példa. Számoljuk ki az

$$\int_{-\infty}^{\infty}\exp\left(-x^2\right)dx$$

integrált!

Először

$$\begin{aligned} I &\stackrel{\circ}{=} \int_0^\infty\int_0^\infty x\exp\left(-x^2\left(1+t^2\right)\right)dxdt= \\ &= \int_0^\infty\left[\frac{\exp\left(-x^2\left(1+t^2\right)\right)}{-2\left(1+t^2\right)}\right]_0^\infty dt= \\ &= \frac{1}{2}\int_0^\infty\frac{1}{1+t^2}dt=\frac{\pi}{4}. \end{aligned}$$

Az $x\exp\left(-x^2\left(1+t^2\right)\right)$ függvény az $x,t\geq 0$ tartományon folytonos, és nem negatív, ezért alkalmazható rá a Fubini-tétel. $u=xt$ helyettesítéssel

$$\begin{aligned} I &= \int_0^\infty\int_0^\infty x\exp\left(-x^2\left(1+t^2\right)\right)dtdx= \\ &= \int_0^\infty x\exp\left(-x^2\right)\int_0^\infty\exp\left(-(xt)^2\right)dtdx= \\ &= \int_0^\infty x\exp\left(-x^2\right)\int_0^\infty\exp\left(-u^2\right)\frac{du}{x}dx=\left(\int_0^\infty\exp\left(-x^2\right)dx\right)^2, \end{aligned}$$

vagyis $\int_0^\infty\exp\left(-x^2\right)dx=\sqrt{\pi}/2$. Az $\exp\left(-x^2\right)$ páros volta alapján

$$\int_{-\infty}^{\infty}\exp\left(-x^2\right)dx=\sqrt{\pi}.$$

□

7.1.6. Példa. Igazoljuk a $\Gamma(1/2)=\sqrt{\pi}$ egyenlőséget.

$u^2 = x$ helyettesítéssel

$$\begin{aligned}\Gamma\left(\frac{1}{2}\right) &\overset{\circ}{=} \int_0^\infty x^{-1/2} \exp(-x)\,dx = \int_0^\infty \left(u^2\right)^{-1/2} \exp\left(-u^2\right) 2u\,du = \\ &= 2\int_0^\infty \exp\left(-u^2\right) du = \int_{-\infty}^\infty \exp\left(-u^2\right) du = \sqrt{\pi}.\end{aligned}$$

Egy másik bizonyítás a gamma és a béta függvények közötti azonosságra épül:

$$\begin{aligned}\left(\Gamma\left(\frac{1}{2}\right)\right)^2 &= \Gamma(1)\,B\left(\frac{1}{2},\frac{1}{2}\right) = 1\cdot B\left(\frac{1}{2},\frac{1}{2}\right) = \\ &= \int_0^1 \frac{1}{\sqrt{x(1-x)}}\,dx.\end{aligned}$$

Az integrálban $x = \sin^2 t$ helyettesítést végezve

$$\frac{dx}{dt} = 2\sin t \cos t$$

$$\int_0^{\pi/2} \frac{1}{\sqrt{\sin^2 t\left(1-\sin^2 t\right)}} 2\sin t\cos t\,dt =$$

$$2\int_0^{\pi/2} 1\,dt = 2\cdot\frac{\pi}{2} = \pi.$$

amiből a $\Gamma(1/2) = \sqrt{\pi}$ már evidens.

□

7.1.7. Példa. Számoljuk ki a standard normális eloszlás várható értékét és szórását.

Az

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

sűrűségfüggvényhez tartozó eloszlást standard normális eloszlásnak mondjuk.

1. $f \geq 0$, így az f sűrűségfüggvény voltához elég megmutatni, hogy az f integrálja a számegyenesen éppen 1. A kézenfekvő $u = x/\sqrt{2}$ helyettesítéssel

$$\begin{aligned}\int_{-\infty}^\infty f(x)\,dx &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \exp\left(-\frac{x^2}{2}\right) dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \exp\left(-u^2\right)\sqrt{2}\,du = \\ &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^\infty \exp\left(-u^2\right) du = 1.\end{aligned}$$

2. A várható érték meghatározása igen egyszerű. Mivel a sűrűségfüggvény páros, így az $xf(x)$ páratlan, és mivel az $xf(x)$ integrálható, ezért

$$\mathbf{E}(\xi) = \int_{-\infty}^{\infty} xf(x)\,dx = 0.$$

3. A szórás kiszámolásához számoljuk ki a második momentumot: Az igen kézenfekvő $u = x^2$ helyettesítéssel

$$\begin{aligned}
\mathbf{E}\left(\xi^2\right) &= \int_{-\infty}^{\infty} x^2 f(x)\,dx = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} x^2 \exp\left(-\frac{x^2}{2}\right)dx = \\
&= \frac{2}{\sqrt{2\pi}}\int_0^{\infty} u \exp\left(-\frac{u}{2}\right)\frac{1}{2\sqrt{u}}du = \\
&= \frac{1}{\sqrt{2\pi}}\int_0^{\infty} u^{1/2}\exp\left(-\frac{u}{2}\right)du = \\
&= \frac{1}{\sqrt{2\pi}}\Gamma\left(\frac{3}{2},\frac{1}{2}\right) = \frac{1}{\sqrt{2\pi}}\frac{\Gamma(3/2)}{(1/2)^{3/2}} = \\
&= \frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{2}}\frac{1}{2}2^{3/2}\Gamma\left(\frac{1}{2}\right) = 1.
\end{aligned}$$

Mivel a várható érték nulla, ezért a szórás $\mathbf{D}(\xi) = \sqrt{\mathbf{E}\left(\xi^2\right) - \mathbf{E}^2(\xi)} = 1.$

□

7.2. Gamma és béta eloszlások

Mivel a gammafüggvény és a bétafüggvény integrandusa nem negatív, az integrálokkal leosztva egyszerűen konstruálhatunk sűrűségfüggvényeket.

7.2.1. Definíció. Ha λ és a pozitív számok, akkor az

$$f(x) \circeq \frac{\lambda^a}{\Gamma(a)}x^{a-1}\exp(-\lambda x), \quad x > 0$$

sűrűségfüggvénnyel rendelkező eloszlást (a,λ) paraméterű gamma eloszlásnak hívjuk és $\Gamma(a,\lambda)$ módon jelöljük[2].

[2]Nem keverendő össze, a hasonlóan jelölt általánosított gammafüggvénnyel. Bár a jelölés azonos, az egyik egy eloszlás, a másik egy függvény. A némiképpen zavaró jelölés magyarázata a két fogalom igen szoros kapcsolata.

7.2.2. Definíció. Ha α és β pozitív paraméterek, akkor az

$$f(x) \stackrel{\circ}{=} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad x \in (0,1)$$

sűrűségfüggvénnyel rendelkező eloszlást (α,β) paraméterű béta eloszlásnak hívjuk és $B(\alpha,\beta)$ módon jelöljük.

7.2.3. Definíció. Ha α és β pozitív paraméterek, akkor a

$$g(x) \stackrel{\circ}{=} \frac{1}{B(\alpha,\beta)} x^{\alpha-1} \frac{1}{(1+x)^{\alpha+\beta}}, \quad x > 0$$

sűrűségfüggvénnyel rendelkező eloszlást általánosított, vagy másodfajú béta eloszlásnak nevezzük. Az eloszlást $\widetilde{B}(\alpha,\beta)$-val fogjuk jelölni. Ha $\alpha = 1$, akkor a

$$\begin{aligned} g(x) &\stackrel{\circ}{=} \frac{1}{B(1,\beta)} \frac{1}{(1+x)^{1+\beta}} = \\ &= \frac{\Gamma(1+\beta)}{\Gamma(1)\Gamma(\beta)} \frac{1}{(1+x)^{1+\beta}} = \\ &= \frac{\beta}{(1+x)^{1+\beta}} \end{aligned}$$

sűrűségfüggvényhez tartozó eloszlást szokás Pareto-eloszlásnak mondani.

A két béta eloszlás kapcsolata egyszerűen megvilágítható:

7.2.4. Állítás. *Ha ξ béta eloszlású, akkor az $\eta \stackrel{\circ}{=} \xi/(1-\xi)$ változó másodfajú béta eloszlású. Ha η másodfajú béta eloszlású, akkor a $\xi \stackrel{\circ}{=} \eta/(1+\eta)$ változó béta eloszlású.*

Bizonyítás: Ha $\varphi(u) \stackrel{\circ}{=} u/(1-u)$, akkor $\varphi^{-1}(x) = x/(1+x)$, és a sűrűségfüggvények transzformációs szabálya szerint

$$\begin{aligned} g(x) &= f\left(\varphi^{-1}(x)\right) \frac{d}{dx} \varphi^{-1}(x) = \\ &= \frac{1}{B(\alpha,\beta)} \left(\frac{x}{1+x}\right)^{\alpha-1} \left(1 - \frac{x}{1+x}\right)^{\beta-1} \frac{1}{(1+x)^2} = \\ &= \frac{1}{B(\alpha,\beta)} x^{\alpha-1} \frac{1}{(1+x)^{\alpha+\beta}}. \end{aligned}$$

A fordított irány igazolása analóg.

□

7.2.5. Példa. Számoljuk ki a gamma eloszlás várható értékét és szórását.

Mivel ismert a sűrűségfüggvény, a várható érték

$$\begin{aligned}
\mathbf{E}(\xi) &= \int_0^\infty x \frac{\lambda^a}{\Gamma(a)} x^{a-1} \exp(-\lambda x)\, dx = \\
&= \frac{\lambda^a}{\Gamma(a)} \int_0^\infty x^a \exp(-\lambda x)\, dx = \\
&= \frac{\lambda^a}{\Gamma(a)} \frac{\Gamma(a+1)}{\lambda^{a+1}} = \frac{a}{\lambda}.
\end{aligned}$$

A második momentum

$$\begin{aligned}
\mathbf{E}\left(\xi^2\right) &= \int_0^\infty x^2 \frac{\lambda^a}{\Gamma(a)} x^{a-1} \exp(-\lambda x)\, dx = \\
&= \frac{\lambda^a}{\Gamma(a)} \int_0^\infty x^{a+1} \exp(-\lambda x)\, dx = \\
&= \frac{\lambda^a}{\Gamma(a)} \frac{\Gamma(a+2)}{\lambda^{a+2}} = \frac{a(a+1)}{\lambda^2}.
\end{aligned}$$

Ebből a szórás

$$\mathbf{D}(\xi) = \sqrt{\frac{a(a+1)}{\lambda^2} - \frac{a^2}{\lambda^2}} = \frac{\sqrt{a}}{\lambda}.$$

□

7.2.6. Példa. Számoljuk ki a béta eloszlás várható értékét és szórását.

A várható érték képlete alapján, felhasználva a gamma és a béta függvények közötti összefüggést,

$$\begin{aligned}
\mathbf{E}(\xi) &= \frac{1}{B(\alpha,\beta)} \int_0^1 x \cdot x^{\alpha-1} (1-x)^{\beta-1}\, dx = \\
&= \frac{1}{B(\alpha,\beta)} B(\alpha+1,\beta) = \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)} = \frac{\alpha}{\alpha+\beta}.
\end{aligned}$$

A második momentum

$$\begin{aligned}
\mathbf{E}\left(\xi^2\right) &= \frac{1}{B(\alpha,\beta)} \int_0^1 x^2 \cdot x^{\alpha-1} (1-x)^{\beta-1}\, dx = \\
&= \frac{1}{B(\alpha,\beta)} B(\alpha+2,\beta) = \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)} = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}.
\end{aligned}$$

Ebből a szórás

$$
\begin{aligned}
\mathbf{D}(\xi) &= \sqrt{\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2} = \\
&= \sqrt{\frac{\alpha(\alpha+1)(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta)^2(\alpha+\beta+1)}} = \\
&= \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}}.
\end{aligned}
$$

□

7.2.7. Példa. Számoljuk ki a másodfajú béta eloszlás momentumait.

A másodfajú béta eloszlás n-edik momentuma

$$\mathbf{E}(\xi^n) = \frac{1}{B(\alpha,\beta)} \int_0^\infty x^{\alpha+n-1} \frac{1}{(1+x)^{\alpha+\beta}} dx.$$

A másodfajú béta eloszlás

$$g(x) \circeq \frac{1}{B(\alpha,\beta)} x^{\alpha-1} \frac{1}{(1+x)^{\alpha+\beta}}$$

sűrűségfüggvénye alapján

$$
\begin{aligned}
\int_0^\infty x^{\alpha+n-1} \frac{1}{(1+x)^{\alpha+\beta}} dx &= B(\alpha+n, \beta+\alpha-\alpha-n) = \\
&= B(\alpha+n, \beta-n).
\end{aligned}
$$

Ebből

$$
\begin{aligned}
\mathbf{E}(\xi^n) &= \frac{B(\alpha+n,\beta-n)}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+n)\Gamma(\beta-n)}{\Gamma(\alpha+\beta)} \frac{\Gamma(a+\beta)}{\Gamma(\alpha)\Gamma(\beta)} = \\
&= \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)} \frac{\Gamma(\beta-n)}{\Gamma(\beta)} = \frac{\alpha(\alpha+1)\cdots(\alpha+n-1)}{(\beta-1)\cdots(\beta-n)}.
\end{aligned}
$$

A levezetésből világos, hogy az $\mathbf{E}(\xi^n)$ momentum pontosan akkor véges, ha $\beta - n > 0$, vagyis $\beta > n$. Egy másik bizonyítás a következő: Miként láttuk, ha a ξ másodfajú béta eloszlású, akkor felírható $\eta/(1-\eta)$ alakban, ahol az η béta eloszlású. Így

$$
\begin{aligned}
\mathbf{E}(\xi^n) &= \mathbf{E}\left(\left(\frac{\eta}{1-\eta}\right)^n\right) = \frac{1}{B(\alpha,\beta)} \int_0^1 \left(\frac{x}{1-x}\right)^n x^{\alpha-1}(1-x)^{\beta-1} dx = \\
&= \frac{B(\alpha+n,\beta-n)}{B(\alpha,\beta)}.
\end{aligned}
$$

□

7.2.8. Példa. Számoljuk ki a másodfajú béta eloszlás várható értékét.

Az előző példa alapján ha $n = 1$, akkor

$$\mathbf{E}(\xi) = \frac{\alpha}{\beta - 1}.$$

A

$$\begin{aligned} B(\alpha+1,\beta) &= \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)} = \frac{\alpha}{\alpha+\beta}\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \\ &= \frac{\alpha}{\alpha+\beta}B(\alpha,\beta) \end{aligned}$$

függvényegyenlet alapján számolva és felhasználva, hogy a béta függvény szimmetrikus,

$$\begin{aligned} \mathbf{E}(\xi) &= \frac{1}{B(a,\beta)}\int_0^1 \frac{x}{x-1}\cdot x^{\alpha-1}(1-x)^{\beta-1}\,dx = \\ &= \frac{B(\alpha+1,\beta-1)}{B(\alpha,\beta)} = \\ &= \frac{\alpha}{\alpha+\beta-1}\frac{B(\alpha,\beta-1)}{B(\alpha,\beta)} = \\ &= \frac{\alpha}{\alpha+\beta-1}\frac{\alpha+\beta-1}{\beta-1}\frac{B(\alpha,\beta-1)}{B(\alpha,\beta-1)} = \\ &= \frac{\alpha}{\beta-1}. \end{aligned}$$

□

7.2.9. Példa. Az exponenciális és a χ_1^2 eloszlások mint gamma eloszlások.

Gamma eloszlásra a legegyszerűbb példa a $\Gamma(1,\lambda)$ eloszlás. Ennek sűrűségfüggvénye

$$f(x) = \frac{\lambda^1}{\Gamma(1)}x^{1-1}\exp(-\lambda x) = \lambda\exp(-\lambda x), \quad x > 0,$$

ami éppen az exponenciális eloszlás sűrűségfüggvénye. Vegyük észre, hogy mivel $a = 1$, ezért ilyenkor a várható értékre és a szórásra kapott képlet éppen visszaadja az exponenciális eloszlásra korábban kapott $1/\lambda$ értéket. További alapvető példa a χ_1^2

eloszlás. Definíció szerint χ_1^2 eloszláson az $N(0,1)$ standard normális eloszlás négyzetének eloszlását értjük. A már többször látott utat követve, ha a ξ standard normális eloszlású, és $x>0$, akkor

$$\begin{aligned}\mathbf{P}\left(\xi^2<x\right) &= \mathbf{P}\left(|\xi|<\sqrt{x}\right)=\frac{1}{\sqrt{2\pi}}\int_{-\sqrt{x}}^{\sqrt{x}}\exp\left(-\frac{t^2}{2}\right)dt= \\ &= \frac{2}{\sqrt{2\pi}}\int_0^{\sqrt{x}}\exp\left(-\frac{t^2}{2}\right)dt.\end{aligned}$$

Ezt deriválva a χ_1^2 eloszlás sűrűségfüggvénye, ha $x>0$,

$$\begin{aligned}f(x) &= \frac{2}{\sqrt{2\pi}}\exp\left(-\frac{x}{2}\right)\frac{1}{2\sqrt{x}}=\frac{(1/2)^{1/2}}{\sqrt{\pi}}x^{-1/2}\exp\left(-\frac{1}{2}x\right)= \\ &= \frac{(1/2)^{1/2}}{\Gamma(1/2)}x^{1/2-1}\exp\left(-\frac{1}{2}x\right).\end{aligned}$$

Ez éppen $\Gamma(1/2,1/2)$ sűrűségfüggvénye. A gamma eloszlás várható értékére kapott a/λ képlet ilyenkor éppen az 1 értéket adja, ami egybeesik avval, hogy a standard normális eloszlás második momentuma éppen a szórása, vagyis 1. Jóval érdekesebb azonban a χ_1^2 második momentuma, ami a standard normális eloszlás negyedik momentuma:

$$\mathbf{E}\left(N(0,1)^4\right)=\mathbf{E}\left(\left(\chi_1^2\right)^2\right)=\frac{\left(1+\frac{1}{2}\right)\frac{1}{2}}{\left(\frac{1}{2}\right)^2}=\frac{\frac{3}{4}}{\frac{1}{4}}=3.$$

□

7.2.10. Példa. A χ_n^2 eloszlás várható értéke és szórása.

A χ_n^2 eloszlást mint n darab független χ_1^2 összegét definiáljuk. Ebből

$$\begin{aligned}\mathbf{E}\left(\chi_n^2\right) &= n\mathbf{E}\left(\chi_1^2\right)=n \\ \mathbf{D}^2\left(\chi_n^2\right) &= n\mathbf{D}^2\left(\chi_1^2\right)=n\left(\mathbf{E}\left(\chi_1^2\right)-\mathbf{E}^2\left(\chi_1^2\right)\right)=2n.\end{aligned}$$

□

A gamma eloszlás tulajdonságai között a legfontosabb, hogy ha a második paraméterük azonos, akkor a gamma eloszlású valószínűségi változók összege is gamma eloszlású.

Mielőtt azonban a pontos állítást ismertetnénk, röviden foglalkozni kell az összeg eloszlásának kérdésével. Később ezt a kérdést egy sokkal általánosabb keretben újra tárgyalni fogjuk. Legyenek ξ és η valószínűségi változók, és az együttes sűrűségfüggvényük legyen $f(x,y)$. Ekkor az együttes sűrűségfüggvény definíciója alapján

$$\mathbf{P}((\xi,\eta)\in B)=\int\int_B f(x,y)\,dxdy.$$

Ezt felhasználva a $\zeta \mathrel{\hat{=}} \xi+\eta$ eloszlásfüggvénye

$$\mathbf{P}(\zeta<z)=\mathbf{P}(\xi+\eta<z)=\mathbf{P}((\xi,\eta)\in B),$$

ahol $B \mathrel{\hat{=}} \{(x,y)\mid x+y<z\}$ egy félsík. Ezt felhasználva

$$\mathbf{P}(\zeta<z)=\int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f(x,y)\,dxdy,$$

ugyanis az $x+y<z$ félsík szerkezete alapján minden fix y esetén az x legfeljebb a $z-y$ értékig futhat. A belső integrálban $x=u-y$ helyettesítéssel[3]

$$\mathbf{P}(\zeta<z)=\int_{-\infty}^{\infty}\int_{-\infty}^{z} f(u-y,y)\,dudy.$$

Mivel az integrandus nem negatív, az integrációs tartomány egy „téglalap” az integrációs határok felcserélhetőek, így

$$\mathbf{P}(\zeta<z)=\int_{-\infty}^{z}\int_{-\infty}^{\infty} f(u-y,y)\,dydu.$$

Ebből a z szerint deriválva az összeg sűrűségfüggvénye

$$g(z)=\int_{-\infty}^{\infty} f(z-y,y)\,dy.$$

Emlékeztetünk a következő fontos definícióra:

7.2.11. Definíció. A ξ és az η változókat függetlennek mondjuk, ha az együttes eloszlásfüggvényük a peremeloszlások szorzata, vagyis minden x és y mellett

$$\mathbf{P}(\xi<x,\eta<y)=\mathbf{P}(\xi<x)\,\mathbf{P}(\eta<y).$$

Ha az együttes eloszlásnak van sűrűségfüggvénye, akkor a

$$\frac{\partial^2 F}{\partial x\partial y}(x,y)=f(x,y)$$

szabályból azonnal következik, hogy az együttes sűrűségfüggvény a peremsűrűségfüggvények szorzata.

[3] A felső határon $z-y=u-y$ egyenlőségből $z=u$.

7.2.12. Példa. Ha a ξ és η függetlenek, továbbá a ξ eloszlása $\Gamma(a,\lambda)$, az η eloszlása pedig $\Gamma(b,\lambda)$, akkor a $\xi+\eta$ összeg eloszlása $\Gamma(a+b,\lambda)$.

Az összeg sűrűségfüggvényének képlete alapján, felhasználva, hogy a változók a nem negatív számokra támaszkodnak, így a negatív számokon a sűrűségfüggvényük nulla, ha a ξ sűrűségfüggvénye $f(x)$ az η sűrűségfüggvénye $g(y)$, akkor az összeg sűrűségfüggvénye

$$\int_{-\infty}^{\infty} f(x-y)\,g(y)\,dy = \int_0^x f(x-y)\,g(y)\,dy =$$
$$= \int_0^x \frac{\lambda^a}{\Gamma(a)}(x-t)^{a-1}\exp(-\lambda(x-t))\frac{\lambda^b}{\Gamma(b)}t^{b-1}\exp(-\lambda t)\,dt =$$
$$= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}\exp(-\lambda x)\int_0^x (x-t)^{a-1}t^{b-1}dt =$$
$$= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}\exp(-\lambda x)\int_0^1 (x-xz)^{a-1}(xz)^{b-1}x\,dz =$$
$$= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}\exp(-\lambda x)x^{a+b-1}\int_0^1 (1-z)^{a-1}z^{b-1}dz =$$
$$= \frac{\lambda^{a+b}}{\Gamma(a+b)}\exp(-\lambda x)x^{a+b-1}.$$

A gondolatmenetből látszik, hogy a gamma eloszlás korlátlanul osztható, ugyanis a $\Gamma(a,\lambda)$ eloszlás tetszőleges n-re felírható n darab független $\Gamma(a/n,\lambda)$ eloszlás összegeként.

□

7.2.13. Példa. Független exponenciális eloszlású változók összege nem exponenciális, hanem amennyiben a λ paraméterük azonos, akkor gamma eloszlású. Két különböző paraméterű exponenciális eloszlású változó összegének eloszlása nem gamma eloszlású.

A konvolúciós képlet alapján az összeg sűrűségfüggvénye

$$\begin{aligned} h(x) &\overset{\circ}{=} \lambda_1\lambda_2\int_0^x \exp(-\lambda_1(x-y))\exp(-\lambda_2 y)\,dy = \\ &= \lambda_1\lambda_2\exp(-\lambda_1 x)\int_0^x \exp(-(\lambda_2-\lambda_1)y)\,dy = \\ &= \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}\left(\exp(-\lambda_1 x)-\exp(-\lambda_2 x)\right). \end{aligned}$$

□

VIII.

ÖSSZEG, SZORZAT ÉS HÁNYADOS ELOSZLÁSA

Az előző fejezet végén már tárgyaltuk az összeg sűrűségfüggvényének képletét. Ebben a fejezetben a szorzat és a hányados sűrűségfüggvényét tárgyaljuk. A szorzat és a hányados sűrűségfüggvényének levezetése az összeghez hasonlóan elvégezhető a Fubini-tétel segítségével is. Ennek ellenére egy általánosabb eljárást fogunk bemutatni és az integráltranszformációs tétel segítségével fogjuk a formulákat levezetni.

8.1. A helyettesítéses integrálás formulája

A számegyenesen az

$$\int_{g(a)}^{g(b)} f(x)\,dx = \int_a^b f(g(y))\,g'(y)\,dy \tag{8.1.1}$$

helyettesítési formula a Newton–Leibniz formula közvetlen folyománya. A formulában a g függvény monotonitását nem kell megkövetelni. Elegendő megkövetelni, hogy az f integrandus folytonos és hogy a g helyettesítő függvény folytonosan deriválható legyen. Ilyenkor a két integrál létezik és mind a két oldalon alkalmazható a Newton–Leibniz formula, amely aztán azonos eredményre vezet. Magasabb dimenzióban a Newton–Leibniz formula ebben az egyszerű alakban nem érvényes, a különböző többdimenziós általánosításai jóval bonyolultabbak, így a helyettesítéses integrálás a (8.1.1) alakban nem várható, hogy igaz legyen. A többdimenziós helyettesítéses integrálási formula számos szempontból eltér az egydimenziós esettől. Ebben az alfejezetben az eltérés okait szeretnénk megvilágítani. A pontos bizonyítás egy sor technikai részletkérdést tartalmaz, amelyek tárgyalását elhagyjuk, így az alábbi gondolatmenet csak az indoklás vázát tartalmazza. A többdimenziós formula tárgyalása kapcsán az első észrevétel, hogy a fenti (8.1.1) alak többdimenzióban formailag is értelmetlen. Például a g' derivált általában egy mátrix és így többdimenzióban a két oldalon már a dimenziók sem egyeznek. További probléma, hogy a többdimenziós integrálokban az integrációs tartományok általában nem intervallumok, pontosabban szólva nem az intervallumoknak megfelelő téglatestek, ezért az integrációs határok jelölése helyett a formulában integrációs tartományokat kell írni. Ha egy B halmaz felett akarunk integrálni, akkor a szokásos jelölés

$$\int_B f(\mathbf{x})\,d\mathbf{x} = \underset{B}{\int \ldots \int} f(x_1,\ldots,x_n)\,dx_1\ldots dx_n,$$

ahol az $\mathbf{x}$ azért van vastagon szedve, hogy ezzel is jelezzük, miszerint az integrálás vektorok felett történik. Ha nem akarjuk az $\mathbf{x}$ változót kiírni, és az integrálás n-dimenzióban történik, akkor szokás az $\int_B f d\lambda_n$ jelölést is használni, de szokás a két jelölést keverni is és az integrálást például $\int_B f(x)\,d\lambda_n(x)$ vagy $\int_B f(\mathbf{x})\,d\lambda_n(\mathbf{x})$ módon jelölni. Az integrációs tartomány új jelölésének bevezetése azonban már önmagában is problémát jelent. Térjünk vissza a fenti (8.1.1) formulához és írjuk át helytelenül,

$$\int_{g(B)} f(x)\,dx = \int_B f(g(y))\,g'(y)\,dy,$$

alakba, ahol nyilván $B=[a,b]$. Ha a g monoton nő, akkor $g([a,b])=[g(a),g(b)]$ alapján

$$\begin{aligned}\int_{g([a,b])} f(x)\,dx &= \int_{g(a)}^{g(b)} f(x)\,dx=\int_a^b f(g(y))\,g'(y)\,dy= \\ &= \int_{[a,b]} f(g(y))\,g'(y)\,dy,\end{aligned}$$

vagyis ilyenkor a két felírási mód ekvivalens. Ha azonban a g monoton csökken, akkor $g([a,b])=[g(b),g(a)]$, így

$$\begin{aligned}\int_{g([a,b])} f(x)\,dx &= \int_{g(b)}^{g(a)} f(x)\,dx=-\int_{g(a)}^{g(b)} f(x)\,dx= \\ &= \int_a^b f(g(y))\left|g'(y)\right|dy= \\ &= \int_{[a,b]} f(g(y))\left|g'(y)\right|dy,\end{aligned}$$

ugyanis ilyenkor a g' negatív, következésképpen az abszolútérték kompenzálja az integrál előtti mínusz jelet. Ugyanakkor érdemes felhívni a figyelmet arra is, hogy az abszolútértékkel kibővített

$$\int_{g([a,b])} f(x)\,dx=\int_{[a,b]} f(g(y))\left|g'(y)\right|dy$$

alak is csak akkor érvényes, ha a g monoton. Ellenpéldaként tekintsük a következőt: A $[0,\pi]$ intervallum képe a $\sin x$ leképezéssel a $[0,1]$ szakasz, így

$$\begin{aligned}1 &= \int_0^1 1dx=\int_{\sin([0,\pi])} 1dx=\int_{[0,\pi]} |\cos x|\,dx=2\int_0^{\pi/2}\cos x dx= \\ &= 2\left[\sin x\right]_0^{\pi/2}=2,\end{aligned}$$

ami nyilvánvalóan nem helyes. Ugyanakkor persze a (8.1.1) képlet szerinti

$$0=\int_{\sin 0}^{\sin\pi} 1dx=\int_0^\pi \cos x dx=0$$

formula helyes. Ezek után nem túl meglepő, hogy $\mathbb{R}^n$ esetén a helyettesítés során fel kell tenni, hogy a helyettesített leképezés invertálható.

8.1.1. Tétel. *Legyen $U\subseteq\mathbb{R}^n$ nyílt halmaz és $g:U\to\mathbb{R}^n$ egy invertálható és folytonosan differenciálható leképezés. Ekkor $B\subseteq U$ esetén*

$$\begin{aligned}\int_{g(B)} f(\mathbf{x})\,d\mathbf{x} &\mathrel{\stackrel{\circ}{=}} \int_{g(B)} f(x)\,d\lambda_n(x)=\int_B f(g(y))\left|\det\left(g'(y)\right)\right|d\lambda_n(y)= \\ &\mathrel{\stackrel{\circ}{=}} \int_B f(g(\mathbf{y}))\left|\det\left(g'(\mathbf{y})\right)\right|d\mathbf{y},\end{aligned}$$

ahol $g'(\mathbf{y})$ *a* g *parciális deriváltjaiból álló Jacobi-mátrix*[1] *az* $\mathbf{y}$ *helyen.*

8.1.2. Példa. A gamma-béta aritmetika és az integráltranszformációs tétel.

Az integráltranszformációs tétel a matematikai analízis egyik alapeszköze. Segítségével egy sor fontos állítás igazolható. Példaként tekintsük a gamma és a béta függvényt összekötő nevezetes képlet egy lehetséges további indoklását. Ha $x, y > 0$, akkor

$$\begin{aligned}\Gamma(x)\Gamma(y) &\mathrel{\overset{\circ}{=}} \int_0^\infty t^{x-1}\exp(-t)\,dt \int_0^\infty s^{y-1}\exp(-s)\,ds = \\ &= \int_0^\infty \int_0^\infty t^{x-1}s^{y-1}\exp(-(s+t))\,dsdt.\end{aligned}$$

Tekintsük az

$$s = uv, t = v(1-u) = v - uv$$

helyettesítést. A helyettesítéskor három lépést kell végrehajtani: Első lépésként ki kell cserélni a változókat. Az integrandusba való behelyettesítéssel

$$\begin{aligned}&(v(1-u))^{x-1}(uv)^{y-1}\exp(-v) = \\ &(1-u)^{x-1}u^{y-1}v^{x+y-2}\exp(-v).\end{aligned}$$

Ezt követően ki kell számolni a Jacobi-determinánst. A deriváltmátrix

$$\begin{pmatrix} \partial s/\partial u & \partial s/\partial v \\ \partial t/\partial u & \partial t/\partial v \end{pmatrix} = \begin{pmatrix} v & u \\ -v & 1-u \end{pmatrix},$$

következésképpen a Jacobi-determináns $v(1-u) + uv = v$. Így helyettesítés után az integrandus, felhasználva, hogy $v \geq 0$,

$$(1-u)^{x-1}u^{y-1}v^{x+y-1}\exp(-v).$$

Harmadik lépésként ki kell számolni az integrációs határokat. Ha a képletben $B \mathrel{\overset{\circ}{=}} (0,1)\times(0,\infty)$, vagyis az (u,v) pár a B halmaz egy tetszőleges eleme, akkor az (s,t) pár a $(0,\infty)\times(0,\infty)$ halmazt futja be. Világos, hogy a leképezés egyértelmű. Ebből következően

$$\Gamma(x)\Gamma(y) = \int_0^\infty \int_0^1 (1-u)^{x-1}u^{y-1}v^{x+y-1}\exp(-v)\,dudv.$$

A belső integrálból a csak v-től függő tagokat kiemelve

$$\Gamma(x)\Gamma(y) = \int_0^\infty v^{x+y-1}\exp(-v)B(x,y)\,dv = \Gamma(x+y)B(x,y),$$

ami éppen a bizonyítandó összefüggés.

[1] A B halmaz általában zárt halmaz, de ennél jóval általánosabb esetek, például a különböző téglatestek is előfordulhatnak. A B pontos karakterizációjával a tárgyalás elnagyolt jellege miatt nem foglalkozunk.

$\square$

Az elmondottakban a leginkább meglepő elem, hogy miként kerül az integráltranszformációs tételbe a $\det(g'(y))$ Jacobi-determináns. A következő alpontban egy némiképpen elnagyolt gondolatmenettel ezt fogjuk indokolni.

A térfogat és a determináns kapcsolata

Mint minden integrál, így az $\int_B f(\mathbf{x})\, d\mathbf{x}$ többdimenziós integrál is lényegében az $f(\mathbf{x})$ integrandus egy súlyozott összege, ahol a súlyokat az $\mathbb{R}^n$ „térfogatelemei" adják. A figyelmes olvasóban azonnal felmerülhet, hogy ez a megközelítés bizonyos értelemben mellébeszélés, ugyanis szemben az egydimenziós esettel távolról sem világos, hogy mit kell tekinteni valamely partíció elemeinek térfogatán. Érdemes észrevenni, hogy már két dimenzióban is a térfogat, ilyenkor persze terület, kiszámolása már önmagában is integrállal történik és legalábbis a nem negatív függvények esetén egy n-dimenziós integrál kiszámolása valójában egy $(n+1)$-dimenziós térfogat meghatározását jelenti. Így a többdimenziós integrálás és a többdimenziós térfogat egymástól nem elválasztható, hanem sokkal inkább egymást kölcsönösen feltételező fogalmak.

8.1.3. Példa. Számoljuk ki az n-dimenziós egységgömb térfogatát.

Jelölje $G_n(r)$ az n-dimenziós r sugarú gömböt, vagyis legyen

$$G_n(r) \overset{\circ}{=} \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq r\},$$

ahol $\|\mathbf{x}\| \overset{\circ}{=} \sqrt{\Sigma_{k=1}^n x_k^2}$. Teljes indukcióval megmutatjuk, hogy

$$\operatorname{vol}(G_n(1)) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

Ha $n = 1$, akkor

$$\frac{\pi^{1/2}}{\Gamma\left(\frac{1}{2}+1\right)} = \frac{\sqrt{\pi}}{\frac{1}{2}\Gamma\left(\frac{1}{2}\right)} = 2\frac{\sqrt{\pi}}{\sqrt{\pi}} = 2,$$

amely valóban a $[-1,1]$ szakasz hossza. Érdemes ellenőrizni az $n = 2$ esetet is. Ilyenkor

$$\frac{\pi}{\Gamma\left(\frac{2}{2}+1\right)} = \frac{\pi}{\Gamma(2)} = \frac{\pi}{1!} = \pi,$$

amely valóban az egységnyi sugarú körlap területe. Tegyük fel, hogy a képletet már egy n-re beláttuk. Az említett eljárás szerint a $(n+1)$-dimenziós térfogat úgy származtatható, hogy egy adott irány szerint venni kell az n-dimenziós szeleteket, ki kell számolni ezen szeletek n-dimenziós térfogatát, ezeket szorozni kell a szeletek magasságával és ezeket értékeket összegezni, majd integrálni kell[2]. Egységnyi sugarú gömb esetén venni kell a $[-1,1]$ szakaszt, azt fel kell osztani az (u_k) pontokkal du_k hosszú szakaszokra. Ezt követően az $I_k = [u_{k-1}, u_k]$ szakaszból venni kell egy ξ_k pontot és venni kell evvel a sugárral egy $G_n\left(\sqrt{1-\xi_k^2}\right)$ sugarú gömböt, ugyanis egy gömb szeletei maguk is alacsonyabb dimenziós gömbök, és ha a szelet sugara r, akkor teljesülni kell az $1 = r^2 + \xi_k^2$ egyenlőségnek. A $G_n\left(\sqrt{1-\xi_k^2}\right) \times I_k$ karikák egyesítése a gömb egy közelítése. Következésképpen

$$\operatorname{vol}(G_{n+1}(1)) = \int_{-1}^{1} \operatorname{vol}\left(G_n\left(\sqrt{1-x^2}\right)\right) dx.$$

Az n-dimenziós térfogat n-edrendben homogén függvény, vagyis

$$\operatorname{vol}(G_n(r)) = r^n \operatorname{vol}(G_n(1)).$$

Ezt és az indukciós feltételt felhasználva

$$\begin{aligned}
\operatorname{vol}(G_{n+1}(1)) &= \int_{-1}^{1} \operatorname{vol}(G_n(1)) \left(\sqrt{1-x^2}\right)^n dx = \\
&= \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} \int_{-1}^{1} \left(\sqrt{1-x^2}\right)^n dx = \\
&= \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} 2 \int_{0}^{1} \left(1-x^2\right)^{n/2} dx.
\end{aligned}$$

$u = x^2$ helyettesítéssel és a gamma és a béta függvények közötti összefüggéssel

$$\begin{aligned}
\operatorname{vol}(G_{n+1}(1)) &= \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} 2 \int_{0}^{1} (1-u)^{n/2} \frac{1}{2\sqrt{u}} dx = \\
&= \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} B\left(\frac{n}{2}+1, \frac{1}{2}\right) = \\
&= \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} \frac{\Gamma\left(\frac{n}{2}+1\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n+1}{2}+1\right)} = \\
&= \frac{\pi^{(n+1)/2}}{\Gamma\left(\frac{n+1}{2}+1\right)},
\end{aligned}$$

ami éppen az igazolandó formula $n+1$ esetén.

□

[2]Gondoljunk arra, hogy egy hagymát karikákra felvágunk. Ezt szokás Cavalieri elvnek is mondani és a már sokat emlegetett Fubini-tétel közvetlen alkalmazását jelenti.

A többdimenziós integrálok kiszámolásakor a legfőbb gondot tehát nem az integrálok definíciója vagy az integrálok kiszámolása jelent, hanem annak igazolása, hogy az n-dimenziós térben egyáltalán létezik egy egyértelműen meghatározott „ésszerű”, a szemléletnek megfelelő térfogatfogalom. A $\mathrm{vol}\,(B)$ térfogatról a következőket szokás feltenni:

1. $\mathrm{vol}\,(B) \geq 0$.
2. Ha a legfeljebb megszámlálható sok halmazból álló (B_i) halmazrendszer elemei páronként diszjunktak, akkor $\mathrm{vol}\,(\cup_i B_i) = \sum_i \mathrm{vol}\,(B_i)$.
3. A $\mathrm{vol}\,(B)$ eltolásinvariáns, vagyis tetszőleges $\mathbf{x}$ vektor esetén $\mathrm{vol}\,(B+\mathbf{x}) = \mathrm{vol}\,(B)$.[3]
4. Az $E \stackrel{\circ}{=} \{\mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}$ egységkocka térfogat éppen 1.

Hangsúlyozni kell, hogy a térfogat fogalma bármennyire is szemléletes, annak igazolása, hogy a fenti tulajdonságokat kielégítő térfogatfogalom létezik és egyértelmű, távolról sem nyilvánvaló. A térfogat fogalma a következő további szemléletes tulajdonságokkal rendelkezik, amelyek már következményei a fenti alaptulajdonságoknak és amelyeket alább röviden indokolni fogunk:

1. A térfogat invariáns a forgatásra és a tükrözésre[4].
2. A térfogat homogén abban az értelemben, hogy tetszőleges $\alpha > 0$ szám esetén ha valamelyik koordinátát α-val megszorzzunk, akkor a térfogat is α-szorosára nő. Így többek között minden $\alpha > 0$ esetén $\mathrm{vol}\,(\alpha B) = \alpha^n \mathrm{vol}\,(B)$, illetve általánosabban, ha Λ egy diagonális mátrix, akkor
$$\mathrm{vol}\,(\Lambda \cdot B) = \lambda_1 \lambda_2 \cdots \lambda_n \mathrm{vol}\,(B),$$
ahol $\lambda_1, \lambda_2, \ldots, \lambda_n$ a Λ diagonálisában levő elemek. Speciálisan tetszőleges téglatest térfogata az oldalak hosszának szorzata.

Természetesen kézenfekvően merül fel a kérdés, hogy n-dimenzióban mit is értünk forgatáson? Forgatáson és tükrözésen olyan invertálható lineáris leképezéseket értünk, amelyekre nézve a tér egységgömbje invariáns. Emlékeztetünk, hogy egy $\mathbf{S}$ mátrixot ortogonálisnak mondunk, ha a transzponáltja éppen az inverze, vagyis $\mathbf{S}\mathbf{S}^* = \mathbf{S}^*\mathbf{S} = \mathbf{I}$, ahol $\mathbf{I}$ az egységmátrix és a $*$ a transzponálás jele. Az ortogonális mátrixok őrzik a távolságot.
$$\|\mathbf{S}\mathbf{x}\|^2 = (\mathbf{S}\mathbf{x}, \mathbf{S}\mathbf{x}) = (\mathbf{x}, \mathbf{S}^*\mathbf{S}\mathbf{x}) = (\mathbf{x}, \mathbf{x}) = \|\mathbf{x}\|^2,$$
így az első tulajdonság miatt ha az $\mathbf{S}$ egy ortogonális mátrix, akkor $\mathrm{vol}\,(\mathbf{S}\cdot B) = \mathrm{vol}\,(B)$.

[3] Ez a tulajdonság azt jelenti, hogy a térfogat fogalma nem függ attól, hogy hol jelöljük ki az origót.

[4] A tulajdonság azt jelenti, hogy a térfogat valóban geometriai fogalom és nem függ attól, hogy a térfogat kiszámolásakor milyen ortonormált koordinátarendszert alkalmazunk.

Az alpont legfontosabb eredményeként megmutatjuk, hogy tetszőleges $\mathbf{A}$ lineáris leképezésre igaz a következő:

$$\operatorname{vol}(\mathbf{A}\cdot B)=|\det(\mathbf{A})|\cdot\operatorname{vol}(B). \tag{8.1.2}$$

Ez a determinánsos képlet triviálisan teljesül, ha az $\mathbf{A}$ oszlopai összefüggnek, ugyanis ilyenkor a determináns nulla, és mivel az $\mathbf{A}\cdot B$ az $\mathbb{R}^n$ egy valódi alterébe esik,

$$\operatorname{vol}(\mathbf{A}\cdot B)=0=|\det(\mathbf{A})|\operatorname{vol}(B).$$

Az igazolandó (8.1.2) egyenlőség egy fontos lineáris algebrai tételen, a komplex számok $z=r\exp(i\varphi)$ felbontásának általánosításának tekinthető poláris felbontás következménye:

8.1.4. Lemma. Ha $\mathbf{A}$ egy invertálható mátrix, akkor az $\mathbf{A}$ felbontható $\mathbf{A}=\mathbf{R}\cdot\mathbf{U}$ alakba, ahol az $\mathbf{R}$ pozitív definit[5], az $\mathbf{U}$ pedig ortogonális.

Bizonyítás: Az $\mathbf{AA}^*$ mátrix szimmetrikus és pozitív definit, így a spektrálfelbontási tétel miatt $\mathbf{O}\Lambda\mathbf{O}^*$ alakba írható, ahol az $\mathbf{O}$ ortogonális mátrix oszlopai az $\mathbf{AA}^*$ sajátvektorai, Λ a sajátértékeket tartalmazó diagonális mátrix. Az $\mathbf{AA}^*$ pozitív szemidefinit, ezért $\Lambda\geq 0$, következésképpen az $\mathbf{R}\stackrel{\circ}{=}\mathbf{O}\sqrt{\Lambda}\mathbf{O}^*$ definíció értelmes. Értelemszerűen a gyökjel a Λ elemeire értendő. Mivel $\mathbf{A}$ invertálható, ezért $\Lambda>0$, így az $\mathbf{R}$ valóban pozitív definit. Az $\mathbf{R}$ nyilván szimmetrikus és ugyanakkor invertálható, ugyanis invertálható mátrixok szorzata. Elegendő igazolni, hogy az $\mathbf{U}\stackrel{\circ}{=}\mathbf{R}^{-1}\mathbf{A}$ egy ortogonális mátrix, ugyanis nyilvánvalóan $\mathbf{A}=\mathbf{RU}$. Ehhez elegendő megmutatni, hogy $\mathbf{R}^2=\mathbf{AA}^*$, ugyanis ezzel

$$\mathbf{UU}^*\stackrel{\circ}{=}\mathbf{R}^{-1}\mathbf{A}\left(\mathbf{R}^{-1}\mathbf{A}\right)^*=\mathbf{R}^{-1}\mathbf{AA}^*\left(\mathbf{R}^{-1}\right)^*=\mathbf{R}^{-1}\mathbf{R}^2\mathbf{R}^{-1}=\mathbf{I}.$$

Ez azonban egyszerű, ugyanis

$$\begin{aligned}\mathbf{R}^2 &= \left(\mathbf{O}\sqrt{\Lambda}\mathbf{O}^*\right)\left(\mathbf{O}\sqrt{\Lambda}\mathbf{O}^*\right)=\\ &= \mathbf{O}\sqrt{\Lambda}\left(\mathbf{O}^*\mathbf{O}\right)\sqrt{\Lambda}\mathbf{O}^*=\\ &= \mathbf{O}\sqrt{\Lambda}\mathbf{I}\sqrt{\Lambda}\mathbf{O}^*=\mathbf{O}\sqrt{\Lambda}\sqrt{\Lambda}\mathbf{O}^*=\\ &= \mathbf{O}\Lambda\mathbf{O}^*=\mathbf{AA}^*.\end{aligned}$$

□

A poláris felbontással és a térfogat említett alaptulajdonságaival az igazolandó (8.1.2) determinánsos formula már egyszerűen belátható. Ha $\mathbf{A}$ invertálható, akkor felírható $\mathbf{A}=\mathbf{RU}$ módon. $\mathbf{R}$ pozitív definit, így a spektrálfelbontási tétel miatt $\mathbf{T}\Lambda\mathbf{T}^*$ alakba

[5] A pozitív definitség fogalmába a szimmetriát is beleértjük.

írható, ahol Λ most az $\mathbf{R}$ sajátértékeiből álló diagonális mátrix. Ennek determinánsa éppen a sajátértékek szorzata, így a forgatásinvariancia és a skálázási tulajdonság miatt

$$\begin{aligned}\operatorname{vol}(\mathbf{A}\cdot B) &= \operatorname{vol}((\mathbf{RU})\cdot B) = \\ &= \operatorname{vol}((\mathbf{T}\Lambda\mathbf{T}^*\mathbf{U})\cdot B) = \operatorname{vol}(\mathbf{T}\cdot((\Lambda\mathbf{T}^*\mathbf{U})\cdot B)) = \\ &= \operatorname{vol}((\Lambda\mathbf{T}^*\mathbf{U})\cdot B) = \left(\prod_{i=1}^{n}\lambda_i\right)\operatorname{vol}(\mathbf{T}^*\mathbf{U}\cdot B) = \\ &= \left(\prod_{i=1}^{n}\lambda_i\right)\operatorname{vol}(B) = \det(\mathbf{R})\operatorname{vol}(B).\end{aligned}$$

Az $\mathbf{U}$ ortogonális mátrix, ezért

$$1 = \det(\mathbf{I}) = \det(\mathbf{UU}^*) = \det(\mathbf{U})\det(\mathbf{U}^*) = \det(\mathbf{U})^2$$

így $\det(\mathbf{U}) = \pm 1$. Ebből következően, felhasználva, hogy $\mathbf{R}$ pozitív definit, így a determinánsa pozitív:

$$\begin{aligned}\operatorname{vol}(\mathbf{A}\cdot B) &= \det(\mathbf{R})\left|\det(\mathbf{U})\right|\operatorname{vol}(B) = \left|\det(\mathbf{R})\det(\mathbf{U})\right|\operatorname{vol}(B) = \\ &= \left|\det(\mathbf{RU})\right|\operatorname{vol}(B) = \left|\det(\mathbf{A})\right|\operatorname{vol}(B).\end{aligned}$$

□

Végezetül térjünk rá a bizonyításban alapvető szerepet játszó tulajdonságok igazolására. Ha $\mathbf{S}$ az egységgömböt önmagára képező invertálható leképezés, akkor a $v(B) \circeq \operatorname{vol}(\mathbf{S}\cdot B)$ is nyilvánvalóan eltolásinvariáns, ugyanis

$$\begin{aligned}v(B+\mathbf{x}) &\circeq \operatorname{vol}(\mathbf{S}\cdot(B+\mathbf{x})) = \operatorname{vol}(\mathbf{S}\cdot B + \mathbf{Sx}) = \\ &= \operatorname{vol}(\mathbf{S}\cdot B) \circeq v(B).\end{aligned}$$

Természetesen nem feltétlenül teljesül, hogy az E egységkocka v-je éppen 1, de ez könnyen elérhető, ha a

$$w(B) \circeq \frac{v(B)}{v(E)}$$

halmazfüggvényt tekintjük. A térfogat feltételezett egyértelműsége miatt

$$w(B) \circeq \frac{v(B)}{v(E)} = \operatorname{vol}(B).$$

Ha most B-nek a G egységgömböt választjuk és kihasználjuk, hogy $\mathbf{S}\cdot G = G$, akkor

$$\operatorname{vol}(G) = w(G) \circeq \frac{v(G)}{v(E)} \circeq \frac{\operatorname{vol}(\mathbf{S}\cdot G)}{v(E)} = \frac{\operatorname{vol}(G)}{v(E)}$$

egyenlőséghez jutunk, következésképpen $v(E) = 1$, így a térfogat, miképpen ezt a bizonyításban felhasználtuk, invariáns az ortogonális mátrixszal való szorzásra nézve. Hasonlóan kell igazolni, a homogenitás tulajdonságát.

A helyettesítéses integrálás formulájának „igazolása”

A determinánsos képlet alapján az integráltranszformációs tétel „indoklása" a következő: Tekintsük a B halmaz egy $\left(V_i^{(k)}\right)$ partícióját. A g invertálhatósága miatt ez a $g(B)$ halmazon egy $U_i^{(k)} \circeq g\left(V_i^{(k)}\right)$ partíciót definiál. Az integrál definíció szerint a végtelenül finomodó partíciókhoz tartozó közelítő összegek határértéke, így közelítőleg

$$\int_{g(B)} f(\mathbf{x})\, d\mathbf{x} \approx \sum_i f(\mathbf{x}_i) \operatorname{vol}\left(U_i^{(k)}\right) = \sum_i f(g(\mathbf{y}_i)) \operatorname{vol}\left(U_i^{(k)}\right),$$

ahol $\mathbf{x}_i \in U_i^{(k)}$ és $\mathbf{x}_i = g(\mathbf{y}_i)$ és $\mathbf{y}_i \in V_i^{(k)}$. A bizonyítás lényege, hogy ha a $V_i^{(k)}$ elég „kicsi", akkor a g a $V_i^{(k)}$ halmazon jól közelíthető a deriváltjával. Ezért, felhasználva az eltolásinvarienciát

$$\begin{aligned}
\operatorname{vol}\left(U_i^{(k)}\right) &= \operatorname{vol}\left(g\left(V_i^{(k)}\right)\right) = \\
&= \operatorname{vol}\left(g(\mathbf{y}_i) + g'(\mathbf{y}_i)\left(V_i^{(k)} - \mathbf{y}_i\right) + o\left(\left\|V_i^{(k)} - \mathbf{y}_i\right\|\right)\right) \approx \\
&\approx \operatorname{vol}\left(g(\mathbf{y}_i) + g'(\mathbf{y}_i)\left(V_i^{(k)} - \mathbf{y}_i\right)\right) = \\
&= \operatorname{vol}\left(g'(\mathbf{y}_i)\left(V_i^{(k)} - \mathbf{y}_i\right)\right) \\
&= \operatorname{vol}\left(g'(\mathbf{y}_i) V_i^{(k)}\right) = \\
&= \left|\det\left(g'(\mathbf{y}_i)\right)\right| \operatorname{vol}\left(V_i^{(k)}\right).
\end{aligned}$$

Így

$$\begin{aligned}
\int_{g(B)} f(\mathbf{x})\, d\mathbf{x} &\approx \sum_i f(g(\mathbf{y}_i)) \left|\det\left(g'(\mathbf{y}_i)\right)\right| \operatorname{vol}\left(V_i^{(k)}\right) \approx \\
&\approx \int_B f(g(\mathbf{y})) \left|\det\left(g'(\mathbf{y})\right)\right| d\mathbf{y}.
\end{aligned}$$

A bizonyítás során garantálni kell, hogy $k \to \infty$, esetén egyrészt a $\left(V_i^{(k)}\right)$ és az $\left(U_i^{(k)}\right)$ partíciók egyszerre legyenek infinitezimálisak, vagyis a partíciókban szereplő halmazok átmérőinek maximuma egyszerre tartson nullához[6], másrészt hogy a másodrendű $\operatorname{vol}\left(o\left(\left\|V_i^{(k)} - \mathbf{y}_i\right\|\right)\right)$ hibák elhagyása során a hibák összege is nullához tartson[7]. A részletek indoklását elhagyjuk.

---

[6] Ehhez ki kell használni hogy a g folytonosan deriválható.

[7] Ehhez azt kell kihasználni, hogy ha a finomítást növeljük, akkor a partícióban levő elemek száma lassabban nő, mint ahogy a $\operatorname{vol}\left(o\left(\left\|V_i^{(k)} - \mathbf{y}_i\right\|\right)\right)$ másodrendű közelítési hibák nagyságai csökkennek. Ha például a felbontás finomságát felére vesszük, akkor 2^n-szeresére nő a felbontásban levő halmazok száma, amelyek átmérője megfeleződik, így a térfogatuk 2^{-n}-szeresénél jobban csökken, ugyanis az átmérők másodrendben kicsik.

8.2. Többdimenziós transzformált változók

Ha $(\xi_k)_{k=1}^n$ az $(\Omega, \mathscr{A}, \mathbf{P})$ téren értelmezett valószínűségi változók, és g az $\mathbb{R}^n$ téren értelmezett folytonos függvény, akkor az

$$\eta \mathrel{\overset{\circ}{=}} g(\xi_1, \ldots, \xi_n)$$

összefüggéssel definiált függvény szintén valószínűségi változó. Hogyan lehet meghatározni az η eloszlását?

8.2.1. Tétel. *Ha f a $(\xi_1, \ldots, \xi_k)$ vektor sűrűségfüggvénye, emellett*

$$(\eta_1, \ldots, \eta_k) \mathrel{\overset{\circ}{=}} \varphi(\xi_1, \xi_2, \ldots, \xi_k),$$

továbbá a $T \mathrel{\overset{\circ}{=}} \varphi^{-1}$ függvény létezik és differenciálható, akkor az $(\eta_1, \ldots, \eta_k)$ sűrűségfüggvénye

$$\begin{aligned} h(x_1, \ldots, x_k) &\mathrel{\overset{\circ}{=}} f\left(\varphi^{-1}(x_1, \ldots, x_n)\right)\left|\det\left(\frac{d}{dx}\varphi^{-1}(x_1, \ldots, x_n)\right)\right| = \\ &= f(T(x_1, \ldots, x_n))\left|\det\left(\frac{d}{dx}T(x_1, \ldots, x_n)\right)\right|. \end{aligned}$$

Bizonyítás: A többdimenziós integráltranszformációs tétel alapján minden B esetén[8]

$$\begin{gathered} \mathbf{P}((\eta_1, \eta_2, \ldots \eta_k) \in B) \mathrel{\overset{\circ}{=}} \\ \mathrel{\overset{\circ}{=}} \mathbf{P}((\varphi_1(\xi_1, \ldots, \xi_k), \ldots, \varphi_k(\xi_1, \ldots, \xi_k)) \in B) = \\ = \mathbf{P}\left((\xi_1, \ldots, \xi_k) \in \varphi^{-1}(B)\right) = \mathbf{P}((\xi_1, \ldots, \xi_k) \in T(B)) = \\ = \int_{T(B)} f \, d\lambda_k = \int_B f(T)\left|\det\left(T'\right)\right| d\lambda_k, \end{gathered}$$

feltéve persze, hogy a tétel feltételei teljesülnek, vagyis a $T \mathrel{\overset{\circ}{=}} \varphi^{-1}$ létezik és differenciálható. Vagyis $f(T)\left|\det(T')\right|$ az (η_i) vektor sűrűségfüggvénye.

□

Az alkalmazásokban az η_i változók száma általában kevesebb, mint a ξ_i változók száma. Ilyenkor a φ függvényt „kiegészítjük“, majd peremeloszlásokra térünk át. Ezt mutatja, a következő példa:

8.2.2. Példa. Határozzuk meg az összeg, a szorzat és a hányados sűrűségfüggvényét.

[8] Akárcsak az integráltranszformációs tételben a B pontos karakterizációját most sem tisztázzuk.

1. Összeg: Ilyenkor

$$\begin{aligned}\varphi(x,y) &= (x+y,y) \overset{\circ}{=} (u,v)\,. \\ \varphi^{-1}(u,v) &= (u-v,v) = (x,y)\,.\end{aligned}$$

Legyen ξ sűrűségfüggvénye h, és η sűrűségfüggvénye pedig legyen g. Ha ξ és η függetlenek, akkor a (ξ,η) pár együttes sűrűségfüggvénye $f(x,y) = h(x)g(y)$. A $(\xi+\eta,\eta)$ sűrűségfüggvénye

$$\begin{aligned}f(T)\left|\det\left(T'\right)\right| &= h(u-v)\,g(v)\left|\det\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\right| = \\ &= h(u-v)\,g(v)\,.\end{aligned}$$

A $\xi+\eta$ eloszlása az első peremeloszlás, vagyis a $\xi+\eta$ sűrűségfüggvénye

$$\int_{-\infty}^{\infty} h(u-v)\,g(v)\,dv.$$

Ha ξ és η nem függetlenek, akkor az összeg sűrűségfüggvénye

$$k(u) = \int_{-\infty}^{\infty} f(u-v,v)\,dv.$$

Azonnal látható, hogy $w = u-v$ helyettesítéssel

$$\int_{-\infty}^{\infty} f(u-v,v)\,dv = \int_{\infty}^{-\infty} f(w,u-w)(-1)\,dw = \int_{-\infty}^{\infty} f(w,u-w)\,dw$$

amely egybevág avval, hogy az összeadás kommutatív.

2. Szorzat. Ilyenkor

$$\begin{aligned}\varphi(x,y) &= (xy,y) \overset{\circ}{=} (u,v)\,. \\ \varphi^{-1}(u,v) &= (u/v,v) = (x,y)\,.\end{aligned}$$

Ha ξ és η függetlenek, akkor

$$\begin{aligned}f(T)\left|\det\left(T'\right)\right| &= h\left(\frac{u}{v}\right) g(v)\left|\det\begin{pmatrix} 1/v & -u/v^2 \\ 0 & 1 \end{pmatrix}\right| = \\ &= h\left(\frac{u}{v}\right) g(v)\,\frac{1}{|v|}.\end{aligned}$$

A $\xi\eta$ eloszlása az első peremeloszlás, vagyis a $\xi\eta$ sűrűségfüggvénye

$$\int_{-\infty}^{\infty} h\left(\frac{u}{v}\right) g(v)\,\frac{1}{|v|}\,dv.$$

Ha a változók nem függetlenek, akkor a szorzat sűrűségfüggvénye

$$\int_{-\infty}^{\infty} f\left(\frac{u}{v},v\right)\frac{1}{|v|}dv.$$

Ilyenkor $w=u/v$ helyettesítéssel

$$v=\frac{u}{w},\frac{dv}{dw}=-\frac{u}{w^2},\Rightarrow dv=-\frac{u}{w^2}dw.$$

Ha u pozitív, akkor a $(0,\infty)$ ráképződik a $(\infty,0)$-ra, illetve a $(-\infty,0)$ a $(0,-\infty)$-re képződik. Ezért a határokat fel kell cserélni. Ilyenkor $u=|u|$ és ezért

$$\begin{aligned}\int_{-\infty}^{\infty} f\left(\frac{u}{v},v\right)\frac{1}{|v|}dv &= -\int_{-\infty}^{\infty} f\left(w,\frac{u}{w}\right)\frac{|w|}{|u|}\left(-\frac{u}{w^2}\right)dw= \\ &= \int_{-\infty}^{\infty} f\left(w,\frac{u}{w}\right)\frac{1}{|w|}dw.\end{aligned}$$

Ha u negatív, akkor a $(0,\infty)$ ráképződik a $(-\infty,0)$-ra, a $(-\infty,0)$ pedig a $(0,\infty)$-re képződik. Ilyenkor $|u|=-u$

$$\begin{aligned}\int_{-\infty}^{\infty} f\left(\frac{u}{v},v\right)\frac{1}{|v|}dv &= \int_{-\infty}^{\infty} f\left(w,\frac{u}{w}\right)\frac{|w|}{|u|}\left(-\frac{u}{w^2}\right)dw= \\ &= \int_{-\infty}^{\infty} f\left(w,\frac{u}{w}\right)\frac{1}{|w|}dw.\end{aligned}$$

A szabály összhangban áll azzal, hogy a szorzat is kommutatív.

3. Hányados. Ilyenkor

$$\begin{aligned}\varphi(x,y) &= (x/y,y)\mathrel{\overset{\circ}{=}}(u,v). \\ \varphi^{-1}(u,v) &= (u\cdot v,v)=(x,y).\end{aligned}$$

Ha ξ és η függetlenek, akkor

$$\begin{aligned}f(T)\left|\det\left(T'\right)\right| &= h(u\cdot v)g(v)\left|\det\begin{pmatrix} v & u \\ 0 & 1\end{pmatrix}\right|= \\ &= h(u\cdot v)g(v)|v|.\end{aligned}$$

A ξ/η eloszlása az első peremeloszlás, vagyis a ξ/η sűrűségfüggvénye

$$\int_{-\infty}^{\infty} h(uv)g(v)|v|dv.$$

Ha nem függetlenek, akkor $\int_{-\infty}^{\infty} f(uv,v)|v|dv$. Ekkor $w=uv$ helyettesítéssel, ha u pozitív, akkor a határokat nem kell módosítani

$$\frac{dw}{dv}=u\Rightarrow dv=\frac{dw}{u},$$

és

$$\int_{-\infty}^{\infty} f(uv, v)\,|v|\,dv = \int_{-\infty}^{\infty} f\left(w, \frac{w}{u}\right)\left|\frac{w}{u}\right|\frac{dw}{u},$$

ami egybevág avval, hogy az osztás nem kommutatív.

□

8.2.3. Példa. Számoljuk ki a Laplace-eloszlás sűrűségfüggvényét.

Ha ξ és η független λ paraméterű exponenciális eloszlású valószínűségi változók, akkor a $\xi - \eta$ eloszlását Laplace-eloszlásnak mondjuk. Ha valamely ζ változó eloszlásfüggvénye $F(x)$, akkor

$$\begin{aligned} \mathbf{P}(-\zeta < x) &= \mathbf{P}(\zeta > -x) = 1 - \mathbf{P}(\zeta \leq -x) = \\ &= 1 - F(-x + 0). \end{aligned}$$

Ezt deriválva ha $f(x)$ a ζ sűrűségfüggvénye, akkor a $-\zeta$ sűrűségfüggvénye $f(-x)$, vagyis a ζ sűrűségfüggvényének az y-tengelyre való tükrözése. A konvolúciós szabály miatt a λ paraméterű Laplace-eloszlás sűrűségfüggvénye

$$\begin{aligned} h(z) &= \lambda^2 \int_{-\infty}^{\infty} \exp(-\lambda y)\,\chi_{(0,\infty)}(y)\exp(\lambda(z-y))\,\chi_{(-\infty,0)}(z-y)\,dy = \\ &= \lambda^2 \int_0^{\infty} \exp(-\lambda y)\exp(\lambda(z-y))\,\chi_{(-\infty,0)}(z-y)\,dy = \\ &= \lambda^2 \exp(\lambda z)\int_0^{\infty} \exp(-2\lambda y)\,\chi_{(-\infty,0)}(z-y)\,dy. \end{aligned}$$

Ha $z > 0$, akkor

$$\begin{aligned} h(z) &= \lambda^2 \exp(\lambda z)\int_z^{\infty} \exp(-2\lambda y)\,dy = \\ &= \lambda^2 \exp(\lambda z)\left[\frac{\exp(-2\lambda y)}{-2\lambda}\right]_z^{\infty} = \\ &= \frac{\lambda}{2}\exp(\lambda z)\exp(-2\lambda z) = \frac{\lambda}{2}\exp(-\lambda z). \end{aligned}$$

Ha $z < 0$, akkor

$$\begin{aligned} h(z) &= \lambda^2 \exp(\lambda z)\int_0^{\infty} \exp(-2\lambda y)\,dy = \\ &= \lambda^2 \exp(\lambda z)\frac{1}{2\lambda} = \frac{\lambda}{2}\exp(\lambda z). \end{aligned}$$

A két oldalt egy képletbe összefogva

$$h(z) = \frac{\lambda}{2}\exp(-\lambda\,|z|).$$

□

8.2.4. Példa. Számoljuk ki az összeg várható értékét.

Ha ξ és η nem feltétlenül függetlenek, akkor az összeg sűrűségfüggvénye $\int_{-\infty}^{\infty} f(u-v,v)\,dv$. Ebből az összeg várható értéke

$$\begin{aligned}
\mathbf{E}(\xi+\eta) &= \int_{-\infty}^{\infty} u \int_{-\infty}^{\infty} f(u-v,v)\,dvdu = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uf(u-v,v)\,dudv = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (v+x) f(x,v)\,dxdv = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,v)\,dxdv + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} vf(x,v)\,dxdv = \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x,v)\,dvdx + \int_{-\infty}^{\infty} v \int_{-\infty}^{\infty} f(x,v)\,dxdv = \\
&= \int_{-\infty}^{\infty} xg(x)\,dx + \int_{-\infty}^{\infty} vh(v)\,dv = \\
&= \mathbf{E}(\xi)+\mathbf{E}(\eta).
\end{aligned}$$

A bizonyításban kulcsszerepet játszott az integrálok felcserélhetősége. Korábban már többször jeleztük, hogy ez mindig megtehető, ha az integrandus nem negatív. Itt általában ez nem teljesül. Közvetett módon azonban a negatív és a pozitív részeket szétválasztva a gondolatmenetet négyszer megismételve és kihasználva, hogy a várható érték létezése azt jelenti, hogy a pozitív és a negatív rész mindegyike végesen integrálható, a gondolatmenet pontosítható[9].

□

8.2.5. Példa. Számoljuk ki a $\xi\eta$ szorzat várható értékét.

A korábbi példából a $\xi\eta$ sűrűségfüggvénye

$$h(u) = \int_{-\infty}^{\infty} f\left(\frac{u}{v}, v\right) \frac{1}{|v|}\,dv.$$

A várható érték képlete alapján

$$\begin{aligned}
\mathbf{E}(\xi\eta) &= \int_{-\infty}^{\infty} u\cdot h(u)\,du = \int_{-\infty}^{\infty} u \int_{-\infty}^{\infty} f\left(\frac{u}{v}, v\right)\frac{1}{|v|}\,dvdu = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uf\left(\frac{u}{v}, v\right)\frac{1}{|v|}\,dvdu = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uf\left(\frac{u}{v}, v\right)\frac{1}{|v|}\,dudv.
\end{aligned}$$

[9] Valójában a bizonyítást fordítva kell olvasni, ugyanis a pontos állítás az, hogy amennyiben létezik a ξ és az η (véges) várható értéke, akkor az összegnek is létezik a (véges) várható értéke, és az a várható értékek összege lesz. Vagyis a várható értékek összevonhatóak. A várható érték azonban elvileg nem feltétlenül szedhető szét.

A belső integrálban $u/v = z$ helyettesítést végezve és kihasználva, hogy $du/dz = v$

$$\begin{aligned}\mathbf{E}(\xi\eta) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} vzf(z,v)\frac{|v|}{|v|}dzdv = \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} vzf(z,v)\,dzdv,\end{aligned}$$

ami éppen a transzformált valószínűségi változók várható értékére vonatkozó képlet. A számolás során kihasználtuk, hogy ha v negatív, akkor a helyettesítés során az integrálás iránya megváltozik, és így változatlan integrációs határok mellett a $du = vdz$ helyettesítés helyett a $du = |v|\,dz$ helyettesítést kell alkalmazni. Vegyük észre, hogy a bizonyítás első felében az integrálokat csak akkor lehet felcserélni, ha a $|\xi\eta|$ változónak létezik várható értéke, ami persze ekvivalens avval, hogy a $\xi\eta$ szorzatnak létezik várható értéke. Ha ξ és η függetlenek, akkor

$$\begin{aligned}\mathbf{E}(\xi\eta) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} vzg(z)\,r(v)\,dzdv = \\ &= \int_{-\infty}^{\infty} vr(v)\int_{-\infty}^{\infty} zg(z)\,dzdv = \\ &= \int_{-\infty}^{\infty} vr(v)\,\mathbf{E}(\xi)\,dv = \mathbf{E}(\xi)\int_{-\infty}^{\infty} vr(v)\,dv = \\ &= \mathbf{E}(\xi)\,\mathbf{E}(\eta).\end{aligned}$$

Diszkrét változók esetén a transzformációs formula igazolása lényegében azonos módon végezhető el. Ilyenkor ha $p_{ij} \mathrel{\overset{\circ}{=}} \mathbf{P}(\xi = i, \eta = j)$, akkor

$$\mathbf{P}(\xi\eta = k) = \sum_l p_{lk/l}$$

ahol értelemszerűen $p_{lk/l} = 0$, ha a k/l nem egész.

$$\begin{aligned}\mathbf{E}(\xi\eta) &= \sum_k k\mathbf{P}(\xi\eta = k) = \sum_k k\sum_l p_{lk/l} = \\ &= \sum_k\sum_l kp_{lk/l} = \sum_i\sum_j ijp_{ij}.\end{aligned}$$

□

IX.

A NORMÁLIS ELOSZLÁS ÉS BARÁTAI

Ebben a fejezetben a standard normális eloszlás és a belőle származtatható eloszlásokat tekintjük át. A standard normális eloszlás sűrűségfüggvénye

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

A standard normális eloszlás várható értéke nulla, a szórása pedig egy. Az $N(\mu, \sigma)$ eloszlást, vagyis a μ várható értékű és σ szórású normális eloszlást az $\eta = \sigma\xi + \mu$ transzformációval kaphatjuk, ahol értelemszerűen a ξ eloszlása $N(0, 1)$. Az $y = \sigma x + \mu$ függvény inverze $x = (y - \mu)/\sigma$. A transzformált változók sűrűségfüggvényének képlete alapján az $N(\mu, \sigma)$ eloszlás sűrűségfüggvénye[1]

$$f(x) = \varphi_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Független normális eloszlású valószínűségi változók összege szintén normális eloszlású, ahol a várható értékek összeadódnak és a szórások a $\sqrt{\sigma_1^2 + \sigma_2^2}$ képlet szerint alakulnak. Ennek igazolása, mármint hogy az összeg normális eloszlású, viszonylag körülményes és a legegyszerűbben a karakterisztikus függvények segítségével végezhető el. Egy viszonylag átlátható közvetlen igazolás a következő: Legyenek ξ_1 és ξ_2 független standard normális eloszlású változók. Be kell látni, hogy a $\sigma_1\xi_1 + \mu_1 + \sigma_2\xi_2 + \mu_2$ eloszlása szintén normális. Ehhez elég belátni, hogy az $\eta \overset{\circ}{=} \sigma_1\xi_1 + \sigma_2\xi_2$ összeg eloszlása normális, ugyanis a konstans hozzáadása nem módosítja a normalitást. A konvolúciós formula direkt használata hosszú számolást eredményez, ezért egy közvetlen utat választunk:

$$\begin{aligned}
\mathbf{P}(\eta < z) &= \mathbf{P}(\sigma_1\xi_1 + \sigma_2\xi_2 < z) = \\
&= \frac{1}{2\pi} \int\int_{\sigma_1 x + \sigma_2 y < z} \exp\left(-\frac{x^2 + y^2}{2}\right) dxdy.
\end{aligned}$$

Az integrandus körszimmetrikus, így az integrál értéke csak attól függ, hogy a $\sigma_1 x + \sigma_2 y = z$ egyenes az origótól milyen távolságra van, ugyanis alkalmas módon elforgatva a síkot, az $\sigma_1 x + \sigma_2 y < z$ félsík egy $y < c$ félsíkba megy át, ahol a c értéke éppen a $\sigma_1 x + \sigma_2 y = z$ egyenesnek az origótól való (előjeles) távolságával egyezik meg. Könnyen

[1] Illik hangsúlyozni, hogy az irodalomban az $N(\mu, \sigma)$ jelölés nem egyértelmű. Gyakran találkozhatunk az $N(\mu, \sigma^2)$ jelöléssel is, amit a többdimenziós eloszlás kapcsán használt formalizmussal való kompatibilitás indokol.

kiszámolható, hogy az egyenes az origótól $c = z/\sqrt{\sigma_1^2+\sigma_2^2}$ távolságra van. Ebből

$$\begin{aligned}
\mathbf{P}(\eta < z) &= \frac{1}{2\pi}\int_{-\infty}^{c}\int_{-\infty}^{\infty}\exp\left(-\frac{x^2+y^2}{2}\right)dxdy = \\
&= \frac{1}{2\pi}\int_{-\infty}^{c}\exp\left(-\frac{y^2}{2}\right)\int_{-\infty}^{\infty}\exp\left(-\frac{x^2}{2}\right)dxdy = \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c}\exp\left(-\frac{y^2}{2}\right)dy = \\
&= \Phi(c) = \Phi\left(\frac{z}{\sqrt{\sigma_1^2+\sigma_2^2}}\right),
\end{aligned}$$

ahol a Φ a standard normális eloszlás eloszlásfüggvénye. Így az összeg $\sigma \circeq \sqrt{\sigma_1^2+\sigma_2^2}$ szórással és nulla várható értékkel rendelkező normális eloszlású változó. A távolság értékére azonban nincs szükség, elég azt tudni, hogy a c távolság a z lineáris függvénye. Ebből a normalitás már következik, és a szórásra vonatkozó képlet pedig már általános megfontolások alapján is nyilvánvaló.

9.1. Normális eloszlású változók szimulálása

Érdemes felidézni, hogy ha ξ eloszlása egyenletes a $[0,1]$ intervallumon, és $\eta \circeq F^{-1}(\xi)$, ahol F egy folytonos, szigorúan monoton növekedő eloszlásfüggvény, akkor

$$\mathbf{P}(\eta < x) \circeq \mathbf{P}\left(F^{-1}(\xi) < x\right) = \mathbf{P}(\xi < F(x)) = F(x),$$

vagyis az η eloszlásfüggvénye F. Megfordítva, ha η eloszlásfüggvénye F és $0 < x < 1$, akkor

$$\mathbf{P}(F(\eta) < x) = \mathbf{P}\left(\eta < F^{-1}(x)\right) = F\left(F^{-1}(x)\right) = x,$$

vagyis ilyenkor $F(\eta)$ eloszlása egyenletes. Ez az összefüggés számos eloszlás generálását teszi lehetővé.

9.1.1. Példa. Exponenciális eloszlású változó generálása.

Legyen $F(x) = 1-\exp(-\lambda x)$. Ebből

$$F^{-1}(y) = -\lambda^{-1}\ln(1-y).$$

Ha ξ egyenletes a $[0,1]$-en, akkor az $1-\xi$ is az, így az

$$\eta \circeq -\lambda^{-1}\ln\xi$$

eloszlása λ paraméterű exponenciális.

□

9.1.2. Példa. Cauchy-eloszlású változó szimulálása.

Ha η egy Cauchy-eloszlású valószínűségi változó, akkor η sűrűségfüggvénye

$$f(x)=\frac{1}{\pi\left(1+x^2\right)}.$$

Ebből az eloszlásfüggvénye

$$\begin{aligned}
F(x) &= \frac{1}{\pi}\int_{-\infty}^{x}\frac{1}{1+t^2}dt=[\arctan]_{-\infty}^{x}=\frac{1}{\pi}\left(\arctan x+\frac{\pi}{2}\right)= \\
&= \frac{1}{\pi}\arctan x+\frac{1}{2}, \\
F^{-1}(y) &= \tan\pi\left(y-\frac{1}{2}\right).
\end{aligned}$$

Ha ξ egyenletes a $[0,1]$ szakaszon, akkor az $\eta \mathrel{\hat{=}} \tan\pi(\xi-1/2)$ Cauchy-eloszlású. □

Mivel a standard normális eloszlás sűrűségfüggvénye

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right),$$

ezért a standard normális eloszlás eloszlásfüggvénye

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right)dt.$$

Az előzőek alapján kézenfekvőnek látszik, hogy az $N(0,1)$ változók szimulálására a $\Phi^{-1}(y)$ függvényt használjuk. Azonban ennek kiszámolása bonyolult, így ritkán használatos. Az alább bemutatott módszert szokás Box–Müller módszernek is mondani.

9.1.3. Tétel. *Ha δ_1,δ_2 független, a $(0,1)$ intervallumon egyenletes eloszlású változók, akkor a*

$$\xi \mathrel{\hat{=}} \sqrt{-2\ln\delta_1}\cos(2\pi\delta_2), \quad \eta \mathrel{\hat{=}} \sqrt{-2\ln\delta_1}\sin(2\pi\delta_2)$$

változók függetlenek, és a ξ és η változók eloszlása $N(0,1)$.

Bizonyítás: A módszer igazolása céljából vegyünk két független $N(0,1)$ változót, és a (ξ,η) párt tekintsük az $\mathbb{R}^2$ sík véletlenül kiválasztott pontjának. Mivel a változók függetlenek, ezért az együttes sűrűségfüggvényük a sűrűségfüggvények szorzata

$$\begin{aligned}
f(x,y) &= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)= \\
&= \frac{1}{2\pi}\exp\left(-\frac{1}{2}\left(x^2+y^2\right)\right).
\end{aligned}$$

Mi történik, ha (ρ,φ) polárkoordinátákra térünk át? Legyen

$$T(\rho,\varphi) \circeq \rho(\cos\varphi,\sin\varphi) = (\xi,\eta)$$

a polárkoordinátákról való visszatérést megadó inverz leképezés. A T az origón kívül, vagyis a $\rho > 0$ tartományon injektív, és az origótól eltekintve teljesíti az integráltranszformációs tétel feltételeit.

$$\left|\det\left(T'\right)\right| = \left|\det\begin{pmatrix} \cos\varphi & -\rho\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{pmatrix}\right| = \rho.$$

Az előző fejezet alapján polárkoordinátákban a sűrűségfüggvény

$$s(r,\varphi) = f(T)\left|\det\left(T'\right)\right| = \frac{1}{2\pi}\exp\left(-\frac{r^2}{2}\right)r = g(r)h(\varphi),$$

ahol

$$\begin{aligned} g(r) &\circeq r\exp\left(-\frac{r^2}{2}\right), \quad r > 0, \\ h(\varphi) &\circeq \frac{1}{2\pi}, \quad \varphi\in(0,2\pi), \end{aligned}$$

vagyis a polárkoordinátákra való áttérés után a ρ sugár és a φ szög függetlenek. A φ szög a $(0,2\pi)$ intervallumban egyenletes eloszlású, a ρ sugár eloszlásának sűrűségfüggvénye pedig

$$g(r) = r\exp\left(-r^2/2\right).$$

A gondolatmenet megfordítható: ha a (ρ,φ) pár eloszlása éppen ilyen, akkor a (ξ,η) két független normális eloszlást definiál. Normális eloszlású változók szimulálása tehát visszavezethető egy egyenletes, és egy $r\exp\left(-r^2/2\right)$ sűrűségfüggvényű változó szimulálására. Ha δ a $(0,1)$ intervallumon egyenletes eloszlású, akkor a $\rho \circeq \sqrt{-2\ln\delta}$ eloszlása

$$\mathbf{P}(\rho < r) = \mathbf{P}\left(\sqrt{-2\ln\delta} < r\right) = \mathbf{P}\left(\delta > \exp\left(-\frac{1}{2}r^2\right)\right) = 1-\exp\left(-\frac{1}{2}r^2\right),$$

amely sűrűségfüggvénye éppen az $r\exp\left(-r^2/2\right)$.

$\square$

9.2. A statisztika néhány eloszlása

A normális eloszlásból egy sor fontos eloszlás származtatható. Ebben az alfejezetben ezeket tekintjük át.

9.2.1. Példa. A χ_n^2 eloszlás.

Az n szabadságfokú χ_n^2 eloszlású változót mint n darab független, standard normális eloszlású változó négyzetének összegét definiáljuk. Ha $n = 1$, akkor, miként már láttuk, a χ_1^2 eloszlása $N(0,1)^2 = \Gamma(1/2, 1/2)$, így sűrűségfüggvénye

$$k_1(x) = \frac{1}{\sqrt{2\pi x}} \exp\left(-\frac{x}{2}\right), \quad x > 0.$$

Az általános esetben a χ_n^2 eloszlása $\Gamma(n/2, n/2)$, és a sűrűségfüggvénye

$$k_n(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} \exp\left(-\frac{x}{2}\right), \quad x > 0,$$

amely a gamma eloszlás additív tulajdonsága miatt evidens. A várható értéke

$$\mathbf{E}\left(\chi_n^2\right) = n\mathbf{E}\left(\chi_1^2\right) = n$$

ugyanis $\mathbf{E}\left(\chi_1^2\right) = \mathbf{E}\left(N^2(0,1)\right) = \mathbf{D}^2(N(0,1))$. A szórás meghatározása a következő:

$$\mathbf{D}^2\left(\chi_n^2\right) = n\mathbf{D}^2\left(\chi_1^2\right) = n(3-1) = 2n$$

ugyanis egy χ_1^2 eloszlású változó négyzete valójában egy $N(0,1)$ változó negyedik hatványa:

$$\begin{aligned}
\mathbf{E}\left(N(0,1)^4\right) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^4 \exp\left(-\frac{x^2}{2}\right) dx = \\
&= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^4 \exp\left(-\frac{x^2}{2}\right) dx = \\
&= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} u^2 \exp\left(-\frac{u}{2}\right) \frac{1}{2\sqrt{u}} du = \\
&= \frac{1}{\sqrt{2\pi}} \Gamma\left(\frac{5}{2}, \frac{1}{2}\right) = \frac{1}{\sqrt{2\pi}} \frac{\Gamma(5/2)}{(1/2)^{5/2}} = \\
&= \frac{1}{\sqrt{2\pi}} 2^{5/2} \frac{3}{2} \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = 3.
\end{aligned}$$

□

9.2.2. Példa. A χ_n eloszlás.

A χ_n eloszlást a χ_n^2 eloszlásból gyökvonással kapjuk. A transzformált valószínűségi változók sűrűségfüggvényének képlete szerint egy (a,λ) paraméterű gamma eloszlású változó gyökének sűrűségfüggvénye

$$\frac{\lambda^a}{\Gamma(a)} x^{2(a-1)} \exp\left(-\lambda x^2\right) 2x$$

amely $\lambda = 1/2$ és $a = n/2$ esetén

$$\begin{aligned} h_n(x) &= k_n\left(x^2\right) 2x = \frac{\left(x^2\right)^{n/2-1} \exp\left(-x^2/2\right)}{2^{n/2}\Gamma(n/2)} 2x = \\ &= \frac{1}{2^{n/2-1}\Gamma(n/2)} x^{n-1} \exp\left(-\frac{x^2}{2}\right). \end{aligned}$$

A várható érték kiszámolásához tekintsük egy gamma eloszlású változó gyökének várható értékét:

$$\begin{aligned} \mathbf{E}(\chi_n) &= \frac{\lambda^a}{\Gamma(a)} \int_0^\infty \sqrt{x} x^{a-1} \exp(-\lambda x)\, dx = \\ &= \frac{\lambda^a}{\Gamma(a)} \int_0^\infty x^{a+1/2-1} \exp(-\lambda x)\, dx = \\ &= \frac{\lambda^a}{\Gamma(a)} \Gamma\left(a + \frac{1}{2}, \lambda\right) = \\ &= \frac{\lambda^a}{\Gamma(a)} \frac{\Gamma\left(a+\frac{1}{2}\right)}{\lambda^{a+1/2}} = \frac{1}{\sqrt{\lambda}} \frac{\Gamma\left(a+\frac{1}{2}\right)}{\Gamma(a)}. \end{aligned}$$

Innen $a = n/2$ és $\lambda = 1/2$ helyettesítéssel

$$\mathbf{E}(\chi_n) = \sqrt{2} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)}.$$

A szórás pedig

$$\mathbf{D}(\chi_n) = \sqrt{\mathbf{E}\left(\chi_n^2\right) - \mathbf{E}(\chi_n)^2} = \sqrt{n - 2\left(\frac{\Gamma((n+1)/2)}{\Gamma(n/2)}\right)^2}.$$

□

9.2.3. Tétel. *Ha ξ eloszlása $\Gamma(a,\lambda)$ és η eloszlása $\Gamma(b,\lambda)$ valamint ξ és η függetlenek, akkor a ξ/η hányados eloszlása $\widetilde{B}(a,b)$, és a $\xi/(\xi+\eta)$ tört eloszlása $B(a,b)$.*

Bizonyítás: A hányados valószínűségi változó sűrűségfüggvényének képletét felírva, és kihasználva a két változó nem negativitását, ha $u > 0$, akkor

$$
\begin{aligned}
h(u) &= \int_0^\infty \frac{\lambda^a}{\Gamma(a)} (uy)^{a-1} \exp(-\lambda uy) \frac{\lambda^b}{\Gamma(b)} y^{b-1} \exp(-\lambda y)\, y dy = \\
&= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} u^{a-1} \int_0^\infty y^{a+b-1} \exp(-\lambda y(u+1))\, dy = \\
&= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} u^{a-1} \Gamma(a+b, \lambda(u+1)) = \\
&= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} u^{a-1} \frac{\Gamma(a+b)}{(\lambda(1+u))^{a+b}} = \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1} \frac{1}{(1+u)^{a+b}} = \frac{1}{B(a,b)} u^{a-1} \frac{1}{(1+u)^{a+b}}.
\end{aligned}
$$

Mivel

$$
\frac{\xi}{\eta+\xi} = \frac{\xi/\eta}{1+\xi/\eta},
$$

ezért elegendő a ξ/η változón $\varphi(x) = x/(1+x)$ transzformációt végezni. Miként már az előző fejezetben láttuk,

$$
\varphi^{-1}(x) = x/(1-x), \quad 0 < x < 1.
$$

A transzformált sűrűségfüggvény így

$$
\begin{aligned}
& \frac{1}{B(a,b)} \left(\frac{x}{1-x}\right)^{a-1} \frac{1}{\left(1+\frac{x}{1-x}\right)^{a+b}} \frac{1}{(1-x)^2} = \\
&= \frac{1}{B(a,b)} \left(\frac{x}{1-x}\right)^{a-1} (1-x)^{a+b} \frac{1}{(1-x)^2} = \\
&= \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}.
\end{aligned}
$$

□

9.2.4. Példa. A Fisher-féle F eloszlás.

Legyenek ξ_i és η_j független $N(0,1)$ eloszlású változók, ahol $1 \leq i \leq m$ és $1 \leq j \leq n$. Legyen

$$
F \circeq \frac{\sum_{i=1}^m \xi_i^2}{\sum_{i=1}^n \eta_i^2}.
$$

A statisztikában m,n szabadságfokú Fisher-féle F eloszláson az

$$
\widetilde{F} \circeq \frac{1/m \sum_{i=1}^m \xi_i^2}{1/n \sum_{i=1}^n \eta_i^2}
$$

változó eloszlását szokás érteni. A továbbiakban csak az F változóval foglalkozunk, az $\widetilde{F}$ változó eloszlását értelemszerű módosítással kaphatjuk. Az F két független χ^2 eloszlás hányadosa, így $\widetilde{B}(m/2,n/2)$ eloszlású, ezért a sűrűségfüggvénye

$$f(x)=\frac{1}{B(m/2,n/2)}x^{m/2-1}\frac{1}{(1+x)^{(m+n)/2}},\quad x>0.$$

A várható érték a korábban látottak alapján

$$\begin{aligned}\mathbf{E}(F) &= \frac{B(\alpha+1,\beta-1)}{B(\alpha,\beta)}= \\ &= \frac{\alpha}{\beta-1}=\frac{m}{n-2},\end{aligned}$$

feltéve, hogy $\beta-1=n/2-1>0$, vagyis $n>2$. Számoljuk ki a második momentumot. Az általánosság kedvéért számoljunk általában α és β paraméterekkel:

$$\mathbf{E}\left(F^2\right)=\frac{1}{B(\alpha,\beta)}\int_0^\infty x^2x^{\alpha-1}\frac{1}{(1+x)^{\alpha+\beta}}dx.$$

A másodfajú béta eloszlás momentumai alapján

$$\begin{aligned}\mathbf{E}\left(F^2\right) &= \frac{1}{B(\alpha,\beta)}\int_0^\infty x^2x^{\alpha-1}\frac{1}{(1+x)^{\alpha+\beta}}dx= \\ &= \frac{B(\alpha+2,\beta-2)}{B(\alpha,\beta)}=\frac{(\alpha+1)\alpha}{(\beta-1)(\beta-2)}= \\ &= \frac{(m/2+1)(m/2)}{(n/2-1)(n/2-2)}= \\ &= \frac{(m+2)m}{(n-2)(n-4)}.\end{aligned}$$

Ebből

$$\mathbf{D}^2(F)=\frac{(m+2)m}{(n-2)(n-4)}-\frac{m^2}{(n-2)^2}=\frac{2m(n+m-2)}{(n-2)^2(n-4)}.$$

□

9.2.5. Példa. Student t_n.

Legyen ξ eloszlása $N(0,1)$ és $t_n \mathrel{\hat{=}} \xi/\chi_n$, ahol a χ_n-ről feltesszük, hogy független a ξ-től és értelemszerűen χ_n eloszlású. A t_n változó eloszlását n szabadságfokú Student-eloszlásnak mondjuk. Tekintsük először a $N(0,1)^2/\chi_n^2$ hányadost, ahol értelemszerűen feltesszük, hogy a számláló független a nevezőtől. Ennek eloszlása $\widetilde{B}(1/2,n/2)$ másodfajú béta eloszlás, így a sűrűségfüggvénye

$$f(x)=\frac{1}{B(1/2,n/2)}x^{1/2-1}\frac{1}{(1+x)^{(n+1)/2}},\quad x>0.$$

Ha gyököt vonunk az eloszlásból, akkor az x^2 inverzzel való, már többször látott transzformációval a sűrűségfüggvény

$$\frac{1}{B(1/2,n/2)}\frac{1}{x}\frac{1}{\left(1+x^2\right)^{(n+1)/2}}2x,\quad x>0$$

módon alakul. Ez nyilván az $|N(0,1)|/\chi_n$ eloszlás sűrűségfüggvénye. Ha η egy szimmetrikus eloszlású, sűrűségfüggvénnyel rendelkező változó, akkor

$$\mathbf{P}(\eta<x)=\frac{1}{2}\begin{cases}(1-\mathbf{P}(|\eta|<-x)) & \text{ha} \quad x\le 0\\ (1+\mathbf{P}(|\eta|<x)) & \text{ha} \quad x>0\end{cases},$$

amiből deriválással könnyen látható, hogy η sűrűségfüggvénye $|\eta|$ sűrűségfüggvényéből tükrözéssel és kettővel való osztással kapható. Így tetszőleges x esetén a t_n sűrűségfüggvénye

$$\frac{1}{B(1/2,n/2)}\frac{1}{\left(1+x^2\right)^{(n+1)/2}}.$$

A statisztikában Student t_n eloszláson gyakran a

$$\widetilde{t}_n=\sqrt{n}t_n=\sqrt{n}\frac{\xi}{\chi_n}=\frac{\xi}{\sqrt{\left(\xi_1^2+\ldots+\xi_n^2\right)/n}}$$

változó eloszlását szokás érteni. A t_n eloszlásából a $\widetilde{t}_n$ eloszlása

$$\widetilde{F}(x)=\mathbf{P}\left(\widetilde{t}_n<x\right)=\mathbf{P}\left(\sqrt{n}t_n<x\right)=F\left(\frac{x}{\sqrt{n}}\right)$$

módon kapható, amiből a $\widetilde{t}_n$ sűrűségfüggvénye

$$\frac{1}{\sqrt{n}B(n/2,1/2)}\frac{1}{\left(1+x^2/n\right)^{(n+1)/2}}.$$

□

9.2.6. Példa. A Cauchy-eloszlás.

Mielőtt a Student-eloszlás várható értékének és a szórásának kiszámolására rátérnénk érdemes megvizsgálni az $n = 1$ esetet. Ha $n = 1$, akkor

$$t_1(x) = \frac{\Gamma(2/2)}{\sqrt{\pi}\Gamma(1/2)} \frac{1}{(1+x^2)} = \frac{1}{\sqrt{\pi}\sqrt{\pi}} \frac{1}{(1+x^2)} = \frac{1}{\pi(1+x^2)}.$$

Vagyis ha ξ és η független $N(0,1)$ eloszlásúak, akkor a $\xi/|\eta|$ hányados Cauchy-eloszlású. Ugyanakkor ha ξ és η standard normális eloszlásúak, akkor a ξ/η sűrűségfüggvénye

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-\frac{(xy)^2}{2}\right) \exp\left(-\frac{y^2}{2}\right) |y|\, dy =$$

$$= \frac{1}{2\pi}\int_{0}^{\infty} y \exp\left(-\frac{y^2(1+x^2)}{2}\right) dy =$$

$$= \frac{1}{2\pi}\left[-\frac{2}{1+x^2}\exp\left(-\frac{y^2(1+x^2)}{2}\right)\right]_0^{\infty} = \frac{1}{\pi}\frac{1}{1+x^2},$$

amely szintén Cauchy-eloszlású.

□

9.2.7. Példa. A t_n eloszlás várható értéke és szórása.

Mivel a sűrűségfüggvény páros, így a várható érték, ha létezik, csak nulla lehet. Ha $n = 1$, akkor nincs várható érték, ha azonban $n > 1$, akkor a várható értéket megadó integrál konvergens, így a várható érték nulla. Térjünk rá a szórásra. Mivel az $n = 1$ esetben a várható érték értelmetlen, ilyenkor a szórás is az. A szórás kiszámolásához ki kell számolni a második momentumot. Miként láttuk a t_n^2 másodfajú béta eloszlású $\alpha = 1/2$ és $\beta = n/2$ paraméterekkel. Legyen a ξ eloszlása (α, β) paraméterű másodfajú béta eloszlás. Számoljuk ki a ξ várható értékét: Miként már az előző fejezetben is láttuk,

$$\mathbf{E}(\xi) = \frac{\alpha}{\beta - 1}.$$

A levezetésből világos, hogy amennyiben $\beta - 1 = n/2 - 1 \leq 0$, vagyis $n < 3$, akkor a szórás végtelen. Ha $n \geq 3$, akkor

$$\begin{aligned} \mathbf{D}^2(t_n) &= \mathbf{E}(\xi) = \frac{\alpha}{\beta - 1} = \\ &= \frac{1/2}{n/2 - 1} = \frac{1}{n-2}. \end{aligned}$$

A már belátottak alapján $\widetilde{t}_n = \sqrt{n} t_n$ esetén világos, hogy ha $n \geq 2$, akkor a várható érték nulla, ha pedig $n \geq 3$, akkor a variancia

$$\mathbf{D}^2(\widetilde{t}_n) = \frac{n}{n-2}.$$

□

9.3. A lognormális eloszlás

Egy ξ változót lognormális eloszlásúnak mondunk, ha a logaritmusa normális eloszlású, vagyis $\xi = \exp(\eta)$ alakú, ahol az η eloszlása $N(\mu, \sigma)$. A definícióból világos, hogy a ξ a nem negatív számokra támaszkodik. A transzformált valószínűségi változók sűrűségfüggvényének képlete alapján ξ sűrűségfüggvénye

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x > 0.$$

A várható érték a transzformált valószínűségi változók várható értékének képlete szerint

$$\begin{aligned}
\mathbf{E}(\xi) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(x) \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2 - 2x\left(\mu + \sigma^2\right) + \mu^2}{2\sigma^2}\right) dx = \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{\left(x - \left(\mu + \sigma^2\right)\right)^2 - \left(\mu + \sigma^2\right)^2 + \mu^2}{2\sigma^2}\right) dx = \\
&= \exp\left(\frac{\left(\mu + \sigma^2\right)^2 - \mu^2}{2\sigma^2}\right) = \exp\left(\frac{2\sigma^2\mu + \sigma^4}{2\sigma^2}\right) = \\
&= \exp\left(\mu + \frac{\sigma^2}{2}\right).
\end{aligned}$$

Vegyük észre, hogy a számolás során kihasználtuk, hogy az

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\left(x - \left(\mu + \sigma^2\right)\right)^2}{2\sigma^2}\right)$$

éppen az $N\left(\mu+\sigma^2,\sigma\right)$ sűrűségfüggvénye, így az integrálja 1. Hasonlóan kell kiszámolni a második momentumot:

$$\begin{aligned}
\mathbf{E}\left(\xi^2\right) &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(2x)\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx = \\
&= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\frac{x^2-2x\left(\mu+2\sigma^2\right)+\mu^2}{2\sigma^2}\right)dx = \\
&= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\frac{\left(x-\left(\mu+2\sigma^2\right)\right)^2-\left(\mu+2\sigma^2\right)^2+\mu^2}{2\sigma^2}\right)dx = \\
&= \exp\left(\frac{\left(\mu+2\sigma^2\right)^2-\mu^2}{2\sigma^2}\right) = \exp\left(\frac{4\sigma^2\mu+4\sigma^4}{2\sigma^2}\right) = \exp\left(2\mu+2\sigma^2\right).
\end{aligned}$$

Ebből

$$\begin{aligned}
\mathbf{D}(\xi) &= \sqrt{\exp\left(2\mu+2\sigma^2\right)-\exp\left(2\mu+\sigma^2\right)} = \\
&= \exp\left(\mu+\frac{\sigma^2}{2}\right)\sqrt{\exp\left(\sigma^2\right)-1} = \\
&= \mathbf{E}(\xi)\sqrt{\exp\left(\sigma^2\right)-1}.
\end{aligned}$$

Természetesen a fordított irányból is számolhatjuk a paramétereket. Ha m és s a ξ várható értéke és szórása, akkor az

$$\begin{aligned}
m &= \exp\left(\mu+\frac{\sigma^2}{2}\right), \\
s &= m\sqrt{\exp\left(\sigma^2\right)-1}.
\end{aligned}$$

egyenleteket a μ és σ értékekre megoldva

$$\begin{aligned}
\sigma^2 &= \ln\frac{s^2+m^2}{m^2} \\
\mu &= \ln m-\frac{\sigma^2}{2} = \ln m-\ln\sqrt{\frac{s^2+m^2}{m^2}} = \\
&= \ln\frac{m}{\sqrt{1+\frac{s^2}{m^2}}}.
\end{aligned}$$

X.

A POISSON-ELOSZLÁS

A normális eloszlás mellett a valószínűségszámítás másik alapvető eloszlása a Poisson-eloszlás. A Poisson-eloszlás a $k = 0, 1, 2, \ldots$ értékekre támaszkodik, vagyis az eloszlás, illetve a mögöttes valószínűségi változó diszkrét. Definíció szerint

$$\mathbf{P}(\xi = k) = \frac{\lambda^k}{k!} \exp(-\lambda), \quad k = 0, 1, 2, \ldots$$

ahol a λ az eloszlás paramétere. Miként már láttuk, a λ paraméter éppen az eloszlás várható értéke és egyúttal az eloszlás szórásnégyzete, vagyis $\mathbf{E}(\xi) = \mathbf{D}^2(\xi) = \lambda$. A Poisson-eloszlást konkrét feladatokban akkor szokás használni, ha valamilyen jelenség, például hibás elemek, darabszámát akarjuk modellezni és az alappopuláció számossága, amelyből az elemek származnak nagy vagy nem ismert. Ez a legtöbb esetben úgy jelentkezik, hogy nincsen megmondva, hogy mekkora populációból származnak a, mondjuk hibás elemek, de ismert a hibás elemek átlagos száma, és arra vagyunk kíváncsiak, hogy az adott átlag mellett például mi annak a valószínűsége, hogy nem találunk hibás elemet, vagy viszonylag kevés hibás elemet találunk[1]. Miként már láttuk, ha egy n elemből álló populációban az egyes elemek egymástól függetlenül valamely p valószínűséggel rendelkeznek valamilyen tulajdonsággal, például miként már említettük, mondjuk hibásak, akkor annak a valószínűsége, hogy éppen k elem fog rendelkezni az adott tulajdonsággal, binomiális eloszlást követ, vagyis

$$\mathbf{P}(\xi = k) \mathrel{\hat{=}} p_k = \binom{n}{k} p^k q^{n-k},$$

ahol értelemszerűen ξ a hibás elemek darabszáma és $q \mathrel{\hat{=}} 1 - p$. A binomiális eloszlás p_k értékét azonban relatíve nehéz kiszámolni, abban az értelemben, hogy ha a p kicsi és az n nagy, akkor az $\binom{n}{k}$ kifejezés sokszor egy nagy szám a p^k pedig egy kicsi szám lesz, és a tényleges valószínűség ezek szorzataként alakul. Például $\binom{1000}{10} = 2{,}634\,1 \times 10^{23}$, ami egy rendkívül nagy érték és ennek arányában a $p^k q^{n-k}$ szorzónak rendkívül kicsinek kell lenni ahhoz, hogy a szorzat értékére a $0 < p_{10} < 1$ teljesüljön. Éppen ezért a binomiális eloszlást érdemes Poisson-eloszlással közelíteni. Ha $\lambda \mathrel{\hat{=}} np$, akkor

$$\begin{aligned} p_k &= \binom{n}{k} p^k q^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n-k} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n} \mathrel{\hat{=}} \\ &\mathrel{\hat{=}} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n} R_n, \end{aligned}$$

[1] További tipikus példa egy adott területen levő festési hibák száma, vagy egy könyvben levő sajtóhibák száma. Szigorúan véve a Poisson-eloszlás feltételezése nem lehet helyes, ugyanis az eloszlás tartója az összes nem negatív egész számok halmaza és például egy könyvben csak véges számú sajtóhiba lehetséges. Ennek ellenére a szöveges példákban, amennyiben az alapeloszlásban nincs explicite megadva a darabszám, mindig Poisson-eloszlást tételezünk fel. (Hasonlóan a kockadobás példánál mindig 1/6 valószínűséget teszünk fel, bár szigorúan véve a valószínűség ettől, például anyaghibák miatt, elvileg eltérhet. Viszont éppen az erre vonatkozó információ hiánya miatt kell $1/6$-dal számolni.)

ahol az R_n módon jelölt korrekciós tényező éppen

$$R_n = \left(1 - \frac{\lambda}{n}\right)^{-k} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n}.$$

Ha most az n elég nagy, akkor, mivel a k fix és az n-hez képest kicsi, közelítőleg $R_n \approx 1$ és $(1-\lambda\backslash n)^n \approx \exp(-\lambda)$, tehát

$$p_k \approx \frac{\lambda^k}{k!} \exp(-\lambda),$$

vagyis a p_k közelítőleg Poisson-eloszlású. A közelítés meglepően pontos. Például, ha $n = 1000$ és $p = 0,005$, vagyis $\lambda = 5$, akkor a Possion-eloszlás és a binomiális eloszlás első tíz eleme

0	0,00673	0,00665
1	0,03368	0,03343
2	0,08422	0,08392
3	0,14037	0,14030
4	0,17546	0,17573
5	0,17546	0,17590
6	0,14622	0,14658
7	0,10444	0,10460
8	0,06527	0,06524
9	0,03626	0,03613

Az első 100 elem között a maximális eltérés 4×10^{-4}, és az átlagos eltérés, vagyis ahol a negatív és a pozitív eltérések kiegyenlítik egymást, 5×10^{-15}. Még ha $n = 100$ és $p = 0,05$ a maximális eltérés az összes elemre akkor is csak $0,0046$.[2] A Poisson-eloszlás azonban nem csak közelítő eloszlásként tekinthető. A fejezet célja annak bemutatása, hogy egy sor feladatban a Poisson-eloszlás igen természetes módon jelentkezik.

10.1. Lévy folyamatok

Sztochasztikus folyamat alatt egy olyan $X(t,\omega)$ kétváltozós függvényt értünk amely az $\omega \in \Omega$ paraméter szerint minden t időpontra valószínűségi változó. Az ω változó rögzítése esetén a $t \mapsto X(t,\omega)$ hozzárendelést a folyamat ω-hoz tartozó trajektóriájának mondjuk. A következő definíció igen természetes:

10.1.1. Definíció. A $t \geq 0$ időtengelyen értelmezett X folyamat Lévy-folyamat, ha

1. $X(0) = 0$,

[2] A számolásokat Matlab segítségével végeztem el. Mind a két eloszlást a Matlab logaritmizálva számolja ki, a faktoriálisokat a gamma függvény logaritmusával határozza meg.

2. az X független és stacionárius növekményű, és
3. a trajektóriák jobbról regulárisak, vagyis jobbról folytonosak, és minden időpontban van bal oldali határértékük.

Értelemszerűen egy X folyamatot független növekményűnek mondunk, ha akárhogyan veszünk egy $t_1 < t_2 < \ldots < t_n$ időpontsorozatot az

$$X(t_2) - X(t_1), X(t_3) - X(t_2), \ldots, X(t_n) - X(t_{n-1})$$

növekmények függetlenek[3]. Egy X folyamat stacionárius növekményű, ha tetszőleges $s > 0$ esetén az $X(t+s) - X(t)$ növekmény eloszlása csak az s-től függ, a t-től pedig nem. Ebből következően ha ξ valamilyen Lévy-folyamat valamilyen növekménye, akkor ξ korlátlanul osztható. Ha például $\xi = X(t) = X(t) - X(0)$, akkor a kt/n pontokban vett n darab $\xi_k \overset{\circ}{=} X(tk/n) - X(t(k-1)/n)$ növekmény éppen a ξ szükséges felbontását megadó teleszkopikus összeg. Bár nem túl egyszerű igazolni, de az állítás megfordítása is igaz: Tetszőleges korlátlanul osztható eloszláshoz létezik olyan X Lévy-folyamat, amelyre az $X(1)$ eloszlása éppen az adott korlátlanul osztható eloszlás. A Lévy-folyamatok és a korlátlanul osztható folyamatok azonosíthatósága indokolja a korlátlanul oszthatóság valószínűségszámításban játszott központi szerepét.

10.1.2. Példa. A legegyszerűbb Lévy-folyamat az azonosan nulla folyamat. Egy konstans értékű folyamat csak akkor Lévy-folyamat, ha a konstans értéke nulla. Minden $X(t) = a \cdot t$ alakú egyszerű lineáris trend Lévy-folyamat. Az $a \cdot t + b$ alakú lineáris függvény ha $b \neq 0$ nem Lévy-folyamat.

A sztochasztikus folyamatok elméletének egyik legfontosabb fogalma a megállási idő. A megállási idő mellett szokás még kilépési szabályokról is beszélni, de egyéb hasonló elnevezésekkel is találkozhatunk. A megállási idők véletlen időpontok, de nem minden véletlen időpont megállási idő. Szándékosan fogalmaztunk úgy, hogy a megállási idők véletlen időpontok és nem azt írtuk, hogy valószínűségi változók. Ennek oka az, hogy a megállási idők felvehetik a végtelen értéket is. A véletlen időpontok és a megállási idők közötti fő eltérés az, hogy egy megállási időről bekövetkezésének időpontjában tudjuk, hogy bekövetkezett. A pontos definíció ismételten messze vezetne, ezért némiképpen pontatlanul a megállási időket azonosítani fogjuk a találati időkkel.

10.1.3. Definíció. Legyen X egy sztochasztikus folyamat és B legyen egy halmaz a

$$\tau \overset{\circ}{=} \inf\{t \mid X(t) \in B\}$$

módon definiált függvényt a B halmaz találati idejének mondjuk. A valószínűségszámításban szokásos jelöléssel a τ az Ω alaptéren van értelmezve és minden $\omega \in \Omega$ kimenetel esetén tekinteni kell a $t \mapsto X(t,\omega)$ trajektóriát és minden ω esetén $\tau(\omega)$ legyen az

[3] Emlékeztetünk, hogy valószínűségi változókat függetlennek mondunk, ha az együttes eloszlásfüggvényük a peremeloszlások szorzataként írható fel.

„első” olyan időpont, ahol a $t \mapsto X(t,\omega)$ trajektória belép a B halmazba. Ha valamely ω kimenetelre a $t \mapsto X(t,\omega)$ trajektória soha nem lép be a B halmazba, akkor az üres halmaz infimumára vonatkozó konvenciónak megfelelően definíció szerint $\tau(\omega) \triangleq \infty$.

A találati időhöz hasonlóan definiálhatjuk az n-edik találati idő, vagyis azt a „első” időpontot, amikor az X folyamat n-edszer lép be a B halmazba. Ezt rekurzióval definiálhatjuk: $\tau_0 \triangleq 0$ és

$$\tau_{n+1} \triangleq \inf\{t > \tau_n \mid X(t) \in B\}.$$

Az n-edik találati idő minden n-re szintén megállási idő. Ugyanakkor az az időpont, amikor az X utoljára lép be a B halmazba nem megállási idő. Hasonlóképpen az az időpont, amikor egy szakaszon a trajektória felveszi a maximumát szintén nem megállási idő, ugyanis csak később derül majd ki, hogy az adott szakaszon mennyi is volt a maximális érték. A Lévy-folyamatok elméletének legfontosabb tétele a következő:

10.1.4. Tétel (Erős Markov-tulajdonság)**.** *Ha $\tau < \infty$ egy tetszőleges véges megállási idő, X pedig egy tetszőleges Lévy-folyamat, akkor az*

$$X^*(t,\omega) \triangleq X(\tau(\omega)+t,\omega) - X(\tau(\omega),\omega), \quad t \geq 0,$$

újraindított folyamat eloszlásban megegyezik az X-szel, és az $\{X^*(t) \mid t \geq 0\}$ változók függetlenek az τ előtt bekövetkezett eseményektől.

A Lévy-folyamatok definíciójából evidens, hogy tetszőleges fix s időpont esetén az s időpontban újraindított $X^*(t) \triangleq X(t+s) - X(s)$ folyamat valószínűségszámítási értelemben megkülönböztethetetlen az eredeti X folyamattól. Vagyis pusztán a növekmény megfigyeléséből nem lehet arra következtetni, hogy például mikor indult el az eredeti folyamat. Másképpen az $X(t)$ állapot egyedül az $X(s)$ állapottól és a folyamat s és a t közötti alakulásától függ, és nem függ attól, hogy milyen módon jutott el a folyamat az $X(s)$ állapotba. Ezt szokás az X Markov-tulajdonságának mondani. Az erős Markov-tulajdonság ennél annyiban erősebb, hogy az s helyébe egy $\tau < \infty$ véges megállási idő is írható. Például a folyamatot akkor indítjuk újra, amikor az belép egy adott B halmazba. Vagyis a folyamat nem csak egy fix időpontig vezető múltat felejti el, hanem azt is ahogyan a B halmazig eljutott. A jövő szempontjából egyedül az érdekes, hogy hol lépett be a B halmazba, de például érdektelen az, hogy mikor lépett oda be, Feltéve persze, hogy 1 valószínűséggel belép a halmazba.

10.2. Poisson-folyamatok

10.2.1. Definíció. Poisson-folyamat alatt, definíció szerint, olyan monoton növekedő trajektóriákkal rendelkező Lévy-folyamatot értünk, amely értékkészlete majdnem minden ω kimenetelre a $\{0,1,2,\cdots\}$ egész számok halmaza. Hangsúlyozni kell, hogy definíció szerint, minden kimenetelre az $\mathbb{N}$ összes eleme felvételre kerül, vagyis nincsenek a folyamatnak 1-nél nagyobb ugrásai. Ennek megfelelően a Poisson-folyamatok éppen a Lévy-típusú számláló folyamatok.[4]

Mivel a folyamat értékei egész számok, és mivel a trajektóriák jobbról folytonosak, továbbá rendelkeznek bal oldali határértékkel, ezért az egyes ugrások közötti szakaszok hossza pozitív. Mivel a folyamat a teljes $t \geq 0$ félegyenesen értelmezve van, és csak véges értéket vehet fel, ezért az ugrások időpontjai nem torlódhatnak véges értékhez. Az értékkészletre tett megkötés alapján a folyamat trajektóriái egységnyi magasságú ugrásokat tartalmaznak.

Az ugrások között eltelt idő exponenciális eloszlású

10.2.2. Állítás. *Legyen X egy Poisson-folyamat. Az első ugrás helyét megadó*

$$\tau_1(\omega) \stackrel{\circ}{=} \inf\{t \mid X(t,\omega) = 1\} = \inf\{t \mid X(t,\omega) > 0\} < \infty$$

találati idő eloszlása exponenciális.

Bizonyítás: A független és stacionárius növekmény feltételét felhasználva

$$\begin{gathered}
\mathbf{P}(\tau_1 > t+s) = \\
= \mathbf{P}(X(t+s) = 0) = \mathbf{P}(X(s) - X(0) = 0, X(t+s) - X(s) = 0) = \\
= \mathbf{P}(X(s) - X(0) = 0)\,\mathbf{P}(X(t+s) - X(s) = 0) = \\
= \mathbf{P}(X(s) = 0)\,\mathbf{P}(X(t) = 0) = \\
= \mathbf{P}(\tau_1 > s)\,\mathbf{P}(\tau_1 > t).
\end{gathered}$$

Ha bevezetjük az $f(t) \stackrel{\circ}{=} \mathbf{P}(\tau_1 > t)$ függvényt, akkor f kielégíti az

$$f(t+s) = f(t)\,f(s)$$

úgynevezett Cauchy-egyenletet. Ha most $a \stackrel{\circ}{=} f(1)$, akkor az egyenlet szerint $f(n) = a^n$, illetve $f(1/n)^n = f(1) = a$, vagyis $f(1/n) = a^{1/n}$. Ebből tetszőleges $p/q \geq 0$ racionális számra $f(p/q) = a^{p/q}$. Az f függvény definíciójából világos, hogy az f jobbról folytonos, így minden $t \geq 0$ esetén $f(t) = a^t$. Az a nem lehet nulla, ugyanis akkor $f(t) \equiv 0$, vagyis minden t-re $\tau_1 \leq t$, vagyis $\tau_1 = 0$, tehát az X azonnal kilép a nulla pontból, ami ellentmond annak, hogy az X jobbról folytonos. Hasonlóan az $a = 1$ azt jelenti, hogy a $\tau_1 = \infty$, ami csak akkor lehetséges, ha az X azonosan nulla, ami szintén lehetetlen, ugyanis az X egy Poisson-folyamat. Mivel $0 < a < 1$, ezért alkalmas $0 < \lambda < \infty$ számra

$$\mathbf{P}(\tau_1 > t) = \mathbf{P}(X(t) = 0) = \exp(-\lambda t).$$

□

[4] Értelemszerűen számláló folyamaton olyan monoton növekedő folyamatokat értünk, amelyek trajektóriáinak értékkészlete az $\mathbb{N} = \{0, 1, 2, \ldots\}$ halmaz.

A λ értéket szokás a folyamat paraméterének vagy intenzitásparaméterének mondani. Az exponenciális eloszlás miatt a τ_1 véges, így alkalmazható az erős Markov-tulajdonság. Az erős Markov-tulajdonság miatt a $X_1^*(t) \stackrel{\circ}{=} X(\tau_1 + t) - X(\tau_1)$ eloszlása azonos a $X(t)$ eloszlásával, így vehetjük a X második ugrásainak helyét megadó

$$\tau_2(\omega) \stackrel{\circ}{=} \inf\{t \mid X(t + \tau_1(\omega), \omega) = 2\} = \inf\{t : X_1^*(t, \omega) > 0\} < \infty$$

megállási időt. Ugyancsak az erős Markov-tulajdonság miatt a τ_1 és a τ_2 függetlenek, és az eloszlásuk azonos. Hasonlóan folytatva kapjuk a következőt:

10.2.3. Következmény. *Egy Poisson-folyamat, esetén az egyes ugrások között eltelt idő exponenciális eloszlású valószínűségi változó. A különböző ugrások között eltelt időszakok hossza független, és az egyes időhosszak eloszlása azonos paraméterű exponenciális eloszlást követ.*

10.2.4. Következmény. *Egy λ paraméterű Poisson-folyamat n-edik ugrásainak időpontja $\Gamma(n, \lambda)$ eloszlású.*

Bizonyítás: Az egyes ugrások közötti időhossz eloszlása $\Gamma(1, \lambda)$ és n darab független $\Gamma(1, \lambda)$ eloszlású változó összegének eloszlása $\Gamma(n, \lambda)$.

□

10.2.5. Tétel. *Ha X egy λ paraméterű Poisson-folyamat, akkor tetszőleges t időpontra*

$$\mathbf{P}(X(t) = n) = \frac{(\lambda t)^n}{n!} \exp(-\lambda t).$$

Bizonyítás: Legyen $\sigma_n \stackrel{\circ}{=} \sum_{k=1}^{n} \tau_k$, ahol τ_k az egyes ugrások között eltelt idő. A σ_{n+1} eloszlása $\Gamma(n+1, \lambda)$.

$$\mathbf{P}(X(t) < n+1) = \mathbf{P}(\sigma_{n+1} > t) = \int_t^\infty \frac{\lambda^{n+1}}{\Gamma(n+1)} x^n \exp(-\lambda x)\, dx =$$

$$= \left[\frac{\lambda^{n+1} x^n}{\Gamma(n+1)} \frac{\exp(-\lambda x)}{-\lambda} \right]_t^\infty + \int_t^\infty n \frac{\lambda^n x^{n-1}}{\Gamma(n+1)} \exp(-\lambda x)\, dx =$$

$$= \frac{(\lambda t)^n}{n!} \exp(-\lambda t) + \mathbf{P}(X(t) < n).$$

Ebből átrendezéssel

$$\mathbf{P}(X(t) = n) = \mathbf{P}(X(t) < n+1) - \mathbf{P}(X(t) < n) = \frac{(\lambda t)^n}{n!} \exp(-\lambda t).$$

Emlékeztetünk, hogy korábban már közvetlen számolással beláttuk, hogy a Poisson-eloszlás korlátlanul osztható. Vegyük észre, hogy ez tulajdonság következik a tételből is, ugyanis a Poisson-eloszlás egy Lévy-folyamathoz tartozó eloszlás.

□

Előrejelezhető megállási idők

10.2.6. Definíció. Egy $\tau > 0$ megállási időt előrejelezhetőnek mondunk, ha létezik megállási idők egy (ρ_n) sorozata, amelyre $\rho_n < \tau$ és $\rho_n \nearrow \tau$. Ez interpretációját tekintve nyilvánvalóan azt jelenti, hogy a (ρ_n) sorozat előrejelzi a τ bekövetkezését.

10.2.7. Tétel. *Egy Poisson-folyamat ugrásainak időpontjai nem előrejelezhetőek.*

Bizonyítás: Ha például az első ugrás időpontját megadó τ_1 előrejelezhető lenne, akkor az

$$X_n^*(t) \circeq X(\rho_n + t) - X(\rho_n)$$

újraindított folyamatok mindegyike Poisson-folyamat lenne, és az eloszlásuk megegyezne az X eloszlásával. Az X_n^* első ugrása éppen a $\tau_1 - \rho_n$ időpontban következik be. De a τ_1-nek, és így mindegyik $\tau_1 - \rho_n$ megállási időnek létezik várható értéke amely az exponenciális eloszlás várható értéke alapján minden n-re éppen $1/\lambda > 0$. Így

$$\frac{1}{\lambda} = \lim_{n\to\infty} \frac{1}{\lambda} = \lim_{n\to\infty} \mathbf{E}(\tau_1 - \rho_n) = \mathbf{E}\left(\lim_{n\to\infty}(\tau_1 - \rho_n)\right) = 0,$$

ami lehetetlen.

□

Poisson-folyamatok és az egyenletes eloszlás

Vegyünk egy n értéket. Jelölje σ_n valamely Poisson-folyamat n-edik ugrásának helyét. Mi lesz a

$$\left(\frac{\sigma_1}{\sigma_n}, \frac{\sigma_2}{\sigma_n}, \ldots, \frac{\sigma_{n-1}}{\sigma_n}\right)$$

véletlenül választott $n-1$ pont eloszlása a $(0,1)$ intervallumban?

10.2.8. Tétel. *Az eloszlás megegyezik a $(0,1)$ intervallumból vett egyenletes eloszlású mintából képzett rendezett minta eloszlásával.*

Bizonyítás: Legyenek $(\xi_k)_{k=1}^n$ független azonos, λ paraméterű exponenciális eloszlású valószínűségi változók. Legyen $\sigma_m \circeq \sum_{k=1}^m \xi_k$. Határozzuk meg az $\eta_k \circeq \sigma_k/\sigma_n$ változók eloszlását.

$$\mathbf{P}(\eta_1 < x) = \mathbf{P}\left(\frac{\xi_1}{\xi_1 + \ldots + \xi_n} < x\right).$$

A ξ_1 eloszlása $\Gamma(1,\lambda)$, a $\sum_{k=2}^n \xi_k$ eloszlása $\Gamma(n-1,\lambda)$. Ebből az η_1 eloszlása $B(1,n-1)$. A $B(1,n-1)$ eloszlás sűrűségfüggvénye

$$f_1(x) \circeq \frac{\Gamma(n)}{\Gamma(1)\Gamma(n-1)} x^{1-1}(1-x)^{n-2}, \quad x \in (0,1).$$

A $\Gamma(n) = (n-1)!$ értéket beírva

$$f_1(x) = (n-1)(1-x)^{n-2} \quad x \in (0,1).$$

Legyenek $(\tau_k)_{k=1}^{n-1}$ a $(0,1)$ intervallumon egyenletes eloszlású változók és jelölje τ_1^* a legkisebb elemet, vagyis $\tau_1^* \circeq \min \tau_k$. $\left\{\tau_1^* < x\right\}$ pontosan akkor, ha legalább egy elem az $(n-1)$-ből kisebb mint x, tehát

$$F_1(x) \circeq \mathbf{P}(\tau_1^* < x) = 1-(1-x)^{n-1}.$$

A τ_1^* sűrűségfüggvénye

$$F_1'(x) = (n-1)(1-x)^{n-2}$$

amely éppen azonos az $f_1(x)$ függvénnyel, vagyis az η_1 eloszlása azonos a τ_1^* eloszlásával. Hasonlóan a $\sum_{i=1}^{k} \xi_i$ eloszlása $\Gamma(k,\lambda)$ a $\sum_{i=k+1}^{n} \xi_i$ eloszlása $\Gamma(n-k,\lambda)$ így az η_k eloszlása $B(k, n-k)$, amely sűrűségfüggvénye

$$\frac{\Gamma(n)}{\Gamma(k)\Gamma(n-k)} x^{k-1}(1-x)^{n-k-1} = (n-1)\binom{n-2}{k-1} x^{k-1}(1-x)^{n-k-1}.$$

Határozzuk meg az τ_k^* eloszlásfüggvényét. Az egyszerűbb jelölés kedvéért legyen először τ_k^* egy egyenletes eloszlásból származó n elemű rendezett minta k-dik eleme. A $\left\{\tau_k^* < x\right\}$ esemény ekvivalens azzal, hogy legalább k változó kisebb mint x. Ebből

$$F_k(x) \circeq \mathbf{P}(\tau_k^* < x) = \sum_{i=k}^{n} \binom{n}{i} x^i (1-x)^{n-i}.$$

A derivált kiszámolásának komplikáltsága miatt a sűrűségfüggvény meghatározása a következő:

$$\frac{F_k(x+h) - F_k(x)}{h} = \frac{\mathbf{P}\left(x \le \tau_k^* < x+h\right)}{h}.$$

Tekintsük a $0 \le x < x+h$ intervallumokat. A $\mathbf{P}\left(x \le \tau_k^* < x+h\right)$ annak a valószínűsége, hogy legfeljebb $(k-1)$ változó kisebb mint x és legalább k változó kisebb mint $x+h$. Annak a valószínűsége, hogy r változó esik az $[x, x+h)$ intervallumba $h^r = o\left(h^{r-1}\right)$ nagyságrendű, így egyedül az $r=0$, illetve az $r=1$ eseteket kell megvizsgálnunk. Ha az $\left\{x \le \tau_k^* < x+h\right\}$ esemény teljesül, akkor az $r=0$ lehetetlen, így a sűrűségfüggvény meghatározásakor egyedül az $r=1$ esetet kell kiszámolnunk. Ilyenkor $k-1$ elem kisebb mint x, egy az $[x, x+h)$ intervallumban van és $n-k$ elem nagyobb mint x, vagyis

$$\begin{aligned}\mathbf{P}(x \le \tau_k^* < x+h) &= \binom{n}{1} \cdot \binom{n-1}{k-1} \cdot h \cdot x^{k-1} \cdot (1-x-h)^{n-k} + \\ &\quad + o(h).\end{aligned}$$

Ebből a sűrűségfüggvény

$$n\binom{n-1}{k-1} x^{k-1}(1-x)^{n-k} = \frac{1}{B(n-k+1,k)} x^{k-1}(1-x)^{n-k}, \tag{10.2.1}$$

vagyis $(n-k+1,k)$ paraméterű béta eloszlást alkot. Ha n helyébe $(n-1)$-et írunk, akkor éppen az η_k sűrűségfüggvényét kapjuk. Érdemes hangsúlyozni, hogy a bizonyítás annyiban egyszerűsíthető, hogy a konstansokat nem szükséges meghatározni, ugyanis abból, hogy

$$\mathbf{P}(x \leq \tau_k^* < x+h) = C \cdot h \cdot x^{k-1} \cdot (1-x-h)^{n-k} + o(h)$$

és

$$\begin{aligned} \frac{F_k(x+h) - F_k(x)}{h} &= \frac{\mathbf{P}\left(x \leq \tau_k^* < x+h\right)}{h} = \\ &= C \cdot x^{k-1} \cdot (1-x-h)^{n-k} + \frac{o(h)}{h}, \end{aligned}$$

ahonnan határátmenettel már következik, hogy a sűrűségfüggvény $Cx^{k-1}(1-x)^{n-k}$ alakú. Ebből pedig a C értéke már evidens.

□

10.2.9. Példa. Legyen N Poisson-eloszlású λ paraméterrel. A $(0,1)$ intervallumba N darab, N-től és egymástól független elhelyezkedésű, egyenletes eloszlású véletlen pontot helyezünk el. Mi lesz a $(0,1/3)$ intervallumba eső pontok számának az eloszlása?

Jelölje ξ a $(0,1/3)$ intervallumba eső pontok számát. Rögzítsük a $(0,1)$-be eső pontok számát: legyen $N=n$. Minden egyes pont $1/3$ valószínűséggel esik a $(0,1/3)$ intervallumba, és ezt a kísérletet n-szer ismételjük. Annak a valószínűsége, hogy az esemény k-szor következik be binomiális eloszlású:

.

$$\mathbf{P}(\xi = k \mid N = n) = \binom{n}{k}\left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{n-k}.$$

Ebből a teljes valószínűség tételével adódik, hogy

$$\begin{aligned} \mathbf{P}(\xi = k) &= \sum_{n=k}^{\infty} \binom{n}{k}\left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{n-k} \frac{\lambda^n}{n!}\exp(-\lambda) = \\ &= \frac{\lambda^k}{k!}\left(\frac{1}{3}\right)^k \exp(-\lambda) \sum_{n=k}^{\infty} \left(\frac{2}{3}\right)^{n-k} \frac{\lambda^{n-k}}{(n-k)!} = \\ &= \frac{(\lambda/3)^k}{k!}\exp(-\lambda)\exp\left(\frac{2}{3}\lambda\right) = \frac{(\lambda/3)^k}{k!}\exp\left(-\frac{\lambda}{3}\right), \end{aligned}$$

vagyis ξ is Poisson eloszlású $\lambda/3$ paraméterrel.

□

Várakozási idők eloszlása

10.2.10. Példa. Egy autóbuszmegállóból úti célunkhoz kétféle autóbusszal juthatunk el. Az egyikre átlagosan 5 percet, a másikra átlagosan 3 percet kell várni. Mennyit kell átlagosan várni, ha mindkét autóbusz jó? (Feltételezzük, hogy a várakozási idők függetlenek, mindkét várakozási idő exponenciális eloszlású.)

Jelöljük a két várakozási időt ξ -vel és η-val. Eloszlásaik paramétere $\lambda = 1/3$ és $\mu = 1/5$. Számítsuk ki a tényleges várakozási idő, a $\min(\xi, \eta)$ eloszlásfüggvényét.

$$\begin{aligned}\mathbf{P}(\min(\xi,\eta) > x) &= \mathbf{P}(\xi > x, \eta > x) = \mathbf{P}(\xi > x)\mathbf{P}(\eta > x) = \\ &= \exp(-\lambda x)\exp(-\mu x) = \exp(-(\lambda+\mu)x),\end{aligned}$$

amiből látható, hogy a várakozási idő eloszlásfüggvénye szintén exponenciális és az átlagos várakozási idő

$$\frac{1}{\lambda+\mu} = \frac{1}{1/3+1/5} = \frac{15}{8}\text{perc.}$$

□

10.2.11. Példa. A Lánchídat éjszakánként 400 izzólámpa világítja meg. Egy napon az összes égőt felújítják. Mennyi ideig marad a kivilágítás tökéletes? (A lámpák élettartama exponenciális eloszlású, és az átlagos élettartam 20000 óra.)

Az előző feladat gondolatmenete alapján

$$\mathbf{P}(\min(\xi_1, \xi_2, \ldots, \xi_{400}) > x) = \exp\left(-\frac{400}{20000}x\right),$$

amiből 50 óra átlagos hibátlan kivilágítási idő adódik.

□

10.2.12. Példa. Tegyük fel, hogy az autógumik élettartama exponenciális eloszlású, 60000 km átlagos élettartammal. Mi a valószínűsége, hogy egy 10000 km-es útra egy pótkerék elegendő?

Mint az előző feladatoknál, itt is az első meghibásodásig megtett út exponenciális eloszlású $60000/4 = 15000$ km átlagos úthosszal. Innen a paraméter $\lambda = 1/15000$. A következő meghibásodásig eltelt idő szintén ugyanilyen eloszlású. Két exponenciális eloszlású valószínűségi változó összegének eloszlása gamma eloszlás, amely sűrűségfüggvénye

$$f(x) = \lambda^2 x \exp(-\lambda x), \quad x > 0.$$

Ebből $a = 10000$ esetén

$$\mathbf{P}(\xi_1 + \xi_2 > a) = \int_a^\infty \lambda^2 x \exp(-\lambda x)\, dx = 0{,}8557.$$

Vegyük észre, hogy a feladat megoldása során kihasználtuk az exponenciális eloszlás úgynevezett örökifjú tulajdonságát. Ez alatt a következőt értjük: a $\xi \geq 0$ valószínűségi változót örökifjúnak mondjuk, ha minden $x, y \geq 0$ számok esetén

$$\mathbf{P}(\xi \geq x + y \mid \xi \geq y) = \mathbf{P}(\xi \geq x),$$

vagyis annak a valószínűsége, hogy a ξ „túléli” az $x+y$ életkort, feltéve, hogy már „túlélte” az y életkort éppen annyi, hogy túléli az x életkort. Mivel

$$\{\xi \geq x+y\} \cap \{\xi \geq y\} = \{\xi \geq x+y\}$$

ezért az örökifjú tulajdonság

$$\begin{aligned} \mathbf{P}(\xi \geq x+y \mid \xi \geq y) &= \frac{\mathbf{P}(\xi \geq x+y)}{\mathbf{P}(\xi \geq x)} = \\ &= \frac{1-F(x+y)}{1-F(y)} = \mathbf{P}(\xi \geq x) = 1-F(x) \end{aligned}$$

egyenlőség teljesülését jelenti. Ez az exponenciális eloszlásra teljesül, ugyanis

$$\mathbf{P}(\xi > x+y \mid \xi > y) = \frac{\exp(-\lambda(x+y))}{\exp(-\lambda y)} = \exp(-\lambda x).$$

Érdemes megjegyezni, hogy ha $G(x) \stackrel{\circ}{=} 1-F(x)$, akkor az örökifjú tulajdonság a $G(x+y) = G(x)G(y)$ egyenletre vezet, amelynek, miként láttuk az eloszlások körében egyedüli megoldása a $G(x) = \exp(-\lambda x)$, vagyis egyedül az exponenciális eloszlás örökifjú.

□

10.2.13. Példa. Az exponenciális eloszlás rendezett mintája.

Legyenek $\xi_1, \xi_2, \ldots, \xi_n$ független azonos λ paraméterhez tartozó exponenciális eloszlású változok. Jelölje $\xi_1^*, \xi_2^*, \ldots, \xi_n^*$ a belőlük készített rendezett mintát. Mi lesz a ξ_k^* eloszlása? A villanykörtés példa alapján a ξ_1^* eloszlása exponenciális lesz és a paramétere $n\lambda$ lesz. A transzformált valószínűségi változók sűrűségfüggvényének képlete alapján evidens, hogy a ξ_1/n éppen $n\lambda$ paraméterű exponenciális eloszlást követ. A villanykörtés példák alapján az elsőt követően a következő lámpa elromlásához, ugyanis már csak $n-1$ darab lámpa maradt, $\lambda(n-1)$ paraméterű exponenciális eloszlás szerint kell várakozni. Vegyük észre, hogy kihasználtuk az exponenciális eloszlás örökifjú tulajdonságát. Ebből következően a ξ_2^* eloszlása a második lámpa meghibásodásának időpontja, amely eloszlása megegyezik a $\xi_1/n + \xi_2/(n-1)$ eloszlásával. Ennek megfelelően a ξ_k^* eloszlása megegyezik a

$$\frac{\xi_1}{n} + \frac{\xi_2}{n-1} + \cdots + \frac{\xi_k}{n-k+1} \tag{10.2.2}$$

összeg eloszlásával. Ez a gondolatmenet azonban némiképpen pontatlan, ugyanis az örökifjú tulajdonságot csak heurisztikusan használhatjuk, mivel a tulajdonság csak fix értékre és nem véletlen időpontokra igazoltuk. Ha valamely $F(x)$ eloszlású független változók (ξ_k) sorozatát rendezzük és az $F(x)$ folytonos és szigorúan növekedő, akkor a sorozat elemeit vehetjük $F^{-1}(\tau_k)$ alakúnak, ahol a τ_k változók egyenletes eloszlásúak a $(0,1)$ intervallumon. Mivel az F^{-1} függvény rendezéstartó, ezért $\xi_k^* = F^{-1}\left(\tau_k^*\right)$. A fenti (10.2.1) sor alapján a τ_k^* sűrűségfüggvénye[5]

$$r_k(x) = \frac{1}{B(n-k+1,k)} x^{k-1}(1-x)^{n-k}.$$

A transzformált változók sűrűségfüggvényének képletét használva a ξ_k^* sűrűségfüggvénye

$$h_k^*(x) = \frac{1}{B(n-k+1,k)} F^{k-1}(x)(1-F(x))^{n-k} f(x),$$

ahol a $f(x)$ a $F(x)$-hez tartozó sűrűségfüggvény. Exponenciális eloszlás esetén

$$\begin{aligned} h_k^*(x) &= \frac{1}{B(n-k+1,k)}(1-\exp(-\lambda x))^{k-1}(\exp(-\lambda x))^{n-k}\lambda\exp(-\lambda x) = \\ &= \frac{\lambda}{B(n-k+1,k)}(1-\exp(-\lambda x))^{k-1}(\exp(-\lambda x))^{n-k+1} = \\ &= \frac{\lambda}{B(n-k+1,k)}(1-\exp(-\lambda x))^{k-1}\exp(-\lambda(n-k+1)x). \end{aligned}$$

Amennyiben az imént vázolt (10.2.2) képlet igaz, akkor a ξ_{k+1}^* sűrűségfüggvénye

$$\int_0^x h_k^*(y)\lambda(n-k)\exp(-\lambda(n-k)(x-y))\,dy.$$

Ez a kitevők összevonása után éppen

$$\frac{\lambda^2(n-k)}{B(n-k+1,k)}(\exp(-\lambda x))^{n-k}\int_0^x(1-\exp(-\lambda y))^{k-1}\exp(-\lambda y)\,dy.$$

Az integrál értéke

$$\frac{1}{\lambda}\frac{(1-\exp(-\lambda y))^k}{k},$$

továbbá

$$\begin{aligned} \frac{n-k}{k\cdot B(n-k+1,k)} &= \frac{n-k}{k}\frac{n!}{(k-1)!(n-k)!} = \\ &= \frac{n!}{k!(n-k-1)!} = \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k)} \\ &= \frac{1}{B(n-k,k+1)}, \end{aligned}$$

amely összefüggéseket visszaírva evidens módon a $h_{k+1}(x)$ képletét kapjuk[6].

□

[5] A legegyszerűbben úgy jegyezhetjük meg, ha tekintjük a σ_k/σ_{n+1} sűrűségfüggvényét.

[6] A teljes indukciós bizonyításhoz szükséges $k=1$ eset igazolását az olvasó könnyen elvégezheti.

10.2.14. Példa. Az autógumis példa újratöltve.

A figyelmes olvasó az autógumis példában is észrevehette, hogy heurisztikusan használtuk az örökifjú tulajdonságot. Ha $\xi_1, \xi_2, \ldots, \xi_5$ az öt autógumi élettartama, akkor a teljes futásteljesítmény $\eta \overset{\circ}{=} \min\left(\xi_2^*, \xi_1^* + \xi_5\right)$, ahol $\left(\xi_k^*\right)$ az első négy gumi élettartamából készített rendezett minta. Mivel ξ_1^* eloszlása megegyezik a $\xi_1/4$ eloszlásával, ξ_2^* eloszlása megegyezik a $\xi_1^* + \xi_2/3$ eloszlásával, ezért η eloszlása

$$\min\left(\xi_1^* + \xi_2/3, \xi_1^* + \xi_5\right) = \xi_1^* + \min\left(\xi_2/3, \xi_5\right).$$

eloszlásával egyezik meg. Ugyanakkor

$$\begin{aligned} \mathbf{P}\left(\min\left(\xi_2/3, \xi_5\right) > x\right) &= \mathbf{P}\left(\xi_2/3 > x\right)\mathbf{P}\left(\xi_5/3 > x\right) = \\ &= \exp\left(-3\lambda x\right)\exp\left(-\lambda x\right) = \exp\left(-4\lambda x\right), \end{aligned}$$

vagyis, miként állítottuk, az η eloszlása két 4λ paraméterű független exponenciális valószínűségi változó összegének eloszlásával azonos.

□

XI.

MOMENTUMOK KISZÁMOLÁSA

A valószínűségszámítás gyakori trükkje, hogy a valószínűségszámítás valamilyen problémáját egy más területről ismert klasszikus matematikai feladat megoldására vezetjük vissza. Ebben a fejezetben a momentumok kiszámolásának problémáját mutatjuk be.

11.0.1. Definíció. Legyen ξ egy valószínűségi változó és jelölje F a ξ eloszlásfüggvényét.

1. A komplex számok[1] alkalmas részhalmazán értelmezett

$$z \mapsto \mathscr{M}(z) \stackrel{\circ}{=} \mathbf{E}(\exp(z \cdot \xi)) = \int_{-\infty}^{\infty} \exp(z \cdot x)\, dF(x)$$

leképezést, ahol z tetszőleges olyan komplex szám, amelyre a várható érték létezik, a ξ, illetve az F komplex momentumgeneráló függvényének mondjuk.

2. A valós számok valamely részhalmazán értelmezett

$$s \mapsto M(s) \stackrel{\circ}{=} \mathbf{E}(\exp(s \cdot \xi)) = \int_{-\infty}^{\infty} \exp(s \cdot x)\, dF(x)$$

függvényt a ξ változó, illetve az F eloszlás (valós) momentumgeneráló függvényének nevezzük, ahol az M értelmezési tartománya az olyan s valós számok halmaza, amelyekre az integrál véges.

3. A komplex momentumgeneráló függvény imaginárius tengely mentén vett

$$\begin{aligned} t \quad \mapsto \quad & \varphi(t) \stackrel{\circ}{=} \mathbf{E}(\exp(it \cdot \xi)) = \int_{-\infty}^{\infty} \exp(it \cdot x)\, dF(x) = \\ = \quad & \int_{-\infty}^{\infty} \cos(t \cdot x)\, dF(x) + i \int_{-\infty}^{\infty} \sin(t \cdot x)\, dF(x) \end{aligned}$$

„szeletét” az eloszlás karakterisztikus függvényének mondjuk. A karakterisztikus függvény, szemben a momentumgeneráló függvénnyel minden t valós szám esetén értelmes.

4. A valószínűségszámításban szokásos momentumgeneráló függvény mellett érdemes definiálni a komplex és valós Laplace-transzformációkat is[2]. A valós Laplace-transzformációt definiáló integrál

$$L(s) \stackrel{\circ}{=} \mathbf{E}(\exp(-s \cdot \xi)) = \int_{-\infty}^{\infty} \exp(-s \cdot x)\, dF(x),$$

[1] A komplex exponenciális függvénnyel kapcsolatos ismereteket röviden a centrális határeloszlás tételéről szóló fejezetben foglaljuk össze. Ebben a fejezetben nincs igazán szükségünk rá, és ezért a tárgyalást nem célszerű megszakítani vele.

[2] Laplace-transzformációt általában akkor használunk, ha ξ nem negatív. Ilyenkor a Laplace-transzformáció kényelmesebb, mint a momentumgeneráló függvény.

a komplex Laplace-transzformációt megadó integrál pedig értelemszerűen

$$\begin{aligned}\mathscr{L}(z) &\doteq \mathbf{E}(\exp(-z\cdot\xi)) = \int_{-\infty}^{\infty}\exp(-z\cdot x)\,dF(x) = \\ &= \int_{-\infty}^{\infty}\exp(-s\cdot x)(\cos(-t\cdot x)+i\sin(-t\cdot x))\,dF(x) = \\ &= \int_{-\infty}^{\infty}\exp(-s\cdot x)(\cos(t\cdot x)-i\sin(t\cdot x))\,dF(x),\end{aligned}$$

ahol az s valós a $z=s+it$ pedig komplex paraméter. A valós számokon értelmezett

$$\begin{aligned}\mathscr{F}(t) &= \mathbf{E}(\exp(-it\cdot\xi)) = \int_{-\infty}^{\infty}\exp(-it\cdot x)\,dF(x) = \\ &= \int_{-\infty}^{\infty}\cos(t\cdot x)-i\sin(t\cdot x)\,dF(x)\end{aligned}$$

függvényt az F, illetve a ξ Fourier-transzformáltjának is szokás mondani.

5. Diszkrét eloszlásokra szokás definiálni az úgynevezett generátorfüggvényt: Ha a ξ változó a $k=0,1,2,\ldots$ számok egy részhalmazán van értelmezve akkor a komplex számok alkalmas részhalmazán értelmezett

$$G(z)=\mathbf{E}\left(z^k\right)=\sum_k\mathbf{P}(\xi=k)\,z^k\doteq\sum_k p_k z^k$$

függvényt a ξ generátorfüggvényének mondjuk.

6. Gyakran hasznos a következő fogalom: Az $M(s)$ momentumgeneráló függvény

$$C(s)\doteq\ln M(s)$$

logaritmusát kumulánsgeneráló függvénynek mondjuk. A

$$C(s)=\sum_{k=1}^{\infty}c_k\frac{s^k}{k!}$$

sorfejtésében szereplő c_k együtthatókat kumulánsoknak[3] nevezzük.

Mielőtt elkezdjük a momentumok vizsgálatát, érdemes az egyes transzformációkat összevetni. A legegyszerűbb transzformáció a generátorfüggvény, amit értelemszerűen csak $k=0,1,\ldots$ értékű változók esetén tudunk használni. A $G(z)$ egy hatványsor, így az analízisben a hatványsorokra elmondottak érvényesek rá. Egy hatványsorral kapcsolatban a legfontosabb kérdés az, hogy mi a konvergenciasugara. Mivel

[3] Mivel $M(0)=1$, ezért $C(0)=0$, vagyis a hatványsor nulladik tagja nulla. Mivel független változók momentumgeneráló függvényei összeszorzódnak, a kumulánsgeneráló függvények, és ennek megfelelően a kumulánsok, összeadódnak. Éppen ez a „kumulatív” tulajdonság indokolja a kumuláns elnevezést.

a $G(z)$ együtthatói eloszlást alkotnak, ezért a $G(z)$ konvergenciasugara legalább 1. Minden hatványsor a konvergenciakörén belül deriválható, de a konvergenciakör határán általában nem tudunk róla mondani semmit. Ugyanakkor vegyük észre, hogy ha $z = \exp(it)$, akkor $|z| = 1$ és ilyenkor a $G(z)$ hatványsora konvergens. Ebből következően a $\varphi(t) = G(\exp(it))$ függvény minden t-re értelmes[4]. A $G(z)$ függvényt elegendő a $[0,1]$ szakaszon ismerni. Ha most tekintjük a $z = \exp(-s)$ transzformációt, akkor az $s \geq 0$ halmazt a $(0,1]$ halmazra képezzük. Így kézenfekvő tekinteni az $L(s) = G(\exp(-s))$ függvényt, amely éppen a Laplace-transzformált. Ha a $z = \exp(s)$ transzformációt tekintjük, akkor az $s \leq 0$ halmazt képezzük a $(0,1]$ intervallumra. Bármelyik transzformációt tekintjük is, egy olyan függvényt kapunk, amely egy az origót nem tartalmazó félegyenesen végtelen sokszor deriválható. A kérdés csak az, hogy mi történik az origóban? Előfordulhat azonban, hogy a $G(z)$ konvergenciasugara nagyobb mint 1. Ilyenkor a $G^{(n)}(1)$ deriváltak mindegyike létezik. Ha a hatványsor konvergenciasugara nagyobb 1-nél, akkor az $M(s)$, illetve az $L(s)$ értelmezve van a 0 pont egy környezetében és ott végtelen sokszor deriválható. Mivel tetszőleges komplex szám felírható $z = r\exp(it)$ alakban, ezért ilyenkor a komplex Laplace-transzformált értelmezve van az origó körüli sávban, ahol a z imaginárius része bármi lehet. Ha a valószínűségi változó nem negatív, akkor a momentumgeneráló függvénye, illetve a Laplace-transzformáltja egy a nullát tartalmazó félegyenesen konvergens és ezen félegyenes belsejében mindig végtelen sokszor deriválható. Azonban azt nem tudjuk, hogy ez a konvergenciatartomány ilyenkor átnyúlik-e az origó másik oldalára vagy sem. A momentumgeneráló függvény elnevezést a következő állítás indokolja:

11.0.2. Állítás. *Ha egy ξ változó momentumgeneráló függvénye értelmezve van a nulla egy I nyílt környezetében, akkor a ξ változónak létezik az összes momentuma, és*

$$\begin{aligned} M^{(n)}(0) &\overset{\circ}{=} \left.\frac{d^n}{ds^n}\mathbf{E}(\exp(s\xi))\right|_{s=0} = \mathbf{E}\left(\frac{d^n}{ds^n}\exp(s\xi)\right)\Big|_{s=0} = \\ &= \mathbf{E}(\xi^n \exp(s\xi))|_{s=0} = \mathbf{E}(\xi^n). \end{aligned}$$

Vagyis ha az $M(s)$ véges a nulla egy nyílt környezetében, akkor az $M(s)$ függvény a nulla pontban végtelen sokszor deriválható és a nulla pontban vett deriváltak éppen a momentumok. A feltétel szerint a komplex momentumgeneráló függvény létezik a $\operatorname{Re}(z) \in I$ *sávban. Ilyenkor a karakterisztikus függvény $s = it$ „helyettesítéssel” számolható ki. Ugyancsak ilyenkor a momentumgeneráló függvény az $s = 0$ pont körül felírható*

$$M(s) = \sum_{k=0}^{\infty} \frac{M^{(k)}(0)}{k!} s^k = \sum_{k=0}^{\infty} \frac{\mathbf{E}\left(\xi^k\right)}{k!} s^k$$

módon.

[4] Vegyük ugyanakkor észre, hogy a $\varphi(t)$ 2π szerint periodikus függvény lesz. Ugyanakkor persze általában a karakterisztikus függvények nem periodikusak!!

11.0.3. Példa. Számoljuk ki az $f(x) \circeq (1/4)\exp\left(-\sqrt{|x|}\right)$ sűrűségfüggvényhez tartozó eloszlás momentumait és a momentumgeneráló függvényét!

Az

$$\begin{aligned}\int_{-\infty}^{\infty} \exp\left(-\sqrt{|x|}\right) dx &= 2\int_{0}^{\infty} \exp\left(-\sqrt{x}\right) dx = 2\int_{0}^{\infty} \exp(-u) \frac{dx}{du} du = \\ &= 2\int_{0}^{\infty} \exp(-u)\, 2u du = 4\Gamma(2) = 4\end{aligned}$$

összefüggés alapján világos, hogy az f sűrűségfüggvény. A függvény páros, ezért a páratlan momentumok, ha léteznek, mindegyike nulla. A páros momentumokat az alábbi számolással határozhatjuk meg:

$$\begin{aligned}\mathbf{E}\left(\xi^{2k}\right) &= \int_{-\infty}^{\infty} x^{2k} \exp\left(-\sqrt{|x|}\right) dx = 2\int_{0}^{\infty} x^{2k} \exp\left(-\sqrt{x}\right) dx = \\ &= 2\int_{0}^{\infty} u^{4k} \exp(-u)\, 2u du = 4\Gamma(4k+2) = 4(4k+1)!.\end{aligned}$$

Világos, hogy a

$$\sum_{k=0}^{\infty} \mathbf{E}\left(\xi^{k}\right) \frac{s^k}{k!} = \sum_{k=0}^{\infty} \frac{4(4k+1)!}{k!} s^k$$

sor konvergenciasugara nulla, és ennek megfelelően ha $s \neq 0$, akkor a momentumgeneráló függvény

$$M(s) = \int_{-\infty}^{\infty} \exp\left(sx - \sqrt{|x|}\right) dx \geq \int_{sx-\sqrt{|x|}\geq 0} 1 dx = \infty.$$

A példából látható, hogy az előző állítás nem fordítható meg, és a momentumok létezéséből még nem következik, hogy a momentumgeneráló függvény értelmezési tartománya nem csak a 0 pontból áll.

□

11.0.4. Példa. Tegyük fel, hogy a ξ momentumgeneráló függvény értelmezve van a 0 valamely nyílt környezetében. Számoljuk ki az első négy kumulánst!

$$\begin{aligned}c_1 &= C'(0) = \frac{M'(0)}{M(0)} = \mathbf{E}(\xi), \\ c_2 &= C''(0) = \frac{M''(0)M(0) - M'(0)^2}{M(0)^2} = \mathbf{E}\left(\xi^2\right) - \mathbf{E}^2(\xi) = \mathbf{D}^2(\xi), \\ c_3 &= C'''(0) = M'''(0) - 3M'(0)M''(0) + 2M'(0)^3 = \\ &= \mathbf{E}\left(\xi^3\right) - 3\mathbf{E}(\xi)\mathbf{E}\left(\xi^2\right) + 2\mathbf{E}(\xi)^3 = \mathbf{E}\left((\xi - \mathbf{E}(\xi))^3\right), \\ c_4 &= C^{(iv)}(0) = \mathbf{E}\left((\xi - \mathbf{E}(\xi))^4\right) - 3\mathbf{D}^4(\xi).\end{aligned}$$

Ennek fontos következménye, hogy egy eloszlás ferdesége[5]

$$\frac{\mathbf{E}\left((\xi-\mathbf{E}(\xi))^3\right)}{\mathbf{D}^3(\xi)}=\frac{c_3}{c_2^{3/2}},$$

illetve kurtózisa, magyarul lapultsága

$$\frac{\mathbf{E}\left((\xi-\mathbf{E}(\xi))^4\right)}{\mathbf{D}^4(\xi)}=\frac{c_4+3\mathbf{D}^4(\xi)}{\mathbf{D}^4(\xi)}=3+\frac{c_4}{c_2^2}.$$

□

11.0.5. Példa. Eloszlások definiálása függvénytranszformációval.

A függvénytranszformációk fontos felhasználása, hogy sok esetben az eloszlást a függvénytranszformációval definiáljuk. Ennek bemutatása igen messze vezetne, és csak egy példát mutatunk meg. Legyen $L(s)=\mathbf{E}(\exp(-s\xi))$ egy nem negatív változó Laplace-transzformáltja. Ilyenkor a deriváltakra minden $s>0$ esetén

$$L^{(n)}(s)=\frac{d^n}{ds^n}\mathbf{E}(\exp(-s\xi))=\mathbf{E}\left(\frac{d^n}{ds^n}\exp(-s\xi)\right)=(-1)^n\mathbf{E}(\xi^n\exp(-s\xi)),$$

amiből

$$(-1)^nL^{(n)}(s)=\mathbf{E}(\xi^n\exp(-s\xi))\geq 0.$$

Többek között $L(s)$ második deriváltja nem negatív, vagyis $L(s)$ konvex. Így például ha $c>0$, akkor az $\exp\left(-c\cdot s^2\right)$ függvény nem lehet egy nem negatív változó Laplace-transzformáltja. De ennél több is igaz. Mivel $L(s)>0$, ezért értelmes a $C(s)=\ln L(s)$ függvény is. A Cauchy-egyenlőtlenség alapján

$$\begin{aligned}L\left(\frac{s+t}{2}\right) &= \mathbf{E}\left(\exp\left(-\frac{s+t}{2}\xi\right)\right)\leq\\ &\leq \mathbf{E}\left(\exp\left(-\frac{s}{2}\xi\right)\exp\left(-\frac{t}{2}\xi\right)\right)=\\ &= \sqrt{\mathbf{E}(\exp(-s\xi))}\sqrt{\mathbf{E}(\exp(-\xi))}.\end{aligned}$$

Logaritmust véve

$$C\left(\frac{s+t}{2}\right)\leq\frac{1}{2}(C(s)+C(t)),$$

vagyis az $L(s)$ Laplace-transzformált nem csak konvex, hanem logkonvex is. Így például ha $\gamma>1$, akkor az

$$U(s)=\exp(-s^{\gamma}),\quad s\geq 0$$

[5] Az első három kumuláns alapján arra gondolhatunk, hogy a kumulánsok éppen a centrális momentumok. Az intő ellenpélda a standard normális eloszlás, amely negyedik centrális momentuma, vagyis a momentuma 3, miközben a negyedik kumulánsa nulla.

függvény nem lehet egy nem negatív változó Laplace-transzformáltja[6]. Analóg állítás igaz a momentumgeneráló függvényekre.

□

11.1. Függvénytranszformációk és az összegzés

11.1.1. Állítás. *Ha a ξ és az η változók függetlenek, akkor*

$$M_{\xi+\eta}(s) = M_{\xi}(s) M_{\eta}(s)$$

feltéve, hogy az s pontban mind a három momentumgeneráló függvény értelmes. Analóg állítások igazak a többi transzformációra.

Bizonyítás: Közvetlen számolással az indoklás evidens, ugyanis független valószínűségi változók szorzatának várható értéke a várható értékek szorzata, így

$$\begin{aligned} M_{\xi+\eta}(s) &\overset{\circ}{=} \mathbf{E}(\exp(s(\xi+\eta))) = \mathbf{E}(\exp(s\xi)\exp(s\eta)) = \\ &= \mathbf{E}(\exp(s\xi))\,\mathbf{E}(\exp(s\eta)) = M_{\xi}(s) M_{\eta}(s). \end{aligned}$$

Vegyük észre, hogy felhasználtuk a függetlenség azon fontos tulajdonságát, hogy ha ξ és η függetlenek, akkor a $h(\xi)$ és a $g(\eta)$ transzformált változók is függetlenek. Érdemes megjegyezni, hogy az állítás akkor is igaz, ha elhagyjuk azt a kitételt, hogy az s paraméter olyan, hogy mind a három transzformált értelmes. Miként a várható érték definíciója során megjegyeztük gyakran hasznos, ha megengedjük, hogy a várható érték végtelen legyen. Ha ezzel a definícióval, vagy inkább konvencióval élünk, akkor az egyenlőség az s paraméterre tett minden megkötés nélkül érvényben marad[7]. Az elemi valószínűségszámítás tárgyalásakor a végtelen értékű várható érték bevezetése esetleg zavaró lehet, így nem feltétlenül szerencsés, így evvel a konvencióval a tárgyalás során nem élünk[8]. Ugyancsak érdemes megjegyezni, hogy ha a három függvénytranszformált közül kettő véges, akkor a harmadiknak is végesnek kell lenni.

□

[6] Egy másik megfontolás a következő: Ha tekintjük az $U'(s) = (-1)\gamma s^{\gamma-1}U(s)$ deriváltat, akkor ha $\gamma > 1$, akkor $U'(0) = 0$, amelyből következne, hogy a várható érték nulla, ami lehetetlen ugyanis az eloszlás a nem negatív számokra koncentrálódik.

[7] Mivel a momentumgeneráló függvény mindig pozitív, ezért nincs szükség a gyakran hasznos nullaszor végtelen egyenlő nulla konvencióra sem.

[8] A végtelen várható érték megengedése egy sor tétel kimondását egyszerűbbé teszi, de az ilyen megfogalmazások elemi szinten több kárral, mint haszonnal járhatnak. Hasonló a helyzet az olyan típusú egyenlőségekkel, amelyek során bizonyos esetekben mind a két oldalon egyszerre állhat értelmetlen kifejezés, és így a tétel kimondása kerekebbé válhat.

11.2. Momentumok és faktoriális momentumok

11.2.1. Definíció. Legyen

$$(x)_n \circeq x(x-1)(x-2)\cdots(x-n+1) = \prod_{k=0}^{n-1}(x-k).$$

Az $\mathbf{E}((\xi)_n)$ értékre mint n-edik faktoriális momentum fogunk hivatkozni.

11.2.2. Példa. Deriváltak és a faktoriális momentumok.

Egy $G(z)$ generátorfüggvény, mint minden hatványsorösszeg, a konvergenciakörén belül mindig végtelen sokszor deriválható és a deriválás tagonként elvégezhető. Mivel (p_k) valószínűségeloszlás, ezért $G(1) = 1$, így a konvergenciasugár mindig legalább 1, vagyis a $(-1,1)$ nyílt intervallumban $G(z)$ végtelen sokszor deriválható. A tagonként való deriválhatóság miatt ha $|z| < 1$, akkor

$$\begin{aligned} G^{(n)}(z) &= \sum_{k=n}^{\infty} k(k-1)\dots(k-n+1)\,p_k z^{k-n} \circeq \\ &\circeq \sum_{k=n}^{\infty}(k)_n\,p_k z^{k-n}. \end{aligned}$$

Tegyük fel, hogy az n-edik momentum véges. Ilyenkor a $0 \le \mathbf{E}((\xi)_n) \le \mathbf{E}(\xi^n)$ n-edik faktoriális momentum is véges, így

$$\mathbf{E}((\xi)_n) = \sum_{k=n}^{\infty} k(k-1)\dots(k-n+1)\,p_k = \sum_{k=n}^{\infty}(k)_n\,p_k < \infty.$$

Mivel a $G^{(n)}(z)$ hatványsor együtthatói nem negatívak, ezért a $G(z)$ hatványsora a Weierstrass-kritérium miatt a teljes $[0,1]$ szakaszon egyenletesen konvergens. Az egyenletesen konvergens függvénysorok tagonként való deriválhatósága alapján[9] a $G^{(n)}(1)$ mint bal oldali derivált létezik és $G^{(n)}(1) = \mathbf{E}((\xi)_n)$.

□

11.2.3. Definíció. Jelölje $\left\{ {n \atop k} \right\}$ egy n elemű halmazból készíthető k elemű partíciók számát. Az $\left\{ {n \atop k} \right\}$ számokat másodfajú Stirling-számoknak mondjuk.

Könnyen látható, hogy teljesül az

$$\left\{ {n+1 \atop k} \right\} = \left\{ {n \atop k-1} \right\} + k\left\{ {n \atop k} \right\} \tag{11.2.1}$$

rekurzió. Ugyanis egy $n+1$ elemből álló halmaz valamely k elemből álló partíciója úgy keletkezhet, hogy az utolsó lépésben kapott $n+1$ sorszámú elemet vagy egy külön még

[9] Emlékeztetünk, hogy a deriváltaknak kell egyenletesen konvergálni.

üres „rekeszbe” tesszük és így a megmaradt n elemből $k-1$ halmazból álló partíciókat készítünk, vagy az $n+1$ sorszámú elemet a már meglevő k darab nem üres halmaz valamelyikébe tesszük be. Definíció szerint

$$\left\{ \begin{matrix} 0 \\ 0 \end{matrix} \right\} \circeq 1, \quad \left\{ \begin{matrix} n \\ 0 \end{matrix} \right\} \circeq \left\{ \begin{matrix} 0 \\ n \end{matrix} \right\} \circeq 0.$$

A rekurziós szabály alapján elemi számolással

$$\left\{ \begin{matrix} 1 \\ 1 \end{matrix} \right\} = 1$$
$$\left\{ \begin{matrix} 2 \\ 1 \end{matrix} \right\} = 1, \left\{ \begin{matrix} 2 \\ 2 \end{matrix} \right\} = 1$$
$$\left\{ \begin{matrix} 3 \\ 1 \end{matrix} \right\} = 1, \left\{ \begin{matrix} 3 \\ 2 \end{matrix} \right\} = 3, \left\{ \begin{matrix} 3 \\ 3 \end{matrix} \right\} = 1$$
$$\left\{ \begin{matrix} 4 \\ 1 \end{matrix} \right\} = 1, \left\{ \begin{matrix} 4 \\ 2 \end{matrix} \right\} = 7, \left\{ \begin{matrix} 4 \\ 3 \end{matrix} \right\} = 6, \left\{ \begin{matrix} 4 \\ 4 \end{matrix} \right\} = 1.$$

11.2.4. Példa. A momentumgeneráló függvény és a generátorfüggvény kapcsolata.

Vegyük észre, hogy ha m és n egész számok, akkor érvényes az

$$m^n = \sum_{\mathscr{P}} (m)_{|\mathscr{P}|} = \sum_{k=1}^{n} \sum_{|\mathscr{P}|=k} (m)_{|\mathscr{P}|} = \sum_{k=1}^{n} \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (m)_k \tag{11.2.2}$$

azonosság, ahol értelemszerűen $\mathscr{P}$ az n elemből álló halmaz összes partícióján fut végig és szintén értelemszerűen $|\mathscr{P}|$ jelöli a partíció elemeinek számát. Valóban: tekintsük egy n elemből álló halmaz összes leképezését egy m elemből álló halmazba. Az így kapott függvények száma nyilván m^n. Ugyanakkor minden f függvény egyértelműen azonosítható az $f^{-1}(y)$ alakú teljes inverzkép-függvényével. Minden ilyen teljes inverzkép-függvény egyértelműen azonosítható egy partícióval[10], illetve a partíció halmazain különböző konstans értékeket felvevő leképezéssel. Ha tehát adott egy k halmazból álló partíció, akkor az első halmazon még a teljes értékkészletből vagyis m elemből választhatunk, a következőn pedig már csak $m-1$ elemből. Az eljárást egész addig folytathatjuk, amíg a partíció elemei el nem fogynak. Miként láttuk

$$\begin{aligned} G^{(n)}(1) &= \sum_k k(k-1)(k-2)\cdots(k-n+1)\,p_k = \\ &\circeq \sum_k (k)_n\, p_k = \mathbf{E}\left((\xi)_n\right). \end{aligned}$$

[10] De persze nem minden partícióhoz tartozik függvény. Ha egy $\mathscr{P}$ partíció halmazainak száma nagyobb mint m, akkor nincs függvény a partícióhoz, és ilyenkor $(m)_{|\mathscr{P}|} = 0$.

A fenti (11.2.2) képletet ξ-re alkalmazva

$$\begin{aligned}\mathbf{E}\left(\xi^{n}\right) &= \sum_{\mathscr{P}}\mathbf{E}\left(\left(\xi\right)_{|\mathscr{P}|}\right)=\sum_{\mathscr{P}}G^{(|\mathscr{P}|)}\left(1\right)= \\ &= \sum_{k=1}^{n}\sum_{|\mathscr{P}|=k}G^{(k)}\left(1\right)=\sum_{k=1}^{n}\left\{ {n \atop k} \right\}G^{(k)}\left(1\right).\end{aligned}$$

□

11.2.5. Példa. Számoljuk ki a Poisson-eloszlás karakterisztikus, momentum- és a kumulánsgeneráló függvényeit! Számoljuk ki az eloszlás momentumait és kumulánsait!

1. A momentumgeneráló függvény

$$\begin{aligned}M\left(s\right) &= \sum_{k=0}^{\infty}\frac{\lambda^{k}}{k!}\exp\left(-\lambda\right)\exp\left(sk\right)=\sum_{k=0}^{\infty}\frac{\left(\lambda\exp\left(s\right)\right)^{k}}{k!}\exp\left(-\lambda\right)= \\ &= \exp\left(\lambda\left(\exp\left(s\right)-1\right)\right),\end{aligned}$$

amiből $C\left(s\right)=\ln M\left(s\right)=\lambda\left(\exp\left(s\right)-1\right)$. Mivel $C\left(s\right)=\lambda\sum_{k=1}^{\infty}s^{k}/k!$, ezért a Poisson-eloszlás összes kumulánsa λ. A kumulánsok felhasználásával a Poisson-eloszlás ferdesége

$$\frac{c_{3}}{c_{2}^{3/2}}=\frac{\lambda}{\lambda^{3/2}}=\frac{1}{\sqrt{\lambda}},$$

a kurtózisa

$$3+\frac{c_{4}}{c_{2}^{2}}=3+\frac{\lambda}{\lambda^{2}}=3+\frac{1}{\lambda}.$$

2. Mivel az $\exp\left(\lambda\left(\exp\left(z\right)-1\right)\right)$ a teljes komplex síkban értelmezve van, ezért $\varphi\left(t\right)=\exp\left(\lambda\left(\exp\left(it\right)-1\right)\right)$.

3. A Poisson-eloszlás esetén a $G\left(z\right)=\exp\left(\lambda\left(z-1\right)\right)$ sorfejtéséből világos, hogy $G^{(k)}\left(1\right)=\lambda^{k}$, következésképpen

$$\mathbf{E}\left(\xi^{n}\right)=\sum_{k=1}^{n}\left\{ {n \atop k} \right\}G^{(1)}\left(1\right)=\sum_{k=1}^{n}\left\{ {n \atop k} \right\}\lambda^{k}.$$

4. Végül a momentumok segítségével is számoljuk ki a kurtózist: Ehhez definíció szerint ki kell számolni a negyedik centrális momentumot.

$$\mathbf{E}\left(\xi-E\left(\xi\right)\right)^{4}=$$

$$\mathbf{E}\left(\xi^{4}\right)-\binom{4}{1}\mathbf{E}\left(\xi^{3}\right)\mathbf{E}\left(\xi\right)+\binom{4}{2}\mathbf{E}\left(\xi^{2}\right)\mathbf{E}^{2}\left(\xi\right)-\binom{4}{3}\mathbf{E}\left(\xi\right)\mathbf{E}^{3}\left(\xi\right)+\mathbf{E}^{4}\left(\xi\right)=$$

$$=\left(\lambda+7\lambda^{2}+6\lambda^{3}+\lambda^{4}\right)-4\left(\lambda^{2}+3\lambda^{3}+\lambda^{4}\right)+6\left(\lambda^{3}+\lambda^{4}\right)-4\lambda^{4}+\lambda^{4}=$$

$$=\lambda+3\lambda^{2}.$$

Ezt osztva a variancia négyzetével, vagyis λ^2-tel, a kurtózis, miként már láttuk, $3 + 1/\lambda$. Mivel az összes kumulánsa λ, ezért a negyedik centrális momentuma

$$\mathbf{E}(\xi - E(\xi))^4 = c_4 + 4\mathbf{D}^4(\xi) = \lambda + 3\lambda^2.$$

$\square$

11.2.6. Példa. Számoljuk ki a geometriai eloszlás momentumait a generátorfüggvénye segítségével.

A geometriai eloszlás generátorfüggvényének deriváltjait viszonylag egyszerű kiszámolni.

$$\begin{aligned}
G(z) &= \sum_{k=1}^{\infty} pq^{k-1}z^k = \frac{pz}{1-qz}, \\
G'(z) &= \frac{p(1-qz)+pzq}{(1-qz)^2} = \frac{p}{(1-qz)^2}, \\
G''(z) &= \frac{2pq(1-qz)}{(1-qz)^4} = \frac{2pq}{(1-qz)^3}, \\
&\vdots \\
G^{(k)}(z) &= \frac{k!pq^{k-1}}{(1-qz)^{k+1}}.
\end{aligned}$$

Ebből

$$G^{(k)}(1) = \frac{k!}{p}\left(\frac{q}{p}\right)^{k-1},$$

amiből

$$\mathbf{E}(\xi^n) = \sum_{k=1}^{n} \left\{ n \atop k \right\} \frac{k!}{p}\left(\frac{q}{p}\right)^{k-1}.$$

Például ha $n = 2$, akkor felhasználva, hogy $\left\{ 2 \atop 1 \right\} = \left\{ 2 \atop 2 \right\} = 1$,

$$\mathbf{E}\left(\xi^2\right) = \frac{1!}{p}\left\{ 2 \atop 1 \right\} + \frac{2!}{p}\frac{q}{p}\left\{ 2 \atop 1 \right\} = \frac{1}{p} + \frac{2q}{p^2} = \frac{1+q}{p^2}.$$

$\square$

11.2.7. Példa. Számoljuk ki a geometriai eloszlás momenmtumgeneráló függvényét, majd ennek segítségével számoljuk ki a kumulánsokat!

A momentumgeneráló függvény definíciója szerint

$$\begin{aligned} M(s) &= \mathbf{E}(\exp(s\xi)) = G(\exp(s)) = \\ &= \frac{p\exp(s)}{1-q\exp(s)}. \end{aligned}$$

A kumulánsgeneráló függvény

$$C(s) = \ln M(s) = \ln p + s - \ln(1-q\exp(s)).$$

Ebből

$$\begin{aligned} c_1 &= \mathbf{E}(\xi) = C'(0) = 1 + \left.\frac{q\exp(s)}{1-q\exp(s)}\right|_{s=0} = \\ &= 1 + \frac{q}{1-q} = \frac{1}{1-q} = \frac{1}{p}. \\ c_2 &= \mathbf{D}^2(\xi) = C''(0) = \\ &= \left.\frac{q\exp(s)(1-q\exp(s)) + q\exp(s)q\exp(s)}{(1-q\exp(s))^2}\right|_{s=0} = \\ &= \left.\frac{q\exp(s)}{(1-q\exp(s))^2}\right|_{s=0} = \frac{q}{(1-q)^2} = \frac{q}{p^2}. \\ c_3 &= \mathbf{E}\left((\xi-\mathbf{E}(\xi))^3\right) = C'''(0) = \\ &= \left.\frac{q\exp(s)(1-q\exp(s))^2 + 2q\exp(s)(1-q\exp(s))q\exp(s)}{(1-q\exp(s))^4}\right|_{s=0} = \\ &= \left.\frac{q\exp(s)(1+q\exp(s))}{(1-q\exp(s))^3}\right|_{s=0} = \frac{q(1+q)}{p^3}. \end{aligned}$$

Ebből a ferdeség

$$\frac{c_3}{c_2^{3/2}} = \frac{q(1+q)}{p^3}\left(\frac{p^2}{q}\right)^{3/2} = \frac{1+q}{\sqrt{q}} = \frac{2-p}{\sqrt{q}}.$$

Ismételten deriválva és az s helyébe nullát téve

$$\begin{gathered} c_4 = C^{(iv)}(0) = \\ = \frac{q(1+2q)(1-q) + 3q(1+q)q}{p^4} = \\ = \frac{q\left(1+4q+q^2\right)}{p^4}. \end{gathered}$$

A kurtózis tehát

$$\begin{aligned} 3+\frac{c_4}{c_2^2} &= 3+\frac{q\left(1+4q+q^2\right)}{p^4}\frac{p^4}{q^2}= \\ &= 3+\frac{1+4q+q^2}{q}= \\ &= 6+\frac{1+q+q^2}{q}= \\ &= 9+\frac{1-2q+q^2}{q}=9+\frac{p^2}{q}. \end{aligned}$$

□

11.2.8. Példa. Számoljuk ki az $N(0,1)$ eloszlás momentumgeneráló és karakterisztikus függvényét, majd határozzuk meg az eloszlás momentumait és kumulánsait!

A momentumgeneráló függvény

$$\begin{aligned} M(s) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(sx)\exp\left(-\frac{x^2}{2}\right)dx= \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\frac{(x-s)^2}{2}\right)\exp\left(\frac{s^2}{2}\right)dx= \\ &= \exp\left(\frac{s^2}{2}\right)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\frac{(x-s)^2}{2}\right)dx= \\ &= \exp\left(\frac{s^2}{2}\right)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\frac{u^2}{2}\right)du=\exp\left(\frac{s^2}{2}\right). \end{aligned}$$

Mivel az $\exp\left(z^2/2\right)$ a teljes komplex síkban értelmezett, ezért

$$\varphi(t)=\mathscr{M}(it)=\exp\left(-\frac{t^2}{2}\right).$$

Az

$$M(s)=\sum_{k=0}^{\infty}\frac{\left(s^2/2\right)^k}{k!}=\sum_{k=0}^{\infty}\frac{1\cdot 3\cdot\ldots\cdot(2k-1)}{(2k)!}s^{2k}=\sum_{k=0}^{\infty}\frac{(2k-1)!!}{(2k)!}s^{2k}$$

sorfejtés alapján a páratlan momentumok mindegyike nulla, a páros momentumok pedig

$$\mathbf{E}\left(\xi^{2k}\right)=(2k-1)!!\stackrel{\circ}{=}(2k-1)(2k-3)\cdots 1.$$

Ebből speciálisan a negyedik momentum 3. Mivel

$$C(s)\stackrel{\circ}{=}\ln\exp\frac{s^2}{2}=\frac{s^2}{2},$$

ezért $c_1=0$, $c_2=1$, a többi kumuláns pedig nulla. Ez alapján a kurtózis $3+c_4/c_2^2=3$.

□

11.2.9. Példa. A centrális határeloszlás tétele és a kumulánsok.

Legyenek (ξ_k) független, azonos eloszlású valószínűségi változók és tegyük fel, hogy a közös eloszlásnak létezik a negyedik momentuma. Jelölje μ a közös várható értéket és σ a közös szórást. Tekintsük az

$$\eta_n \overset{\circ}{=} \frac{\sum_{k=1}^{n} \xi_k - n\mu}{\sigma\sqrt{n}}$$

normalizált és centralizált összeget. A centralizáció miatt az η_n várható értéke nulla, a normalizáció miatt a szórása pedig 1. De mi történik a ferdeséggel és a kurtózissal? Ha egy ξ valószínűségi változó momentumgeneráló függvénye $M(s)$, akkor az $a\xi + b$ momentumgeneráló függvénye

$$\mathbf{E}(\exp(s(a\xi + b))) = \exp(sb)M(as).$$

Ebből az $a\xi + b$ kumulánsgeneráló függvénye $sb + C(as)$. Az összetett függvény deriválási szabálya miatt a k-adik kumuláns kiszámolásakor egy a^k szorzót kell alkalmazni. Az egyszerűség kedvéért tegyük fel, hogy $\mu = 0$ és $\sigma = 1$. Ha a ξ_k változók közös eloszlásának kumulánsai c_k, akkor az η_n ferdesége az $c_3/c_2^{3/2}$ képlet szerint

$$\frac{(nc_3)\left(\frac{1}{\sqrt{n}}\right)^3}{\left(nc_2\left(\frac{1}{\sqrt{n}}\right)^2\right)^{3/2}} = \frac{c_3\frac{1}{\sqrt{n}}}{c_2^{3/2}}$$

amely kifejezés $1/\sqrt{n}$ nagyságrenddel nullához tart. A kurtózis pedig a $3 + c_4/c_2^2$ képlet szerint

$$3 + \frac{nc_4\left(\frac{1}{\sqrt{n}}\right)^4}{\left(nc_2\left(\frac{1}{\sqrt{n}}\right)^2\right)^2} = 3 + \frac{c_4/n}{c_2^2}$$

amely pedig $1/n$ sebességgel 3-hoz tart. Vagyis a centralizált normalizált összeg ferdesége és kurtózisa a standard normális eloszlás ferdeségéhez és kurtózisához tart.

□

11.2.10. Példa. Az első négy kumuláns és az eloszlás normalitása.

Kézenfekvően merül fel a kérdés, miszerint, ha egy centralizált és normalizált eloszlás ferdesége nulla és kurtózisa 3, akkor az eloszlás normális-e? A válasz: nem. Legyen ξ egy λ paraméterű Poisson-eloszlású változó, és tükrözzük $\sqrt{\xi}$ eloszlását az y-tengelyre. Vagyis $\sqrt{\xi}$ eloszlását mérjük fel szimmetrikus módon az x-tengelyre, majd a megfelelő valószínűségeket osszuk el 2-vel. Az így kapott eloszlás szimmetrikus, így a várható értéke és a ferdesége nulla. Könnyen látható, hogy a szórása éppen a ξ várható értéke, λ, a negyedik momentuma pedig a ξ második momentuma, vagyis $\lambda+\lambda^2$. Ebből a kurtózis éppen $\left(\lambda+\lambda^2\right)/\lambda^2 = 1+1/\lambda$. Így ha $\lambda = 1/2$, akkor a kurtózis három. A kurtózis nem függ a skálázástól, a variancia pedig négyzetesen szorzódik, így a ξ helyett a $\sqrt{2}\xi$ változót véve a variancia is 1 lesz.

□

11.2.11. Példa. Számoljuk ki az exponenciális eloszlás momentumgeneráló és karakterisztikus függvényét, és határozzuk meg a momentumait és a kumulánsokat!

$$M(s) = \int_0^\infty \exp(sx)\,\lambda\exp(-\lambda x)\,dx = \lambda\int_0^\infty \exp((s-\lambda)x)\,dx = \frac{\lambda}{\lambda - s}.$$

A függvény értelmezve van az $s<\lambda$ halmazon és az

$$\frac{\lambda}{\lambda-s} = \frac{1}{1-s/\lambda} = \sum_{k=0}^{\infty}\frac{s^k}{\lambda^k}$$

sorfejtés alapján a k-dik momentum $\mathbf{E}\left(\xi^k\right) = k!/\lambda^k$. A kumulánsgeneráló függvény az

$$\ln(1+x) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k}$$

egyenlőség alapján

$$C(s) = -\ln(1-s/\lambda) = \sum_{k=1}^{\infty}\frac{1}{k\lambda^k}s^k,$$

amiből $c_k = (k-1)!/\lambda^k$. Ebből az exponenciális eloszlás ferdesége

$$\frac{c_3}{c_2^{3/2}} = \frac{2!}{\lambda^3}\cdot\frac{\left(\lambda^2\right)^{3/2}}{1} = 2.$$

A kurtózis pedig

$$3+\frac{c_4}{c_2^2} = 3+\frac{3!}{\lambda^4}\left(\frac{\lambda^2}{1!}\right)^2 = 3+6 = 9.$$

Az exponenciális eloszlás momentumgeneráló függvénye értelmezve van a 0 $(-\infty,\lambda)$ környezetében, ami alapján

$$\varphi(t) = \frac{1}{1-it/\lambda}.$$

□

11.2.12. Példa. Számoljuk ki az (a,λ) paraméterű gamma eloszlás momentumgeneráló függvényét, karakterisztikus függvényét és határozzuk meg a momentumait és a kumulánsokat!

Az (a,λ) paraméterű gamma eloszlás sűrűségfüggvénye

$$\frac{\lambda^a}{\Gamma(a)}x^{a-1}\exp(-\lambda x), \quad x>0.$$

Ez alapján $\lambda - s > 0$ esetén

$$\begin{aligned}
M(s) &= \int_0^\infty \exp(sx)\frac{\lambda^a}{\Gamma(a)}x^{a-1}\exp(-\lambda x)\,dx = \\
&= \frac{\lambda^a}{\Gamma(a)}\int_0^\infty x^{a-1}\exp(-(\lambda - s)x)\,dx = \\
&= \frac{\lambda^a}{\Gamma(a)}\Gamma(a,\lambda - s) = \frac{\lambda^a}{\Gamma(a)}\frac{\Gamma(a)}{(\lambda - s)^a} = \\
&= \left(\frac{\lambda}{\lambda - s}\right)^a = \frac{1}{(1-s/\lambda)^a}.
\end{aligned}$$

A $(-\lambda,\lambda)$ intervallumban a binomiális sor összegképlete alapján

$$\left(1-\frac{s}{\lambda}\right)^{-a} = \sum_{k=0}^\infty \binom{-a}{k}\left(-\frac{s}{\lambda}\right)^k = \sum_{k=0}^\infty \frac{s^k}{k!}\frac{a(a+1)\dots(a+k-1)}{\lambda^k},$$

amiből a k-dik momentum

$$\mathbf{E}\left(\xi^k\right) = \frac{\prod_{j=0}^{k-1}(a+j)}{\lambda^k}.$$

Érdemes ezt közvetlenül is kiszámolni:

$$\begin{aligned}
\mathbf{E}\left(\xi^k\right) &= \frac{\lambda^a}{\Gamma(a)}\int_0^\infty x^k x^{a-1}\exp(-\lambda x)\,dx = \\
&= \frac{\lambda^a}{\Gamma(a)}\Gamma(a+k,\lambda) = \frac{\lambda^a}{\Gamma(a)}\frac{\Gamma(a+k)}{\lambda^{a+k}} = \\
&= \frac{\prod_{j=0}^{k-1}(a+j)}{\lambda^k}.
\end{aligned}$$

A kumulánsgeneráló függvény

$$C(s) = \ln\left(\frac{1}{(1-s/\lambda)^a}\right) = -a\ln\left(1-\frac{s}{\lambda}\right) = a\sum_{k=1}^\infty \frac{1}{k\lambda^k}s^k,$$

amiből $c_k = a(k-1)!/\lambda^k$. Ebből a ferdeség

$$\frac{c_3}{c_2^{3/2}} = \frac{a \cdot 2!}{\lambda^3} \frac{\left(\lambda^2\right)^{2/3}}{a^{3/2}} = \frac{2}{\sqrt{a}}.$$

A kurtózis

$$3 + \frac{c_4}{c_2^2} = 3 + \frac{a \cdot 3!}{\lambda^4} \left(\frac{\lambda^2}{a}\right)^2 = 3 + \frac{6}{a}.$$

A karakterisztikus függvény pedig

$$\varphi(t) = \left(1 - \frac{it}{\lambda}\right)^{-a}.$$

$\square$

XII.

EGY HITELKOCKÁZATI MODELL

Ebben a fejezetben egy összefoglaló alkalmazást mutatunk be és az összetett, vagy másnéven kevert eloszlások problémáit tárgyaljuk. Matematikai szempontból a fejezet a Poisson-eloszlás és a geometriai eloszlás kapcsolatát tárgyalja. A fejezet egyik fő eredménye, hogy a geometriai eloszlás is korlátlanul osztható. Ha egy ξ valószínűségi változó eloszlása korlátlanul osztható, akkor tetszőleges a konstans esetén a $\xi+a$ eloszlása is korlátlanul osztható. Így a geometriai eloszlás helyett érdemesebb a pq^n alakú elsőrendű Pascal-eloszlásokkal foglalkozni. Ennek előnye, hogy az elsőrendű Pascal-eloszlás tartója, megegyezik a Poisson-eloszlás tartójával. Miként az exponenciális eloszlás gamma eloszlások összegére bomlik, az elsőrendű Pascal-eloszlás negatív binomiális eloszlásokra bontható szét. A két eloszláscsalád kapcsolatát a kevert eloszlások fogalma biztosítja. Az elsőrendű Pascal-eloszlás Poisson-eloszlás exponenciális eloszlással való keveréseként állítható elő, a negatív binomiális eloszlás pedig Poisson-eloszlás gamma eloszlással való keveréseként vezethető be.

12.1. Összetett eloszlások

12.1.1. Definíció. Ha N egy a $0,1,2,\ldots$ értékeket felvevő valószínűségi változó és $\xi_1,\xi_2,\ldots$ valószínűségi változók egy sorozata, akkor az $\eta \triangleq \Sigma_{k=1}^{N}\xi_k$ véletlen tagszámú összeget összetett valószínűségi változónak mondjuk. Általában eloszlások keverésén azt az eljárást értjük, amikor egy valószínűségi változó értékét két lépésben határozzuk meg. Az első lépésben kiválasztunk egy valószínűségi változót, majd ennek felhasználásával egy másik változót veszünk, amely értéke lesz a kevert változó értéke. Legtöbbször a keverés a keverendő eloszlás valamelyik paraméterének randomizációjával történik. A kevert eloszlások speciális esete az összetett eloszlások, amikor a darabszámot randomizáljuk.

12.1.2. Állítás. *Ha a (ξ_k) változók azonos eloszlásúak és egymástól és az N változótól is függetlenek, akkor az $\eta \triangleq \Sigma_{k=1}^{N}\xi_k$ véletlen tagszámú összeg várható értéke*

$$\mathbf{E}(\eta)=\mathbf{E}(\xi)\cdot\mathbf{E}(N), \tag{12.1.1}$$

szórása pedig

$$\mathbf{D}(\eta)=\sqrt{\mathbf{D}^2(\xi)\,\mathbf{E}(N)+\mathbf{E}^2(\xi)\,\mathbf{D}^2(N)}, \tag{12.1.2}$$

ahol ξ a (ξ_k) változók bármelyike. Az azonosságok felírásakor értelemszerűen feltettük, hogy a képletekben szereplő várható értékek léteznek.

Bizonyítás: Az első egyenlőség kiszámolása a teljes várható érték tétele alapján igen egyszerű:

$$
\begin{aligned}
\mathbf{E}(\eta) &= \sum_{j=0}^{\infty} \mathbf{E}(\eta \mid N=j)\,\mathbf{P}(N=j) = \\
&= \sum_{j=0}^{\infty} \mathbf{E}\left(\sum_{k=1}^{N} \xi_k \mid N=j\right)\mathbf{P}(N=j) = \\
&= \sum_{j=0}^{\infty} \mathbf{E}\left(\sum_{k=1}^{j} \xi_k\right)\mathbf{P}(N=j) = \sum_{j=0}^{\infty}\left(\sum_{k=1}^{j}\mathbf{E}(\xi_k)\right)\mathbf{P}(N=j) = \\
&= \sum_{j=0}^{\infty} j\mathbf{E}(\xi)\,\mathbf{P}(N=j) = \mathbf{E}(\xi)\sum_{j=0}^{\infty} j\mathbf{P}(N=j) = \mathbf{E}(\xi)\,\mathbf{E}(N).
\end{aligned}
$$

A második momentum kiszámolása a következő:

$$
\begin{aligned}
\mathbf{E}\left(\eta^2\right) &= \sum_{j=0}^{\infty} \mathbf{E}\left(\eta^2 \mid N=j\right)\mathbf{P}(N=j) = \\
&= \sum_{j=0}^{\infty} \mathbf{E}\left(\left(\sum_{k=1}^{N}\xi_k\right)^2 \mid N=j\right)\mathbf{P}(N=j) = \\
&= \sum_{j=0}^{\infty} \mathbf{E}\left(\left(\sum_{k=1}^{j}\xi_k\right)^2\right)\mathbf{P}(N=j) = \\
&= \sum_{j=0}^{\infty}\left(\mathbf{D}^2\left(\sum_{k=1}^{j}\xi_k\right)+\mathbf{E}^2\left(\sum_{k=1}^{j}\xi_k\right)\right)\mathbf{P}(N=j) = \\
&= \sum_{j=0}^{\infty}\left(j\mathbf{D}^2(\xi)+\left(\sum_{k=1}^{j}\mathbf{E}(\xi_k)\right)^2\right)\mathbf{P}(N=j) = \\
&= \sum_{j=0}^{\infty}\left(j\mathbf{D}^2(\xi)+(j\mathbf{E}(\xi))^2\right)\mathbf{P}(N=j) = \\
&= \sum_{j=0}^{\infty}\left(j\mathbf{D}^2(\xi)+j^2\mathbf{E}^2(\xi)\right)\mathbf{P}(N=j) = \\
&= \mathbf{E}(N)\,\mathbf{D}^2(\xi)+\mathbf{E}^2(\xi)\,\mathbf{E}\left(N^2\right).
\end{aligned}
$$

Ebből a szórás definíciója szerint

$$\begin{aligned}\mathbf{D}^2(\eta) &= \mathbf{E}\left(\eta^2\right) - \mathbf{E}^2(\eta) = \\ &= \mathbf{E}(N)\mathbf{D}^2(\xi) + \mathbf{E}^2(\xi)\mathbf{E}\left(N^2\right) - (\mathbf{E}(\xi)\mathbf{E}(N))^2 = \\ &= \mathbf{E}(N)\mathbf{D}^2(\xi) + \mathbf{E}^2(\xi)\left(\mathbf{E}\left(N^2\right) - \mathbf{E}^2(N)\right),\end{aligned}$$

amiből a második (12.1.2) egyenlőség már evidens. A számolások során használtuk a

$$\mathbf{E}\left(\sum_{k=1}^{N}\xi_k \mid N=j\right) = \mathbf{E}\left(\sum_{k=1}^{j}\xi_k\right)$$

egyenlőséget, amely során az N véletlen tagszám helyébe betettük az $N = j$ feltételt és elhagytuk a feltételt. Hasonló összefüggést használtunk a négyzetes tagokra is. Vegyük észre, hogy $\mathbf{P}(N=j) > 0$, ugyanis ellenkező esetben ezt a j értéket nem használjuk. A feltételes eloszlás, illetve a feltételes várható érték definíciója szerint

$$\mathbf{E}\left(\sum_{k=1}^{N}\xi_k \mid N=j\right) \circeq \frac{\mathbf{E}\left(\chi(N=j)\sum_{k=1}^{N}\xi_k\right)}{\mathbf{P}(N=j)}.$$

Ilyenkor triviálisan

$$\frac{\mathbf{E}\left(\chi(N=j)\sum_{k=1}^{N}\xi_k\right)}{\mathbf{P}(N=j)} = \frac{\mathbf{E}\left(\chi(N=j)\sum_{k=1}^{j}\xi_k\right)}{\mathbf{P}(N=j)}.$$

Az N és a (ξ_k) változók függetlensége miatt a szorzat várható értéke a várható értékek szorzata, így

$$\frac{\mathbf{E}\left(\chi(N=j)\sum_{k=1}^{j}\xi_k\right)}{\mathbf{P}(N=j)} = \frac{\mathbf{P}(N=j)\mathbf{E}\left(\sum_{k=1}^{j}\xi_k\right)}{\mathbf{P}(N=j)},$$

amiből az egyenlőség már evidens.

$\square$

12.1.3. Példa. Poisson-eloszlás szerint összetett eloszlás szórása és várható értéke.

Legyen N Poisson-eloszlású λ paraméterrel. Ekkor $\mathbf{E}(N) = \mathbf{D}^2(N) = \lambda$, és ilyenkor

$$\begin{aligned} \mathbf{E}(\eta) &= \lambda \mathbf{E}(\xi), \\ \mathbf{D}(\eta) &= \sqrt{\lambda}\sqrt{\mathbf{D}^2(\xi) + \mathbf{E}^2(\xi)} = \sqrt{\lambda \mathbf{E}\left(\xi^2\right)}. \end{aligned}$$

Érdemes megjegyezni, hogy ha az N egy Poisson-folyamat szerint alakul, akkor az $X(t) \stackrel{\circ}{=} \sum_{k=1}^{N(t)} \xi_k$ összetett folyamat várható értéke és a varianciája az idő szerint lineárisan változik.

□

12.2. Veszteségek eloszlása

Közgazdasági szempontból az alapprobléma a következő: Tekintsünk egy $[0,T]$ időszakot. Mekkora veszteség érhet minket az adott időszak alatt? A teljes veszteség nagysága két dologtól függ:

1. Az egyedi veszteségek nagyságától, és
2. a veszteségek darabszámától.

Feltesszük, hogy

1. a t időpontig bekövetkezett veszteségek $N(t)$ darabszáma egy nem negatív egészekből álló folyamatot alkot,
2. a (ξ_k) egyedi veszteségek nagysága egy az N-től független, azonos eloszlású, független valószínűségi változókból álló sorozat.

A megközelítés nagyrészt a biztosításmatematikából került át a pénzügyekbe és elsősorban az operációs kockázatok modellezésére szokás használni. A nehézség az, hogy a teljes veszteséget leíró így kapott bonyolult valószínűségi változó eloszlása nehezen kezelhető analitikusan. A T időhorizont általában egy év. Szemben a piaci kockázatokkal a veszteségek nem, vagy csak részben „gyógyulnak vissza”. Ez azt jelenti, hogy a (ξ_k) egyedi veszteségekről feltesszük, hogy nem negatívak. A valószínűségi változó kockázatát általában a valószínűségi változó szórásával szokás jellemezni. A pénzügyi alkalmazások során a probléma nagyrészt abból ered, hogy a veszteségeket megadó összetett valószínűségi változó nagyon elnyúló, a sűrűségfüggvénye lassan konvergál nullához. Ennek következtében sem az összetett veszteség várható értéke, sem a szórása pénzügyi szempontból nem releváns[1]. Ezért más kockázati mértéket érdemes bevezetni. Az irodalomban számos kockázati mutató létezik. A leggyakrabban használt a következő:

[1]Nem beszélve arról, hogy a szórás lehet végtelen is, így megadásának információtartalma nulla.

12.2.1. Definíció. Egy adott α valószínűséghez tartozó kockáztatott érték az a legkisebb X_α érték, amelyre annak a valószínűsége, hogy a *Loss* módon jelölt veszteség nagyobb lesz mint X_α kisebb vagy egyenlő mint α. Képlettel

$$Var_\alpha(Loss) \triangleq X_\alpha \triangleq \inf\{x \mid \mathbf{P}(Loss > x) \le \alpha\}.$$

A $[0,T]$ időtartomány során keletkezett teljes veszteség

$$Loss = \sum_{k=1}^{\infty} \chi(\sigma_k \le T)\,\xi_k = \sum_{k=1}^{N(T)} \xi_k,$$

ahol (σ_k) az N folyamatban levő ugrások időpontja[2]. Az alábbi számítások célja adott α esetén a kockáztatott érték meghatározása. Ez a feladat matematikai szempontból a *Loss* változóhoz tartozó eloszlásfüggvény inverzének a kiszámolását jelenti. Sajnos a *Loss* változó eloszlásának közvetlen kiszámolása nem egyszerű feladat. Ezért egy kerülőutat választunk. A kulcsészrevétel az, hogy mivel a *Loss* változó értelemszerűen nem negatív, ezért az

$$L_{Loss}(s) \triangleq \mathbf{E}(\exp(-s \cdot Loss))$$

Laplace-transzformáltja[3] minden $s \ge 0$ esetén véges, és ilyenkor a Laplace-transzformált egyértelműen meghatározza az eloszlást[4]. Mivel T ismert, az egyszerűség kedvéért vezessük be az $N \triangleq N(T)$ jelölést. Ha $\mathbf{P}(N = k) \triangleq p_k$, akkor a teljes várható érték már látott tétele alapján

$$\begin{aligned}
L_{Loss}(s) &\triangleq \mathbf{E}(\exp(-s \cdot Loss)) = \mathbf{E}(\mathbf{E}(\exp(-s \cdot Loss) \mid N)) = \\
&= \sum_{k=0}^{\infty} \mathbf{E}(\exp(-s \cdot Loss) \mid N = k) \cdot p_k = \\
&= \sum_{k=0}^{\infty} \mathbf{E}\left(\exp\left(-s \cdot \sum_{i=1}^{k} \xi_i\right)\right) \cdot p_k = \\
&= \sum_{k=0}^{\infty} (\mathbf{E}(\exp(-s \cdot \xi)))^k p_k = G_N\left(L_\xi(s)\right)
\end{aligned}$$

ahol $G_N(z) \triangleq \sum_{k=0}^{\infty} z^k p_k$ az ugrások számának generátorfüggvénye és $L_\xi(s)$ az egyedi veszteségek eloszlásának közös Laplace-transzformáltja. Ha a ξ_k egyedi veszteségek

[2] Miként látni fogjuk, ha nem Poisson-folyamatról van szó, akkor egyszerre több ugrás is bekövetkezhet.

[3] A momentumgeneráló függvény és a Laplace-transzformált lényegében azonos, egyedül a kitevő előjele más. Nem negatív változók esetén a Laplace-transzformált talán kényelmesebb.

[4] Általában a Laplace-transzformált vagy a momentumgeneráló függvény nem határozza meg egyértelműen az eloszlást. Ha azonban a változó nem negatív, akkor igen. Így valójában két matematikai feladatot kellene megoldani: Egyrészt matematikailag igazolni kellene, hogy nem negatív változók esetén a Laplace-transzformált egyértelműen jellemzi az eloszlást, másrészt egy hatékony algoritmus kellene megalkotni, amely segítségével a Laplace-transzformáltból visszaszámolható az eloszlásfüggvény. A továbbiakban a feladat ilyen általánosságban való vizsgálatától eltekintünk.

eloszlása diszkrét, akkor az L_ξ Laplace-transzformált helyett vehetjük a ξ_k változók közös G_ξ generátorfüggvényét is, és a gondolatmenet értelemszerű megismétlésével a $G_{Loss}(z) = G\left(G_\xi(z)\right)$ képletet kapjuk.

12.2.2. Példa. Ha a veszteségek N darabszáma Poisson-eloszlást követ, akkor

$$\begin{aligned} G(z) &\overset{\circ}{=} \sum_{k=0}^{\infty} z^k \frac{\lambda^k}{k!} \exp(-\lambda) = \exp(-\lambda) \sum_{k=0}^{\infty} \frac{(z\lambda)^k}{k!} = \\ &= \exp(-\lambda) \exp(z\lambda) = \exp(\lambda(z-1)). \end{aligned}$$

Ha a ξ_k egyedi veszteségek nagysága exponenciális eloszlást követ μ paraméterrel:

$$\begin{aligned} L_\xi(s) &= \int_0^\infty \mu \exp(-\mu x) \exp(-sx)\, dx = \mu \int_0^\infty \exp(-(\mu+s)x)\, dx = \\ &= \mu \left[-\frac{\exp(-\mu+s)x}{\mu+s} \right]_0^\infty = \frac{\mu}{\mu+s} \\ L_{Loss}(s) &= \exp\left(\lambda\left(\frac{\mu}{\mu+s} - 1\right)\right) = \exp\left(\frac{-\lambda s}{\mu+s}\right). \end{aligned}$$

□

12.3. Negatív binomiális eloszlás

A Poisson-folyamat λ paramétere számos külső tényezőtől függ. Így nem tudjuk a λ partaméter értékét. Tegyük fel, hogy csak a darabszám feltételes eloszlása lesz Poisson:

$$\mathbf{P}(N = n \mid \lambda) = \frac{\lambda^n}{n!} \exp(-\lambda).$$

Tegyük fel, hogy a λ eloszlása (α, β) paraméterű gamma eloszlást alkot, ahol (α, β) a külső tényezők eloszlásának függvénye. Mi lesz az N eloszlása?

$$\mathbf{P}(N = n) = \mathbf{E}(\mathbf{P}(N = n \mid \lambda)) = \frac{1}{n!} (\mathbf{E}(\lambda^n \exp(-\lambda))).$$

A gamma eloszlás sűrűségfüggvényének képlete alapján

$$\begin{aligned} \mathbf{E}(\lambda^n \exp(-\lambda)) &= \int_0^\infty x^n \exp(-x) \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x)\, dx = \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+n-1} \exp(-(\beta+1)x)\, dx = \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \Gamma(\alpha+n.\beta+1) = \frac{\beta^\alpha \Gamma(\alpha+n)}{\Gamma(\alpha)(\beta+1)^{\alpha+n}}, \end{aligned}$$

Mivel $\Gamma(x+1) = x\Gamma(x)$, ezért

$$\begin{aligned}
\mathbf{P}(N=n) &= \frac{\Gamma(\alpha+n)}{n!\Gamma(\alpha)} \frac{\beta^\alpha}{(1+\beta)^{\alpha+n}} = \\
&= \frac{(\alpha+n-1)\Gamma(\alpha+n-1)}{n!\Gamma(\alpha)} \frac{\beta^\alpha}{(1+\beta)^{\alpha+n}} = \cdots = \\
&= \frac{\prod_{k=0}^{n-1}(\alpha+k)\Gamma(\alpha)}{n!\Gamma(\alpha)} \left(\frac{\beta}{\beta+1}\right)^\alpha \left(\frac{1}{\beta+1}\right)^n \overset{\circ}{=} \\
&\overset{\circ}{=} \binom{\alpha+n-1}{n} p^\alpha q^n.
\end{aligned}$$

ahol $p \overset{\circ}{=} \beta/(1+\beta)$, $q \overset{\circ}{=} 1/(1+\beta)$.

12.3.1. Definíció. A

$$p_n \overset{\circ}{=} \binom{\alpha+n-1}{n} p^\alpha q^n, \quad n = 0, 1, \ldots$$

eloszlást (α, p) paraméterű negatív binomiális eloszlásnak szokás nevezni.

12.3.2. Példa. A geometriai és a negatív binomiális eloszlás kapcsolata.

Ha $\alpha = 1$, akkor a gamma eloszlás exponenciális. Ekkor

$$\mathbf{P}(N=n) = pq^n, \quad n = 0, 1, 2, \ldots.$$

Emlékeztetünk, hogy ezt az eloszlást, hogy ne keverjük a geometriai eloszlással, szokás elsőrendű Pascal-eloszlásnak mondani.

□

Ha n egy természetes szám, akkor $\Gamma(n) = (n-1)!$ Így definíció szerint

$$\binom{\alpha+n-1}{n} \overset{\circ}{=} \frac{\Gamma(\alpha+n)}{n!\Gamma(\alpha)} = \frac{\prod_{k=0}^{n-1}(\alpha+k)}{n!}.$$

Más oldalról

$$\binom{-\alpha}{n} \overset{\circ}{=} \frac{-\alpha(-\alpha-1)\cdots(-\alpha-n-1)}{n!} = (-1)^n \frac{\prod_{k=0}^{n-1}(\alpha+k)}{n!}.$$

Következésképpen

$$\mathbf{P}(N=n) = (-1)^n \binom{-\alpha}{n} p^\alpha q^n,$$

amely kifejezés indokolja a negatív binomiális elnevezést. A Poisson-eloszlás generátorfüggvényét és a gamma eloszlás sűrűségfüggvényét használva a negatív binomiális eloszlás generátorfüggvénye

$$
\begin{aligned}
G(z) &\mathrel{\hat{=}} \mathbf{E}\left(z^N\right) = \mathbf{E}\left(\mathbf{E}\left(z^N \mid \lambda\right)\right) \mathrel{\hat{=}} \mathbf{E}\left(G_\lambda(z)\right) = \\
&= \int_0^\infty G_\lambda(z) \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\beta\lambda)\, d\lambda = \\
&= \int_0^\infty \exp(\lambda(z-1)) \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\beta\lambda)\, d\lambda = \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \exp(-\lambda(\beta+1-z))\, d\lambda = \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \Gamma(a, \beta+1-z) = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\beta+1-z)^\alpha} = \left(\frac{\beta}{\beta+1-z}\right)^\alpha .
\end{aligned}
$$

12.3.3. Példa. Ha a veszteségek nagysága exponenciális eloszlású μ paraméterrel, akkor

$$
L_{Loss}(s) = G\left(L_\xi(z)\right) = \left(\frac{\beta}{\beta+1-\frac{\mu}{\mu+s}}\right)^\alpha .
$$

12.3.4. Példa. A negatív binomiális eloszlás korlátlanul osztható.

Emlékeztetünk, hogy egy eloszlást korlátlanul oszthatónak mondunk, ha tetszőleges n-re az eloszlás felbontható n darab azonos eloszlású független változó összegére. Mivel független diszkrét eloszlások összegének generátorfüggvénye összeszorzódik, ezért elég megjegyezni, hogy n darab $(\alpha/n, \beta)$ paraméterű független negatív binomiális eloszlású változó összege (α, β) paraméterű negatív binomiális eloszlású változó. Miként a Lévy-folyamatok kapcsán említettük, igazolható egy fontos tétel, amely szerint minden korlátlanul osztható eloszlás beágyazható egy Lévy-folyamatba. Ez alatt azt érjük, hogy minden korlátlanul osztható eloszláshoz van egy olyan $X(t)$ Lévy-folyamat, hogy az $X(1)$ eloszlása éppen az adott korlátlanul osztható eloszlás[5]. Ennek megfelelően beszélhetünk negatív binomiális folyamatról. A negatív binomiális folyamat egész értékeket felvevő folyamat, így joggal merülhet fel, hogy mennyiben különbözik a Poisson-folyamattól. Viszonylag könnyen megmutatható, hogy minden egész értékeket felvevő Lévy-folyamat összetett Poisson-folyamat[6], vagyis alkalmas $N(t)$ Poisson-folyamattal $X(t) = \Sigma_{k=1}^{N(t)} \xi_k$ alakú véletlen összeg, ahol a ξ_k változók független azonos eloszlásból

[5] Érdemes hangsúlyozni, hogy általában egy sztochasztikus folyamat eloszlásán az összes lehetséges $X(t_1), \ldots, X(t_n)$ véges sorozat együttes eloszlásának rendszerét értjük. Lévy-folyamatok esetén azonban elegendő megadni az eloszlást egyetlen $t > 0$ időpontban, például az $X(1)$, eloszlását.

[6] Általában a Lévy-folyamatokra nem igaz, hogy összetett Poisson-folyamatok. A Lévy-folyamatok struktúrája ennél sokkal bonyolultabb.

vett ugrásnagyságok. Ennek igazolásához elegendő először venni az első ugrás helyét megadó

$$\tau_1 = \inf\{t \mid |X(t)| \geq 1\}$$

megállási időt. A Poisson-folyamathoz hasonlóan a Lévy-folyamatok jobbról való folytonossága miatt $\mathbf{P}(\tau_1 \geq t) = \mathbf{P}(X(t) = 0)$. Ebből a Poisson-folyamatnál követett módon azonnal látható, hogy a τ_1 eloszlása exponenciális. Legyen $\xi_1 \stackrel{\circ}{=} X(\tau_1)$. Az X Lévy-folyamat, így érvényes rá az erős Markov-tulajdonság. τ_2 legyen az X^* újraindított folyamat első ugrásának időpontja és $\xi_2 \stackrel{\circ}{=} X^*(\tau_2)$ stb. A (τ_k) sorozat egy Poisson-folyamatot generál, és az X éppen a ξ_k ugrásokkal vett összetett Poisson-folyamat. Alább megmutatjuk, hogy a negatív binomiális eloszlás esetén az ugrások nagyságát megadó változók eloszlása úgynevezett logaritmikus eloszlás, amelyre

$$p_n = \frac{1}{-\ln(1-p)} \frac{p^n}{n}, \quad n = 1, 2, \ldots$$

Vagyis a negatív binomiális folyamat bár egész értékeket felvevő, monoton növekedő trajektóriákkal rendelkező folyamat, nem számláló folyamat, ugyanis az ugrásainak nagysága lehet 1-nél nagyobb. A logaritmikus eloszlás várható értéke

$$\begin{aligned} M &= \sum_{n=1}^{\infty} n p_n = \frac{1}{-\ln(1-p)} \sum_{n=1}^{\infty} p^n = \\ &= \frac{1}{-\ln(1-p)} \frac{p}{1-p} > 1, \end{aligned}$$

ugyanis ha $x > 0$ és $x \neq 1$, akkor $x - 1 > \ln x$, ezért

$$\frac{p}{1-p} = \frac{1}{1-p} - 1 > \ln \frac{1}{1-p}.$$

A logaritmikus eloszlás generátorfüggvénye

$$G(z) = \frac{1}{-\ln(1-p)} \sum_{k=1}^{\infty} \frac{p^n}{n} z^n = \frac{\ln(1-pz)}{\ln(1-p)}.$$

Miként megjegyeztük az összetett eloszlás generátorfüggvényét úgy kapjuk, hogy a Poisson-eloszlás $\exp(\lambda(z-1))$ generátorfüggvényébe betesszük a logaritmikus eloszlás generátorfüggvényét

$$\begin{gathered} \exp\left(\lambda\left(\frac{\ln(1-pz)}{\ln(1-p)} - 1\right)\right) = \\ = \exp\left(\frac{\lambda}{\ln(1-p)}\left(\ln(1-pz) - \ln(1-p)\right)\right) \stackrel{\circ}{=} \\ \stackrel{\circ}{=} \left(\frac{1-p}{1-pz}\right)^{\alpha} = \left(\frac{1/p-1}{1/p-z}\right)^{\alpha} \stackrel{\circ}{=} \left(\frac{\beta}{\beta+1-z}\right)^{\alpha}. \end{gathered}$$

Mivel a generátorfüggvény egyértelműen megadja az eloszlást a Poisson-eloszlás szerint kevert logaritmikus eloszlás negatív binomiális eloszlású[7].

□

12.3.5. Példa. A negatív binomiális folyamat várható értéke és szórása.

A fenti $X(t)$ Lévy-folyamatot szokás negatív binomiális folyamatnak nevezni[8]. Jelölje λ a Poisson-folyamat paraméterét és legyen p a logaritmikus eloszlás paramétere. Ekkor a fenti (12.1.1) képlet alapján

$$\mathbf{E}(X(t)) = \lambda t \cdot \frac{1}{-\ln(1-p)} \frac{p}{1-p}.$$

A varianciára vonatkozó (12.1.2) képlet alkalmazásakor vegyük figyelembe, hogy $\mathbf{E}(N(t)) = \mathbf{D}^2(N(t)) = \lambda t$, így

$$\mathbf{D}^2(X(t)) = \lambda t \cdot \left(\mathbf{D}^2(\xi) + \mathbf{E}^2(\xi)\right) = \lambda t \cdot \mathbf{E}\left(\xi^2\right).$$

A logaritmikus eloszlás második momentuma

$$\begin{aligned}
\mathbf{E}\left(\xi^2\right) &= \sum_{n=1}^{\infty} n^2 p_n = \frac{1}{-\ln(1-p)} \sum_{n=1}^{\infty} np^n = \\
&= \frac{1}{-\ln(1-p)} \sum_{n=1}^{\infty} np^n = \frac{p}{-\ln(1-p)} \sum_{n=1}^{\infty} np^{n-1} = \\
&= \frac{p}{-\ln(1-p)} \left(\sum_{n=0}^{\infty} p^n\right)' = \frac{p}{-\ln(1-p)} \left(\frac{1}{1-p}\right)' = \\
&= \frac{1}{-\ln(1-p)} \frac{p}{(1-p)^2}.
\end{aligned}$$

Vagyis szemben a Poisson-folyamattal a variancia mindig nagyobb mint a várható érték.

□

12.3.6. Példa. A negatív binomiális eloszlás várható értéke és szórása.

A várható érték a generátorfüggvény első deriváltja a $z = 1$ helyen:

$$\frac{d}{dz}\left(\frac{\beta}{\beta+1-z}\right)^{\alpha} = \alpha \left(\frac{\beta}{\beta+1-z}\right)^{\alpha-1} \frac{\beta}{(\beta+1-z)^2},$$

amiből a várható érték éppen

$$\mathbf{E}(\xi) = \alpha \left(\frac{\beta}{\beta}\right)^{\alpha-1} \frac{\beta}{(\beta)^2} = \frac{\alpha}{\beta} = \alpha \frac{q}{p}$$

[7] Az α és a β paramétereknek pozitívnak kell lenni, ami teljesül!

[8] Helyesebb lenne a negatív binomiális Lévy-folyamat elnevezés.

ahol $p \mathrel{\hat{=}} \beta/(1+\beta)$, $q \mathrel{\hat{=}} 1/(1+\beta)$. A logaritmikus eloszlás bevezetésekor a $\beta \mathrel{\hat{=}} 1/p-1$ jelöléssel éltünk[9]. Ebből $p = 1/(1+\beta)$, amit viszont most q-val jelöltünk. A negatív binomiális folyamat várható értékének és szórásnak hányadosa az ott használt jelöléssel $1-p$, ami a most használt jelöléssel éppen p. Vagyis a negatív binomiális eloszlás varianciája

$$
\begin{aligned}
\mathbf{D}^2(\xi) &= \frac{\mathbf{E}(\xi)}{p} = \alpha\frac{q}{p^2} = \alpha\frac{1/(1+\beta)}{(\beta/(1+\beta))^2} = \\
&= \alpha\frac{1}{1+\beta}\frac{(1+\beta)^2}{\beta^2} = \alpha\frac{1+\beta}{\beta^2} = \\
&= \frac{\alpha}{\beta} + \frac{\alpha}{\beta^2} = \mathbf{E}(\xi) + \frac{1}{\alpha}\mathbf{E}^2(\xi).
\end{aligned}
$$

□

12.3.7. Példa. Az első ugrás időpontja a negatív binomiális folyamatban.

Kézenfekvően merül fel a kérdés, hogy mi lesz a negatív binomiális folyamatban a várakozási idők eloszlása. Mivel az ugrások nagysága lehet 1-nél nagyobb, ezért a $\tau \mathrel{\hat{=}} \inf\{t \mid X(t) \geq 1\}$ találati idő eloszlására vagyunk kíváncsiak. Legyen (α, β) adott és legyen $X(t)$ egy olyan negatív binomiális Lévy-folyamat, amelyre $X(1)$ eloszlása (α, β) paraméterű negatív binomiális eloszlású változó. Mi lesz az $X(t)$ eloszlása? Az $X(1)$ generátorfüggvénye

$$G_1(z) = \left(\frac{\beta}{\beta+1-z}\right)^{\alpha}.$$

Ebből a korlátlanul való oszthatóság miatt tetszőleges r racionális számra

$$G_r(z) = \left(\frac{\beta}{\beta+1-z}\right)^{r\alpha},$$

ami folytonosan kiterjeszthető minden valós t-re. A $\mathbf{P}(X(t) = 0)$ valószínűség éppen a generátorfüggvény értéke a $z = 0$ helyen, amiből

$$
\begin{aligned}
\mathbf{P}(\tau \geq t) &= \mathbf{P}(X(t) = 0) = \left(\frac{\beta}{\beta+1}\right)^{t\alpha} = \\
&= \exp\left(-\alpha \ln\left(\frac{\beta+1}{\beta}\right) t\right) \mathrel{\hat{=}} \exp(-\lambda t),
\end{aligned}
$$

[9] Vegyük észre, hogy a negatív binomiális eloszlás esetén az n a q kitevőjében szerepelt, a logaritmikus eloszlás esetén a p kitevőjében. Ennek oka, hogy a negatív binomiális eloszlásban, akárcsak a geometriai eloszlásban, az úgymond sikertelen próbálkozásokat számoljuk, ezért jelöljük q-val.

vagyis az exponenciális várakozási idő paramétere

$$\lambda \stackrel{\circ}{=} \alpha \ln\left(\frac{\beta+1}{\beta}\right) > 0.$$

Ha most a logaritmikus eloszlásnál használt $p = 1/(1+\beta)$ konvenciót használjuk, akkor

$$\lambda = -\alpha \ln(1-p).$$

Ezzel a jelöléssel a negatív binomiális folyamat várható értéke[10]

$$\mathbf{E}(X(t)) = \lambda t \cdot \frac{1}{-\ln(1-p)} \frac{p}{1-p} = t\alpha \frac{p}{q},$$

varianciája

$$\mathbf{D}^2(X(t)) = t\alpha \frac{p}{q^2}.$$

A p és a q felcserélésével és a $t = 1$ behelyettesítéssel kapjuk a negatív binomiális eloszlás várható értékének és szórásának már ismert képletét.

□

12.3.8. Példa. Poisson-folyamat megállítása.

Az imént bemutatott keverés egy másik lehetséges interpretációja a következő: Legyen $N(t)$ egy Poisson-folyamat és legyen $\tau < \infty$ egy véletlen időpont. Mi lesz az

$$N(\tau)(\omega) \stackrel{\circ}{=} N(\tau(\omega), \omega)$$

megállított változó eloszlása? Mivel

$$\mathbf{P}(N(\tau) = n \mid \tau = t) = \frac{(\lambda t)^n}{n!} \exp(-\lambda t),$$

ezért, ha a τ gamma eloszlású, akkor a megállított változó eloszlása negatív binomiális lesz. Vegyük észre, hogy a τ változót nem a λ, hanem a t helyébe tettük. Ha az $N(t)$ Poisson-folyamat paramétere λ, a τ változó paraméterei pedig (α, β) akkor az $N(\tau)$ tekinthető egy olyan keverésnek, amikor a keverő változó $\lambda\tau$. A τ sűrűségfüggvénye

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x),$$

amiből a transzformált változók sűrűségfüggvényére vonatkozó képlet alapján a $\lambda\tau$ sűrűségfüggvénye

$$\begin{aligned} g(x) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{x}{\lambda}\right)^{\alpha-1} \exp\left(-\beta \frac{x}{\lambda}\right) \frac{1}{\lambda} = \\ &= \frac{(\beta/\lambda)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\beta \frac{x}{\lambda}\right), \end{aligned}$$

ami gamma eloszlású, de a β paraméter helyébe β/λ írandó.

[10] Ismét ügyeljünk a p és a q két eltérő használatára.

□

12.3.9. Példa. Szubordinált, vagy más néven alárendelt Poisson-folyamatok.

Lévy-folyamatok konstruálásának jól ismert és a gyakorlatban előszeretettel használt módszere a szubordináció, vagy alárendelt folyamatok konstruálása. A szubordináció alapgondolata az, hogy a megfigyelt folyamat két komponensből áll. Egyrészt a folyamatnak van egy véletlen „belső" órája, ami egy monoton növekedő Lévy-folyamattal írható le, másrészt van egy másik Lévy-folyamat, amely a rendszer reakcióit írja le a determinisztikusan változó időben[11]. Az első folyamatot szokás alárendelt folyamatnak mondani. A két Lévy-folyamat kompozíciója adja a megfigyelt folyamatot, amelyről megmutatható, hogy szintén Lévy-folyamat. Mivel a gamma eloszlás korlátlanul osztható, ezért létezik gamma folyamat, vagyis olyan $X(t)$ Lévy-folyamat, amelyre az $X(1)$ eloszlása adott (α, β) paraméterű gamma eloszlást követ. Mivel az $X(t) - X(s)$ minden $t > s$ esetén gamma eloszlást követ $(\alpha(t-s), \beta)$ paraméterekkel, ezért az $X(t) - X(s)$ egy valószínűséggel pozitív, így a gamma folyamat használható alárendelt időtranszformáltnak. Ha most $N(t)$ egy Poisson-folyamat, akkor az $Y(t) = N(X(t))$ folyamatot szubordinált, vagy alárendelt Poisson-folyamatnak mondjuk. Mivel tetszőleges t esetén az $X(t)$ gamma eloszlású, az elmondottak alapján az $Y(t)$ eloszlása negatív binomiális lesz. Ha a Poisson-folyamat paramétere λ, akkor az $Y(t)$ eloszlásának paraméterei $(\alpha t, \beta/\lambda)$ lesznek. Megmutatható, hogy egy független növekményű folyamat csak akkor lehet folytonos, ha minden időpontban az eloszlása normális, ezért egy $X(t)$ szubordinátor, az $X(t) = t$ triviális esettől eltekintve[12], nem lehet folytonos. Szemléletesen az $X(t)$ szubordinátor egy ugró folyamat, vagyis hosszabb-rövidebb szakaszokon az értéke konstans, majd egy kisebb-nagyobb ugrással az értéke megváltozik stb. Ha az időugrás elég nagy, akkor az ugráshoz tartozó időszak alatt az $N(t)$ Poisson-folyamat több ugrása is lezajlik, így az összetett folyamatban egyszerre több ugrást is megfigyelhetünk. Célszerű hangsúlyozni, hogy mivel az alárendelt Lévy-folyamatok szintén Lévy-folyamatok, ezért a gamma folyamattal szubordinált Poisson-folyamat, a paraméterek egyezése esetén, eloszlásban megegyezik a korábban konstruált negatív binomiális Lévy-folyamattal.

□

[11] Ezt a módszert előszeretettel szokás használni az árfolyamok időben való alakulásának modellezésére. A szokásos interpretáció, hogy például a piacon az aktivitás időben változik. Amikor a kereskedés aktívabb, a „belső" idő gyorsabban pörög.

[12] Az t értéket nulla szórással rendelkező normális eloszlású változónak tekintjük.

12.4. Kockáztatott érték kiszámolása rekurzióval

Tegyük fel, hogy a lehetséges egyedi veszteségeloszlásokat egy diszkrét eloszlással közelítjük[13]. Jelölje $\left(r_j\right)_{j=0}^{M}$ a közelítő valószínűségeket.

$$\begin{aligned} G(z) &\mathrel{\hat{=}} \sum_{n=0}^{\infty} \mathbf{P}(L=n)\, z^n \mathrel{\hat{=}} \sum_{n=0}^{\infty} A_n z^n = \mathbf{E}\left(z^L\right) = \mathbf{E}\left(\mathbf{E}\left(z^L\right) \mid N=k\right) = \\ &= \mathbf{E}\left(\mathbf{E}\left(z^{\sum_{k=1}^{N} \xi_k}\right) \mid N=k\right) = \sum_k \mathbf{E}\left(z^{\sum_{i=1}^{k} \xi_i}\right) \cdot p_k = \\ &= \sum_k \left(\mathbf{E}\left(z^{\xi}\right)\right)^k \cdot p_k = G_N\left(P(z)\right), \end{aligned}$$

ahol $P(z) \mathrel{\hat{=}} \sum_{j=1}^{M} z^j r_j$ veszteség generátorfüggvénye és G_N a veszteség frekvenciáinak generátorfüggvénye.

12.4.1. Példa. Ha a lehetséges veszteségek $1,2$ és 4 millió és ezek bekövetkezési valószínűségei $r_1 = 0,1$, $r_2 = 0,4$ és $r_4 = 0,5$, akkor

$$P(z) = 0,1z + 0,4z^2 + 0z^3 + 0,5z^4.$$

Ha a bekövetkezések száma Poisson-eloszlást alkot, akkor

$$G_N(z) \mathrel{\hat{=}} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \exp(-\lambda)\, z^k = \exp(\lambda(z-1)).$$

Vezessük be a $\mu_j \mathrel{\hat{=}} \lambda r_j$ jelölést. Ekkor

$$P(z) = \sum_{j=1}^{M} r_j z^j \mathrel{\hat{=}} \sum_{j=1}^{M} \frac{\mu_j}{\lambda} z^j.$$

$$\widehat{P}(z) \mathrel{\hat{=}} \lambda P(z) = \sum_{j=1}^{M} \mu_j z^j$$

12.4.2. Lemma. A szorzatfüggvény n-edik deriváltjának képlete $(uv)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} u^{(k)} v^{(n-k)}$.

[13] Ez az alkalmazásokban közvetlenül adódik.

Bizonyítás: Az azonosságot teljes indukcióval lehet igazolni. Ha $n = 0$ vagy $n = 1$, akkor a formula közismerten igaz. Az n-ről $n+1$-re való lépés bizonyítása pedig

$$
\begin{aligned}
(uv)^{(n+1)} &= \sum_{k=0}^{n} \binom{n}{k} \left(u^{(k)} v^{(n-k)}\right)' = \\
&= \sum_{k=0}^{n} \binom{n}{k} \left(u^{(k+1)} v^{(n-k)} + u^{(k)} v^{(n+1-k)}\right) = \\
&= \binom{n}{0} uv^{(n+1)} + \sum_{k=1}^{n} u^{(k)} v^{(n+1-k)} \left(\binom{n}{k} + \binom{n}{k-1}\right) + \binom{n}{n} u^{(n+1)} v = \\
&= \binom{n+1}{0} uv^{(n+1)} + \sum_{k=1}^{n} u^{(k)} v^{(n+1-k)} \binom{n+1}{k} + \binom{n+1}{n+1} u^{(n+1)} v = \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} u^{(k)} v^{(n+1-k)}.
\end{aligned}
$$

□

A Poisson-eloszlás képletét használva ha

$$
A_n \circeq \mathbf{P}(Loss = n) = \frac{1}{n!} G^{(n)}(0),
$$

akkor a $z = 0$ helyen

$$
\begin{aligned}
A_n &= \frac{1}{n!} \frac{d^{n-1}}{dz^{n-1}} \left(\frac{d}{dz} G(z)\right) = \frac{1}{n!} \frac{d^{n-1}}{dz^{n-1}} \left(\lambda G(z) \frac{d}{dz} P(z)\right) = \\
&= \frac{1}{n!} \frac{d^{n-1}}{dz^{n-1}} \left(G(z) \frac{d}{dz} \lambda P(z)\right) = \frac{1}{n!} \frac{d^{n-1}}{dz^{n-1}} \left(G(z) \frac{d}{dz} \widehat{P}(z)\right) = \\
&= \frac{1}{n!} \sum_{k=0}^{n-1} \binom{n-1}{k} G^{(n-k-1)}(0) \widehat{P}^{(k+1)}(0) = \\
&= \sum_{k=0}^{n-1} \frac{1}{n!} \binom{n-1}{k} A_{n-k-1} (n-k-1)! (k+1)! \mu_{k+1} = \\
&= \sum_{k=0}^{n-1} \frac{k+1}{n} A_{n-k-1} \mu_{k+1},
\end{aligned}
$$

ugyanis egyszerűen számolva

$$
\begin{aligned}
&\frac{1}{n!} \binom{n-1}{k} (n-k-1)! (k+1)! = \\
&= \frac{1}{n!} \frac{(n-1)!}{k! (n-1-k)!} (n-k-1)! (k+1)! = \\
&= \frac{k+1}{n}.
\end{aligned}
$$

Az iteráció induló értéke $G(0) = \exp(\lambda(P(0)-1))$. Mivel $P(0) = 0$, ugyanis a P valóban csak a veszteségeket tartalmazza, ezért $G(0) = \exp(-\lambda)$, ami annak a valószínűsége, hogy nem lesz ugrás a Poisson-folyamatban.

Mi történik akkor, ha a darabszámok eloszlása nem Poisson? Tegyük fel, hogy a veszteségek darabszámának G_N generátorfüggvényére teljesül a

$$\frac{d}{dz}\ln(G_N(z)) = \frac{A(z)}{B(z)} = \frac{\sum_{k=0}^{r} a_k z^k}{\sum_{k=0}^{s} b_k z^k}$$

szabály, vagyis

$$B(z)\frac{d}{dz}G_N(z) = G_N(z)A(z).$$

12.4.3. Példa. A negatív binomiális eloszlással számolva

$$\frac{d}{dz}\ln\left(\frac{\beta}{\beta+1-P(z)}\right)^{\alpha} = \frac{\frac{d}{dz}\left(\frac{\beta}{\beta+1-P(z)}\right)^{\alpha}}{\left(\frac{\beta}{\beta+1-P(z)}\right)^{\alpha}} = \frac{\frac{d}{dz}(\beta+1-P(z))^{-\alpha}}{(\beta+1-P(z))^{-\alpha}}$$
$$= \frac{-\alpha(\beta+1-P(z))^{-\alpha-1}P'(z)}{(\beta+1-P(z))^{-\alpha}} = \frac{-\alpha P'(z)}{\beta+1-P(z)} \mathrel{\mathring{=}} \frac{A(z)}{B(z)}.$$

G_N helyébe a hatványsorát beírva,

$$\left(\sum_{k=0}^{s} b_k z^k\right)\left(\sum_{n=0}^{\infty}(n+1)A_{n+1}z^n\right) = \sum_{k=0}^{r} a_k z^k\left(\sum_{n=0}^{\infty} A_n z^n\right).$$

z^n-hez tartozó tagok a két oldalon egyenlőek

$$\sum_{j=0}^{\min(s,n)} b_j(n+1-j)A_{n+1-j} = \sum_{i=0}^{\min(r,n)} a_i A_{n-i},$$

amiből

$$b_0(n+1)A_{n+1} = \sum_{i=0}^{\min(r,n)} a_i A_{n-i} - \sum_{j=1}^{\min(s,n)} b_j(n+1-j)A_{n+1-j}$$

ami egy rekurzió az (A_n) sorozatra.

XIII.

A CENTRÁLIS HATÁRELOSZLÁS TÉTELE

Ebben a fejezetben a centrális határeloszlás tételét vizsgáljuk meg. A centrális határeloszlás tétele a normális eloszlás szerepét indokolja. A tételnek számos alakja van, mi ezek közül csak a legegyszerűbbet tárgyaljuk: Legyenek $\xi_1, \xi_2, \ldots$ független, azonos eloszlású valószínűségi változók, és tekintsük az

$$\eta_n \overset{\circ}{=} \frac{\sum_{k=1}^{n} \xi_k - n \cdot \mathbf{E}(\xi)}{\sqrt{n}\mathbf{D}(\xi)}$$

centralizált és normalizált sorozatot, ahol ξ a ξ_k változók bármelyike. A tétel szerint minden x-re

$$\lim_{n \to \infty} \mathbf{P}(\eta_n < x) = \Phi(x),$$

ahol

$$\Phi(x) \overset{\circ}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt$$

a standard normális eloszlás eloszlásfüggvénye[1].

13.1. A karakterisztikus függvény módszer

A tétel bizonyításának számos alakja ismert, de talán a leginkább kézenfekvő a karakterisztikus függvényekre (Fourier-transzformáltakra) épülő bizonyítás. A módszer lényege, hogy az eloszlásfüggvények konvergenciája helyett a megfelelő karakterisztikus függvények konvergenciáját mutatjuk meg. Ehhez azonban némi magyarázatot kell adni.

1. A leginkább kézenfekvő kérdés az, hogy miért használjuk a függvénytranszformáltakat és miért nem közvetlenül foglalkozunk az eloszlásfüggvényekkel? A válasz viszonylag egyszerű: A tételben szereplő összeg eloszlásfüggvényét nehéz kiszámolni, miközben a karakterisztikus függvényüket egyszerű, ugyanis miként már említettük ha ξ és η függetlenek, akkor a $\xi + \eta$ összeg karakterisztikus függvénye a tagok karakterisztikus függvényeinek szorzata:

$$\begin{aligned} \varphi_{\xi+\eta}(t) &\overset{\circ}{=} \mathbf{E}(\exp(it(\xi+\eta))) = \mathbf{E}(\exp(it\xi)\exp(it\eta)) = \\ &= \mathbf{E}(\exp(it\xi))\mathbf{E}(\exp(it\eta)) = \varphi_\xi(t)\varphi_\eta(t). \end{aligned}$$

Vegyük észre, hogy kihasználtuk, hogy a függetlenség tulajdonsága a transzformált változókra való áttéréskor is megmarad. Az összeggel rokon és gyakran használt szabály a

$$\varphi_{a\xi}(t) = \mathbf{E}(\exp(it(a\xi))) = \mathbf{E}(\exp(i(ta)\xi)) = \varphi_\xi(at)$$

képlet, amely a konstanssal való szorzás karakterisztikus függvényét adja meg.

[1] Feltéve persze, hogy az η_n változók értelmesek. Evvel kizárjuk a konstans sorozatokat és azokat, amelyeknek nincs véges szórásuk.

2. Miért kell áttérni komplex számokra, miért használjuk a karakterisztikus függvényt és nem a momentumgeneráló függvényt? A válasz ismét kézenfekvő: A momentumgeneráló függvény nem jellemzi az eloszlást. Miként láttuk, az is előfordulhat, hogy egy eloszlásnak létezik az összes momentuma, de a momentumgeneráló függvénye az $s = 0$ ponton kívül mindenhol végtelen. Ha a valószínűségi változó nem negatív, akkor a momentumgeneráló függvény egyértelműen jellemzi az eloszlást, de a tételben ezzel a megkötéssel nem tudunk élni. A karakterisztikus függvény azonban minden t értékre értelmes, ugyanis az

$$\exp(it) = e^{it} = \cos t + i \sin t$$

Euler-képlet miatt

$$\varphi(t) \stackrel{\circ}{=} \mathbf{E}(\exp(it\xi)) = \mathbf{E}(\cos t\xi) + i\mathbf{E}(\sin t\xi),$$

ahol a két várható érték mindig véges. Ugyan nem teljesen evidens, de megmutatható[2], hogy a karakterisztikus függvény egyértelműen jellemzi az eloszlást, vagyis ha két valószínűségi változónak különböző az eloszlása, akkor különböző lesz a karakterisztikus függvényük is. Természetesen némi ellenérzést válthat ki a komplex számok használata, de könnyen látható, hogy csak az

$$\exp(z_1 + z_2) = \exp(z_1)\exp(z_2)$$

azonosságra van szükségünk, amely indoklása igen egyszerű. Emlékeztetünk, hogy definíció szerint tetszőleges z komplex számra

$$\exp(z) \stackrel{\circ}{=} \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Egy hatványsort természetes módon a komplex síkban érdemes definiálni. Az $\exp(z)$ kifejezést definiáló komplex hatványsor minden z komplex számra abszolút konvergens, így az alábbi számítás során a sor átrendezése megengedett, így alkalmazhatjuk a hatványsorokra vonatkozó Cauchy-szorzást:

$$\begin{aligned}
\exp(z_1)\exp(z_2) &= \sum_{k=0}^{\infty} \frac{z_1^k}{k!} \cdot \sum_{l=0}^{\infty} \frac{z_2^l}{l!} = \sum_{k=0}^{\infty} \sum_{l=0}^{k} \frac{z_1^k}{l!} \frac{z_2^{k-l}}{(k-l)!} = \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{l=0}^{k} \binom{k}{l} z_1^k z_2^{k-l} = \sum_{k=0}^{\infty} \frac{1}{k!} (z_1 + z_2)^k = \\
&= \exp(z_1 + z_2).
\end{aligned}$$

Az komplex exponenciális függvényre vonatkozó addíciós tétel miatt tetszőleges $z = x + iy$ esetén $\exp(z) = \exp(x)\exp(iy)$, ahol $\exp(x)$ a közönséges valós exponenciális

[2] V.ö.: B.2.3. Állítás, 242. oldal.

függvény. Elegendő tehát csak a karakterisztikus függvényben szereplő $\exp(it)$ függvénnyel foglalkoznunk. Az is könnyen látható, hogy $\exp(\bar{z}) = \overline{\exp(z)}$, ugyanis a komplex számok konjugálása felcserélhető az algebrai műveletekkel és folytonos transzformáció. Ebből kifolyólag

$$\begin{aligned} 1 &= \exp(0) = \exp(it - it) = \\ &= \exp(it)\exp(-it) = \exp(it)\overline{\exp(it)} = \\ &= |\exp(it)|^2 \end{aligned}$$

vagyis az $\exp(it)$ függvény a valós számokat a komplex egységkörre képezi. Több módon is igazolható, hogy ilyenkor a fent említett Euler-képlet teljesül. Tekintsük például a következőt. Jelölje $\cos t$ és $\sin t$ az $\exp(it)$ pont koordinátáit. Ekkor definíció szerint

$$\exp(it) \stackrel{\circ}{=} \cos t + i \sin t.$$

Nem világos azonban, hogy az így definiált két függvény azonos-e a középiskolában bevezetett trigonometrikus függvényekkel. A komplex exponenciális függvény komplex értelemben deriválható. Az addíciós képlet miatt

$$\frac{\exp(z+h) - \exp(z)}{h} = \exp(z) \frac{\exp(h) - 1}{h}.$$

A második tagba a definíciót betéve és az algebrai műveleteket elvégezve azonnal látható, hogy a második tag a teljes komplex síkban konvergens

$$1 + \frac{h}{2!} + \frac{h^2}{3!} + \ldots$$

sor, így mindenhol folytonos, tehát

$$\lim_{h \to 0} \frac{\exp(h) - 1}{h} = 1,$$

vagyis komplex értelemben az $\exp(z)$ függvény deriváltja önmaga. Hasonlóan igazolható az $(\exp(az))' = a \exp(z)$ szabály, ahol a egy tetszőleges komplex konstans. Ebből következően

$$\frac{d}{dt} \exp(it) = i \exp(it) = i \cos t - \sin t.$$

Vagyis a $\cos t$ deriváltja $-\sin t$ és a $\sin t$ deriváltja $\cos t$. Ugyanakkor a köríven az $(1,0)$ pont és az $(\cos t, \sin t)$ közötti ív hossza

$$\begin{aligned} \int_0^t \sqrt{\left(\frac{d \cos s}{ds}\right)^2 + \left(\frac{d \sin s}{ds}\right)^2} ds &= \int_0^t \sqrt{(\sin s)^2 + (\cos s)^2} ds = \\ &= \int_0^t |\exp(is)| \, ds = \int_0^t 1 ds = t, \end{aligned}$$

vagyis a $\cos t$ és a $\sin t$ éppen a t hosszú ívhez tartozó két koordináta[3].

[3] Ez éppen a középiskolai definíció.

3. A momentumgeneráló függvény legfontosabb tulajdonsága, hogy ha az értelmezési tartománya tartalmazza az origó egy nyílt környezetét, akkor segítségével a momentumok meghatározhatóak. Karakterisztikus függvény esetén azonban nem kell aggódnunk az értelmezési tartomány miatt. Ha létezik az n-edik momentum, akkor a karakterisztikus függvény n-szer deriválható és

$$\varphi^{(n)}(0) \stackrel{\circ}{=} \frac{d^n\varphi}{dt^n}(0) = i^n\mathbf{E}(\xi^n).$$

Ugyanis miként már említettük, komplex esetben is az $\exp(az)$ deriváltja $a\exp(az)$, így például az első deriváltakra

$$\varphi'(t) = \frac{d}{dt}\mathbf{E}(\exp(it\xi)) = \mathbf{E}\left(\frac{d}{dt}\exp(it\xi)\right) = \mathbf{E}(i\xi\exp(it\xi)),$$

amiből $t = 0$ helyettesítéssel $\varphi'(0) = i\mathbf{E}(\xi)$. Hasonlóan

$$\varphi''(t) = \frac{d^2}{dt^2}\mathbf{E}(\exp(it\xi)) = \mathbf{E}\left(\frac{d^2}{dt^2}\exp(it\xi)\right) = \mathbf{E}\left(i^2\xi^2\exp(it\xi)\right),$$

amiből ha $t = 0$, akkor $\varphi''(0) = i^2\mathbf{E}\left(\xi^2\right) = -\mathbf{E}\left(\xi^2\right)$, és így tovább.

4. A karakterisztikus függvény módszer kulcstétele az úgynevezett Lévy-féle konvergenciatétel[4]. A tétel szerint, ha a karakterisztikus függvényekből álló $\varphi_n(t)$ sorozat minden t-re egy $\varphi(t)$ karakterisztikus függvényhez tart, akkor a $\varphi_n(t)$-hez tartozó $F_n(x)$ eloszlásfüggvények sorozata a $\varphi(t)$-hez tartozó $F(x)$ eloszlásfüggvényhez tart gyengén. Gyenge konvergencia alatt azt értjük, hogy az $F_n(x)$ sorozat $F(x)$-hez tart de csak akkor, ha az x az $F(x)$ határeloszlásnak folytonossági pontja. Mivel a centrális határeloszlás tételében $F(x) = \Phi(x)$ és a $\Phi(x)$ folytonos, ezért ilyenkor az $F_n(x) \to F(x)$ konvergencia minden pontban teljesül.

5. Szükségünk lesz még a standard normális eloszlás karakterisztikus függvényére. Miként korábban a momentumgeneráló függvények tárgyalása kapcsán megjegyeztük, ha a momentumgeneráló függvény véges az $s = 0$ egy nyílt környezetében, akkor elegendő $s = it$ helyettesítést alkalmazni. Miként már láttuk, tetszőleges s esetén

$$\begin{aligned} M(s) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(sx)\exp\left(-\frac{x^2}{2}\right)dx = \\ &= \exp\left(\frac{s^2}{2}\right)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\frac{(x-s)^2}{2}\right)dx = \\ &= \exp\left(\frac{s^2}{2}\right), \end{aligned}$$

[4] V.ö.: B.3.15. Állítás, 256. oldal.

ugyanis az integrál az $1/\sqrt{2\pi}$ konstanssal együtt éppen az s várható értékű, 1 szórású normális eloszlás sűrűségfüggvényének integrálja. Ha most $s = it$ helyettesítést végzünk, akkor $\varphi_{N(0,1)}(t) = \exp\left(-t^2/2\right)$. Mivel igen fontos állításról van szó, egy másik bizonyítást is bemutatunk.

$$\begin{aligned}\varphi(t) &\stackrel{\circ}{=} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(itx)\exp\left(-\frac{x^2}{2}\right)dx = \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\cos(tx)\exp\left(-\frac{x^2}{2}\right)dx,\end{aligned}$$

ugyanis a második, a szinuszos tag integrálja nulla, hiszen a szinusz függvény páratlan. Az integrál mögé deriválva, majd parciálisan integrálva

$$\begin{aligned}\varphi'(t) &= -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin(tx)\,x\exp\left(-\frac{x^2}{2}\right)dx = \\ &= \frac{1}{\sqrt{2\pi}}\left(\left[\exp\left(-\frac{x^2}{2}\right)\sin(tx)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\cos(tx)\,t\exp\left(-\frac{x^2}{2}\right)dx\right) = \\ &= \frac{1}{\sqrt{2\pi}}\left(0 - t\int_{-\infty}^{\infty}\cos(tx)\exp\left(-\frac{x^2}{2}\right)dx\right) = -t\varphi(t).\end{aligned}$$

Mivel a φ karakterisztikus függvény, ezért $\varphi(0) = 1$, így a differenciálegyenlet egyetlen megoldása a $\varphi(t) = \exp\left(-t^2/2\right)$. Mielőtt tovább megyünk, érdemes egy megjegyzést tenni. Mivel a független változók összegének karakterisztikus függvényei összeszorzódnak, ezért a normális eloszlás korlátlanul osztható, ugyanis ha a $\xi_1, \ldots, \xi_n$ független változók eloszlása $N(0, 1/\sqrt{n})$, akkor az összeg eloszlása $N(0,1)$. Ugyanakkor van itt egy további érdekesség. Az $N(0, 1/\sqrt{n})$ szintén normális eloszlású, de ami fontosabb az eloszlás típusa is megegyezik az $N(0,1)$ eloszlással. Az eloszlások körében a típus fogalmát gyakran pontatlanul, lényegében stilisztikai tartalommal szokás használni. Így például szokás Poisson-típusú eloszlásról, vagy valószínűségi változóról beszélni, ami alatt azt szokás érteni, hogy valamilyen λ paraméterrel rendelkező Poisson-eloszlású változóról van szó. Ez valójában egy kicsit pontatlan. Ugyanis definíció szerint két eloszlást azonos típusúnak mondunk, ha a mögöttük levő valószínűségi változók egymásból $ax + b$ alakú lineáris transzformációval származtathatóak. Így két különböző paraméterű Poisson-eloszlás típusa nem azonos[5], de két normális esetén a típus azonos. Egy korlátlamul osztható eloszlást stabil eloszlásnak mondunk, ha a felbontásban szereplő eloszlások típusa megegyezik az eredeti eloszlás típusával. Távolról sem triviális, de az egyetlen olyan eloszlás, amely stabil és van szórása, az a normális eloszlás. Így például a Poisson-eloszlás korlátlanul osztható, de nem stabil. Egy másik hasznos

[5] Ha ξ Poisson-eloszlású, akkor az $a\xi + b$ általában nem lesz Poisson-eloszlású, ugyanis a tartója nem lesz a $0, 1, \ldots$ halmaz.

és rendkívül gyakran említett és korábban már igazolt összefüggés, hogy független normális eloszlású változók összege normális. Ez következik abból is, hogy az

$$\exp\left(-\left(\sigma_1 t\right)^2\right)\exp\left(it\mu_1\right)\exp\left(-\left(\sigma_2 t\right)^2\right)\exp\left(it\mu_2\right)$$

szorzat felírható

$$\exp\left(-\left(t\sqrt{\sigma_1^2+\sigma_2^2}\right)^2\right)\exp\left(it\left(\mu_1+\mu_2\right)\right)$$

módon, ami az $N\left(\mu_1+\mu_2,\sqrt{\sigma_1^2+\sigma_2^2}\right)$ eloszlás karakterisztikus függvénye.

6. Ennyi előkészítés után a centrális határeloszlás tétel igazolása már viszonylag egyszerű. A tétel állítása szerint a (ξ_k) változók közös eloszlásának van várható értéke és szórása. Mivel a centralizált és normalizált összeg határeloszlását kell vizsgálni, ezért feltehető, hogy a közös várható érték nulla, a szórás pedig 1. Mivel az első két momentum létezik, a $\varphi(t)$ kétszer deriválható, vagyis értelmes az

$$\begin{aligned}1-\frac{t^2}{2} &= \frac{1}{0!}+\frac{i\mathbf{E}(\xi)}{1!}t+\frac{i^2\mathbf{E}\left(\xi^2\right)}{2!}= \\ &= \frac{\varphi(0)}{0!}+\frac{\varphi'(0)}{1!}t+\frac{\varphi''(0)}{2!}t^2\end{aligned}$$

Taylor-polinom. A L'Hospital szabály miatt

$$\lim_{t\to 0}\frac{\varphi(t)-\left(\frac{\varphi(0)}{0!}+\frac{\varphi'(0)}{1!}t+\frac{\varphi''(0)}{2!}t^2\right)}{t^2}=0,$$

így $\varphi(t)=1-t^2/2+o\left(t^2\right)$, ahol a $o\left(t^2\right)$ a közelítés maradéktagja, és csak annyit tudunk róla, hogy t^2-tel osztva nullához tart. A $\Sigma_{k=1}^n\xi_k$ összeg karakterisztikus függvénye a szorzási szabály miatt

$$\varphi^n(t)=\left(1-\frac{t^2}{2}+o\left(t^2\right)\right)^n.$$

De mivel az összeget el kell osztani $\sqrt{n}$-nel, ezért a már említett

$$\varphi_{a\xi}(t)=\mathbf{E}\left(\exp\left(it\left(a\xi\right)\right)\right)=\mathbf{E}\left(\exp\left(i\left(ta\right)\xi\right)\right)=\varphi_\xi(at)$$

szabály miatt ha az η_n karakterisztikus függvénye $\varphi_n(t)$, akkor

$$\begin{aligned}\varphi_n(t) &= \left(1-\frac{\left(t/\sqrt{n}\right)^2}{2}+o\left(\left(t/\sqrt{n}\right)^2\right)\right)^n= \\ &= \left(1-\frac{t^2}{2n}+o\left(\frac{t^2}{n}\right)\right)^n.\end{aligned}$$

Közismerten

$$\lim_{n\to\infty}\left(1-\frac{t^2}{2n}\right)^2=\exp\left(-\frac{t^2}{2}\right),$$

de meg kell mutatni, hogy az $o\left(t^2/n\right)$ tag sem okoz gondot. Ha $n\to\infty$, akkor

$$\begin{aligned}\left|\left(1-\frac{t^2}{2n}+o\left(\frac{t^2}{n}\right)\right)^n-\left(1-\frac{t^2}{2n}\right)^n\right| &\circeq \\ \circeq |a^n-b^n|=|a-b|\left|\sum_{k=0}^{n-1}a^kb^{n-k-1}\right| &\leq \\ \leq |a-b|\sum_{k=0}^{n-1}\left|a^kb^{n-k-1}\right|\leq n\,|a-b|=no\left(\frac{t^2}{n}\right) &= \\ =t^2\left(\frac{o\left(\frac{t^2}{n}\right)}{t^2/n}\right)&\to 0,\end{aligned}$$

ahol kihasználtuk, hogy ha n elég nagy akkor

$$\begin{aligned}|a| &\circeq |\varphi(t)|=|\mathbf{E}(\exp(it\xi))|\leq\mathbf{E}(|\exp(it\xi)|)=\mathbf{E}(1)=1\\ |b| &\circeq \left|\left(1-\frac{t^2}{2n}\right)\right|\leq 1.\end{aligned}$$

Ezzel a centrális határeloszlás tételének igazolását befejeztük.

□

13.2. A közelítés sebessége

Közkeletű elképzelés, hogy a centrális határeloszlás tételében a konvergencia sebessége gyors. Ezt azonban a szokásos matematikai tételek nem igen igazolják. A konvergencia sebességére vonatkozó legfontosabb tétel a következő:

13.2.1. Tétel (Berry–Esseen)**.** *A korábban tárgyalt feltételeken kívül tegyük fel, hogy a* (ξ_n) *sorozat tagjai közös eloszlásának véges a* $\rho\circeq\mathbf{E}\left(|\xi|^3\right)$ *harmadik abszolút momentuma, ahol* ξ *a* ξ_n *változók bármelyike. Tegyük fel továbbá, hogy* $\mathbf{E}(\xi)=0$, *de nem tesszük fel, hogy a* $\sigma\circeq\mathbf{D}(\xi)$ *szórás egyenlő volna* 1*-gyel. Ilyenkor létezik olyan* $C<0,5$ *konstans, hogy minden n-re és x-re*

$$|F_n(x)-\Phi(x)|\leq C\frac{\rho}{\sigma^3\sqrt{n}}.$$

A tétel szerint a konvergencia sebessége $O(1/\sqrt{n})$ nagyságrendű. Ez azt jelenti, hogy $1/\sqrt{n}$-nel osztva, vagyis $\sqrt{n}$-nel szorozva a konvergencia sebességét megadó kifejezés korlátos. Az úgynevezett Hölder-egyenlőtlenséggel megmutatható, hogy tetszőleges $n \leq m$ esetén $\sqrt[n]{\mathbf{E}(|\xi|^n)} \leq \sqrt[m]{\mathbf{E}(|\xi|^m)}$, amiből $\sigma^3 \leq \rho$, vagyis $\rho/\sigma^3 \geq 1$.

13.2.2. Példa. Számoljuk ki a Berry–Esseen tételben szereplő konstanst az egyenletes eloszlás esetében.

Mivel a várható értéknek nullának kell lenni, ezért egy 0 közepű intervallumon kell az eloszlást tekinteni. Az egyszerűség kedvéért ez legyen a $(-1/2, 1/2)$ intervallum. Ekkor

$$\rho = \int_{-1/2}^{1/2} |x|^3 \, dx = 2\int_0^{1/2} x^3 dx = \frac{1}{32}.$$

$\sigma^3 = \left(1/\sqrt{12}\right)^3$. Következésképpen

$$\frac{\rho}{\sigma^3} = \frac{1/32}{\left(1/\sqrt{12}\right)^3} = 1,299.$$

Így tétel szerint a konvergencia sebessége

$$C\frac{\rho}{\sigma^3\sqrt{n}} \leq \frac{0,5 \cdot 1,3}{\sqrt{n}} = \frac{0,65}{\sqrt{n}}.$$

Az egyenlőtlenség minden n-re érvényes. Ha $n = 1$, akkor a $(-1/2.1/2)$ intervallumon $F_1(x) = x + 1/2$. Ilyenkor

$$\sup_x |F_1(x) - \Phi(x)| \approx 0,3085 < 0,65.$$

□

Az alábbi némiképpen heurisztikus gondolatmenet célja annak vizsgálata, hogy miért csak négyzetgyökös a konvergencia a centrális határeloszlás tételében. Lehet-e javítani a konvergencia sebességén, ha nem csak a harmadik momentum létezését, hanem ennél magasabb momentumok létezését is feltesszük? A konvergencia sebességével függ össze az úgynevezett Edgeworth-kifejtés. A kifejtésben központi szerepet játszanak az úgynevezett Hermite-polinomok. A Hermite-polinomok definíciója az irodalomban nem egységes, de valószínűségszámítási környezetben az $\exp\left(-x^2/2\right)$ deriváltjaival definiáljuk őket[6].

[6] Időnként, például a Matlab programcsomag, a Hermite-polinomokat az $\exp\left(-x^2\right)$ deriváltjaival definiálják. Így a Matlab Hermite-polinomjaira a $2^{-n/2}H_n\left(x/\sqrt{2}\right)$ helyettesítést kell használni.

13.2.3. Definíció. Ha $\varphi(x) \circeq 1/\sqrt{2\pi}\exp\left(-x^2/2\right)$ a standard normális eloszlás sűrűségfüggvénye, akkor a

$$\varphi^{(n)}(x) \circeq \frac{d^n}{dx^n}\varphi(x) = (-1)^n H_n(x)\varphi(x)$$

összefüggésben szereplő $H_n(x)$ polinomot az n-edik Hermite-polinomnak mondjuk.

13.2.4. Példa. Rekurzív formulák.

A polinomok számolásakor nem szükséges a deriválásokat elvégezni. A definiáló egyenletet deriválva

$$\begin{aligned}
\varphi^{(n+1)}(x) &= \frac{d}{dx}\left((-1)^n H_n(x)\varphi(x)\right) = \\
&= (-1)^n\left(H_n'(x)\varphi(x) + H_n(x)\varphi'(x)\right) = \\
&= (-1)^n\left(H_n'(x)\varphi(x) - xH_n(x)\varphi(x)\right) = \\
&= (-1)^{n+1}\left(xH_n(x) - H_n'(x)\right)\varphi(x),
\end{aligned}$$

amiből

$$H_{n+1}(x) = xH_n(x) - H_n'(x).$$

Indukcióval könnyen látható, hogy $H_n'(x) = nH_{n-1}(x)$. Valóban $H_0(x) = 1$ és $H_1(x) = x$, így az $n = 1$ eset evidens. Az indukciós feltételt használva

$$\begin{aligned}
H_{n+1}'(x) &= H_n(x) + xH_n'(x) - nH_{n-1}'(x) = \\
&= H_n(x) + xnH_{n-1}(x) - nH_{n-1}'(x) = \\
&= H_n(x) + n\left(xH_{n-1}(x) - H_{n-1}'(x)\right) = \\
&= H_n(x) + nH_n(x) = (n+1)H_n(x).
\end{aligned}$$

Tehát

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}.$$

Az első hat Hermite-polinom a következő:

$$\begin{aligned}
H_0(x) &= 1, \\
H_1(x) &= x, \\
H_2(x) &= x^2 - 1, \\
H_3(x) &= x^3 - 3x, \\
H_4(x) &= x^4 - 6x^2 + 3, \\
H_5(x) &= x^5 - 10x^3 + 15x, \\
H_6(x) &= x^6 - 15x^4 + 45x^2 - 15.
\end{aligned}$$

□

Az Edgeworth-kifejtésben a centralizált és normalizált összeg sűrűség- és eloszlásfüggvényének Hermite-polinomokkal való közelítését adjuk meg. A közelítésben szereplő kifejezések bár nem sűrűségfüggvények, ugyanis az értékük lehet negatív és az integráljuk sem feltétlenül egy, de jól reprezentálják a konvergencia sebességében szerepet játszó mutatók jelentőségét. Az egyszerűbb jelölés céljából tegyük fel, hogy a ξ_k változók közös eloszlásának várható értéke nulla, a σ szórása pedig legyen egy. Ugyancsak az egyszerűbb jelölés miatt csak a másodrendű Edgeworth-közelítéseket adjuk meg. Ilyenkor a negyedik momentum létezését tételezzük fel. A kifejés lényege, hogy a centrális határeloszlás tételében az eloszlásfüggvények konvergenciája ekvivalens a karakterisztikus függvény konvergenciájával. A karakterisztikus függvény egy komplex értékű függvény, amelynek nincs feltétlenül klasszikus értelemben vett logaritmusa. Ennek ellenére definiálhatjuk a kumulánsait, és a kumulánsok segítségével megpróbáljuk sorba fejteni a karakterisztikus függvényt. Miként korábban megjegyeztük, hiába tart a kurtózis $1/n$ nagyságrendben a normális eloszlás kurtózisához, a ferdeség csak $1/\sqrt{n}$-nel tart nullához, így a konvergencia sebessége is csak négyzetgyökös lehet.

1. Legyen C_j a ξ változó j-edik kumulánsa és vezessük be a $\rho_j \stackrel{\circ}{=} C_j/\sigma^j$ hányadosokat. Az általános Edgeworth-kifejtésben a ρ_j konstansok szerepelnek. Például

$$\rho_3 \stackrel{\circ}{=} \frac{C_3}{\sigma^3} = \frac{\mathbf{E}\left((\xi - \mathbf{E}(\xi))^3\right)}{\mathbf{D}(\xi)^3}$$

a ferdeség, illetve a

$$\rho_4 = \frac{C_4}{\sigma^4} = \frac{\mathbf{E}\left((\xi - \mathbf{E}(\xi))^4\right) - 3\mathbf{D}^4(\xi)}{\mathbf{D}^4(\xi)} = \frac{\mathbf{E}\left((\xi - \mathbf{E}(\xi))^4\right)}{\mathbf{D}^4(\xi)} - 3$$

az úgynevezett többletkurtózis, vagyis a kurtózis hárommal csökkentett értéke. Mivel csak a másodrendű közelítéssel fogunk foglalkozni, ezért csak erre a két mutatóra lesz szükségünk.

2. Miként megjegyeztük, a karakterisztikus függvény egyértelműen jellemzi az eloszlást. Az eloszlásfüggvény és a karakterisztikus függvény közvetlen kapcsolatát leíró úgynevezett inverziós formula viszonylag bonyolult képlet. Bizonyos esetekben azonban egy $\psi(t)$ karakterisztikus függvény és a hozzá tartozó $f(x)$ sűrűségfüggvény között egy szép és egyszerű reláció írható fel[7].

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-itx)\,\psi(t)\,dt.$$

Figyeljünk az $1/(2\pi)$ konstansra és a kitevőben szereplő negatív előjelre[8]! Példaként vegyük a standard normális eloszlás esetét. A t és az x szerepét felcserélve, a normális

[7] Reméljük, hogy nem zavaró, hogy most $\varphi(x)$ a standard normális eloszlás sűrűségfüggvénye, bár legtöbbször $\varphi(t)$ egy karakterisztikus függvény. Ezért jelöljük $\psi(t)$-vel a karakterisztikus függvényt.

[8] A formula érvényességéhez szükséges és elegendő, hogy a $\psi(t)$ karakterisztikus függvény abszolút integrálható legyen. Ez egy elegáns, de viszonylag erős feltétel.

eloszlás karakterisztikus függvényének képlete segítségével

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-itx\right)\exp\left(-\frac{t^2}{2}\right)dt = \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-itx\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)dt = \\
&= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\left(-x\right)^2}{2}\right) = \\
&= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) \stackrel{\circ}{=} \varphi\left(x\right).
\end{aligned}$$

A két oldalt n-szer x szerint deriválva

$$\begin{aligned}
\left(-1\right)^n\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(it\right)^n\exp\left(-itx\right)\exp\left(-\frac{t^2}{2}\right)dt &= \varphi^{(n)}\left(x\right) = \\
&= \left(-1\right)^n H_n\left(x\right)\varphi\left(x\right),
\end{aligned}$$

amiből

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(it\right)^n\exp\left(-itx\right)\exp\left(-\frac{t^2}{2}\right)dt = H_n\left(x\right)\varphi\left(x\right).$$

3. Az Edgeworth-közelítés alapgondolata az, hogy megpróbáljuk a ξ_k változók közös karakterisztikus függvényét magasabb rendben sorba fejteni. Tekintsük tehát a normalizált és centralizált közelítő összegben szereplő ξ változó $\psi\left(t\right)$ karakterisztikus függvényének negyedrendű sorba fejtését. Felhasználva, hogy a várható érték nulla, a szórás pedig 1:

$$\psi\left(t\right) = 1 - \frac{t^2}{2!} + \frac{\left(it\right)^3}{3!}\mathbf{E}\left(\xi^3\right) + \frac{\left(it\right)^4}{4!}\mathbf{E}\left(\xi^4\right) + M_4\left(t\right).$$

A t helyébe a $\sqrt{n}$ -nel való normalizálás miatt $t/\sqrt{n}$-et téve a már korábban látott módon

$$\begin{aligned}
\psi\left(\frac{t}{\sqrt{n}}\right) &= 1 - \frac{t^2}{2!n} + \frac{\left(it\right)^3}{3!n^{3/2}}\mathbf{E}\left(\xi^3\right) + \frac{\left(it\right)^4}{4!n^2}\mathbf{E}\left(\xi^4\right) + M_4\left(\frac{t}{\sqrt{n}}\right) = \\
&= 1 - \frac{t^2}{2n} + \frac{\left(it\right)^3}{6n^{3/2}}\rho_3 + \frac{\left(it\right)^4}{24n^2}\left(\rho_4 + 3\right) + M_4\left(\frac{t}{\sqrt{n}}\right).
\end{aligned}$$

Mivel a karakterisztikus függvény négyszer deriválható, ezért a Taylor-sorfejtésben a L'Hospital szabály szerint a maradéktag $o\left(t^4\right)$ nagyságrendű[9], így

$$M_4\left(\frac{t}{\sqrt{n}}\right) = o\left(\frac{1}{n^2}\right).$$

[9] Figyeljünk megint a jelölésre. A $o\left(t^4\right)$ azt jelenti, hogy t^4-nel osztva a kifejezés nullához tart!

Ezt követően vegyük az n-edik hatványt, majd az $1/n$-nél magasabb rendű tagokat vonjuk össze egy $o(1/n)$ hibatagba

$$\psi\left(\frac{t}{\sqrt{n}}\right)^n = \left(1-\frac{t^2}{2n}\right)^n +$$
$$+\binom{n}{1}\left(1-\frac{t^2}{2n}\right)^{n-1}\left(\frac{(it)^3}{6n^{3/2}}\rho_3+\frac{(it)^4}{24n^2}(\rho_4+3)\right)+$$
$$+\binom{n}{2}\left(1-\frac{t^2}{2n}\right)^{n-2}\frac{(it)^6\rho_3^2}{36n^3}+o\left(\frac{1}{n}\right).$$

4. A L'Hospital szabály segítségével könnyen látható, hogy tetszőleges x esetén[10]

$$\left(1+\frac{x}{n}\right)^n = \exp(x)\left(1-\frac{x^2}{2n}\right)+o\left(\frac{1}{n}\right).$$

Ezeket behelyettesítve és kihasználva, hogy az alacsonyabb hatványok felszorzása csak további $o(1/n)$ nagyságú tagokat eredményez:

$$\psi\left(\frac{t}{\sqrt{n}}\right)^n = \exp\left(-\frac{t^2}{2}\right)\left(1-\frac{t^4}{8n}\right)+$$
$$+\exp\left(-\frac{t^2}{2}\right)\left(\frac{(it)^3}{6n^{1/2}}\rho_3+\frac{(it)^4}{24n}(\rho_4+3)\right)+$$
$$+\exp\left(-\frac{t^2}{2}\right)\frac{(it)^6\rho_3^2}{72n}+o\left(\frac{1}{n}\right).$$

Ebből a $t^4/8n$ tagot a ρ_4-hez tartozó kifejezéssel összevonva

$$\psi\left(\frac{t}{\sqrt{n}}\right)^n - \exp\left(-\frac{t^2}{2}\right) =$$
$$=\exp\left(-\frac{t^2}{2}\right)\left(\frac{(it)^3}{6\sqrt{n}}\rho_3+\frac{(it)^4}{24n}\rho_4+\frac{(it)^6\rho_3^2}{72n}\right)+o\left(\frac{1}{n}\right).$$

5. Ha $f_n(x)$ jelöli az η_n normalizált és centralizált összeg sűrűségfüggvényét, akkor az imént ismertetett inverziós formulát alkalmazva[11]

$$f_n(x)-\varphi(x)\approx\varphi(x)\left(\frac{\rho_3}{6\sqrt{n}}H_3(x)+\frac{\rho_4}{24n}H_4(x)+\frac{\rho_3^2}{72n}H_6(x)\right).$$

[10] A két oldal különbségét kell venni, majd osztani kell n-nel, s ezt követően meg kell mutatni, hogy a határérték nulla. Ez a lépés helyettesíti a karakterisztikus függvény logaritmusának vételét, ami komplex esetben nem magától értetődő.

[11] Vegyük észre, hogy az $o(1/n)$ tagot elhagytuk. Ennek oka, hogy nem tudjuk, hogy a közelítés egyenletese. Ez az egyik oka annak, hogy az Edgeworth-előállítás heurisztikus.

Az x változó szerint integrálva[12]

$$\Phi(x) - F_n(x) \approx \varphi(x)\left(\frac{\rho_3}{6\sqrt{n}}H_2(x) + \frac{\rho_4}{24n}H_3(x) + \frac{\rho_3^2}{72n}H_5(x)\right).$$

Látható, hogy az eltérés a többletkurtózissal $O(1/n)$ nagyságrendben, de a ferdeséggel csak $O(1/\sqrt{n})$ nagyságrendben tart nullához[13].

13.2.5. Példa. Moivre–Laplace képlet.

A Moivre–Laplace képlet a binomiális eloszlás normálissal való közelítését adja meg. A $B(n.p)$ binomiális eloszlás n-darab indikátorváltozó összege, így

$$\lim_{n\to\infty} \mathbf{P}\left(a \le \frac{B(n,p) - np}{\sqrt{npq}} \le b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b \exp\left(-\frac{x^2}{2}\right)dx.$$

A képlet sok szempontból érdekes. Egyrészt az

$$\eta_n \stackrel{\circ}{=} \frac{B(n,p) - np}{\sqrt{npq}}$$

változók értékkészletének egyesítése egy U megszámlálható halmazt alkot. Az U halmazba esés valószínűsége minden n-re 1, de a határeloszlás esetén, ugyanis a normális eloszlás folytonos, 0. Vagyis a centrális határeloszlás tétele nem élesíthető úgy, hogy a közelítő összegek eloszlása eseményenként tart a határeloszláshoz[14]. A képlet sok bizonyítása ismert. A bizonyítások előnye, hogy lehetőség van a konvergencia sebességének közvetlen becslésére is. Nem túl meglepő módon ezek szintén a $O(1/\sqrt{n})$ képletet eredményezik. Ugyanakkor közvetlen numerikus szimulációkra hivatkozva az irodalomban szokás az $npq > 18$ úgynevezett Feller-szabályt javasolni. Ha például $p = 0,1$, akkor

$$n \ge \frac{18}{0,9 \cdot 0,1} = 200.$$

A legjobb esetet akkor kapjuk, ha $p = 0,5$. Ilyenkor

$$n \ge \frac{18}{0,5 \cdot 0,5} = 72,$$

vagyis a binomiális eloszlás normálissal való közelítése, feltéve, hogy a p nem nagyon kicsi már százas nagyságrendű minta esetén indokolt. Ugyanakkor a Feller-szabály csak

[12]Ügyeljünk az előjelre és a Hermite-polinomok indexére!

[13]Mivel a ferdeség okozza a nagyobb problémát szokás a $n > 25\rho_3^2$ ökölszabályt használni. Az exponenciális eloszlás ferdesége $\rho_3 = 2$. Ebből következően a közelítés során $n \ge 100$. A lognormális eloszlás ferdesége $\rho_3 = \left(e^{\sigma^2} + 2\right)\sqrt{e^{\sigma^2} - 1}$. Ha például $\sigma = 1$, akkor $\rho_3 = 6,1849$. Ebből körülbelül $n \ge 1000$ adódik. Ha azonban $\sigma = 2$, akkor $\rho_3 = 414,36$ és ilyenkor $n \ge 4,3 \cdot 10^6$ mintaelemszám adódik. Ugyanakkor a részvények hozamának becslésekor $\sigma = 0,2$-del, vagyis 20%-kal szokás számolni. Ilyenkor $n \ge 10$ adódik.

[14]Emlékeztetünk, hogy eloszláson a $\mu(B) = \mathbf{P}(\xi \in B)$ halmazfüggvényt értjük.

egy heurisztikus észrevétel, amely használhatóságára semmilyen matematikai garancia nincs.

□

13.2.6. Példa. Binomiális eloszlás közelítése Poisson-eloszlással.

Ha a p értéke nagyon kicsi, mondjuk 10^{-4}, akkor

$$n \geq \frac{18}{10^{-4} \cdot \left(1 - 10^{-4}\right)} > 1,8 \cdot 10^5,$$

ami nagyon nagy elemszámot jelent. Mivel igen ritka eseményről van szó és ilyenkor a keresett esemény előfordulásának darabszámát is célszerű kicsinek venni, így célszerű a Poisson-közelítést választani. A binomiális eloszlás Poisson-eloszlással való közelítése kapcsán felmerülhet, hogy lehet-e a Poisson-eloszlást közelíteni normálissal: A válasz igen, ha $\lambda \to \infty$. A legegyszerűbb ismét a karakterisztikus függvény módszert használni. A Poisson-eloszlás karakterisztikus függvénye

$$\begin{aligned} \varphi(t) &= \mathbf{E}(\exp(it\xi)) = \sum_{k=0}^{\infty} \exp(itk) \frac{\lambda^k}{k!} \exp(-\lambda) = \\ &= \sum_{k=0}^{\infty} \frac{(\lambda \exp(it))^k}{k!} \exp(-\lambda) = \exp(\lambda(\exp(it) - 1)). \end{aligned}$$

Mivel a normálissal való közelítéshez ismét centralizálni és normalizálni kell a sorozatot, ezért a $(\xi - \lambda)/\sqrt{\lambda}$ sorozat határértékét kell $\lambda \to \infty$ esetén venni. Karakterisztikus függvényekre áttérve

$$\begin{aligned} &\exp\left(\lambda\left(\exp\left(\frac{it}{\sqrt{\lambda}}\right) - 1\right) - \lambda \frac{it}{\sqrt{\lambda}}\right) = \\ &= \exp\left(\lambda\left(\exp\left(\frac{it}{\sqrt{\lambda}}\right) - 1 - \frac{it}{\sqrt{\lambda}}\right)\right) = \\ &= \exp\left(\lambda\left(\frac{1}{2}\left(\frac{it}{\sqrt{\lambda}}\right)^2 + o\left(\left(\frac{it}{\sqrt{\lambda}}\right)^2\right)\right)\right) = \\ &= \exp\left(-\frac{t^2}{2} + \lambda o\left(\left(\frac{it}{\sqrt{\lambda}}\right)^2\right)\right) \to \exp\left(-\frac{t^2}{2}\right). \end{aligned}$$

□

13.2.7. Példa. Binomiális eloszlás közelítése Poisson-eloszlással.

A binomiális eloszlás karakterisztikus függvénye $\varphi(t) = (q + p\exp(it))^n$. Ha bevezetjük a $\lambda \mathrel{\overset{\circ}{=}} np$ jelölést, akkor

$$\begin{aligned}\varphi(t) &= \left(\left(1-\frac{\lambda}{n}\right)+\frac{\lambda}{n}\exp(it)\right)^n = \\ &= \left(1+\frac{\lambda(\exp(it)-1)}{n}\right)^n \to \exp(\lambda(\exp(it)-1)),\end{aligned}$$

amely éppen a λ paraméterű Poisson-eloszlás karakterisztikus függvénye.

□

XIV.

A NAGY SZÁMOK TÖRVÉNYEI

A nagy számok különböző törvényei a centrális határeloszlás tételek mellett a másik alapvető határeloszlás tétel típus. Akárcsak a centrális határeloszlás tételben most is legyenek $\xi_1, \xi_2, \ldots$ független, azonos eloszlású valószínűségi változók, és tekintsük az

$$\eta_n \stackrel{\circ}{=} \frac{\sum_{k=1}^{n} \xi_k}{n}$$

hányadost. A lényeges eltérés, hogy most más osztót alkalmazunk. A nagy számok különböző törvényei szerint az η_n sorozat valamilyen értelemben az $\mathbf{E}(\xi)$ határértékhez tart, ahol ξ a ξ_k változók bármelyike. A törvény szerint a számtani átlag éppen a várható értékhez tart. Bizonyos szempontból a centrális határeloszlás tétele és a nagy számok törvénye rokon állítások: Mind a kettőben független azonos eloszlású változók összege szerepel. Két lényeges eltérés azonban van: Egyrészt más az osztó, de sokkal fontosabb, hogy más a konvergencia módja. A centrális határeloszlás tételében, miként azt a neve is tartalmazza csak az eloszlásfüggvények, konvergálnak, a nagy számok törvényében azonban maguk az η_n változók konvergálnak. Valószínűségi változók esetén a konvergenciát számtalan módon definiálhatjuk. Bár vannak más alakok is, a konvergencia módja alapján a nagy számok törvényének két alakját szokás megkülönböztetni: Beszélhetünk a nagy számok erős törvényéről, amikor a konvergencia 1 valószínűséggel teljesül, és beszélhetünk a nagy számok gyenge törvényéről, amikor a konvergencia csak sztochasztikus konvergencia értelemben teljesül. Megmutatható, hogy az 1 valószínűséggel való konvergencia implikálja a sztochasztikus konvergenciát. Kézenfekvő tehát a kérdés: Ha igaz a nagy számok erős törvénye, akkor miért érdekes a gyenge törvénye? Ennek oka igen egyszerű: egyrészt a gyenge törvényt jóval egyszerűbb igazolni, másrészt, és ez a fontosabb, a gyenge törvény esetén nem csak a konvergencia tényét tudjuk igazolni, hanem a konvergencia sebességére is tudunk becslést mutatni.

14.1. Sztochasztikus és erős konvergencia

A valószínűségszámításban a valószínűségi változók körében a legtöbbet használt két konvergencia típus az erős és sztochasztikus konvergencia:

14.1.1. Definíció. Legyen (ξ_n) változók egy sorozata és ξ egy további valószínűségi változó.

1. Azt mondjuk, hogy a (ξ_n) sorozat erős értelemben, vagy más néven 1 valószínűséggel tart a ξ-hez, ha az olyan ω kimenetelek halmazának valószínűsége, amelyre a $\xi_n(\omega) \to \xi(\omega)$ konvergencia teljesül éppen 1. Ezt szokás a $\mathbf{P}(\xi_n \to \xi) = 1$, vagy $\mathbf{P}(\lim_{n\to\infty} \xi_n = \xi) = 1$ módon is írni. Az erős konvergencia egy másik gyakran használt elnevezése a majdnem mindenhol való konvergencia. Erre utal a gyakran használt $\xi_n \stackrel{m.m.}{\to} \xi$ jelölési mód.

2. Azt mondjuk, hogy a (ξ_n) sorozat sztochasztikus értelemben, vagy más néven valószínűségben tart a ξ-hez, ha tetszőleges $\varepsilon > 0$ szám esetén

$$\mathbf{P}(|\xi_n - \xi| > \varepsilon) \to 0.$$

Ezt másképpen úgy is kifejezhetjük, hogy tetszőleges $\varepsilon > 0$ és $\delta > 0$ számokhoz van olyan N küszöbindex, hogy ha $n \geq N$, akkor

$$\mathbf{P}(|\xi_n - \xi| > \varepsilon) < \delta.$$

A sztochasztikus konvergenciát szokás $\xi_n \xrightarrow{P} \xi$ módon is jelölni.

14.1.2. Állítás. *Ha $\xi_n \overset{m.m.}{\to} \xi$, akkor $\xi_n \xrightarrow{P} \xi$. Vagyis az 1 valószínűséggel való konvergencia implikálja a sztochasztikus konvergenciát.*

Fontos hangsúlyozni, hogy az 1 valószínűséggel való konvergencia általában nem implikálja a várható értékek konvergenciáját. Vagyis általában a várható érték operátora és a pontonkénti határérték operátora nem cserélhető fel.

14.1.3. Tétel (Kolmogorov)**.** *Legyen (ξ_n) azonos eloszlású, független valószínűségi változók egy sorozata és legyen ξ a ξ_n változók bármelyike. Valamely M konstansra az*

$$\eta_n \mathrel{\hat{=}} \frac{\sum_{k=1}^{n} \xi_k}{n} \overset{m.m.}{\to} M$$

konvergencia pontosan akkor teljesül, ha az $\mathbf{E}(\xi)$ várható érték véges. Ilyenkor $M = \mathbf{E}(\xi)$ és

$$\lim_{n\to\infty} \mathbf{E}(|\eta_n - \mathbf{E}(\xi)|) \to 0.$$

A gyenge törvények igazolásához első lépésként két egyszerű egyenlőtlenséget igazolunk.

14.1.4. Lemma (Markov-egyenlőtlenség)**.** Ha $\xi \geq 0$ egy valószínűségi változó és $a \geq 0$ egy szám, akkor

$$a\mathbf{P}(\xi \geq a) \leq \mathbf{E}(\xi).$$

Bizonyítás: A feltételek és a várható érték definíciója alapján ha $F(x)$ a ξ eloszlásfüggvénye, akkor

$$\begin{aligned} \mathbf{E}(\xi) &\mathrel{\hat{=}} \int_0^\infty x dF(x) \geq \int_a^\infty x dF(x) \geq \int_a^\infty a dF(x) = \\ &= a\int_a^\infty 1 dF(x) = a\mathbf{P}(\xi \geq a). \end{aligned}$$

Vegyük észre, hogy az egyenlőtlenség akkor is igaz, ha $\mathbf{E}(\xi)$ végtelen. Ilyenkor persze nincs mit igazolni.

$\square$

14.1.5. Lemma (Csebisev-egyenlőtlenség)**.** Tetszőleges ξ véges szórású változóra és minden $k>0$ számra

$$\mathbf{P}(|\xi-\mathbf{E}(\xi)|\geq\varepsilon)\leq\frac{\mathbf{D}^2(\xi)}{\varepsilon^2}.$$

Bizonyítás: A Markov-egyenlőtlenség alapján

$$\begin{aligned}\mathbf{P}(|\xi-\mathbf{E}(\xi)|\geq\varepsilon) &= \mathbf{P}\left((\xi-\mathbf{E}(\xi))^2\geq\varepsilon^2\right)\leq \\ &\leq \frac{\mathbf{E}\left((\xi-\mathbf{E}(\xi))^2\right)}{\varepsilon^2}=\frac{\mathbf{D}^2(\xi)}{\varepsilon^2}.\end{aligned}$$

□

A gyenge törvények közül a legismertebb az alábbi, amely a Csebisev-egyenlőtlenség közvetlen következménye.

14.1.6. Tétel (Nagy számok gyenge törvénye)**.** *Ha a* (ξ_k) *változók korrelálatlanok, és a* $(\mathbf{D}(\xi_k))$ *sorozat létezik és korlátos, akkor*

$$\sum_{k=1}^{n}\frac{\xi_k-\mathbf{E}(\xi_k)}{n}\xrightarrow{P}0.$$

Bizonyítás: Az általánosság megszorítása nélkül feltehető, hogy $\mathbf{E}(\xi_k)=0$. Jelölje K a szórások egy korlátját. A Csebisev-egyenlőtlenség alapján

$$\begin{aligned}\mathbf{P}\left(\left|\frac{\sum_{k=1}^n\xi_k}{n}\right|\geq\varepsilon\right) &\leq \frac{1}{\varepsilon^2}\mathbf{D}^2\left(\frac{\sum_{k=1}^n\xi_k}{n}\right)= \\ &= \frac{1}{\varepsilon^2n^2}\sum_{k=1}^{n}\mathbf{D}^2(\xi_k)\leq\frac{nK^2}{n^2}\to0.\end{aligned}$$

□

Az állítást a leggyakrabban a következő alakban szokás idézni:

14.1.7. Tétel (Bernoulli)**.** *Ha a* (ξ_k) *változók azonos eloszlásúak és függetlenek, és a közös eloszlásnak létezik szórása, akkor tetszőleges* $\varepsilon>0$ *esetén*

$$\mathbf{P}\left(\left|\frac{\sum_{k=1}^n\xi_k}{n}-M\right|>\varepsilon\right)\leq\frac{\mathbf{D}^2(\xi)}{\varepsilon^2n}\to0$$

ahol M *a közös eloszlás várható értéke.*

A Bernoulli-féle nagy számok törvénye nem egy szükséges és elégséges tétel. Éppen ezért rendkívül érdekes a következő tétel:

14.1.8. Tétel (Ehrenfeucht–Fisz)**.** *Legyen M egy konstans. Ha (ξ_n) független, azonos eloszlású változók sorozata, akkor a*

$$\frac{\sum_{k=1}^{n}\xi_k}{n} \xrightarrow{P} M \qquad (14.1.1)$$

konvergencia szükséges és elegendő feltétele, hogy a közös eloszlás $\varphi(t)$ karakterisztikus függvénye deriválható legyen a $t=0$ pontban. Ilyenkor

$$iM = \varphi'(0).$$

14.1.9. Példa. A Cauchy-eloszlás karakterisztikus függvénye nem deriválható a $t=0$ pontban, és nem is teljesül a nagy számok gyenge törvénye.

A (standard) Cauchy-eloszlás karakterisztikus függvénye $\varphi(t) = \exp(-|t|)$. Egy n-elemű (standard) Cauchy-eloszlásból vett minta számtani átlagának karakterisztikus függvénye

$$\exp\left(-\left|\frac{t}{n}\right|\right)^n = \exp(-|t|).$$

Vagyis a számtani átlag karakterisztikus függvénye szintén (standard) Cauchy eloszlású. Vegyük észre, hogy a $\varphi(t) = \exp(-|t|)$ képlet alapján a Cauchy-eloszlás korlátlanul osztható, ugyanis ha a ξ eloszlása Cauchy, akkor a ξ/n karakterisztikus függvénye $\exp(-|t|/n)$. Ebből az is következik, hogy a normális eloszlás mellett a Cauchy-eloszlás is stabil eloszlás. Ugyancsak, két független Cauchy-eloszlású változó összegének eloszlása szintén Cauchy-eloszlású. Valóban, legyenek ξ_1 és ξ_2 független standard Cauchy-eloszlású változók. Emlékeztetünk rá, hogy Cauchy-eloszláson a standard Cauchy-eloszlás lineáris transzformációit értjük. Könnyen látható tehát, hogy az általános Cauchy-eloszlás karakterisztikus függvénye

$$\exp(-|at|)\exp(itb).$$

Az $a_1\xi_1 + b_1$ és az $a_2\xi_2 + b_2$ független Cauchy-eloszlású változók összegének karakterisztikus függvénye

$$\exp(-|a_1 t|)\exp(ib_1 t)\exp(-|a_2 t|)\exp(ib_2 t),$$

ami éppen

$$\exp(-(|a_1| + |a_2|)|t|)\exp(it(b_1+b_2)),$$

amely szintén Cauchy-eloszlású.

□

14.2. A konvergencia sebessége

A nagy számok törvényének legfontosabb alkalmazása a relatív gyakoriság valószínűséghez való konvergálásának igazolása, illetve a konvergencia sebességének becslése.

Legyen adva egy A esemény, amely bekövetkezésének valószínűségét jelölje p. Végezzünk el n kísérletet, és tegyük fel, hogy az A esemény r_n-szer következett be. Ha a ξ_k 0 vagy 1 értékű változók jelölik az A esemény bekövetkezését, illetve be nem következését, akkor $r_n = \Sigma_{k=1}^{n} \xi_k$. Mivel $\mathbf{E}(\xi_k) = p$ és $\mathbf{D}^2(\xi_k) = pq$ ezért a nagy számok Bernoulli-féle törvénye alapján tetszőleges $\varepsilon > 0$ esetén

$$\mathbf{P}\left(\left|\frac{r_n}{n} - p\right| > \varepsilon\right) \leq \frac{pq}{n\varepsilon^2}.$$

Ha a p értékét nem ismerjük, akkor a $pq = p(1-p) \leq 1/4$ alapján még a legrosszabb esetben is a

$$\mathbf{P}\left(\left|\frac{r_n}{n} - p\right|\right) \leq \frac{1}{4n\varepsilon^2}$$

becsléssel élhetünk. Ez a konvergencia sebességére egy viszonylag gyenge becslést ad. Jóval erősebb becslés azonban a következő:

14.2.1. Tétel (Bernstein)**.** *A nagy számok gyenge törvényében, ha ε elég kicsi, nevezetesen ha $\varepsilon \leq \min(p,q)$, akkor*

$$\begin{aligned} \mathbf{P}\left(\left|\frac{r_n}{n} - p\right| \geq \varepsilon\right) &\leq 2\exp\left(-\frac{n\varepsilon^2}{2pq}\right) \leq \\ &\leq 2\exp\left(-2n\varepsilon^2\right), \end{aligned} \tag{14.2.1}$$

ahol r_n/n az első n kísérletből származó relatív gyakoriság.

A konvergencia sebességét megadó mindkét képletben ε^2 szerepel. Ez a centrális határeloszlás tételében látott négyzetgyökös konvergencia megfelelője. Ha az ε helyébe $\varepsilon/2$-t írunk, akkor ahhoz, hogy a közelítés valószínűségének becslése ne változzon, a kísérletek n számát a négyszeresére kell emelni. A sebesség becslése kapcsán hangsúlyozni kell, hogy a képletekben szereplő három változó az ε, a δ és az n közül kettőt rögzítve meghatározhatjuk a harmadikat. Vagyis adott n és ε esetén megmondhatjuk, hogy mennyi a δ, vagy például ε és δ megadásakor mekkora kell, hogy legyen az n ahhoz, hogy teljesüljön a $\mathbf{P}(|r_n/n - p| \geq \varepsilon) < \delta$ egyenlőtlenség.

14.2.2. Példa. A Csebisev- és Bernstein-egyenlőtlenségek összehasonlítása.

A Csebisev-egyenlőtlenségre alapuló Bernoulli-tétel szerint a $\mathbf{P}(|r_n/n - p| \geq \varepsilon) \leq \delta$ teljesüléséhez több mint

$$n_1(\delta) \circeq \left[\frac{1}{4\varepsilon^2\delta}\right] + 1 = \left[\frac{1/(2\delta)}{2\varepsilon^2}\right] + 1$$

kísérletre van szükség, ahol a $[\cdot]$ az egészrész operáció. A (14.2.1) alapján a Bernstein-féle alakban már

$$n_2(\delta) \circeq \left[\frac{\ln(2/\delta)}{2\varepsilon^2}\right] + 1$$

elem is elégséges. Az [ ] egészérték jelet elhagyva, ha $\delta \to 0$

$$\frac{n_1(\delta)}{n_2(\delta)} = \frac{1/2\delta}{\ln(2/\delta)} \to \infty,$$

vagyis kis δ-kra a Bernstein-féle becslés sokkal jobb mint a Csebisev-féle becslés. Ha például $\delta = \varepsilon = 10^{-2}$, akkor

$$\begin{aligned} n_1(\delta) &= 1/\left(4 \cdot 10^{-6}\right) = 250000, \\ n_2(\delta) &= \ln(200)/\left(2 \cdot 10^{-4}\right) = 26492, \end{aligned}$$

vagyis körülbelül tizede annyi elem kell a Bernstein-féle becslés szerint. Érdemes jelezni, hogy az $\varepsilon \le \min(p,q)$ feltétel miatt a Bernstein-féle becslés csak akkor használható, ha az esemény sem nem túl ritka, sem nem túl gyakori.

□

14.2.3. Példa. Nagy számok törvénye és a centrális határeloszlás tétele.

Ha se túl ritka, se túl gyakori eseményekről van szó, vagyis ha sem a p, sem a q értéke nem túl kicsi, akkor az $npq > 18$ Feller-szabály alapján már elegendően kicsi n esetén is a binomiális eloszlás jól közelíthető normális eloszlással. Az r_n változó binomiális eloszlással bír, így

$$\begin{aligned} \mathbf{P}\left(\left|\frac{r_n}{n} - p\right| \ge \varepsilon\right) &= \mathbf{P}\left(\left|\frac{r_n - np}{n}\right| \ge \varepsilon\right) = \\ &= \mathbf{P}\left(\left|\frac{r_n - np}{\sqrt{npq}}\right| \ge \frac{\sqrt{n}}{\sqrt{pq}}\varepsilon\right) \approx \\ &\approx \mathbf{P}\left(|N(0,1)| \ge \frac{\sqrt{n}}{\sqrt{pq}}\varepsilon\right) \le \\ &\le \mathbf{P}\left(|N(0,1)| \ge \frac{\sqrt{n}}{\sqrt{1/4}}\varepsilon\right) = \\ &= 2\left(1 - \Phi\left(2\sqrt{n}\varepsilon\right)\right) = \delta. \end{aligned}$$

Miként az előző példában, legyen $\varepsilon = \delta = 0,01$. Ezekkel

$$\Phi\left(2\sqrt{n}\varepsilon\right) = 1 - \frac{\delta}{2} = 0,995,$$

amiből

$$2\sqrt{n}\varepsilon \ge \Phi^{-1}(0,995) = 2,5758.$$

Tehát

$$n \ge \left(\frac{2,5758}{2 \cdot 0,01}\right)^2 = 16587,$$

amely jóval kevesebb, mint a Bernstein-féle becslés. Ha a δ értékét 5%-ra enyhítjük, akkor

$$\Phi\left(2\sqrt{n}\varepsilon\right) = 1 - \frac{\delta}{2} = 1 - \frac{0,05}{2} = 0,975.$$

Mivel $\Phi^{-1}(0,975) = 1,96$, ezért ilyenkor

$$n \geq \left(\frac{1,96}{2 \cdot 0,01}\right)^2 = 9604,$$

ami lényegében egy 10000-es mintának felel meg. Ha most $\varepsilon = \delta = 0,05$, akkor

$$n \geq \left(\frac{1,96}{2 \cdot 5 \cdot 10^{-2}}\right)^2 \approx 385.$$

Ha pedig $\varepsilon = 0,05$ és $\delta = 0,01$, akkor

$$n \geq \left(\frac{2,5758}{2 \cdot 5 \cdot 10^{-2}}\right)^2 \approx 664.$$

De gondolkodhatunk fordítva is. Tegyük fel, hogy $\varepsilon = 0,01$ és $n = 5000$. Mekkora lesz a tévedés valószínűsége? Ilyenkor

$$1 - \frac{\delta}{2} = \Phi\left(2\sqrt{n}\varepsilon\right) = \Phi\left(2\sqrt{5000} \cdot 0.01\right) = \Phi(1,1442) = 0,9214$$

ami $\delta = 15,5\%$-os tévedésvalószínűségnek felel meg. Ugyanakkor már

$$\Phi\left(2\sqrt{5000} \cdot 0,02\right) = 0,9977$$

ami körülbelül $\delta = 0,5\%$ tévedésvalószínűségnek felel meg.

□

14.2.4. Példa. A Brexit közvéleménykutatások és a nagy számok törvényének becslése a centrális határeloszlás tételével.

A Brexit-közvéleménykutatás igen szoros eredményt jósolt, vagyis a p értéke $0,5$ körül alakult. Tegyük fel, hogy $\varepsilon = 0,01$ hibával akarjuk a p értékét megbecsülni és a tévedés valószínűségét is $\delta = 1\%$ körül akarjuk tartani. Hány választót kell megkérdezni? Az előző példa alapján körülbelül 16500 személyt. A Financial Times összegyűjtötte a népszavazással kapcsolatos közvéleménykutatásokat. Ezek közül egyetlen egy volt, amelyben húszezer embert kérdeztek meg, és a közvéleménykutatás 2014 január 20-án történt. Ekkor 41% gondolta, hogy ki kell lépni és ugyanennyi gondolta, hogy maradni kell. Egy másik, 2015 december 5-én végzett mintavételben 10015 személyt kérdeztek meg, itt 42% gondolta, hogy ki kell lépni és 40%, hogy maradni kell. Az utolsó, egy nappal a szavazás előtt, 2016 június 22-én tartott közvéleménykutatás során 4700 személyt kérdeztek meg, és 55% mondta, hogy maradni akar, és 45% gondolta, hogy ki kell lépni. Tegyük fel, hogy a tévedés valószínűségét vagyis a δ értékét 1%-on akarjuk tartani. Ilyenkor az

$$2,5758 = 2\sqrt{n}\varepsilon$$

szabály alapján ha $\delta = 0,01$, akkor

$$\varepsilon = \frac{2,5758}{2\sqrt{n}} = \frac{2,5758}{2\sqrt{4700}} = 1,8786 \times 10^{-2} \approx 2\%$$

Azaz a lehetséges hiba 2% alatt lesz, vagyis a minta alapján esetleg gondolhatták, hogy a Brexit-szavazás eredményeként Nagy Britannia nem fog kilépni az EU-ból. Persze másképpen is gondolkodhatunk. Mivel a p becslését nagyon jól el akarjuk végezni, ezért az $\varepsilon = 0,01$ értéket tartani kell. Ilyenkor a tévedés valószínűsége

$$\delta = 2\left(1 - \Phi\left(2\varepsilon\sqrt{n}\right)\right) = 0,1703,$$

ami túlságosan nagy. Ugyanezen a napon egy másik közvéleménykutató 1032 embert kérdezett meg. Ők 48% maradást 42% kilépést és 11% bizonytalan szavazót mértek. Az ő esetükben a pontosság

$$\varepsilon = \frac{2,5758}{2\sqrt{n}} = \frac{2.5758}{2\sqrt{1032}} = 4,0091 \cdot 10^{-2} \approx 4\%$$

hibahatárt jelent. Ha most a 11% fele-fele arányban oszlik meg, akkor 53,5% a bennmaradás valószínűsége, de a 4%-os pontosság miatt előfordulhat, hogy $\delta = 0,01$ mellett is a kilépésre szavaz a többség. Mint tudjuk, végül 48% voksolt csak a maradásra Így végült több mint 5% volt az $|r_n/n - p|$. Ugyanakkor a tévedés valószínűsége $\varepsilon = 0,01$ és $n = 1032$ esetén $\delta = 0,52\%$ volt.

□

14.2.5. Példa. A Brexitet megelőző közvéleménykutatások és a Bernstein-egyenlőtlenség.

A centrális határeloszlás tételében a konvergencia sebességét talán túlértékeltük. Próbáljuk meg a

$$\mathbf{P}\left(\left|\frac{r_n}{n} - p\right| \geq \varepsilon\right) \leq 2\exp\left(-2n\varepsilon^2\right)$$

Bersntein-becslést alkalmazni. Mivel p nagyon közel van az $1/2$ értékhez, ezért az $\varepsilon = 0,01$ indokolt. Ha $n = 4700$, akkor a hibás becslés valószínűsége

$$2\exp\left(-2 \cdot 4700 \cdot (0.01)^2\right) = 0,78126,$$

ami elképesztően nagy. Ha $n = 1000$, akkor

$$2\exp\left(-2 \cdot 1000 \cdot (0.01)^2\right) = 1,6375,$$

ami már semmilyen információt nem hordoz.

□

XV.

ELLENŐRZŐ KÉRDÉSEK

Az alábbi állítások közül melyek igazak és melyek hamisak?

1. Az $f(x) = \sin x$ a $(0, \pi)$ intervallumon sűrűségfüggvény.
2. Ha ξ eloszlása a $(0, \pi/2)$ szakaszra koncentrálódik, és a sűrűségfüggvénye $A \sin x$ alakú, akkor $A = 1/2$.
3. Ha ξ egyenletes eloszlású a $(0, \pi)$ intervallumon, akkor az $\eta = \sin \xi$ változó várható értéke 2.
4. Jelölje $\left\{ {n \atop k} \right\}$ a másodfajú Stirling-számokat. $\left\{ {5 \atop 2} \right\} = 15$.
5. Jelölje $\left\{ {n \atop k} \right\}$ a másodfajú Stirling-számokat. $\left\{ {8 \atop 2} \right\} = 127$.
6. Ha ξ gamma eloszlású (a, λ) paraméterekkel, akkor a ξ változó n-edik momentuma
$$\frac{\lambda^a \Gamma(a+n)}{\Gamma(a+1) \lambda^n}.$$
7. Ha ξ és η független exponenciális eloszlású valószínűségi változók $\lambda = 2$ és $\mu = 5$ paraméterekkel, akkor a $\xi + \eta$ varianciája kisebb, mint $1/2$.
8. Legyen ξ Poisson-eloszlású $\lambda = 5$ paraméterrel, és legyen η az első olyan időpont, amikor hatost dobunk. $\mathbf{E}(\xi) < \mathbf{E}(\eta)$.
9. A Poisson-folyamat harmadik ugrásának időpontja χ_6^2 eloszlást követ.
10. A Poisson-eloszlás korlátlanul osztható abban az értelemben, hogy ha ξ Poisson-eloszlású, akkor tetszőleges n-re vannak olyan független, azonos eloszlású $\xi_1, \xi_2, \ldots, \xi_n$ változók, hogy a $\xi_1 + \xi_2 + \cdots + \xi_n$ eloszlása megegyezik a ξ eloszlásával.
11. A teljes valószínűség tétele szerint, ha (B_n) egy teljes eseményrendszer, akkor minden A esemény esetén
$$\mathbf{P}(A) = \sum_k \mathbf{P}(A \mid B_k) \mathbf{P}(B_k).$$
12. Legyen ξ eloszlása $N(0,1)$. Ha $F(x)$ az $\eta = \exp(\xi)$ változó eloszlásfüggvénye, akkor $F(1) = 1/2$.
13. Ha ξ és η binomiális eloszlású valószínűségi változók, akkor a $\xi + \eta$ is binomiális eloszlású.
14. Ha egy süteményben átlag két szem mazsola van, akkor annak a valószínűsége, hogy egy véletlenül kiválasztott süteményben lesz mazsola, $1 - \exp(-2)$.
15. Ha egy könyvben oldalanként átlagban két hiba van, akkor annak a valószínűsége, hogy egy oldalon nem lesz hiba, $1 - \exp(-2)$.

16. Ha ξ és η tetszőleges valószínűségi változók, akkor $\mathbf{E}\left(\xi^2\eta\right) \leq \mathbf{E}^2(\eta\xi)$.

17. Ha ξ normális eloszlású, akkor a ξ^2 gamma eloszlású.

18. Ha ξ geometriai eloszlású $p = 1/2$ paraméterrel, akkor $\mathbf{E}(\xi) = \sqrt{2}/2$.

19. Független gamma eloszlású változók szorzata Cauchy-eloszlást követ.

20. Ha ξ eloszlása (α, β) paraméterű másodfajú béta eloszlás, akkor $\mathbf{E}(\xi) = \alpha/(\beta - 1)$.

21. Legyenek ξ és η független valószínűség változók és jelölje f illetve g a sűrűségfüggvényüket. Ekkor a ξ/η hányadosnak is van sűrűségfüggvénye amely a
$$h(y) = \int_{-\infty}^{\infty} f(xy)\, g(x)\, |x|\, dx$$
integrállal számolható.

22. Ha ξ és η független valószínűségi változóknak van szórása, akkor a $\xi\eta$ szorzatnak is van szórása.

23. Ha ξ standard normális eloszlású, akkor $|\xi|$ sűrűségfüggvénye
$$f(x) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{x^2}{2}\right).$$

24. Ha a ξ változó sűrűségfüggvénye $f(x)$, akkor $1/\xi$ sűrűségfüggvénye
$$g(x) = f(1/x)\frac{1}{x^2}.$$

25. Ha a (ξ, η) pár együttes sűrűségfüggvénye $f(x, y)$, akkor az $(1/\xi, 1/\eta)$ pár együttes sűrűségfüggvénye
$$f(1/x, 1/y)\frac{1}{x^2}\frac{1}{y^2}.$$

26. Ha ξ eloszlása χ_n^2, akkor $\mathbf{E}(\xi) = n^2$.

27. Ha a (ξ, η) pár eloszlása a $T = [-1, 2] \times [0, 1]$ halmazra koncentrálódik, és az együttes eloszlásfüggvénye a T halmazon $A(1+x)^3 y^2$ alakú, akkor $A = 1/8$.

28. Ha a (ξ, η) pár eloszlása a $T = [0, 1] \times [0, 1]$ halmazra koncentrálódik, és az együttes eloszlásfüggvénye a T halmazon Ax^3y^2 alakú, akkor a $\xi\eta$ várható értéke kisebb, mint $1/2$.

29. Az exponenciális eloszlás Laplace-transzformációjának és momentumgeneráló függvényének értelmezési tartományainak közös része egy valódi intervallum.

30. Legyen ξ binomiális eloszlású, η pedig Poisson-eloszlású. Ha $\mathbf{E}(\xi)=\mathbf{E}(\eta)$, akkor $\mathbf{D}(\xi)\leq\mathbf{D}(\eta)$.

31. Legyen ξ binomiális eloszlású, η pedig exponenciális eloszlású. Ha $\mathbf{E}(\xi)=\mathbf{E}(\eta)$, akkor $\mathbf{D}(\xi)\leq\mathbf{D}(\eta)$.

32. Legyen ξ binomiális eloszlású, η pedig geometriai eloszlású. Ha $\mathbf{E}(\xi)=\mathbf{E}(\eta)$, akkor $\mathbf{D}(\xi)\leq\mathbf{D}(\eta)$.

33. Legyen ξ exponenciális eloszlású, az η pedig Poisson-eloszlású. Ha $\mathbf{E}(\xi)=\mathbf{E}(\eta)$, akkor $\mathbf{D}(\xi)\leq\mathbf{D}(\eta)$.

34. A standard normális eloszlás hatodik momentuma nagyobb, mint 10.

35. Az (a,λ) paraméterű gamma eloszlás várható értéke a/λ.

36. Ha ξ Cauchy-eloszlású, akkor ξ^2-nek van várható értéke.

37. Ha ξ Cauchy-eloszlású, akkor $\sqrt{|\xi|}$-nek van várható értéke.

38. Ha ξ Cauchy-eloszlású, akkor $\sqrt[4]{|\xi|}$-nek van várható értéke.

39. Ha ξ Cauchy-eloszlású, akkor $1/\xi$ is Cauchy-eloszlású.

40. Ha a ξ változó korlátos, vagyis ha van olyan K konstans, hogy $|\xi|\leq K$, akkor a ξ karakterisztikus függvénye végtelen sokszor deriválható.

41. $\exp(it)=\sin t+i\cos t$.

42. Ha $B(x,y)$ jelöli a béta függvényt, akkor $B(1,1/2)$ értéke racionális szám.

43. Ha $B(x,y)$ jelöli a béta függvényt, akkor $B(1,3/2)$ értéke racionális szám.

44. Ha $B(x,y)$ jelöli a béta függvényt, akkor $B(3,1/2)$ értéke racionális szám.

45. Ha $\xi\eta=0$, akkor $\mathbf{E}(\xi)\mathbf{E}(\eta)=0$.

46. Ha $\xi\eta=0$, akkor $|\mathbf{E}(\xi)\mathbf{E}(\eta)|\leq\mathbf{D}(\xi)\mathbf{D}(\eta)$.

47. A centrális határeloszlás tételében szereplő változók eloszlásfüggvényei a számegyenes minden pontjában konvergálnak.

48. Ha ξ és η függetlenek, akkor a korrelációs együtthatójuk nulla.

49. Ha ξ standard normális eloszlású, akkor ξ^2 sűrűségfüggvénye

$$f(x)=\frac{1}{\sqrt{2\pi x}}\exp\left(-\frac{x}{2}\right),\quad x>0.$$

50. Ha ξ_1 és ξ_2 független normális eloszlású valószínűségi változók, akkor $\xi_1 + 2\xi_2$ eloszlása szintén normális.

51. Ha a (ξ, η) pár eloszlása a $[0,1] \times [0,1]$ négyzetre koncentrálódik, akkor a négyzeten az xy lehet a közös eloszlásfüggvényük.

52. Ha a (ξ, η) pár eloszlása a $[0,1] \times [0,1]$ négyzetre koncentrálódik, akkor a négyzeten az $x + y$ lehet a közös eloszlásfüggvényük.

53. Ha a (ξ, η) pár eloszlása a $T = [0,1] \times [0,1]$ négyzetre koncentrálódik és az együttes eloszlásfüggvény a T halmazon $(xy)^3$, akkor ξ várható értéke $3/4$.

54. Az exponenciális eloszlás Laplace-transzformáltjának és momentumgeneráló függvényének deriváltjai megegyeznek.

55. A Poisson-eloszlás k-dik faktoriális momentuma λ.

56. A Poisson-eloszlás k-dik kumulánsa λ^k.

57. Ha valamely ξ valószínűségi változó karakterisztikus függvényének létezik az összes deriváltja, akkor létezik az összes momentuma is.

58. A $[0,1]$ szakaszon egyenletes eloszlású valószínűségi változó szórása $1/\sqrt{12}$.

59. $\exp(i\pi) + 1 = 0$.

60. Az exponenciális eloszlás várható értéke megegyezik a harmadik momentumával.

61. $\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t)\, dt$.

62. A Markov-egyenlőtlenség szerint ha $\xi \geq 0$, akkor $\mathbf{P}(\xi \leq a) \leq \mathbf{E}(\xi)/a$.

63. Ha $\xi \geqq a$, akkor $\mathbf{D}(\xi) \geq a$.

64. Ha ξ és η függetlenek, akkor $\mathbf{E}\left(\xi^k \eta^l\right) = \mathbf{E}\left(\xi^k\right) \mathbf{E}\left(\eta^l\right)$.

65. Ha $f(x)$ és $g(x)$ két sűrűségfüggvény, akkor a $h(x) = f(x) g(x)$ és az $u(x) = (f(x) + g(x))/2$ is sűrűségfüggvény.

66. Ha az $f(x)$ és a $g(y)$ a $[0,2]$ szakaszra koncentrálódó két sűrűségfüggvény, akkor a $h(x,y) = f(x) g(y)$ és az $u(x,y) = (f(x) + g(y))/2$ is egy-egy a $[0,2] \times [0,2]$ négyzetre koncentrálódó kétdimenziós sűrűségfüggvény.

67. $\Gamma(x, \lambda) = \Gamma(x)/\lambda^x$.

68. Ha ξ egy (a,b) paraméterű béta eloszlású változó, akkor $\mathbf{E}(\xi) = a/(a+b)$.

69. Legyen ξ Poisson-eloszlású változó $\lambda = 5$ paraméterrel, és legyen η egy $\lambda = 5$ paraméterű exponenciális eloszlású változó. Ekkor $\mathbf{E}(\xi) = \mathbf{D}^2(\eta)$.

70. Ha ξ és η független Poisson-eloszlású változók λ és μ paraméterekkel, akkor $\mathbf{P}(\xi+\eta=0)=1-\exp(-(\lambda+\mu))$.

71. Ha egy ξ változó eloszlása szimmetrikus, vagyis $\mathbf{P}(\xi\in B)=\mathbf{P}(\xi\in -B)$, akkor a ξ karakterisztikus függvénye mindig valós értékű.

72. Ha ξ eloszlása χ_n^2, akkor $\mathbf{D}\left(\xi^2\right)=2n$.

73. Ha valamely ξ diszkrét változó generátorfüggvénye $G(z)$, akkor $\mathbf{E}(\xi)=G'(1)$.

74. Legyenek ξ és η Cauchy-eloszlású változók. Ha $\zeta=\xi^2/\left(\xi^2+\eta^2\right)$, akkor a ζ várható értéke véges.

75. A $[0,\pi/2]$ szakaszra koncentrálódó ξ változó eloszlásfüggvénye ezen a szakaszon legyen
$$F(x)=1-\cos x.$$
Ekkor ξ^2 eloszlásfüggvénye az $x=\pi$ pontban $1-\cos\sqrt{\pi}$.

76. Ha az A és B események diszjunktak, akkor nem lehetnek függetlenek.

77. Ha ξ karakterisztikus függvénye $\varphi(t)$, akkor $-\xi$ karakterisztikus függvénye $\varphi(-t)$.

78. Ha valamely változó második momentuma véges, akkor van szórása.

79. A második kumuláns éppen a variancia.

80. A binomiális eloszlás várható értéke mindig nagyobb, mint a szórása.

81. Lehet-e eloszlásfüggvény az
$$F(x)=\begin{cases} 0 & \text{ha} \quad x<0 \\ \frac{2}{\pi}\arctan x & \text{ha} \quad x\geq 0 \end{cases}.$$

82. Ha ξ^2 egyenletes eloszlású a $(0,1)$ intervallumon, akkor $\mathbf{E}(\xi)<2/3$.

83. Ha ξ normális eloszlású, akkor $1/\xi$ várható értéke véges.

84. Ha ξ exponenciális eloszlású λ paraméterrel, akkor ξ momentumgeneráló függvénye $M(s)=\lambda(\lambda-s)$, ha $s<\lambda$.

85. Legyenek ξ_1 és ξ_2 független standard normális eloszlású változók. Ekkor $\xi_1^2+\xi_2^2$ exponenciális eloszlású.

86. Legyen ξ Poisson-eloszlású $\lambda=5$ paraméterrel. Ekkor ξ második momentuma 30.

87. Legyen ξ geometriai eloszlású $p=1/3$ valószínűséggel. Ekkor $\mathbf{E}(\xi)=3$.

88. Ha ξ_1 és ξ_2 független Poisson-eloszlású változók $\lambda_1 = 1$ és $\lambda_2 = 2$ paraméterekkel, akkor a $\xi_1 \xi_2$ várható értéke kisebb, mint a $\xi_1 + \xi_2$ várható értéke.

89. A Bayes-tétel szerint ha (B_n) egy teljes eseményrendszer, akkor

$$\mathbf{P}(B_n \mid A) = \frac{\mathbf{P}(A \mid B_n)\mathbf{P}(A)}{\sum_k \mathbf{P}(A \mid B_k)\mathbf{P}(B_k)}.$$

90. Ha ξ_1 és ξ_2 független standard normális eloszlású változók, akkor $|\xi_1 \xi_2|$ várható értéke kisebb, mint 2.

91. Legyenek ξ_1 és ξ_2 a $(0,1)$ intervallumon egyenletes eloszlású valószínűségi változók. Ha $\eta = \min(\xi_1, \xi_2)$, akkor $\mathbf{E}(\eta) < 1/3$.

92. Egy megyében átlag öt benzinkút van. Annak a valószínűsége, hogy egy megyében nincs benzinkút, $\exp(-5)$.

93. Legyen (ξ_1, ξ_2) egyenletes eloszlású az $x_1, x_2 \geq 0, x_1^2 + x_2^2 \leq 1$ halmazon, és legyen (η_1, η_2) egyenletes eloszlású a $x_1, x_2 \geq 0, x_1 + x_2 \leq 1$ halmazon. A ξ_1 és a ξ_2 korrelációja nagyobb, mint az η_1 és az η_2 korrelációja.

94. Ha a (ξ_n) sorozat gyengén tart egy standard normális eloszlású változóhoz, akkor a ξ_n változók F_n eloszlásfüggvényei minden pontban tartanak a standard normális eloszlás eloszlásfüggvényéhez.

95. A nagy számok erős törvénye szerint független azonos eloszlású valószínűségi változok számtani átlaga egy 0 valószínűségű esemény kimeneteleitől eltekintve egy véges vagy végtelen határértékhez tart.

96. Valószínűségi változók szorzatának várható értéke csak akkor egyezik meg a várható értékek szorzatával, ha van korrelációs együtthatójuk és a korrelációs együttható nulla.

97. A χ_n^2 eloszlás második momentuma $n^2 + 2n$.

98. Legyenek ξ és η független változók, és jelölje f és g a sűrűségfüggvényeiket. Ekkor a $\xi \cdot \eta$ szorzatnak is van sűrűségfüggvénye amely

$$h(z) = \int_{-\infty}^{\infty} f(z-x)\, g(x)\, |x|\, dx.$$

99. A Csebisev-egyenlőtlenség szerint tetszőleges ξ változóra

$$\mathbf{P}(|\xi - \mathbf{E}(\xi)| > a) \leq \frac{\mathbf{D}^2(\xi)}{a^2}.$$

100. A $\varphi(t) = \cos t$ függvény egy eloszlás karakterisztikus függvénye.

101. Van olyan eloszlás, amely karakterisztikus függvénye megegyezik a momentumgeneráló függvényével.

102. A $\varphi(t) = \sin t$ függvény egy eloszlás karakterisztikus függvénye.

103. A $\varphi(t) = \cos t + \sin t$ függvény egy eloszlás karakterisztikus függvénye.

104. A $\varphi(t) = \exp\left(-t^3\right)$ függvény egy eloszlás karakterisztikus függvénye.

105. A $\varphi(t) = \exp(-|t|)$ függvény egy eloszlás momentumgeneráló függvénye.

106. Van olyan nem azonosan nulla valószínűségi változó, amely Laplace-transzformáltja megegyezik a momentumgeneráló függvényével.

107. Ha $F(x)$ és $G(x)$ két eloszlásfüggvény, akkor

$$H(x) = F(x)G(x)$$

és

$$U(x) = \frac{F(x) + G(x)}{2}$$

is eloszlásfüggvény.

108. Ha $F(x)$ és $G(y)$ két eloszlásfüggvény, akkor

$$H(x,y) = F(x)G(y)$$

és

$$U(x,y) = \frac{F(x) + G(y)}{2}$$

is eloszlásfüggvény.

109. Létezik olyan $\xi \neq 0$ valószínűségi változó, amelyre $\mathbf{E}(\xi) = \mathbf{E}\left(\xi^2\right) = \mathbf{D}(\xi)$.

110. Ha ξ gamma eloszlású, akkor $\xi/(1+\xi)$ béta eloszlású.

111. Ha ξ és η független valószínűségi változók és ξ eloszlása exponenciális $\lambda = 1$ paraméterrel, η pedig gamma eloszlású $a = 2$ és $\lambda = 2$ paraméterekkel, akkor a $\xi + \eta$ varianciája kisebb mint 1.

112. Ha ξ és η független Poisson-eloszlású változók, akkor $\xi\eta$ is Poisson-eloszlású.

113. Ha ξ és η független Poisson-eloszlású változók, akkor $(\xi + \eta)/2$ is Poisson-eloszlású.

114. Ha ξ és η együttes sűrűségfüggvénye $f(x,y)$, akkor $\xi\eta$ sűrűségfüggvénye

$$h(x) = \int_{-\infty}^{\infty} f(x/y, y)\,\frac{1}{|y|}\,dy.$$

115. Ha ξ exponenciális eloszlású $\lambda = 1$ paraméterrel, akkor ξ^2 eloszlásfüggvénye az $x = 1$ helyen kisebb mint $1/2$.
116. Ha ξ és η független gamma eloszlású változók, akkor $\xi + \eta$ is gamma eloszlású.
117. Ha valamely változóra $\mathbf{E}\left(\xi^2\right) \leq \mathbf{E}^2(\xi)$, akkor ξ konstans.
118. $\Gamma(3/2) = \sqrt{\pi}/2$.
119. Ha $B(x,y)$ a bétafüggvény, akkor
$$B(x,y) = \frac{\Gamma(x+y)}{\Gamma(y)\Gamma(x)}.$$
120. Egy normális eloszlású változó Laplace-transzformáltja megegyezik a momentumgeneráló függvényével.
121. Van olyan eloszlás, amely momentumgeneráló függvénye $M(s) = \exp\left(-s^2/2\right)$.
122. Ha ξ binomiális eloszlású n és p paraméterekkel, akkor $\eta = -2\xi$ szórása $2\sqrt{npq}$.
123. Ha ξ és η függetlenek, akkor ξ^2 és η^3 korrelálatlan.
124. $M(s) = 1/\left(1+s^2\right)$ egy eloszlás momentumgeneráló függvénye.
125. $\varphi(t) = 1/\left(1+t^2\right)$ egy eloszlás karakterisztikus függvénye.
126. $\int_0^{2\pi} \cos x d \cos x = 0$.
127. A binomiális eloszlás korlátlanul osztható.
128. Ha ξ gamma eloszlású $\alpha > 0$ és $\beta > 0$ paraméterekkel, és $\lambda > 0$, akkor $\lambda\xi$ szintén gamma eloszlású α/λ és β/λ paraméterekkel.
129. A negatív binomiális eloszlás várható értéke mindig kisebb mint a varianciája.
130. Van olyan összetett Poisson-eloszlás, amely varianciája kisebb, mint a várható értéke.
131. Van olyan egész értékeket felvevő összetett Poisson-eloszlás, amely varianciája kisebb mint a várható értéke.
132. Van olyan összetett Poisson-eloszlás, amely nem korlátlanul osztható.
133. Két független korlátlanul osztható eloszlás összege is korlátlanul osztható.
134. A Moivre–Laplace-képletben a konvergencia sebessége $o(1/n)$.
135. A Moivre–Laplace-képletben a konvergencia sebessége $o(1/\sqrt{n})$.

136. Ha ξ egyenletes eloszlású a $[0,1]$ szakaszon, akkor $\exp(\xi)$ várható értéke nagyobb, mint $3/2$.

137. Ha ξ egyenletes eloszlású a $[0,1]$ szakaszon, akkor $\exp(-\xi)$ várható értéke nagyobb, mint $1/2$.

138. Ha $\sqrt{\xi}$ egyenletes eloszlású a $[0,1]$ szakaszon, akkor ξ^2 várható értéke nagyobb, mint $1/2$.

139. Ha $1/\xi$ egyenletes eloszlású a $[0,1]$ szakaszon, akkor ξ^2 várható értéke nagyobb, mint 1.

140. Ha ξ_1 és ξ_2 függetlenek, és az eloszlásfüggvényük F_1 és F_2, akkor $\eta \circeq \max(\xi_1,\xi_2)$ eloszlásfüggvénye F_1F_2.

141. Ha ξ_1 és ξ_2 függetlenek, és az eloszlásfüggvényük F_1 és F_2, akkor $\eta \circeq \min(\xi_1,\xi_2)$ eloszlásfüggvénye $(1-F_1)(1-F_2)$.

142. Ha ξ gamma eloszlású (a,λ) paraméterekkel, akkor a ξ Laplace-transzformáltja

$$L(s)=\left(1-\frac{t}{a}\right)^{-\lambda}.$$

143. A ξ valószínűségi változót szimmetrikus gamma eloszlásúnak mondjuk (a,λ) paraméterekkel, ha felírható két független (a,λ) paraméterű gamma eloszlású változó különbségeként, vagyis $\xi=\xi_1-\xi_2$, ahol ξ_1 és ξ_2 független (a,λ) paraméterű gamma eloszlású valószínűségi változók. Ha ξ szimmetrikus gamma eloszlású, akkor karakterisztikus függvénye $\varphi(t)=\left(\lambda^2/\left(\lambda^2+t^2\right)\right)^a$.

144. A ξ változót inverz gamma eloszlásúnak mondjuk, ha $1/\xi$ gamma eloszlású. Az inverz gamma eloszlásnak mindig van várható értéke.

145. Az $F(x,y)=\sqrt{x^2+y^2-1}$ ha $x,y\in[0,1]$ egy kopula eloszlásfüggvénye.

146. Ha ξ_1,ξ_2,ξ_3 és ξ_4 független $\lambda=1$ paraméterű exponenciális eloszlású változók és $\xi_1^*\le\xi_2^*\le\xi_3^*\le\xi_4^*$ a rendezett elemek, akkor a ξ_2^* szórása nagyobb, mint $1/2$.

147. A normális eloszlás lapultsága 3.

148. A λ paraméterű Poisson-eloszlás negyedik centrális momentuma $\lambda+3\lambda^2$.

149. Két független, azonos paraméterű exponenciális eloszlású valószínűségi változó hányadosa másodfajú béta eloszlást követ.

150. Ha ξ lognormális eloszlású és $\mathbf{E}(\xi)\le 1$, akkor $\mathbf{E}(\log\xi)\le 0$.

Függelékek

A.

NÉHÁNY TOVÁBBI MEGJEGYZÉS

Egy ilyen jegyzet kapcsán felmerülhet a kérdés, minek? Minek ír valaki egy újabb jegyzetet a valószínűségszámítás témakörében, amikor korábban már tízezer számra írtak hasonló jegyzeteket. Ennek nagyon egyszerű oka van. Mivel a témát tanítani kell, a hallgatók számára a legkényelmesebb, ha van egy írott anyag, amiből fel lehet készülni. De az oktató számára is ez a legegyszerűbb. Itt van, ezt kell tudni, se többet se kevesebbet. De ez nem válaszolja meg a kérdést. Minek? Miért nem választunk egy már meglevőt, és miért nem azt használjuk? A válasz tulajdonképpen jóval bonyolultabb, mint aminek gondolnánk. Az elemi valószínűségszámítás oktatásának legfőbb gondja a következő: A valószínűségszámítás axiómái lényegében annak deklarálásából állnak, hogy a tárgykör nem érthető meg a mértékelmélet ismerete nélkül. A valószínűségszámítás Kolmogorov-féle axiómái pontosan azt állítják, hogy az $(\Omega, \mathscr{A}, \mathbf{P})$ egy olyan mértéktér, amelyben az Ω alaphalmaz mértéke 1, vagyis az $(\Omega, \mathscr{A}, \mathbf{P})$ egy 1-re normált mértéktér. Később a valószínűségi változók bevezetésekor pontosan azt mondjuk, hogy a valószínűségi változók mérhető függvények, pontosabban azok nulla valószínűségű események erejéig vett ekvivalenciaosztályai, a várható érték pedig a változó $\mathbf{P}$ mérték szerint vett integrálja. A mértékelmélet nem tartozik a matematika legnehezebb fejezetei közé, de mégis némi matematikai érettséget kíván. Ez a bevezető valószínűségszámítási kurzusok során általában, inkább mindig, a hallgatók esetében nincs meg. Különösen nem várható el az alapszintű közgazdász hallgatóktól. Akkor mégis mit lehet csinálni? Hogyan lehet megvalósítani a lehetetlent? A bevezető valószínűségszámítás egész oktatása világszerte fából vaskarika. Minden, amit csinálunk, szembemegy az axiómákkal. Az általános eljárás az, hogy az elemi valószínűségszámítási kurzusok során túlzottan nagy hangsúlyt helyeznek az olyan feladatok megoldására, amelyeket elemi eszközökkel meg lehet oldani. Speciálisan az elemi kombinatorikára, illetve azokra a feladatokra, amelyek megoldása időnként nem csekély trükkös, vagy körmönfont gondolatmenetet igényelnek. Ezzel némiképpen, talán nem is némiképpen, hanem igencsak becsapják a hallgatót. A jegyzetben és az általam tanított kurzus során megpróbálok szakítani evvel a gyakorlattal és azokra a tételekre koncentrálni, amelyeket tényleg fontosnak tartok. Ennél sokkal, de sokkal többet kellene megtanítani, de hát ennyi fér bele tizennégy hétbe. Annak ellenére, hogy matematikai szempontból a tárgyalás nem hézagmentes, azt gondolom, hogy egy sor fontos tételt bemutatok. Másoldalról a mértékelmélet oktatásának legfőbb nehézsége, hogy a hallgatók nem értik, hogy miért van szükség a tárgykörre. Az általános vélemény, hogy szép-szép dolog a mértékelmélet, de mégis minek? Amit én a kurzus tanítása során megértettem, hogy minden igyekezett ellenére a mértékelmélet nélkül a legegyszerűbb állításokat sem lehet igazolni. Természetesen, hiszen a kérdés már az axiómák szintjén eldőlt. Minden elemi erőfeszítés csak félrevezető módon replikálhatja az általános esetet. A leggyakrabban használt valószínűségszámítási állítás a várható érték additivitása, de igazolása a megközelítés elemi szintjén lényegében lehetetlen. Lehet mellébeszélni, álruhában belátni az absztrakt tételt, be lehet látni diszkrét esetben, illetve ha van folytonos sűrűségfüggvény, de ha jobban megnézzük ezeket a bizonyításokat, akkor azok valójában túlzottan speciálisak és ami még fontosabb: a szükségesnél bonyolultabbak. Ha az érdeklődő és figyelmes hallgató ezt megértette, akkor talán segítettem a mértékelmélet elfogadását. A matematika tanulása nem egy lineáris folyamat. Maga a matematika, még ha így is próbáljuk gyakran eladni, nem egy lineáris ismeretanyag, hanem egy háló, amelyben minden mindennel összefügg. Az, hogy az axiómákból indulunk ki, az nem egy szükségszerű, hanem egy kényelmi lépés. Mint minden kényelmi eszközt, ezt is addig kell használni, amíg hasznos. A lineáris haladás helyett a csigalépcső a jobb hasonlat. Mindig ugyanazt tanuljuk, de egyre jobban körüljárjuk a szükséges ismereteket. A csigalépcső első fordulója az elemi analízis, a második az elemi valószínűségszámítás, majd jönnek a további lépcsőfokok. Ki-ki annyi lépcsősort mászik meg, amennyire ideje, kedve és lehetősége van. A valószínűségszámítás tárgyalása során a mértékelmélet tételeit több helyen is használni kell. A matematika tételei, legalábbis a valószínűségszámításban, két részre oszthatók. Egyrészt vannak az elvi kérdések, másrészt vannak a konkrét számítások, konkrét képletek. A valószínűségszámításban az elvi kérdések közül szinte semmi, vagy inkább egyáltalán semmi sem igazolható mértékelméleti ismeretek nélkül. De gyakran hasonló a helyzet a konkrét számolások során is. Ennek megfelelően a jegyzet helyenként pontatlan vagy inkább hézagos. Választani kell: Vagy pontosak vagyunk és zavaró módon barkácsolunk, vagy a valódi bizonyítást mutatjuk be, de elnagyoljuk a gondolatmenetet. Hogy mégis némi segítséget nyújtsak, néhány gyakran használt tételt itt felsorolok. Az állítások igen speciálisak és csak folytonos függvényekre és Stieltjes-integrálokra mondom ki őket, de ezek segítségével a könyvben szereplő számolások nagyrészt tisztázhatók és pontosíthatók. A tételek mindegyike az integrálás és a határérték felcserélésére vonatkozik. A dolog iróniája az, hogy a mértékelmélet absztrakt fogalmainak legfőbb haszna, hogy ezeket a tételeket be lehet a segítségükkel bizonyítani. Vagyis az absztrakció legfőbb haszna a gyakran használt eszközök egyszerű bizonyítása. Nyomatékosan hangsúlyozni kell, hogy általában az állítások jóval absztraktabb körülmények között is teljesülnek, de az anyag megértése szempontjából elegendő, ha minden olyan függvény, amely az integrálok alatt áll folytonos, a súlyfüggvények monotonok, balról folytonosak és az integrálokat abszolút konvergens improprius Stieltjes-integráloknak kell elgondolni. Érdemes hangsúlyozni, hogy a Stieltjes-integrálok kiszámolásakor a súlyfüggvény értéke az integrációs intervallum belsejében levő szakadási pontokban érdektelen. Ez nyilván nem vonatkozik az intervallumok végpontjaira. Ez a tény különösen nehézzé teszi a véges pontokban vett improprius integrálok kiszámolását. Ennek megfelelően a súlyfüggvényeket regularizálni kell, ami alatt azt értjük, hogy vagy balról vagy jobbról folytonossá kell tenni őket[1]. Valószínűságszámítási környezetben balról folytonos súlyfüggvényekkel szokás dolgozni, de például a sztochasztikus folyamatok elméletében a jobbról folytonos regularizálást szokás feltételezni.

A.0.1. Tétel (Monoton konvergencia tétele). *Ha $0 \leq f_n \nearrow f$ pontonként, akkor*

$$\lim_{n \to \infty} \int_{\mathbb{R}} f_n(x)\, dF(x) = \int_{\mathbb{R}} \lim_{n \to \infty} f_n(x)\, dF(x) = \int_{\mathbb{R}} f(x)\, dF(x).$$

[1] A mértékelméletben ez úgy jelenik meg, hogy a súlyfüggvénynek az intervallumokon mértéknek kell lenni. A mértékelmélet megértésének egyik kulcsa annak megértése, hogy a Stieltjes-integrál milyen korlátokkal bír.

Az állításban az $\mathbb{R}$ helyébe tetszőleges I intervallum is írható. A határérték, illetve a jobb oldali integrál egyidejüleg lehet végtelen is.

A.0.2. Tétel (Majorált konvergencia tétele)**.** *Ha $f_n \to f$ pontonként, valamint létezik olyan g, amelyre*

$$|f_n| \leq g$$

és

$$\int_{\mathbb{R}} g(x)\, dF(x) < \infty,$$

akkor

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)\, dF(x) = \int_{\mathbb{R}} \lim_{n\to\infty} f_n(x)\, dF(x) = \int_{\mathbb{R}} f(x)\, dF(x).$$

Az állításban az $\mathbb{R}$ helyébe tetszőleges I intervallum is írható.

A.0.3. Tétel (Fubini tétele)**.** *Ha $f(x,y) \geq 0$, akkor*

$$\int_{\mathbf{R}} \int_{\mathbf{R}} f(x,y)\, dF(x)\, dG(y) = \int_{\mathbf{R}} \int_{\mathbf{R}} f(x,y)\, dG(y)\, dF(x).$$

Ilyenkor a két oldal lehet végtelen is. Ha $f(x,y)$-ra a nem negativitási feltétel nem teljesül, de $|f(x,y)|$ kettős integrálja véges, akkor az integrál továbbra is felcserélhető. Ilyenkor a két oldalon biztosan véges értékek állnak. Az állításban az $\mathbb{R}$ helyébe tetszőleges I és J intervallumok is írhatók.

A.0.4. Tétel (Integrál mögé deriválás)**.** *Ha $f(x,y)$ az x változó szerint deriválható, valamint létezik olyan $g(y)$ függvény, amelyre $\int_{\mathbb{R}} g(y)\, dF(y) < \infty$, és*

$$\left| \frac{\partial f}{\partial x}(x,y) \right| \leq g(y),$$

akkor

$$\frac{d}{dx} \int_{\mathbf{R}} f(x,y)\, dF(y) = \int_{\mathbf{R}} \frac{\partial}{\partial x} f(x,y)\, dF(y).$$

Az állításban az $\mathbb{R}$ helyébe tetszőleges I intervallum is írható.

A mértékelmélet azonban nem csak néhány absztrakt elvi kérdés és a fenti tételek, illetve számolási szabályok bizonyításához szükséges. A Stieltjes-integrálás igen hasznos és egyszerű fogalom, de nem folytonos függvényeket nem tudunk segítségével integrálni. Így például általában nem tudjuk integrálni az intervallumok χ_I indikátorfüggvényeit sem. Nyilvánvaló módon ha F egy eloszlásfüggvény, akkor $\int_a^b 1 dF(x) = F(b) - F(a)$, amely az $[a,b)$ szakaszba esés valószínűsége. Ugyanakkor az $\mathbf{E}(\chi_A) = \mathbf{P}(A)$ képletnek megfelelő

$$\int_{-\infty}^{\infty} \chi_{[a,b)}(x)\, dF(x) = F(b) - F(a)$$

képlet nem feltétlenül igaz, ugyanis a Stieltjes-integrál adott esetben nem lesz feltétlenül értelmes. Ezért az integrálást ki kell terjeszteni oly módon, hogy az integrandus nem folytonos függvény is lehessen.

A.0.5. Példa. Momentumok és deriváltak.

A generátorfüggvények tárgyalásakor említettük, hogy a $G(z)$ generátorfüggvény a $[0,1)$ intervallumon deriválható, és ha az eloszlásnak van várható értéke, akkor a $G'(1)$ balról vett derivált létezik és az értéke éppen a $\Sigma_n n p_n$ várható érték. Hasonló állítás igaz a magasabb faktoriális momentumokra. Érdemes megjegyezni, hogy analóg tétel igaz a nem negatív változók momentumgeneráló függvényére és Laplace-transzformáltjára. A Laplace-transzformációs jelölést használva és felhasználva az integrál alatt való deriválhatóságról szóló tételt, ha egy nem negatív változónak létezik az n-edik momentuma, akkor

$$L^{(n)}(0) = (-1)^n \int_0^\infty x^n dF(x),$$

ahol a nulla pontban vett deriváltat a Laplace-transzformált esetén jobbról kell venni. Vegyük ugyanis észre, hogy az $\exp(-sx)$ függvény s változó szerinti n-edik parciális deriváltja

$$\frac{\partial^n}{\partial s^n} \exp(-sx) = (-1)^n x^n \exp(-sx).$$

Mivel a változó nem negatív, ezért csak az $x \geq 0$ esettel kell foglalkozni és ilyenkor ha $s \geq 0$, akkor

$$\left| \frac{\partial^n}{\partial s^n} \exp(-sx) \right| = x^n \exp(-sx) \leq x^n.$$

Mivel a momentum létezik, ezért az $\int_0^\infty x^n dF(x)$ integrál véges, vagyis a $g(x) = x^n$ függvény a parciális deriváltnak integrálható majoránsa, és miként említettük, ilyenkor alkalmazható az integrál mögött való deriválásról szóló szabály. Vegyük továbbá észre, hogy tetszőleges $\varepsilon > 0$ esetén van olyan K, hogy minden $x \geq 0$ esetén

$$x^n \leq K \exp(\varepsilon x).$$

Ha $s > 0$, $\varepsilon \overset{\circ}{=} s/2$ akkor, felhasználva, hogy $x \geq 0$,

$$\begin{aligned} x^n \exp(-sx) &\leq K \exp\left(\frac{s}{2}x\right) \exp(-sx) = \\ &= K \exp\left(-\frac{s}{2}x\right) \leq K, \end{aligned}$$

vagyis a parciális derivált egyenletesen korlátos, ezért ha $s > 0$, akkor a deriválás az F szerinti integrál alatt mindig elvégezhető. Vagyis ha $\xi \geq 0$, akkor az $L(s)$ minden $s > 0$ esetén végtelen sokszor deriválható, és érvényes rá az

$$L^{(n)}(s) = (-1)^n \int_0^\infty x^n \exp(-sx) dF(x)$$

képlet.

□

A.0.6. Példa. Abel és Tauber tétele, valamint a momentumok létezése.

A $G(z)$ generátorfüggvény a $[0,1)$ intervallumon deriválható, és ha az eloszlásnak van várható értéke, akkor a tagonként vett deriváltakból álló hatványsor a teljes $[0,1]$ szakaszon egyenletesen konvergens, így a $\lim_{z\nearrow 1} G'(z) \overset{\circ}{=} A$ határérték létezik, és az értéke éppen a $\Sigma_{n=0}^{\infty} np_n$ várható érték. Ezt szokás nevezni Abel-tételnek[2]. Hasonló állítás igaz a magasabb faktoriális momentumokra. Felmerül azonban a kérdés: Mi van akkor, ha csak azt tudjuk, hogy a fenti A határérték létezik? Ilyenkor létezik-e a várható érték, amely ez esetben, Abel-tétele alapján, szükségszerűen éppen az A határérték lesz? Az Abel-típusú tételek ilyen fajta megfordításait szokás Tauber-féle tételeknek mondani. A Tauber-típusú tételek irodalma kötegekre rúg és a válasz általában nemleges[3]. Abból, hogy egy hatványsor a konvergenciakörének egy pontjára folytonosan kiterjeszthető, még nem biztos, hogy ott a hatványsor konvergens. Generátorfüggvény esetén, felhasználva, hogy az alábbi (a_n) együtthatók nem negatívak, a monoton konvergencia tétel miatt

$$\begin{aligned}
\lim_{z\nearrow 1} G^{(m)}(z) &= \lim_{z\nearrow 1}\sum_{n=0}^{\infty}(n)_m p_n z^{n-m} \overset{\circ}{=} \lim_{z\nearrow 1}\sum_{n=0}^{\infty} a_n z^n = \\
&= \lim_{z\nearrow 1}\int_0^{\infty} z^n dA(n) = \int_0^{\infty}\lim_{z\nearrow 1} z^n dA(n) = \int_0^{\infty} 1 dA(n) = \\
&= \sum_{n=0}^{\infty} a_n \overset{\circ}{=} \sum_{n=0}^{\infty}(n)_m p_n.
\end{aligned}$$

Itt a két oldal egyszerre véges vagy végtelen és $A(n)$ értelemszerűen az (a_n) sorozat monoton eloszlásfüggvénye. Érdemes megjegyezni, hogy hasonló tétel igaz a nem negatív változók Laplace-transzformációjára és momentumgeneráló függvényére. A Laplace-transzformációs jelölést használva, és felhasználva az integrál alatt való deriválhatóságról szóló tételt, amely alapján minden $s > 0$ esetén a deriválás az integrál mögött elvégezhető,

$$\begin{aligned}
\lim_{s\searrow 0} L^{(n)}(s) &= (-1)^n \lim_{s\searrow 0}\int_0^{\infty} x^n \exp(-sx)\, dF(x) = \\
&= (-1)^n \int_0^{\infty}\lim_{s\searrow 0} x^n \exp(-sx)\, dF(x) = \\
&= (-1)^n \int_0^{\infty} x^n dF(x).
\end{aligned}$$

Következésképpen, ha nem negatív változók esetén $L^{(n)}(s)$ folytonosan kiterjeszthető az $s = 0$ pontra, akkor létezik az n-edik momentum. Az egyenlőségsort fordítva olvas-

[2]Emlékeztetünk, hogy Abel tétele szerint ha egy a valós számokon értelmezett hatványsor a konvergenciatartomány határán is konvergens, akkor a hatványsor összege, beleértve ezt a végpontot is, folytonos. (Az, hogy az összeg a konvergenciasugáron belül folytonos, az nyilvánvaló. Amit igazolni kell, az az, hogy a végpontban is folytonos lesz. A nehézség abból áll, hogy a végpontban a konvergencia lehet feltételes. Ez a generátorfüggvény esetén nem áll fent, így ilyenkor az indoklás valamivel egyszerűbb.)

[3]Gondoljunk a $\Sigma_{n=0}^{\infty}(-1)^n x^n = 1/(1+x)$ függvényre, amelyre a határérték létezik az $x = 1$ helyen, de a $\Sigma_{n=0}^{\infty}(-1)^n$ összeg nem létezik.

va, ha létezik az n-edik momentum, akkor az $L^{(n)}(s)$ folytonosan kiterjeszthető az $s=0$ pontra. Ennél erősebb állítás nem várható, ugyanis például a lognormális eloszlásnak létezik az összes momentuma, de a Laplace-transzformáltja csak az $s \geq 0$ halmazon konvergens. Végezetül tegyük fel, hogy a Laplace-transzformált jobbról deriválható a nulla pontban. Mivel L első deriváltja negatív, a második deriváltja pedig pozitív, ezért L egy csökkenő konvex függvény. Mivel a harmadik derivált ismét negatív, ezért az $L'(s)$ derivált egy növekedő, konkáv függvény. Konvex függvények deriváltjai nőnek, ezért az $L'(0)$ az $L'(s)$ egy alsó korlátja. Ha $s \searrow 0$, akkor az $L'(s)$ szintén csökken, így az $\lim_{s\searrow 0} L'(s)$ pontosan akkor véges, ha az $L'(0)$ is véges. Vagyis nem negatív változók esetén pontosan akkor létezik a várható érték, ha a Laplace-transzformált a nulla pontban jobbról deriválható. Analóg módon kezelhetőek a magasabb momentumok. Érdemes megjegyezni, hogy amennyiben a változókra nem teljesül a nem negativitás, akkor hasonló állítás a karakterisztikus függvények esetében nem igazolható. Karakterisztikus függvények esetében csak a páros számú deriváltak létezéséből következik a páros számú momentum létezése, ezzel szemben páratlanokra, így a várható értékre az összefüggés nem érvényes.

□

A.0.7. Példa. Független valószínűségi változók összegének várható értéke.

A várható érték additivitása kapcsán érdemes hangsúlyozni, hogy az

$$\mathbf{E}(\xi+\eta)=\mathbf{E}(\xi)+\mathbf{E}(\eta)$$

egyenlőség valójában jobbról balra olvasandó. Ha létezik a ξ és az η várható értéke, akkor a $\xi+\eta$ várható értéke is létezik és értéke a külön vett várható értékek összege. A másik irány általában nem igaz. A legegyszerűbb példa, ha ξ egy Cauchy-eloszlású változó és $\eta=-\xi$. Ilyenkor $\xi+\eta=0$, vagyis $\mathbf{E}(\xi+\eta)=0$, de az $\mathbf{E}(\xi)+\mathbf{E}(\eta)$ értelmetlen. Vagyis a várható érték additivitása azt jelenti, hogy a várható értékeket össze lehet vonni, de csak akkor lehet szétszedni, ha tudjuk, hogy a szétszedett változók várható értéke értelmes. Persze, ha $\xi \geq 0$ és $\eta \geq 0$, akkor az egyenlőség két oldala egyszerre véges, vagy végtelen, de ez az általános esetben nem sokat segít. Miként láttuk, a Fubini-tétel az egyik leghasznosabb számítási eszköz. Ennek segítségével belátjuk a következő igen hasznos egyenlőséget: Ha a ξ és az η függetlenek és a $\xi+\eta$ várható értéke véges, akkor a ξ és az η várható értéke is véges. Valóban: az általános a transzformált valószínűségi változók várható értékére vonatkozó formula alapján

$$\mathbf{E}(g(\xi_1,\xi_2,\ldots,\xi_n))=\int_{\mathbb{R}^n} g(x_1,x_2,\ldots x_n)\,dF(x_1,\ldots,x_n),$$

ahol F az együttes eloszlásfüggvény és az integrál az általunk nem tárgyalt többdimenziós integrál. Ilyenkor a fenti képlet azonosság, vagyis a két oldal egyszerre létezik vagy nem létezik. Az összeg esetén

$$\mathbf{E}(\xi+\eta)=\int_{\mathbb{R}^2} x+y\,dF(x,y).$$

Ha a megfelelő peremeloszlásfüggvények H és G, akkor, felhasználva, hogy a függetlenség miatt az F együttes eloszlásfüggvény a peremeloszlások szorzata, és hogy Fubini-tétele miatt a kettős integrál kétszeres integrálként írható:

$$\mathbf{E}(\xi+\eta)=\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}(x+y)\,dH(x)\right)dG(y).$$

Mivel a kettős integrál véges, ezért a belső integrál is, legalább egy y pontban véges, vagyis van olyan y, hogy az

$$\int_{-\infty}^{\infty}x+ydH(x)=\int_{-\infty}^{\infty}xdH(x)+y$$

egyenlőség fennáll. Ezért az $\mathbf{E}(\xi)$ létezik, és minden y-ra

$$\int_{-\infty}^{\infty}x+ydH(x)=\mathbf{E}(\xi)+y.$$

Innen pedig

$$\begin{aligned}\mathbf{E}(\xi+\eta) &= \int_{-\infty}^{\infty}\mathbf{E}(\xi)+ydG(y)= \\ &= \mathbf{E}(\xi)+\int_{-\infty}^{\infty}ydG(y),\end{aligned}$$

amiből az $\mathbf{E}(\eta)$ létezése is következik.

$\square$

B.

ALAPESZKÖZÖK

Ebben a függelékben néhány a valószínűségszámítás alapjait érintő tételt mutatok be. A mértékelmélet ismeretét feltételezem, így a mértékelméleti tételek tárgyalásától eltekintek. Az itt leírtak értelemszerűen nem részei a bevezető valószínűségszámítás tananyagnak.

B.1. Alapfogalmak

B.1.1. Definíció. Valamely $\mathscr{R}$ halmazrendszert algebrának szokás mondani, ha az $\mathscr{R}$ zárt a három halmazműveletre, vagyis ha $R \in \mathscr{R}$, akkor $R^c \in \mathscr{R}$, és ha $R_1, R_2 \in \mathscr{R}$, akkor $R_1 \cup R_2 \in \mathscr{R}$, és ilyenkor persze az $R_1 \cap R_2 \in \mathscr{R}$ is teljesül. A valószínűségszámítás axiómái szerint a lehetséges események $\mathscr{A}$ halmaza nem pusztán algebra, hanem úgynevezett σ-algebra, ahol a σ jelző arra utal, hogy az egyesítés és a metszet művelete megszámlálható számosságú halmazrendszerek esetén is végrehajtható. Ha adott egy X alaphalmaz, és egy az X részhalmazaiból álló $\mathscr{F}$ σ-algebra, akkor az $(X, \mathscr{F})$ párost mérhető térnek, valószínűségszámítási környezetben eseménytérnek szokás mondani.

B.1.2. Definíció. Ha $(X, \mathscr{F})$ és $(Y, \mathscr{G})$ két mérhető tér, akkor egy $f : X \to Y$ függvényt mérhetőnek mondunk, ha minden $G \in \mathscr{G}$ esetén $f^{-1}(G) \in \mathscr{F}$. Speciálisan ha $Y = \mathbb{R}$ és nem teszünk semmi további megkötést, akkor a $\mathscr{G}$ a nyílt halmazok által generált σ-algebra, vagyis az úgynevezett Borel-halmazok családja. Miként ismert, ilyenkor a mérhető függvények megegyeznek a lépcsős függvények pontonkénti határértékeként előálló függvényekkel. Ugyancsak ismert, hogy a nem negatív függvények esetén a közelítő lépcsős függvények sorozata választható nem negatívnak és monoton növekedőnek.

B.1.3. Definíció. Ha adott egy $(X, \mathscr{F})$ mérhető tér, akkor az $\mathscr{F}$ halmazrendszer $F \in \mathscr{F}$ elemein értelmezett $\mu(F)$ halmazfüggvényt mértéknek mondjuk, ha

1. minden $F \in \mathscr{F}$ halmaz esetén $\mu(F) \geq 0$, megengedve az esetleges $\mu(F) = +\infty$ értéket is,
2. a μ σ-additív, vagyis ha az (F_n) megszámlálható darab $\mathscr{F}$-ből vett halmazból álló páronként diszjunkt halmazrendszer, akkor
$$\mu(\cup_n F_n) = \sum_n \mu(F_n),$$
3. illetve a trivialitások elkerülése véget $\mu(\emptyset) = 0$.

$(X, \mathscr{F}, \mu)$ hármast szokás mértéktérnek mondani. Következésképpen minden $(\Omega, \mathscr{A}, \mathbf{P})$ valószínűségi mező egy egyre normált mértéktér.

B.1.4. Definíció. Az $(\Omega, \mathscr{A}, \mathbf{P})$ valószínűségi mezőn értelmezett mérhető valós függvényeket valószínűségi változónak mondjuk. Két valószínűségi változót ekvivalensnek mondunk, ha csak a **P** mérték szerint nullmértékű halmazon különböznek. Vagyis a valószínűségi változók mérhető függvények ekvivalenciaosztályai. A valószínűségi változók **P** szerinti integráljait várható értéknek mondjuk, vagyis ha ξ egy valószínűségi változó, akkor definíció szerint

$$\mathbf{E}(\xi) \mathrel{\stackrel{\circ}{=}} \int_{\Omega} \xi \, d\mathbf{P}.$$

Miként a mértékelméletből ismert, az integrál lineáris operáció, így a definíció fontos következménye, hogy a várható érték lineáris művelet, vagyis például $\mathbf{E}(\xi+\eta) = \mathbf{E}(\xi)+\mathbf{E}(\eta)$ feltéve, hogy mind a két oldal értelmes. A valószínűségszámítás központi fogalma az eloszlás:

B.1.5. Definíció. Ha $(X, \mathscr{F}, \mu)$ egy mértéktér és $(Y, \mathscr{G})$ egy mérhető tér valamint $f : X \to Y$, egy mérhető függvény, akkor a

$$\nu(G) \mathrel{\stackrel{\circ}{=}} \mu\left(f^{-1}(G)\right), \quad G \in \mathscr{G}$$

módon definiált halmazfüggvényt az f eloszlásának mondjuk. Annak igazolását, hogy az $(Y, \mathscr{G}, \nu)$ egy mértéktér az olvasó könnyen ellenőrizheti.

A transzformált valószínűségi változók várható értékére vonatkozó formula a következő állítás speciális esete.

B.1.6. Állítás (Helyettesítéses integrálás formulája)**.** *Az előző definíció jelölését használva, minden $h : (Y, \mathscr{G}) \to \mathbb{R}$ mérhető függvény esetén*

$$\int_X h(f(x))\, d\mu(x) = \int_Y h(y)\, d\nu(y),$$

ahol a két oldal egyidejüleg véges, végtelen, illetve értelmes vagy értelmetlen. Speciálisan ha $h(y) = g(y)\chi_C(y)$, akkor

$$\int_{f^{-1}(C)} g(f(x))\, d\mu(x) = \int_C g(y)\, d\nu(y). \tag{B.1.1}$$

Bizonyítás: A bizonyítás a mértékelméletben megszokott módon végezhető el. Először legyen $h(y) = \chi_G$, ahol $G \in \mathscr{G}$. Ekkor $h(f(x)) = \chi_{f^{-1}(G)}$ és

$$\begin{aligned} \int_X h(f(x))\, d\mu(x) &= \int_X \chi_{f^{-1}(G)} d\mu(x) = \\ &= \mu\left(f^{-1}(G)\right) = \nu(G) = \int_Y h(y)\, d\nu(y). \end{aligned}$$

Az integrál additivitása miatt, ha $h(y) = \Sigma_k c_k \chi_{G_k}$ alakú lépcsős függvény, akkor az egyenlőség továbbra is teljesülni fog. Ha $h \geq 0$, akkor a h nem negatív lépcsős függvények monoton határértéke és így a monoton konvergencia tétele miatt az egyenlőség érvényben marad. Az általános eset az integrál definíciója miatt evidens, ugyanis az integrál definíció szerint az integrandus pozitív és negatív részek integráljának különbsége.

□

B.2. A monoton osztály tétele

A mérhető függvényekre vonatkozó legfontosabb tétel a monoton osztály tétele. Ennek kimondásához szükséges néhány fogalom bevezetése:

B.2.1. Definíció. A valós értékű függvényekből álló $\mathscr{P}$ családot π-osztálynak mondjuk, ha $f_1, f_2 \in \mathscr{P}$-ből következik, hogy $f_1 f_2 \in \mathscr{P}$. Az ugyancsak valós függvényekből álló $\mathscr{L}$ családot λ-rendszernek mondjuk, ha $\mathscr{L}$ olyan lineáris tér, amely zárt a monoton konvergenciára nézve, vagyis ha $0 \leq f_n \nearrow f$ pontonként és minden n indexre $f_n \in \mathscr{L}$, akkor az $f \in \mathscr{L}$ is teljesül.

B.2.2. Állítás (Monoton osztály tétele)**.** *Legyen $\mathscr{P}$ egy π-rendszer. Ha $\mathscr{L}$ egy korlátos függvényekből álló λ-rendszer és $\mathscr{P} \subset \mathscr{L}$, akkor az $\mathscr{L}$ tartalmazza a $\mathscr{P}$ által generált legszűkebb σ-algebra szerint mérhető korlátos függvények halmazát.*

Az állítás számtalan következménnyel rendelkezik. Ezek közül az egyik legfontosabb a következő:

B.2.3. Állítás (Egyértelműségi tétel)**.** *Minden eloszlást egyértelműen jellemez a karakterisztikus függvénye. Ha az eloszlás a nem negatív számokra koncentrálódik, akkor az eloszlást a Laplace-transzformáltja is egyértelműen meghatározza.*

Bizonyítás: Ha F_1 és F_2 két eloszlás, amely karakterisztikus függvénye megegyezik, akkor az olyan korlátos és mérhető függvények halmaza, amelyre

$$\int_{\mathbb{R}} u(x)\, dF_1(x) = \int_{\mathbb{R}} u(x)\, dF_2(x)$$

tartalmazza a $\sin tx$ és $\cos tx$ trigonometrikus függvényeket. Az integrál linearitása és a monoton konvergencia tétel miatt az ilyen u függvények halmaza λ-rendszert alkot. Mivel a trigonometrikus polinomok halmaza π-rendszer, ezért alkalmazható a monoton osztály tétel. A tétel miatt az $u(x)$ helyébe beírható az összes olyan G halmaz χ_G indikátorfüggvénye amelyre G eleme a trigonometrikus polinomok által generált σ-algebrának. Elegendő azt megmutatni, hogy ez a halmaz tartalmazza a Borel-mérhető halmazokat. Ha

$$f_n(x) \overset{\circ}{=} n \sin x/n,$$

akkor f_n nyilván eleme a keresett mérhető függvények családjának, így az adott családra nézve mérhető a

$$\lim_{n\to\infty} f_n(x) = x \lim_{n\to\infty} \frac{\sin x/n}{x/n} = x$$

függvény is, amely által generált σ-algebra triviálisan éppen a Borel-mérhető függvények halmaza. Hasonlóan igazolható az állítás másik része, csak ott az $\exp(-sx)$ alakú függvényeket kell tekinteni az $x \geq 0$ halmazon, ahol ezek a függvények korlátosak.

$\square$

B.2.4. Állítás. *Az eloszlást egyértelműen jellemzi az eloszlásfüggvény.*

Bizonyítás: A bizonyítás lényegében azonos az előzővel. Legyen $(\xi_1,\ldots,\xi_n)$ egy n-dimenziós változó. Definíció szerint eloszláson a Borel-halmazokon értelmezett

$$\mu(B) = \mathbf{P}((\xi_1,\ldots,\xi_n) \in B)$$

mértéket értjük. Eloszlásfüggvényen az

$$F(x_1,x_2,\ldots,x_n) = \mathbf{P}(\xi_1 < x_1,\ldots,\xi_n < x_n)$$

függvényt értjük. Az előző bizonyítást elegendő megismételni az

$$I \stackrel{\circ}{=} \{y_1 < x_1,\ldots,y_n < x_n\}$$

halmazok χ_I karakterisztikus/indikátor függvényeiből álló π-rendszerrel.

□

B.2.5. Példa. A momentumok nem jellemzik az eloszlást.

A monoton osztály tételben fontos szerepet játszik a függvények korlátossága. Ezért alkalmazható az egyértelműségi tétel a trigonometrikus polinomokra, de nem alkalmazható az ugyancsak szorzatzárt rendszert alkotó közönséges polinomokra. Ennek fontos következménye, hogy előfordulhat, hogy két eloszlás összes momentuma véges és megegyezik, de az eloszlások mégsem azonosak. Az ellenpélda a következő: Legyen

$$f(x) = \frac{1}{x\sqrt{2\pi}} \exp\left(-\frac{(\ln x)^2}{2}\right), \quad x > 0$$

a standard lognormális eloszlás sűrűségfüggvénye. A

$$g(x) = f(x)\left(1 + \sin(2\pi \ln x)\right), \quad x > 0$$

függvény nem negatív. Az alábbi számításból evidens, hogy g sűrűségfüggvény és a momentumai megegyeznek a lognormális eloszlás momentumaival. Minden $k \geq 0$

egészre $u = \ln x - k$ helyettesítéssel

$$
\begin{aligned}
&\int_0^\infty x^k \left(g(x) - f(x)\right) dx = \\
&\frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{1}{x} x^k \exp\left(-\frac{(\ln x)^2}{2}\right) \sin(2\pi \ln x)\, dx = \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{1}{x} \exp\left(-\frac{(\ln x)^2 - 2k \ln x}{2}\right) \sin(2\pi \ln x)\, dx = \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{1}{x} \exp\left(-\frac{(\ln x - k)^2 - k^2}{2}\right) \sin(2\pi \ln x)\, dx = \\
&= \frac{\exp\left(k^2/2\right)}{\sqrt{2\pi}} \int_0^\infty \frac{1}{x} \exp\left(-\frac{(\ln x - k)^2}{2}\right) \sin(2\pi \ln x)\, dx = \\
&= \frac{\exp\left(k^2/2\right)}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(-\frac{u^2}{2}\right) \sin(2\pi (u + k))\, du = \\
&= \frac{\exp\left(k^2/2\right)}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(-\frac{u^2}{2}\right) \sin(2\pi u)\, du = 0.
\end{aligned}
$$

□

Miként azt már a valószínűségi mezők bevezetésekor is jeleztük, a három halmazműveletehez tartozó valószínűségek kiszámolásakor egyedül a metszet valószínűségének kiszámolása nem magától érthetődő. Például a peremeloszlásokból ezért nem tudjuk előállítani az együttes eloszlásokat. Ezt a problémát kerüli meg a függetlenség fogalma.

B.2.6. Definíció. Legyen adva az $(\Omega, \mathscr{A}, \mathbf{P})$ mező, és Γ legyen tetszőleges indexhalmaz.

1. Ha $\left\{\mathscr{M}_\gamma\right\}_{\gamma \in \Gamma}$ az $\mathscr{A}$ részhalmazaiból álló halmazok családja, akkor az $\left\{\mathscr{M}_\gamma\right\}_{\gamma \in \Gamma}$ halmazcsaládokat függetlennek mondjuk, ha bárhogyan is veszünk ki a Γ halmazból véges számú $\gamma_1, \gamma_2, \ldots, \gamma_n$ indexeket és az $\mathscr{M}_{\gamma_1}, \mathscr{M}_{\gamma_2}, \ldots, \mathscr{M}_{\gamma_n}$ családokból
$$A_{\gamma_1} \in \mathscr{M}_{\gamma_1}, A_{\gamma_2} \in \mathscr{M}_{\gamma_2}, \ldots, A_{\gamma_n} \in \mathscr{M}_{\gamma_n}$$
halmazokat, akkor
$$\mathbf{P}\left(A_{\gamma_1} \cap A_{\gamma_2} \cap \ldots \cap A_{\gamma_n}\right) = \mathbf{P}\left(A_{\gamma_1}\right) \cdot \mathbf{P}\left(A_{\gamma_2}\right) \cdot \ldots \cdot \mathbf{P}\left(A_{\gamma_n}\right). \tag{B.2.1}$$

2. Ha $\mathscr{X} = \left\{\mathscr{X}_\gamma\right\}_{\gamma \in \Gamma}$ valószínűségi változók halmazaiból álló család, akkor az $\mathscr{X}$ családot függetlennek mondjuk, ha a generált σ-algebrákból álló $\left\{\sigma\left(\mathscr{X}_\gamma\right)\right\}_{\gamma \in \Gamma}$

családok függetlenek. Speciálisan, a $\left(\xi_\gamma\right)_{\gamma\in\Gamma}$ változókat függetlennek mondjuk, ha az általuk generált $\left(\sigma\left(\xi_\gamma\right)\right)_{\gamma\in\Gamma}$ σ-algebrák függetlenek.

B.2.7. Állítás. *Ha az $\left(\mathscr{M}_\gamma\right)_{\gamma\in\Gamma}$ halmazrendszerek mindegyike π-rendszer és a halmazrendszerek függetlenek, akkor függetlenek a $\left(\sigma\left(\mathscr{M}_\gamma\right)\right)_{\gamma\in\Gamma}$ generált σ-algebrák is. Speciálisan a $\left(\xi_\gamma\right)_{\gamma\in\Gamma}$ változók függetlenségének szükséges és elegendő feltétele, hogy a nívóhalmazaikból álló halmazok függetlenek legyenek, vagyis hogy bármely véges részhalmaz esetén az együttes eloszlásfüggvények a peremeloszlások szorzata legyen.*

Bizonyítás: Mivel a függetlenség definíciójában mindig csak véges számú halmaz szerepel, feltehetjük, hogy a Γ halmaz véges. Rögzítsük az $(A_k)_{k=2}^n$ elemeket ahol az A_k halmaz az $\mathscr{M}_k$ összesség egy eleme, vagy Ω. Legyen $B \circeq \cap_{k=2}^n A_k$, és a $\sigma\left(\mathscr{M}_1\right)$ halmazon definiáljuk egyrészt a $\mu(A) \circeq \mathbf{P}(A\cap B)$, másrészt a $\nu(A) \circeq \mathbf{P}(A)\mathbf{P}(B)$ mértékeket. Az $\mathscr{M}_1$ a feltétel szerint π rendszer, és a $\mathscr{M}_k$ halmazrendszerek feltételezett függetlensége miatt a két mérték a $\mathscr{M}_1$ rendszeren megegyezik. Tekintsük az olyan korlátos és $\sigma\left(\mathscr{M}_1\right)$-mérhető függvények halmazát, amelyekre a μ és a ν szerinti integrálok megegyeznek. Az ilyen függvények halmaza λ-rendszer és mivel a $\mathscr{M}_1$ a feltétel szerint π-rendszer, így a χ_A, $A\in\mathscr{M}_1$ egy π-rendszer. Így a monoton osztály tétel miatt ha $A\in\sigma\left(\mathscr{M}_1\right)$, akkor

$$\begin{aligned}\mathbf{P}(A\cap B) &\circeq \mu(A)=\int_\Omega \chi_A d\mu=\int_\Omega \chi_A d\nu= \\ &= \nu(A)\circeq \mathbf{P}(A)\mathbf{P}(B)=\mathbf{P}(A)\mathbf{P}(A_2)\ldots\mathbf{P}(A_n),\end{aligned}$$

tehát a

$$\sigma\left(\mathscr{M}_1\right),\mathscr{M}_2,\ldots,\mathscr{M}_n \tag{B.2.2}$$

halmazok függetlenek. A gondolatmenetet megismételve a (B.2.2) és a $\sigma\left(\mathscr{M}_2\right)$ halmazokra belátható, hogy a $\sigma\left(\mathscr{M}_1\right),\sigma\left(\mathscr{M}_2\right),\ldots,\mathscr{M}_n$ halmazok is függetlenek. Az eljárást folytatva az összes generált σ-algebra függetlenségét igazolhatjuk. A második állítás egyszerű következménye[1] annak, hogy az $\{\xi\mathscr{R}\lambda\}$, $\mathscr{R}$ lehet $\leq$, $\geq$, $<$, $>$, alakú halmazok π-rendszert alkotnak, és az általuk generált σ-algebra éppen a ξ által generált σ-algebra.

□

B.2.8. Állítás. *Ha a $\left(\xi_\gamma\right)_{\gamma\in\Gamma}$ változók függetlennek, és $(\Gamma_\alpha)_{\alpha\in A}$ a Γ indexhalmaz partíciója, akkor a*

$$\sigma\left(\xi_\gamma \mid \gamma\in\Gamma_\alpha\right),\quad \alpha\in A$$

σ-algebrák függetlenek.

[1] V.ö.: B.2.7. állítás, 245. oldal.

Bizonyítás: Legyen $\mathscr{M}_\alpha$ a $\xi_\gamma^{-1}(B)$ alakú halmazok véges metszeteiből álló halmaz, ahol B tetszőleges Borel-halmaz, és $\gamma \in \Gamma_\alpha$. A függetlenség definíciója alapján a különböző α indexekhez tartozó $\mathscr{M}_\alpha$ halmazok függetlenek. Mivel az $\mathscr{M}_\alpha$ halmazok evidens módon π-rendszert alkotnak, ezért az állítás közvetlen következménye az előző B.2.7. állításnak.

□

A függetlenség számos fontos és meghökkentő tételben játszik kulcsszerepet. Jó példa a következő:

B.2.9. Tétel. *Ha a* (ξ_k) *sorozat tagjai függetlenek, akkor az*

$$A \stackrel{\circ}{=} \left\{ \omega \mid \sum_{k=1}^{\infty} \xi_k(\omega) \text{ sor konvergens} \right\}$$

halmaz valószínűsége 0*, vagy* 1.

A tétel speciális esete egy általánosabb elvnek. Legyen (η_k) változók egy sorozata. Az

$$\mathscr{F}_k \stackrel{\circ}{=} \sigma\left((\eta_i)_{i=k}^{\infty}\right)$$

σ-algebra olyan eseményeket tartalmaz, amelyek megfogalmazásához nincs szükség az első $k-1$ elemre. Az

$$\mathscr{F} \stackrel{\circ}{=} \cap_{k=1}^{\infty} \mathscr{F}_k$$

σ-algebra olyan eseményeket tartalmaz, amelyek megfogalmazásakor a sorozat bármely véges számú tagja érdektelen.

B.2.10. Definíció. Az $\mathscr{F}$ σ-algebrát az (η_k) sorozat farok σ-algebrájának mondjuk.

Triviális módon az

$$F \stackrel{\circ}{=} \{\omega \mid \text{az } \eta_k(\omega) \text{ sorozat konvergens}\}$$

esemény független a sorozat elejétől, vagyis az F minden n-re $\sigma\left((\eta_k)_{k=n}^{\infty}\right)$ mérhető, tehát eleme az (η_k) sorozat farok σ-algebrájának. Ennek megfelelően a fenti B.2.9. tétel nyilvánvaló következménye a következő alapvető tételnek:

B.2.11. Tétel (Nulla vagy egy törvény)**.** *Ha az* (η_k) *változók függetlenek, akkor az* (η_k) *sorozat* $\mathscr{F}$ *farok* σ*-algebrájának minden eleme* 0 *vagy* 1 *valószínűségű.*

Bizonyítás: A bizonyítás fő gondolata, hogy az $\mathscr{F}$ minden eleme független önmagától, vagyis ha $A \in \mathscr{F}$, akkor $\mathbf{P}(A) = \mathbf{P}(A \cap A) = \mathbf{P}^2(A)$, amiből evidens módon $\mathbf{P}(A)$ értéke 0, vagy 1. Ehhez elegendő belátni, hogy az A független az (η_k) sorozat által generált teljes $\sigma\left((\eta_i)_{i=1}^{\infty}\right)$ σ-algebrától. Vagyis ha (η_k) független változók, akkor a generált és a farok σ-algebrák függetlenek egymástól. Mivel evidens módon $\mathscr{F} \subset \sigma\left((\eta_i)_{i=1}^{\infty}\right)$, ezért az $A \in \mathscr{F}$ független önmagától. Tekintsük tehát a

$$\mathscr{G} \doteq \cup_{k=1}^{\infty} \sigma\left((\eta_i)_{i=1}^{k}\right)$$

halmazt. Megmutatjuk, hogy egyrészt a $\mathscr{G}$ π-rendszer, másrészt generálja a $\sigma\left((\eta_i)_{i=1}^{\infty}\right)$ σ-algebrát. Ha $B, C \in \mathscr{G}$, és $B \in \sigma\left((\eta_i)_{i=1}^{j}\right)$, $C \in \sigma\left((\eta_i)_{i=1}^{k}\right)$ és ha $n \doteq \max(j,k)$, akkor $B, C \in \sigma\left((\eta_i)_{i=1}^{n}\right)$, amiből $B \cap C \in \mathscr{G}$, következésképpen a $\mathscr{G}$ valóban π-rendszer[2]. Az összes η_k $\mathscr{G}$-mérhető, ugyanis az η_k nívóhalmazai $\mathscr{G}$-ben vannak. Mivel

$$\mathscr{G} \subset \sigma\left((\eta_i)_{i=1}^{\infty}\right),$$

ezért

$$\sigma\left((\eta_i)_{i=1}^{\infty}\right) \subset \sigma(\mathscr{G}) \subset \sigma\left((\eta_i)_{i=1}^{\infty}\right),$$

így $\sigma\left((\eta_i)_{i=1}^{\infty}\right) = \sigma(\mathscr{G})$. Ha $A \in \mathscr{F}$, akkor minden k-ra $A \in \sigma\left((\eta_i)_{i=k}^{\infty}\right)$. Ha $B \in \sigma\left((\eta_i)_{i=1}^{k-1}\right)$, akkor a valószínűségi változók feltételezett függetlensége alapján[3] A és B függetlenek, ami alapján A független $\mathscr{G}$ minden elemétől. Mivel $\mathscr{G}$ π-rendszer, ezért a generált σ-algebrák is függetlenek, így A független a $\sigma(\mathscr{G}) = \sigma\left((\eta_i)_{i=1}^{\infty}\right)$ minden elemétől.

□

B.2.12. Példa. Ha w Wiener-folyamat[4], akkor

$$\limsup_{t\to\infty} w(t) \overset{m.m.}{=} \infty, \quad \liminf_{t\to\infty} w(t) \overset{m.m.}{=} -\infty.$$

Tetszőleges $m > 0$ számra, ha Φ jelöli az $N(0,1)$ eloszlásfüggvényét, akkor

$$\lim_{n\to\infty} \mathbf{P}(w(n) \geq m) = 1 - \lim_{n\to\infty} \Phi\left(\frac{m}{\sqrt{n}}\right) = \frac{1}{2}.$$

Ha A_m jelöli azokat a kimeneteleket, amelyekre végtelen sokszor $w(n) \geq m$, akkor elég nagy k-ra és n-re

$$\begin{aligned}\mathbf{P}(A_m) &= \mathbf{P}\left(\cap_{k=1}^{\infty}\cup_{n=k}^{\infty}\{w(n)\geq m\}\right) = \lim_{k\to\infty}\mathbf{P}\left(\cup_{n=k}^{\infty}\{w(n)\geq m\}\right) \geq \\ &\geq \mathbf{P}\left(\cup_{n=k}^{\infty}\{w(n)\geq m\}\right) - \frac{1}{8} \geq \mathbf{P}(w(n)\geq m) - \frac{1}{8} \geq \frac{1}{4}.\end{aligned}$$

[2] A $\mathscr{G}$ evidens módon valójában algebra.

[3] V.ö.: B.2.8. Állítás, 245. oldal.

[4] A példa megértéséhez a Wiener-folyamatról elegendő annyit tudni, hogy $w(n) = \sum_{k=1}^{n} \xi_k$, ahol a (ξ_k) független $N(0,1)$ eloszlású változók sorozata. A gondolatmenet csekély módosítással érvényben marad a diszkrét idejű bolyongásra is, vagyis amikor (ξ_k) $1/2$ valószínűséggel ± 1 értéket felvevő, független változók sorozata.

Mivel $A_{m+1} \subseteq A_m$, ezért

$$\mathbf{P}\left(\limsup_{n\to\infty} w(n) = \infty\right) = \mathbf{P}\left(\cap_m A_m\right) \geq \frac{1}{4}.$$

A nulla vagy egy törvény miatt

$$\mathbf{P}\left(\limsup_{n\to\infty} w(n) = \infty\right) = 1.$$

Szimmetria okok miatt a másik egyenlőség is teljesül.

□

B.3. Lévy-féle folytonossági tétel

A következő bemutatandó eszköz a valószínűségszámításban gyakran használt úgynevezett folytonossági tétel. Ez valójában egy az eloszlásokra vonatkozó kompaktsági tétel. A valószínűségszámításban használt fogalmak nagyrészt a matematika egyéb területein használt fogalmak közvetlen átültetései és szimpla átnevezései. A tételek kerek voltát általában az biztosítja, hogy az alapul vett mérték végessége legegyszerűsíti a tárgyalást. Az eloszlások konvergenciája esetén azonban a valószínűségszámítás nem csak némiképpen más terminológiát használ, hanem a matematikai tartalom is csekély mértékben eltér. Mértékek konvergenciájának természetes definíciója a dualitás koncepciójára épül. Az alapvető felismerés, mely a modern absztrakt analízis egyik legfontosabb tétele, hogy a mértékek a folytonos függvények duális tereinek elemeiként foghatók fel, vagyis igen általános körülmények között a folytonos függvényeken értelmezett folytonos funkcionálok mindig alkalmas mérték szerinti integrálként írhatók fel. A legegyszerűbb és számunkra most a legfontosabb eset, amikor a számegyenesen értelmezett valószínűségi mértékeket vizsgáljuk. Ha $C_0(\mathbb{R})$ jelöli azokat a számegyenesen értelmezett folytonos függvényeket, amelyek határértéke mind a két végtelenben nulla, akkor ezen a téren értelmezett folytonos lineáris funkcionálok tere azonosítható a véges (előjeles) mértékek halmazával. Érdemes felidézni, hogy véges mérték esetén nem mindig kötjük meg azt, hogy a halmazok mértéke nem negatív legyen, hanem, miközben kizárjuk a végtelen értékeket, időnként megengedjük, hogy akár negatív is legyen egy halmaz mértéke. Ennek az a fő előnye, hogy így a már említett dualitási tétel kerek módon kimondható és az összes folytonos lineáris funkcionál azonosítható valamilyen, esetleg előjeles, mértékkel. Ennek következtében az absztrakt, gyakran algebrai ihletésű matematikai tételek alkalmazhatók a mértékekre, illetve a hozzájuk kapcsolódó súlyfüggvényekre. A számunkra most legfontosabb állítás a duális terek egységgömbjének kompaktságát kimondó igen absztrakt tétel[5]. Ezt felhasználva könnyen belátható, hogy ha adott a számegyenesen valószínűségi eloszlások egy (F_n) sorozata, akkor ennek

[5] Az úgynevezett Banach–Alaoglu tétel.

van olyan (F_{n_k}) részsorozata és egy olyan F súlyfüggvény, amelyre minden $f \in C_0(\mathbb{R})$ esetén

$$\lim_{k\to\infty}\int_{\mathbb{R}} f(x)\,dF_{n_k}(x) = \int_{\mathbb{R}} f(x)\,dF(x).$$

A kulcs észrevétel, hogy az F nem lesz feltétlenül eloszlásfüggvény, vagyis az F által indukált mérték ugyan nem negatív marad, de nem lesz a teljes egyenes mértéke 1. Ezt a jelenséget nevezzük mértékszivárgásnak. Hogy megint az absztrakt analízis terminológiáját használjuk a nem negatív mértékek halmaza ugyan zárt, de ezen belül a valószínűségi mértékek halmaza a $C_0(\mathbb{R})$ által generált halvány topológiában nem zárt.

B.3.1. Definíció. Azt mondjuk, hogy a súlyfüggvények (F_n) sorozata halványan tart az F súlyfüggvényhez ha minden $f \in C_0(\mathbb{R})$ esetén[6]

$$\lim_{n\to\infty}\int_{\mathbb{R}} f(x)\,dF_n(x) = \int_{\mathbb{R}} f(x)\,dF(x).$$

A konvergenciát szokás $F_n \xrightarrow{v} F$ módon is jelölni.

Az elmondottakat a következő állításban foglalhatjuk össze:

B.3.2. Állítás. *Ha (F_n) monoton növekedő, balról folytonos súlyfüggvények egy egyenletesen korlátos sorozata, akkor a sorozatnak van olyan részsorozata amely halványan tart egy F súlyfüggvényhez. Speciálisan ha az (F_n) eloszlásfüggvények egy sorozata, akkor van olyan F nem negatív, monoton növekedő, balról folytonos súlyfüggvény, hogy az (F_n) egy alkalmas részsorozata halványan tart az F-hez.*

B.3.3. Példa. Mértékszivárgás.

Mértékszivárgásra a legkézenfekvőbb példa az n pontra egységnyi súlyt helyező

$$F_n(x) \circeq \begin{cases} 0 & \text{ha} \quad x \le n \\ 1 & \text{ha} \quad x > n \end{cases}$$

eloszlásfüggvények sorozata. Ilyenkor minden részsorozatra ha $f \in C_0(\mathbb{R})$, akkor

$$\lim_{k\to\infty}\int_{\mathbb{R}} f(x)\,dF_{n_k}(x) = \lim_{k\to\infty} f(n_k) = 0,$$

vagyis az (F_n) sorozat halvány konvergenciában vett határértéke az $F = 0$ súlyfüggvény. Vegyük észre, hogy $F_n(x) \to F(x) \equiv 0$. Talán nem érdektelen megjegyezni, hogy az egy pontra koncentrálódó eloszlások esetén a kompaktsági tétel a következő miatt igaz: Ha az eloszlások tartó pontjainak halmaza korlátos, akkor a mértékek a konvergens részsorozata éppen a pontok konvergens részsorozatához tartozik. Ha azonban a pontoknak nincs konvergens részsorozata, akkor van végtelenbe tartó részsorozat, amely a halvány konvergencia sajátossága miatt szintén konvergens mértéksorozatot generál.

[6] A funkcionálanalízis terminológiája szerint az (F_n) sorozat gyenge-* konvergenciában konvergál az F-hez.

$\square$

Mivel a valószínűségszámítás tárgya a valószínűségi változók eloszlásainak vizsgálata, kézenfekvő, ha a konvergenciát az eloszlásfüggvények szempontjából közelítjük meg[7].

B.3.4. Definíció. Azt mondjuk, hogy az (F_n) sorozat a folytonossági pontokban konvergens ha az F határérték minden x folytonossági pontjában $F_n(x) \to F(x)$. A konvergencia szokásos jelölése $F_n \xrightarrow{c} F$.

B.3.5. Definíció. Ha a számegyenesen értelmezett minden f korlátos folytonos függvény esetén

$$\lim_{n\to\infty} \int_{\mathbb{R}} f(x)\, dF_n(x) = \int_{\mathbb{R}} f(x)\, dF(x), \tag{B.3.1}$$

akkor azt mondjuk, hogy az (F_n) sorozat gyengén tart az F-hez. Ennek szokásos jelölése $F_n \xrightarrow{w} F$.

A már bemutatott példa alapján világos, hogy a folytonosági pontokban való konvergencia esetén előfordulhat a mértékszivárgás, de mivel az $f(x) \equiv 1$ konstans függvény egy korlátos folytonos függvény a gyenge konvergencia esetén a mértékszivárgás nem fordulhat elő. A különböző konvergenciakoncepciók összehasonlítása arra épül, hogy a folytonos függvények melyik osztályán teljesül az integrálok (B.3.1) konvergenciája. Mivel az integrál létezését értelemszerűen meg akarjuk követelni, ezért a legbővebb számbajöhető osztály a gyenge konvergencia mögött álló korlátos folytonos függvények halmaza.

B.3.6. Állítás. *Jelölje $C\left(\overline{\mathbb{R}}\right)$ az olyan folytonos függvényeket, amelyeknek a $\pm\infty$ pontokban van határértéke. Ha (F_n) és F véges, növekedő és balról folytonos súlyfüggvények, és minden $g \in C\left(\overline{\mathbb{R}}\right)$ függvényre*

$$\lim_{n\to\infty} \int_{\mathbb{R}} g\,dF_n = \int_{\mathbb{R}} g\,dF,$$

akkor az (F_n) sorozat a folytonossági pontokban tart az F-hez.

Bizonyítás: Érdemes megjegyezni, hogy az állításban szereplő integrálok minden $g \in C\left(\overline{\mathbb{R}}\right)$ esetén léteznek ugyanis minden ilyen függvény korlátos. Tegyük fel, hogy x az F folytonossági pontja, és ε legyen tetszőleges pozitív szám. Tekintsük a

$$g(y) \circeq \begin{cases} 1 & \text{ha} \quad y < x \\ l(y) & \text{ha} \quad y \in [x, x+\varepsilon] \\ 0 & \text{ha} \quad y > x+\varepsilon \end{cases}$$

[7] Érdemes megjegyezni, hogy az eloszlásfüggvények minden pontban való konvergenciája túl erős megkötés lenne. Ha például $x_n \nearrow x$ valamilyen (x_n) pontsorozatra, akkor az x_n pontokra koncentrálódó eloszlások eloszlásfüggvényei az x pontban mind egy értéket vesznek fel, de az x pontra koncentrálódó eloszlás eloszlásfüggvénye a balról való folytonosság miatt nulla.

folytonos függvényt, ahol l az $(x,1)$ és az $(x+\varepsilon,0)$ pontokat összekötő lineáris függvény. Világos, hogy

$$\chi_{(-\infty,x)} \leq g,$$

következésképpen,

$$F_n((-\infty,x)) = F_n(x) \leq \int_{\mathbb{R}} g dF_n.$$

Hasonlóan, mivel

$$g \leq \chi_{(-\infty,x+\varepsilon)},$$

ezért

$$\int_{\mathbb{R}} g dF \leq F((-\infty,x+\varepsilon)) = F(x+\varepsilon),$$

ahol remélhetőleg nem félreérthető módon az eloszlásfüggvényt és az általa generált mértéket azonos módon jelöltük. Mivel $g \in C\left(\overline{\mathbb{R}}\right)$, ezért a feltétel miatt

$$\limsup_{n\to\infty} F_n(x) \leq \limsup_{n\to\infty} \int_{\mathbb{R}} g dF_n = \int_{\mathbb{R}} g dF \leq F(x+\varepsilon).$$

Mivel x az F folytonossági pontja, ezért az F jobbról folytonos, tehát ha $\varepsilon \searrow 0$, akkor

$$\limsup_{n\to\infty} F_n(x) \leq F(x).$$

Analóg módon ha

$$g(y) \stackrel{\circ}{=} \begin{cases} 1 & \text{ha} \quad y < x-\varepsilon \\ l(y) & \text{ha} \quad y \in [x-\varepsilon,x] \\ 0 & \text{ha} \quad y > x \end{cases}$$

akkor

$$\liminf_{n\to\infty} F_n(x) \geq \liminf_{n\to\infty} \int_{\mathbb{R}} g dF_n = \int_{\mathbb{R}} g dF \geq F(x-\varepsilon).$$

Az F, mint minden eloszlásfüggvény, az x pontban balról folytonos, tehát

$$\liminf_{n\to\infty} F_n(x) \geq F(x),$$

amiből

$$\lim_{n\to\infty} F_n(x) = F(x).$$

□

B.3.7. Lemma. Ha (F_n) és F súlyfüggvények és $F_n \xrightarrow{c} F$, továbbá a és b az F eloszlásfüggvény folytonossági pontjai, és $g \in C([a,b])$, akkor

$$\lim_{n\to\infty} \int_a^b g dF_n = \int_a^b g dF.$$

Speciálisan ha $C_c(\mathbb{R})$ jelöli az olyan folytonos függvényeket, amelyek egy alkalmas, a függvénytől függő $[a,b]$ zárt intervallumon kívül mindenhol nullák, akkor ha a folytonossági pontokban az (F_n) sorozat F-hez tart, akkor minden $g \in C_c(\mathbb{R})$ esetén

$$\lim_{n\to\infty} \int_{\mathbb{R}} g dF_n = \int_{\mathbb{R}} g dF.$$

Bizonyítás: A második állítás egyszerűen következik az elsőből, elég az a és b pontokat úgy megválasztani, hogy ott a g már nulla legyen és a és b folytonossági pontjai legyenek F-nek. Az első állítás egyszerű következménye annak, hogy korlátos és zárt intervallumon értelmezett folytonos függvény az intervallumon egyenletesen is folytonos. Az egyenletes folytonosság miatt tetszőleges $\varepsilon > 0$ értékhez van olyan $\delta > 0$, hogy az $[a,b]$ intervallum minden δ-nál rövidebb részintervallumokból álló $a = t_0, t_1, \ldots, t_n = b$ felbontására az egyes intervallumokon g ingadozása kisebb lesz mint ε. A t_i pontokat bármely sűrű halmazból választhatjuk, vagyis feltehetjük, hogy minden t_i az F függvény folytonossági pontja. Tekintsünk egy ilyen felosztást, és minden egyes részintervallumból vegyünk ki egy x_i reprezentánst. Ha

$$s_\varepsilon(y) \overset{\circ}{=} \sum_i g(x_i)\,\chi_{[t_i,t_{i-1})}(y),$$

akkor világos, hogy

$$|g(y) - s_\varepsilon(y)| < \varepsilon, \quad y \in [a,b].$$

Vagyis ha $\varepsilon \to 0$, akkor az s_ε egyenletesen tart a g-hez. Mivel a feltételek szerint

$$F_n(b) - F_n(a) \to F(b) - F(a) < \infty,$$

ezért

$$\begin{aligned}
\left|\int_a^b g\,dF_n - \int_a^b s_\varepsilon\,dF_n\right| &\leq \int_a^b |g - s_\varepsilon|\,dF_n \leq \\
&\leq \varepsilon\left[F_n(b) - F_n(a)\right] \leq K \cdot \varepsilon,
\end{aligned}$$

tehát az ε szerinti konvergencia az n szerint egyenletes. A t_i pontok választása miatt

$$\begin{aligned}
\lim_{n\to\infty}\int_a^b s_\varepsilon\,dF_n &= \lim_{n\to\infty}\sum_i g(x_i)\left[F_n(t_i) - F_n(t_{i-1})\right] = \\
&= \sum_i g(x_i)\left[F(t_i) - F(t_{i-1})\right] = \int_a^b s_\varepsilon\,dF,
\end{aligned}$$

amiből

$$\begin{aligned}
\lim_{n\to\infty}\int_a^b g\,dF_n &= \lim_{n\to\infty}\int_a^b \lim_{\varepsilon\to 0} s_\varepsilon\,dF_n = \lim_{n\to\infty}\lim_{\varepsilon\to 0}\int_a^b s_\varepsilon\,dF_n = \\
&= \lim_{\varepsilon\to 0}\lim_{n\to\infty}\int_a^b s_\varepsilon\,dF_n = \lim_{\varepsilon\to 0}\int_a^b s_\varepsilon\,dF = \\
&= \int_a^b g\,dF,
\end{aligned}$$

ugyanis egyrészt az egyenletes konvergencia és a véges mérték szerinti integrálás felcserélhető, másrészt mivel az ε szerinti konvergencia az n szerint egyenletes, ezért az ε és az n szerinti határérték is felcserélhető.

□

B.3.8. Állítás. *Ha (F_n) egyenletesen korlátos, akkor a folytonossági pontokban való $F_n \xrightarrow{c} F$ konvergencia implikálja az $F_n \xrightarrow{v} F$ halvány konvergenciát.*

Bizonyítás: Jelölje K az (F_n) sorozat egy korlátját. Legyen $g \in C_0(\mathbb{R})$ tetszőleges, és legyenek a és b olyan folytonossági pontjai F-nek, amelyekre $|g(x)| < \varepsilon/K$, ha $x \notin [a,b)$. Evidens módon

$$\begin{aligned} \left| \int_{\mathbb{R}} g dF_n - \int_a^b g dF_n \right| &\leq \varepsilon, \quad (n = 1, 2, \ldots) \\ \left| \int_{\mathbb{R}} g dF - \int_a^b g dF \right| &\leq \varepsilon. \end{aligned}$$

Ha $F_n \xrightarrow{c} F$, akkor az előző lemma szerint

$$\lim_{n \to \infty} \int_a^b g dF_n = \int_a^b g dF,$$

ezért ha n elég nagy, akkor

$$\begin{aligned} & \left| \int_{\mathbb{R}} g dF_n - \int_{\mathbb{R}} g dF \right| \leq \\ \leq \quad & \left| \int_{\mathbb{R}} g dF_n - \int_a^b g dF_n \right| + \left| \int_a^b g dF_n - \int_a^b g dF \right| + \left| \int_a^b g dF - \int_{\mathbb{R}} g dF \right| \leq 3\varepsilon. \end{aligned}$$

Mivel az ε tetszőleges, ezért

$$\limsup_{n \to \infty} \left| \int_{\mathbb{R}} g dF_n - \int_{\mathbb{R}} g dF \right| = 0,$$

vagyis (F_n) halványan tart az F-hez.

□

Most térjünk rá a gyenge konvergencia vizsgálatára. Miként jeleztük a legnagyobb gondot a mértékszívárás jelenti. Ezt zárja ki a következő definícióban szereplő megkötés:

B.3.9. Definíció. Eloszlásfüggvények $(F_\alpha)_{\alpha \in \Gamma}$ összességet szorosnak mondjuk, ha minden $\varepsilon > 0$ számhoz léteznek olyan $a_\varepsilon < b_\varepsilon$ pontok, hogy minden $\alpha \in \Gamma$ indexre

$$F_\alpha(b_\varepsilon) - F_\alpha(a_\varepsilon) \geq 1 - \varepsilon. \tag{B.3.2}$$

B.3.10. Példa. Ha az eloszlásfüggvényekből álló (F_n) sorozat a folytonossági pontokban tart az F-hez, és az F is eloszlásfüggvény, akkor az (F_n) sorozat szoros.

Az eloszlásfüggvény definíciója alapján eloszlásfüggvényekből álló minden véges halmaz szoros. Nyilván az F folytonossági pontjai sűrűn vannak, így, kihasználva, hogy az F eloszlásfüggvény bármely ε-ra vannak olyan a_ε és b_ε pontok amelyre

$$F(b_\varepsilon) - F(a_\varepsilon) \geq 1 - \frac{\varepsilon}{2}.$$

Ebből és a folytonossági pontokban való konvergencia feltételéből következően az előírt (B.3.2) egyenlőtlenség véges sok indextől eltekintve már teljesül. Ebből a példa igazolása már evidens.

□

B.3.11. Állítás. *Ha eloszlásfüggvények egy (F_n) sorozata szoros, akkor a gyenge konvergencia ekvivalens a folytonossági pontokban való konvergenciával. Ebből következően ha az összes F_n és az F is eloszlásfüggvény, akkor az $F_n \xrightarrow{w} F$ gyenge konvergencia és az $F_n \xrightarrow{c} F$ folytonossági pontokokban való konvergencia ekvivalens.*

Bizonyítás: Miként láttuk a gyenge konvergencia implikálja a folytonossági pontokban való konvergenciát. Fordítva általában nem lehet igaz a tétel, mert gyenge konvergencia esetén nincs mértékszívárgás, de a folytonossági pontokban való konvergencia esetén a mértékszívárgás előfordulhat. Tegyük fel tehát, hogy az (F_n) szoros és az F folytonossági pontjaiban $F_n(x) \to F(x)$. Legyen $\varepsilon > 0$, és legyenek α_ε és b_ε az ε-hoz tartozó számok. Feltehetjük, hogy ezek a pontok folytonossági pontjai F-nek. A szorosság definíciója és a folytonossági pontokban való konvergencia miatt a fenti (B.3.2) egyenlőtlenség F-re is teljesül. Legyen f korlátos folytonos függvény.

$$\int_{\mathbb{R}} f dF_n = \int_{-\infty}^{a_\varepsilon} f dF_n + \int_{a_\varepsilon}^{b_\varepsilon} f dF_n + \int_{b_\varepsilon}^{\infty} f dF_n.$$

A korábban belátott lemma miatt

$$\int_{a_\varepsilon}^{b_\varepsilon} f dF_n \to \int_{a_\varepsilon}^{b_\varepsilon} f dF.$$

Ugyanakkor a szorosság miatt ha K az f egy korlátja

$$\left| \int_{-\infty}^{a_\varepsilon} f dF_n - \int_{-\infty}^{a_\varepsilon} f dF \right| \leq \int_{-\infty}^{a_\varepsilon} |f| dF_n + \int_{-\infty}^{a_\varepsilon} |f| dF \leq 2K\varepsilon.$$

Hasonlóan becsülhető a másik intervallum is. Ebből $\int_{\mathbb{R}} f dF_n \to \int_{\mathbb{R}} f dF$.

□

B.3.12. Állítás (Prohorov)**.** *Ha eloszlásfüggvények egy (F_n) sorozata szoros, akkor a sorozatnak létezik olyan (F_{n_k}) részsorozata, amely gyengén konvergens.*

Bizonyítás: Legyen (r_k) a racionális számok egy felsorolása. Az $(F_n(r_1))_n \subseteq [0,1]$ korlátos sorozatnak létezik $(F_{n_k}(r_1))_k$ konvergens részsorozata. Ezt követően tekintsük az

$(F_{n_k}(r_2))_k$ sorozatot, amely ismételten megritkítható úgy, hogy konvergens legyen. Az eljárást folytatva, menjünk végig az összes racionális számon, majd átlós eljárással vegyük azt az $(F_{n_k})_k$ részsorozat, amely minden $r \in \mathbb{Q}$ racionális szám esetén konvergens. Legyen

$$G(r) \circeq \lim_{k\to\infty} F_{n_k}(r).$$

Mivel az összes F_{n_k} monoton nő, ezért a G is monoton növekedő lesz. Legyen

$$F(x) \circeq \sup\{G(r) \mid r < x,\ r \in \mathbb{Q}\}.$$

A definícióból világos, hogy az F továbbra is növekedő függvény. Legyenek $x \in \mathbb{R}$ tetszőleges és $\varepsilon > 0$. A definícióból világos, hogy létezik olyan $r \in \mathbb{Q}$, hogy $r < x$ és $G(r) > F(x) - \varepsilon$. Ha most az (x_n) sorozat balról az x ponthoz tart, akkor elég nagy n indexre $x_n > r$, vagyis az F definíciója alapján

$$F(x) \geq F(x_n) \geq G(r) > F(x) - \varepsilon,$$

ami alapján az F balról folytonos[8]. Belátjuk, hogy az F minden x folytonossági pontjában

$$\lim_{k\to\infty} F_{n_k}(x) = F(x). \tag{B.3.3}$$

Ha $r > x$ racionális szám, akkor F_{n_k} monotonítása alapján

$$\limsup_{k\to\infty} F_{n_k}(x) \leq \limsup_{k\to\infty} F_{n_k}(r) = \lim_{k\to\infty} F_{n_k}(r) \circeq G(r).$$

Hasonlóan, ha $s < x$ racionális, akkor

$$\liminf_{k\to\infty} F_{n_k}(x) \geq \liminf_{k\to\infty} F_{n_k}(s) = \lim_{k\to\infty} F_{n_k}(s) \circeq G(s).$$

F definíciójából G monotonítása alapján

$$G(s) \geq \sup\{G(r) : r < s,\ r \in \mathbb{Q}\} \circeq F(s),$$

tehát

$$F(s) \leq G(s) \leq \liminf_{k\to\infty} F_{n_k}(x) \leq \limsup_{k\to\infty} F_{n_k}(x) \leq G(r),$$

amiből, ha $s \nearrow x$, akkor az F már belátott balról való folytonossága alapján

$$F(x) \leq \liminf_{k\to\infty} F_{n_k}(x) \leq \limsup_{k\to\infty} F_{n_k}(x) \leq G(r).$$

Legyen $\varepsilon > 0$ tetszőleges, és indirekt módon tegyük fel, hogy van racionális számoknak egy olyan $r_m \searrow x$, sorozata, amelyre $G(r_m) \geq F(x) + \varepsilon$. Ilyenkor minden $y > x$ pontra,

[8] Az x^n, $x \in [0,1]$ példája mutatja, hogy a G nem lesz feltétlenül balról folytonos. Ezért van szükség a bizonyításban szereplő korrekcióra.

ha m elég nagy, $y > r_m \geq x$, ezért $F(y) \geq G(r_m) \geq F(x) + \varepsilon$, ami ellentmond annak, hogy az x folytonossági pontja az F-nek. Ez alapján tehát ha r elég közel van az x ponthoz akkor

$$F(x) \leq \liminf_{k\to\infty} F_{n_k}(x) \leq \limsup_{k\to\infty} F_{n_k}(x) \leq G(r) \leq F(x) + \varepsilon,$$

tehát miként állítottuk teljesül a fenti (B.3.3). A feltétel alapján az (F_n) sorozat szoros, ezért létezik olyan N szám, hogy minden $x < -N$ esetén $F_n(x) < \varepsilon$. Ez alapján minden $r < -N$ racionális számra $G(r) \leq \varepsilon$, amiből minden $x < -N$ esetén $F(x) \leq \varepsilon$, vagyis $\lim_{x\to-\infty} F(x) = 0$. Hasonlóan látható, hogy $\lim_{x\to\infty} F(x) = 1$.

□

A bizonyításból világos, hogy a szorosság feltételére csak azért volt szükség, hogy biztosítani tudjuk, hogy az F valószínűségi eloszlás marad. Ha a szorosság feltételét elhagyjuk, akkor a következő állítást igazoltuk:

B.3.13. Állítás (Helly-féle kiválasztási tétel)**.** *Ha (F_n) balról folytonos és monoton növekedő súlyfüggvények egyenletesen korlátos sorozata, akkor létezik olyan (F_{n_k}) részsorozat, és F balról folytonos, monoton növekedő súlyfüggvény, hogy $F_{n_k} \xrightarrow{c} F$.*

Végezetül a dualitás által definiált konvergenciafogalom és a folytonossági pontokban való konvergencia összevetése szempontjából nem érdektelen a következő:

B.3.14. Példa. Halványan konvergens sorozat, amely nem gyengén konvergens.

Legyen $\xi_{2n} = n$. $\xi_{2n+1} = -n$. Ha $f \in C_0(\mathbb{R})$, akkor $\lim_{n\to 0} f(\xi_n) = 0$, így triviálisan a megfelelő eloszlások a halvány konvergenciában nullához tartanak, de az eloszlásfüggvények egyetlen pontban sem konvergálnak. Ugyanakkor persze a sorozat nem szoros, de a Helly-féle kiválasztási tétel érvényben marad, ugyanis például a (ξ_{2n}) sorozat a folytonossági pontokban való konvergencia értelmében az azonosan nulla eloszláshoz tart.

□

A karakterisztikus függvény módszer kiemelkedően fontos eszköze a következő:

B.3.15. Állítás (Lévy-féle folytonossági tétel)**.** *Ha a számegyenesen értelmezett eloszlásfüggvényekből álló (F_n) sorozat gyengén tart az F eloszlásfüggvényhez, akkor a karakterisztikus függvények megfelelő (φ_n) sorozata pontonként tart az F eloszlásfüggvény φ karakterisztikus függvényéhez. Megfordítva, legyen (φ_n) karakterisztikus függvények egy sorozata, és tegyük fel, hogy minden $t \in \mathbb{R}$ esetén létezik a $\lim_{n\to\infty} \varphi_n(t) \stackrel{\circ}{=} \varphi(t)$ határérték, és a φ függvény a $t = 0$ pontban folytonos. Ekkor a φ egy F eloszlás karakterisztikus függvénye, és a φ_n-hez tartozó F_n eloszlásfüggvények gyengén tartanak az F eloszlásfüggvényhez.*

Bizonyítás: Valóban, a gyenge konvergencia definíciója alapján minden $t \in \mathbb{R}$ esetén a korlátos $x \mapsto \sin tx$ és $x \mapsto \cos tx$ függvényekre

$$\begin{aligned}
\lim_{n\to\infty} \int_{\mathbb{R}} \sin tx \, dF_n(x) &= \int_{\mathbb{R}} \sin tx \, dF(x), \\
\lim_{n\to\infty} \int_{\mathbb{R}} \cos tx \, dF_n(x) &= \int_{\mathbb{R}} \cos tx \, dF(x),
\end{aligned}$$

ami alapján

$$\lim_{n\to\infty} \int_{\mathbb{R}} \exp(it)\, dF_n(x) = \int_{\mathbb{R}} \exp(itx)\, dF(x).$$

A fordított irányt a Prohorov-tétel segítségével fogjuk belátni. Megmutatjuk, hogy az (F_n) sorozat szoros. Ehhez szükségünk lesz a következő lemmára:

B.3.16. Lemma. Ha F tetszőleges eloszlásfüggvény, és φ jelöli az F karakterisztikus függvényét, akkor minden $\tau > 0$ esetén

$$1 - F\left(\frac{2}{\tau}\right) + F\left(-\frac{2}{\tau}\right) \leq 2 - \frac{1}{\tau}\left|\int_{-\tau}^{\tau} \varphi(t)\, dt\right|. \tag{B.3.4}$$

A lemma bizonyítása: A karakterisztikus függvény definícióját felírva és a Fubini-tétel alapján felcserélve az integrálokat a következő módon számolhatunk:

$$\begin{aligned}
\left|\frac{1}{2\tau}\int_{-\tau}^{\tau} \varphi(t)\, dt\right| &= \left|\frac{1}{2\tau}\int_{-\tau}^{\tau}\int_{\mathbb{R}} \exp(itx)\, dF(x)\, dt\right| = \\
&= \left|\frac{1}{2\tau}\int_{\mathbb{R}}\int_{-\tau}^{\tau} \exp(itx)\, dt dF(x)\right| = \\
&= \left|\frac{1}{2\tau}\int_{\mathbb{R}}\int_{-\tau}^{\tau} \cos tx + i \sin tx \, dt dF(x)\right| = \\
&= \left|\frac{1}{2\tau}\int_{\mathbb{R}}\int_{-\tau}^{\tau} \cos tx dt dF(x)\right| = \\
&= \left|\frac{1}{2\tau}\int_{\mathbb{R}}\left[\frac{\sin tx}{x}\right]_{-\tau}^{\tau} dF(x)\right| = \left|\int_{\mathbb{R}} \frac{\sin \tau x}{\tau x} dF(x)\right| \leq \\
&\leq \int_{\mathbb{R}} \left|\frac{\sin \tau x}{\tau x}\right| dF(x).
\end{aligned}$$

Mivel $|\sin(\tau x)/\tau x| \leq 1$, ezért az egyenlőtlenség tetszőleges $a > 0$ esetén az alábbi módon folytatható:

$$\begin{aligned}
\left|\frac{1}{2\tau}\int_{-\tau}^{\tau}\varphi(t)\,dt\right| &\leq \int_{-a}^{a}\left|\frac{\sin\tau x}{\tau x}\right| dF(x) + \frac{1}{\tau a}\int_{[-a,a]^c}|\sin\tau x|\,dF(x) \leq \\
&\leq \int_{-a}^{a} dF(x) + \frac{1}{\tau a}\int_{[-a,a]^c} dF(x) = \\
&= F(a) - F(-a) + \frac{1}{\tau a}\left(1-\left(F(a)-F(-a)\right)\right) = \\
&= \left(F(a)-F(-a)\right)\left(1-\frac{1}{\tau a}\right) + \frac{1}{\tau a}.
\end{aligned}$$

Ha $a = 2/\tau$, akkor

$$\left|\frac{1}{2\tau}\int_{-\tau}^{\tau}\varphi(t)\,dt\right| \leq \left(F\left(\frac{2}{\tau}\right) - F\left(-\frac{2}{\tau}\right)\right)\frac{1}{2} + \frac{1}{2},$$

vagyis

$$\left|\frac{1}{\tau}\int_{-\tau}^{\tau}\varphi(t)\,dt\right| \leq F\left(\frac{2}{\tau}\right) - F\left(-\frac{2}{\tau}\right) + 1.$$

Ha mind a két oldalhoz hozzáadunk 1-et, és átrendezzük az egyenlőtlenséget, akkor

$$1 - F\left(\frac{2}{\tau}\right) + F\left(-\frac{2}{\tau}\right) \leq 2 - \left|\frac{1}{\tau}\int_{-\tau}^{\tau}\varphi(t)\,dt\right|.$$

□

Az állítás bizonyítása: Mivel a feltétel szerint a φ folytonos a $t = 0$ pontban tetszőleges $\varepsilon > 0$ esetén létezik olyan $\tau > 0$, hogy ha $|t| < \tau$ akkor

$$|1 - \varphi(t)| < \frac{\varepsilon}{4}.$$

Az egyenlőtlenséget a $[-\tau, \tau]$ intervallumon integrálva, és figyelembe véve, hogy

$$||a| - |b|| \leq |a - b|$$

$$\begin{aligned}
\frac{\varepsilon}{4} &\geq \frac{1}{2\tau}\int_{-\tau}^{\tau}|1-\varphi(t)|\,dt \geq \frac{1}{2\tau}\left|\int_{-\tau}^{\tau}1-\varphi(t)\,dt\right| = \\
&= \frac{1}{2\tau}\left|\int_{-\tau}^{\tau}1dt - \int_{-\tau}^{\tau}\varphi(t)\,dt\right| \geq 1 - \frac{1}{2\tau}\left|\int_{-\tau}^{\tau}\varphi(t)\,dt\right|.
\end{aligned} \tag{B.3.5}$$

Mivel a karakterisztikus függvények φ_n sorozata egyenletesen korlátos, a majorált konvergencia tétel alapján

$$\lim_{n\to\infty}\int_{-\tau}^{\tau}\varphi_n(t)\,dt = \int_{-\tau}^{\tau}\varphi(t)\,dt.$$

Ebből a (B.3.5) alapján elég nagy n indexre

$$1-\left|\frac{1}{2\tau}\int_{-\tau}^{\tau}\varphi_n(t)\,dt\right|\leq\frac{\varepsilon}{2}.$$

De ekkor a (B.3.4) egyenlőtlenségből

$$1-\left(F_n\left(\frac{2}{\tau}\right)-F_n\left(-\frac{2}{\tau}\right)\right)\leq 2-\left|\frac{1}{\tau}\int_{-\tau}^{\tau}\varphi_n(t)\,dt\right|\leq\varepsilon,$$

ami alapján az (F_n) sorozat szoros, tehát gyengén sorozatkompakt. Így létezik olyan F eloszlás, és (F_{n_k}) részsorozat, hogy gyenge konvergenciában $F_{n_k}\to F$. Ez minden részsorozatra érvényes. Mivel a részsorozat határértékének karakterisztikus függvénye azonos, ezért az egyértelműségi tétel miatt az F nem függ a részsorozattól. Innen az állítás bizonyítása már evidens, hiszen az $F_n\to F$ is teljesül.

□

Laplace-transzformáltak esetén az analóg állítás a következő:

B.3.17. Állítás. *Ha a nem negatív számokon értelmezett eloszlásfüggvényekből álló (F_n) sorozat gyengén tart az F eloszlásfüggvényhez, akkor a Laplace-transzformáltak megfelelő (L_n) sorozata az $s\geq 0$ halmazon pontonként tart az F eloszlásfüggvény L Laplace-transzformáltjához. Megfordítva, legyen (L_n) a nem negatív számokra támaszkodó eloszlások Laplace-transzformáltjainak egy sorozata, és tegyük fel, hogy minden $s\geq 0$ esetén létezik a $\lim_{n\to\infty}L_n(s)\circeq L(s)$ határérték, és az L függvény az $s=0$ pontban folytonos. Ekkor az L egy F eloszlás Laplace-transzformáltja, és az L_n-hez tartozó F_n eloszlásfüggvények gyengén tartanak az F eloszlásfüggvényhez.*

Bizonyítás: Az egyik irány ismételten evidens, ugyanis ha $s\geq 0$, akkor az $x\geq 0$ félegyenesen az $\exp(-sx)$ függvények korlátos folytonos függvények. A fordított irány igazolásához ismét elég belátni, hogy a megadott feltételek esetén az (F_n) sorozat szoros. A fenti (B.3.4) becslés megfelelője most a következő: Ha L egy a nem negatív számokra koncentrálódó F eloszlás Laplace-transzformáltja, akkor minden $r>0$ esetén

$$1-F\left(\frac{1}{r}\right)\leq 2(1-L(r)).$$

Ennek igazolása igen egyszerű, ugyanis, ha $x\geq 1$, akkor $\exp(-x)\leq 1/e<1/2$, amiből

$$\begin{aligned}2(1-L(r)) &= 2\int_0^\infty 1-\exp(-rx)\,dF(x)\geq 2\int_{1/r}^\infty 1-\exp(-rx)\,dF(x)\geq\\ &\geq 2\int_{1/r}^\infty 1/2\,dF(x)=1-F\left(\frac{1}{r}\right).\end{aligned}$$

Legyen $\varepsilon>0$ tetszőleges. Mivel az $L(s)$ határérték folytonos és $L(0)=\lim_n F_n(0)=1$, azért ha r elég kicsik, akkor $2(1-L(r))\leq\varepsilon/2$. Mivel $L_n(r)\to L(r)$, ezért elég nagy

n-re $2(1-L_n(r)) \leq \varepsilon$. Mivel véges számú index nem befolyásolja a szorosságot, ezért az (F_n) szoros. Ettől a ponttól kezdve a bizonyítás már a bemutatott módon befejezhető.

□

A folytonossági tételnek számos alkalmazása ismert. Tekintsük a következőt:

B.3.18. Példa. Ha az F korlátlanul osztható, akkor az F φ karakterisztikus függvénye sehol sem nulla.

A φ helyett tekintsük a $\phi \overset{\circ}{=} |\varphi|^2$ függvényt. A ϕ szintén karakterisztikus függvény ugyanis a $\varphi(-t) = \overline{\varphi(t)}$ szintén egy karakterisztikus függvény és karakterisztikus függvények szorzata szintén karakterisztikus függvény. Mivel $\phi(0) = 1$ és a ϕ folytonos, ezért van az origónak olyan V környezete, ahol $\phi(t) \neq 0$. Ha $\varphi = \varphi_n^n$ és $\phi_n \overset{\circ}{=} |\varphi_n|^2$, akkor $\phi = \phi_n^n$, vagyis a ϕ mögötti eloszlás is korlátlanul osztható. Mivel $\phi(t)$ és $\phi_n(t)$ nem negatív és a nem negatív számok körében a gyökvonás egyértelmű, ezért

$$\phi_n(t) = \sqrt[n]{\phi(t)}.$$

Ennek megfelelően

$$\phi_\infty(t) \overset{\circ}{=} \lim_{n\to\infty} \phi_n(t) = \begin{cases} 1 & \text{ha} \quad \phi(t) > 0 \\ 0 & \text{ha} \quad \phi(t) = 0 \end{cases}.$$

Mivel a $\phi(t)$ az origó környezetében pozitív, ezért a $\phi_\infty(t)$ karakterisztikus függvények olyan pontonkénti határértéke, amelyik az origóban folytonos, tehát maga is karakterisztikus függvény. Ha valamely t_0 pontban $\varphi(t_0) = 0$, akkor $\phi_\infty(t_0) = 0$, ami lehetetlen, hiszen minden karakterisztikus függvény folytonos.

□

Egy fontos további alkalmazás a következő:

B.3.19. Példa. Korlátlanul osztható változók gyenge konvergenciában vett határértéke szintén korlátlanul osztható.

Legyen (ξ_n) korlátlanul osztható változók sorozata és legyen ξ a sorozat gyenge konvergenciában vett határértéke. A megfelelő karakterisztikus függvényekre áttérve $\varphi_n(t) \to \varphi(t)$. Meg kell mutatni, hogy a ξ, pontosabban a ξ eloszlása, is korlátlanul osztható. Legyen $m > 1$ tetszőleges természetes szám. A korlátlanul való oszthatóság definíciója miatt mindegyik ξ_n felírható m darab független és azonos eloszlású $\xi_n^{(k)}$ változó összegeként, vagyis $\xi_n = \Sigma_{k=1}^m \xi_n^{(k)}$. Ahhoz, hogy a ξ határérték is korlátlanul osztható legyen elég megmutatni, minden k-ra a $\xi_n^{(k)}$ sorozatnak van gyenge konvergenciában vett konvergens részsorozata. Ezt m-szer alkalmazva és $\xi^{(k)}$-val jelölve a határértéket és $\varphi^{(k)}(t)$-vel a megfelelő karakterisztikus függvényt

$$\begin{aligned}\varphi(t) &= \lim_{n\to\infty}\varphi_n(t) = \lim_{n\to\infty}\prod_{k=1}^{m}\varphi_n^{(k)}(t) = \\ &= \prod_{k=1}^{m}\varphi^{(k)}(t) \overset{\circ}{=} \psi^m(t).\end{aligned}$$

Elegendő belátni, hogy mindegyik $\xi_n^{(k)}$ sorozat szoros. A függetlenség és az azonos eloszlás feltétele miatt

$$\begin{aligned}\mathbf{P}\left(\xi_n^{(k)} \geq a\right)^m &= \mathbf{P}\left(\xi_n^{(i)} \geq a, i = 1,2,\ldots,m\right) \leq \\ &\leq \mathbf{P}\left(\Sigma_{i=1}^{m}\xi_n^{(i)} \geq ma\right) = \mathbf{P}(\xi_n \geq ma).\end{aligned}$$

Mivel a (ξ_n) sorozat gyengén konvergens, ezért szoros, így ha a elég nagy, akkor $\mathbf{P}\left(\xi_n^{(k)} \geq a\right) \leq \varepsilon/2$. Hasonlóan érvelve kaphatjuk a szorossághoz tartozó másik becslést.

□

B.4. Feltételes várható érték

A mértékelmélet és a valószínűségszámítás szoros és megkerülhetetlen kapcsolata leginkább a feltételes várható érték elmélete kapcsán érzékelhető. Mielőtt rátérnénk a definícióra, néhány megjegyzést teszünk:

1. A feltételes valószínűség speciális esete a feltételes várható értéknek, és a továbbiak során az általánosabb fogalmat, a feltételes várható értéket, tárgyaljuk.

2. Az általános esetben a feltétel nem valószínűségi változó, hanem σ-algebra. Ez a megközelítés lehetővé teszi, hogy a feltétel esetleg több változóval legyen megfogalmazva. Ha a feltételt az $(\eta_\alpha)_{\alpha\in A}$ változókkal akarjuk megfogalmazni, akkor az $\mathscr{F}$ feltételi σ-algebra éppen a $\sigma\left((\eta_\alpha)_{\alpha\in A}\right)$ generált σ-algebra. Minden valószínűségi változóhoz hozzárendelhető az általa generált σ-algebra, de természetesen nem minden σ-algebra fogalmazható meg valószínűségi változóval. Természetesen a valószínűségi változó, mint feltétel, nem egy a speciális esetek közül, hanem a legfontosabb speciális eset, ezért a tárgyalás során különös figyelemmel fogjuk kezelni.

3. Az általános σ-algebrára vonatkozó feltételes várható érték az Ω alaptéren értelmezett valószínűségi változó, amely csak egy nullmértékű halmaz erejéig értelmezett. Bár ennek jelentősége első ránézésre nem világos mégis döntően befolyásolja a valószínűségszámítás elméletét.

4. Rögzítsük az $(\Omega, \mathscr{A}, \mathbf{P})$ valószínűségi mezőt, legyen $\mathscr{F}$ az $\mathscr{A}$ rész σ-algebrája, és legyen adva egy $\xi : \Omega \to \mathbb{R}$ valószínűségi változó. A ξ általában $\mathscr{A}$ mérhető, de a feltételes várható érték egy $\mathscr{F}$-mérhető valószínűségi változó. A feltételes várható érték tárgyalása során mindig feltesszük, hogy a ξ-nek létezik $\mathbf{E}(\xi)$ véges várható értéke. Vagyis a feltételes várható érték az $(\Omega, \mathscr{A}, \mathbf{P})$ téren értelmezett integrálható függvények osztályát az $(\Omega, \mathscr{F}, \mathbf{P})$ téren értelmezett integrálható függvények osztályába képezi. Miként be fogjuk látni a leképezés lineáris, folytonos és projekció, ami alatt azt értjük, hogy az operációt másodszor is végrehajtva a második operáció azonos eredményre vezet mint az első.

B.4.1. Definíció. Az $(\Omega, \mathscr{A}, \mathbf{P})$ téren értelmezett ξ változó $\mathscr{F}$-re vonatkozó feltételes várható értékén azt az $\mathbf{E}(\xi \mid \mathscr{F})$-fel jelölt és az $(\Omega, \mathscr{F}, \mathbf{P})$ téren értelmezett valószínűségi változót értjük, amelyre

$$\int_F \xi \, d\mathbf{P} = \int_F \mathbf{E}(\xi \mid \mathscr{F}) \, d\mathbf{P}, \quad \forall F \in \mathscr{F}. \tag{B.4.1}$$

Ha $\mathscr{F} \stackrel{\circ}{=} \sigma\left((\eta_\alpha)_{\alpha \in A}\right)$, akkor az $\mathbf{E}(\xi \mid \mathscr{F})$ feltételes várható értéket az egyszerűség kedvéért $\mathbf{E}(\xi \mid \eta_\alpha, \alpha \in A)$ módon szokás jelölni. Speciálisan, ha $\xi = \chi_A$, akkor az $\mathbf{E}(\xi \mid \mathscr{F})$ feltételes várható értéket $\mathbf{P}(A \mid \mathscr{F})$ módon jelöljük, és az A $\mathscr{F}$-re vonatkozó feltételes valószínűségének mondjuk. Ilyenkor a (B.4.1) definiáló egyenlet értelemszerűen

$$\mathbf{P}(A \cap F) = \int_F \mathbf{P}(A \mid \mathscr{F}) \, d\mathbf{P}, \quad \forall F \in \mathscr{F}.$$

B.4.2. Állítás (Feltételes várható érték létezése)**.** *Legyen ξ valószínűségi változó, $\mathscr{F}$ feltételi σ-algebra.*

1. *Ha ξ-nek létezik várható értéke, akkor létezik az $\mathbf{E}(\xi \mid \mathscr{F})$ feltételes várható érték.*
2. *Ha ξ-nek létezik várható értéke, akkor az $\mathbf{E}(\xi \mid \mathscr{F})$ $\mathbf{P}$-majdnem mindenhol véges.*
3. *Ha $\xi \geq 0$, akkor $\mathbf{P}$-majdnem mindenhol $\mathbf{E}(\xi \mid \mathscr{F}) \geq 0$.*
4. *Az $\mathbf{E}(\xi \mid \mathscr{F})$ egy ekvivalenciaosztály és ezért csak egy $(\Omega, \mathscr{F}, \mathbf{P})$-nulla halmaz erejéig egyértelműen meghatározott.*

Bizonyítás: Vegyük észre, hogy az állítás bizonyos elemeit már a definíció során felhasználtuk. Ha ξ-nek van véges várható értéke, akkor a

$$\nu(C) \stackrel{\circ}{=} \int_C \xi \, d\mathbf{P}, \quad C \in \mathscr{F}$$

valós szám értékű előjeles mérték. Ha $\xi \geq 0$, akkor $\nu \geq 0$. Ha μ jelöli a $\mathbf{P}$ leszűkítését az $\mathscr{F}$-re, akkor $\nu \ll \mu$, hiszen ha

$$\mu(C) = 0, \quad C \in \mathscr{F},$$

akkor $\mathbf{P}(C)=0$, amiből $\nu(C)=0$. Az $(\Omega,\mathscr{F})$ mérhető téren vegyük a $d\nu/d\mu$ Radon–Nikodym-féle deriváltat. Legyen

$$\mathbf{E}(\xi \mid \mathscr{F}) \stackrel{\circ}{=} \frac{d\nu}{d\mu}.$$

Ekkor a derivált definíciója alapján érvényes a (B.4.1). Az egyértelműség a $d\nu/d\mu$ egyértelműsége alapján evidens. Ha $\nu \geq 0$, akkor $d\nu/d\mu \geq 0$.

□

B.4.3. Következmény. *Ha $A \in \mathscr{A}$ tetszőleges esemény és $\mathscr{F}$ tetszőleges σ-algebra, akkor létezik, mégpedig nullmértékű halmaztól eltekintve egyetlen $\mathbf{P}(A \mid \mathscr{F})$ feltételes valószínűség*[9].

Térjünk rá a $\mathbf{P}(A \mid \eta=y)$ módon jelölt regresszív, feltételes valószínűségre és az $\mathbf{E}(\xi \mid \eta=y)$ regressziós függvényre. Vegyük észre, hogy a $\mathbf{P}(A \mid \eta=y)$ és a $\mathbf{P}(A \mid \eta)$, illetve az $\mathbf{E}(\xi \mid \eta=y)$ és az $\mathbf{E}(\xi \mid \eta)$ jelölések különböző fogalmakat takarnak. Az első kifejezések az $\mathbb{R}$ egyenesen, a két második kifejezés pedig az Ω absztrakt téren van értelmezve. Kapcsolatukat tisztázza az alábbi definíció, amely kimondása előtt tekintsük a következő egyszerű, bár igen fontos, lemmát:

B.4.4. Lemma (Doob)**.** Legyenek X tetszőleges halmaz, $(Y,\mathscr{Y})$ tetszőleges mérhető tér és $f : X \to Y$ tetszőleges leképezés. Ha $g : X \to \mathbb{R}^n$ $\sigma(f)$-mérhető, akkor létezik[10] $h : (Y,\mathscr{Y}) \to \mathbb{R}^n$ mérhető leképezés, amelyre $g = h \circ f$.

Bizonyítás: Mivel a $\sigma(f)$ éppen az $f^{-1}(A)$, $A \in \mathscr{B}(\mathbb{R}^n)$ alakú halmazokból áll, ezért ha g $\sigma(f)$-mérhető lépcsős függvény, akkor $g = \sum_{k=1}^{n} c_k \chi_{B_k}$, ahol a B_k halmazok mindegyike $f^{-1}(A_k)$, $A_k \in \mathscr{B}(\mathbb{R}^n)$ alakú. Ilyenkor ha $h \stackrel{\circ}{=} \sum_{k=1}^{n} c_k \chi_{A_k}$, akkor

$$h(f(x)) = \sum_{k=1}^{n} c_k \chi_{A_k}(f(x)) = \sum_{k=1}^{n} c_k \chi_{f^{-1}(A_k)}(x) = \sum_{k=1}^{n} c_k \chi_{B_k}(x) = g(x).$$

Mivel minden mérhető függvény lépcsősek határértéke, ezért

$$g = \lim_{n\to\infty} g_n = \lim_{n\to\infty} h_n(f),$$

ahol h_n mérhető. Csekély technikai problémát jelent, hogy a (h_n) sorozata nem feltétlenül konverges. Világos, hogy ha $y \in f(X)$, akkor a $(h_n(y))$ konvergens. Sajnos azonban egy mérhető függvény képhalmaza nem feltétlenül mérhető. Az ebből eredő nehézségeket azonban a $h(y) \stackrel{\circ}{=} \limsup_{n\to\infty} h_n(y)$ definícióval könnyen megkerülhetjük. A h mérhető, és

$$g(x) = \lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} h_n(f(x)) = h(f(x)).$$

□

[9] A feltételes valószínűség ekvivalenciaosztályként egyértelmű.

[10] A h nem egyértelmű.

B.4.5. Definíció. Ha η tetszőleges, akkor az $\mathbf{E}(\xi \mid \eta)$ minden reprezentánsa egy $\sigma(\eta)$ mérhető függvény, így a Doob-lemma alapján alkalmas g függvénnyel $\mathbf{E}(\xi \mid \eta) = g(\eta)$. Ezt a g függvényt $\mathbf{E}(\xi \mid \eta = y)$ módon fogjuk jelölni, és a ξ η-ra vonatkozó regressziós függvényének, vagy a ξ η-ra vonatkozó regresszív, feltételes várható értékének fogjuk mondani. Mivel sem az $\mathbf{E}(\xi \mid \eta)$ sem a Doob-lemmában szereplő függvény nem egyértelmű, ezért az $\mathbf{E}(\xi \mid \eta = y)$, illetve analóg módon a $\mathbf{P}(\xi \mid \eta = y)$ egy függvényosztály.

A feltételes várható érték definíciója miatt

$$\int_F g(\eta)\, d\mathbf{P} \stackrel{\circ}{=} \int_F \mathbf{E}(\xi \mid \eta)\, d\mathbf{P} = \int_F \xi d\mathbf{P}, \quad F \in \sigma(\eta).$$

A generált σ-algebra struktúrája miatt az $F \in \sigma(\eta)$ felírható $F = \eta^{-1}(C)$ módon, ahol $C \in \mathscr{B}(\mathbb{R})$. Ha az η eloszlása G, akkor az általános helyettesítéses integrálás formulája miatt[11] minden $C \in \mathscr{B}(\mathbb{R})$ esetén

$$\begin{aligned} \int_C \mathbf{E}(\xi \mid \eta = y)\, dG(y) &= \int_{\eta^{-1}(C)} \xi d\mathbf{P} = \\ &= \mathbf{E}(\xi \cdot \chi(\eta \in C)). \end{aligned}$$

Könnyen látható, hogy az $\mathbf{E}(\xi \mid \eta = y)$ kifejezés az η eloszlásában nullmértékű halmaz erejéig egyértelműen meghatározott[12].

B.4.6. Állítás (A feltételes várható érték tulajdonságai)**.** *A feltételes várható értékre teljesülnek a következő összefüggések:*

1. *Ha ξ és $\mathscr{F}$ függetlenek, akkor $\mathbf{E}(\xi \mid \mathscr{F}) \stackrel{m.m.}{=} \mathbf{E}(\xi)$, ahol a majdnem mindenhol reláció az $\mathscr{F}$-mérhető függvények körében értendő. Speciálisan, ha $\xi \equiv c$ egy konstans, akkor $\mathbf{E}(\xi \mid \mathscr{F}) \stackrel{m.m.}{=} \mathbf{E}(\xi) = c$. Ha $\mathscr{F} = \{\emptyset, \Omega\}$, akkor $\mathbf{E}(\xi \mid \mathscr{F}) = \mathbf{E}(\xi)$.*

2. *Ha ξ már $\mathscr{F}$-mérhető, akkor $\mathbf{E}(\xi \mid \mathscr{F}) \stackrel{m.m.}{=} \xi$.*

3. *Ha $\xi \leq \eta$, akkor*
$$\mathbf{E}(\xi \mid \mathscr{F}) \stackrel{m.m.}{\leq} \mathbf{E}(\eta \mid \mathscr{F}),$$
speciálisan
$$0 \stackrel{m.m.}{\leq} \mathbf{P}(A \mid \mathscr{F}) \stackrel{m.m.}{\leq} 1.$$

4. *Tetszőleges a konstansra $\mathbf{E}(a\xi \mid \mathscr{F}) \stackrel{m.m.}{=} a\mathbf{E}(\xi \mid \mathscr{F})$.*

[11] V.ö.: (B.1.1) sor, 241. oldal.
[12] V.ö.: 6.1.4. állítás, 91. oldal.

5. *Ha az* $\mathbf{E}(\xi)+\mathbf{E}(\eta)$ *kifejezés értelmes, vagyis* ξ*-nek és* η*-nak van véges várható értéke, akkor tetszőleges* $\mathscr{F}$ σ*-algebrára nézve a* $\xi+\eta$*-nak is van feltételes várható értéke és*

$$\mathbf{E}(\xi+\eta \mid \mathscr{F}) \overset{m.m.}{=} \mathbf{E}(\xi \mid \mathscr{F})+\mathbf{E}(\eta \mid \mathscr{F}).$$

Speciálisan, ha A *és* B *diszjunkt halmazok, akkor*

$$\mathbf{P}(A\cup B \mid \mathscr{F}) \overset{m.m.}{=} \mathbf{P}(A \mid \mathscr{F})+\mathbf{P}(B \mid \mathscr{F}),$$

amiből, ha $A\subseteq B$, *akkor*

$$\mathbf{P}(B\setminus A \mid \mathscr{F}) \overset{m.m.}{=} \mathbf{P}(B \mid \mathscr{F})-\mathbf{P}(A \mid \mathscr{F}).$$

6. $|\mathbf{E}(\xi \mid \mathscr{F})| \overset{m.m.}{\leq} \mathbf{E}(|\xi| \mid \mathscr{F})$.

7. *Teljesül a teljes várható érték tétel, vagyis ha* $\mathscr{G}\subseteq\mathscr{F}$, *akkor*

$$\mathbf{E}(\mathbf{E}(\xi \mid \mathscr{F}) \mid \mathscr{G}) \overset{m.m.}{=} \mathbf{E}(\xi \mid \mathscr{G}).$$

Speciálisan

$$\mathbf{E}(\mathbf{E}(\xi \mid \mathscr{F}))=\mathbf{E}(\xi).$$

8. *Ha* $\mathscr{G}\subseteq\mathscr{F}$, *akkor*

$$\mathbf{E}(\mathbf{E}(\xi \mid \mathscr{G}) \mid \mathscr{F}) \overset{m.m.}{=} \mathbf{E}(\xi \mid \mathscr{G}).$$

9. *A feltételes várható értékre teljesül a monoton konvergencia tétel, vagyis ha*

$$0\leq \xi_n \nearrow \xi,$$

akkor majdnem mindenhol

$$0 \overset{m.m.}{\leq} \mathbf{E}(\xi_n \mid \mathscr{F}) \overset{m.m.}{\nearrow} \mathbf{E}(\xi \mid \mathscr{F}).$$

Speciálisan, ha $A_n \nearrow A$, *akkor*

$$\mathbf{P}(A_n \mid \mathscr{F}) \overset{m.m.}{\nearrow} \mathbf{P}(A \mid \mathscr{F}),$$

illetve ha $\xi_n \geq 0$, *akkor*

$$\sum_{n=1}^{\infty}\mathbf{E}(\xi_n \mid \mathscr{F}) \overset{m.m.}{=} \mathbf{E}\left(\sum_{n=1}^{\infty}\xi_n \mid \mathscr{F}\right),$$

vagyis ha az A_n *halmazok páronként diszjunktak és* $A=\cup_n A_n$, *akkor*

$$\sum_{n=1}^{\infty}\mathbf{P}(A_n \mid \mathscr{F}) \overset{m.m.}{=} \mathbf{P}(A \mid \mathscr{F}).$$

10. *Teljesül a majorált konvergencia tétel. Ha*

$$\lim_{n\to\infty} \xi_n = \xi, \quad \textit{és} \quad |\xi_n| \leq \eta,$$

ahol $\mathbf{E}(\eta) < \infty$*, akkor*

$$\lim_{n\to\infty} \mathbf{E}(|\xi_n - \xi| \mid \mathscr{F}) \overset{m.m.}{=} 0, \quad \lim_{n\to\infty} \mathbf{E}(\xi_n \mid \mathscr{F}) \overset{m.m.}{=} \mathbf{E}(\xi \mid \mathscr{F}).$$

Speciálisan, ha $A_n \searrow A$*, akkor*

$$\lim_{n\to\infty} \mathbf{P}(A_n \mid \mathscr{F}) \overset{m.m.}{=} \mathbf{P}(A \mid \mathscr{F}).$$

11. *Érvényes a kiemelési szabály, vagyis ha az* η *változó* $\mathscr{F}$*-mérhető, és léteznek az* $\mathbf{E}(\xi)$ *és az* $\mathbf{E}(\eta\xi)$*, véges várható értékek, akkor*

$$\mathbf{E}(\eta\xi \mid \mathscr{F}) \overset{m.m.}{=} \eta\mathbf{E}(\xi \mid \mathscr{F}).$$

Speciálisan az η *korlátos és* $\mathscr{F}$*-mérhető és az* $\mathbf{E}(\xi)$ *véges, akkor az* η *kiemelhető a feltételes várható értékből.*

Bizonyítás: A bizonyítás általában a definíció és a megfelelő „klasszikus” állítás közvetlen következménye.

1. Az $\mathbf{E}(\xi)$ konstans függvény $\mathscr{F}$-mérhető, ugyanakkor a ξ és az $\mathscr{F}$ függetlensége miatt minden $F \in \mathscr{F}$ halmazra triviálian

$$\int_F \xi d\mathbf{P} = \mathbf{E}(\chi_F \xi) = \mathbf{E}(\chi_F)\mathbf{E}(\xi) = \int_F \mathbf{E}(\xi) d\mathbf{P}.$$

2. Egyrészt evidens módon a ξ jó $\mathbf{E}(\xi \mid \mathscr{F})$-nek, másrészt az $\mathbf{E}(\xi \mid \mathscr{F})$ egy nulla halmaz erejéig egyértelmű.

3. Ha $\xi \leq \eta$ és $F \in \mathscr{F}$, akkor

$$\int_F \mathbf{E}(\xi \mid \mathscr{F}) d\mathbf{P} = \int_F \xi d\mathbf{P} \leq \int_F \eta d\mathbf{P} = \int_F \mathbf{E}(\eta \mid \mathscr{F}) d\mathbf{P},$$

ami alapján $\mathbf{E}(\xi \mid \mathscr{F}) \overset{m.m.}{\leq} \mathbf{E}(\eta \mid \mathscr{F})$.

4. Ha $F \in \mathscr{F}$, akkor

$$\int_F a\mathbf{E}(\xi \mid \mathscr{F})\,d\mathbf{P} = a\int_F \mathbf{E}(\xi \mid \mathscr{F})\,d\mathbf{P} = a\int_F \xi\,d\mathbf{P} = \int_F a\xi\,d\mathbf{P}.$$

Az $a\mathbf{E}(\xi \mid \mathscr{F})$ $\mathscr{F}$-mérhető, ezért az $a\xi$-nek is létezik $\mathscr{F}$-feltételes várható értéke, ami értelemszerüen éppen az $a\mathbf{E}(\xi \mid \mathscr{F})$, következésképpen

$$a\mathbf{E}(\xi \mid \mathscr{F}) = \mathbf{E}(a\xi \mid \mathscr{F}).$$

5. Felhasználva, hogy a feltételes várható érték pontosan olyan értelemben integrálható az $F \in \mathscr{F}$ halmazokon mint az eredeti változó, az integrál additívása miatt

$$\begin{aligned}
\int_F (\xi+\eta)\,d\mathbf{P} &= \int_F \xi\,d\mathbf{P} + \int_F \eta\,d\mathbf{P} = \\
&= \int_F \mathbf{E}(\xi \mid \mathscr{F})\,d\mathbf{P} + \int_F \mathbf{E}(\eta \mid \mathscr{F})\,d\mathbf{P} = \\
&= \int_F (\mathbf{E}(\xi \mid \mathscr{F}) + \mathbf{E}(\eta \mid \mathscr{F}))\,d\mathbf{P},
\end{aligned}$$

amiből, mivel az $\mathbf{E}(\xi \mid \mathscr{F}) + \mathbf{E}(\eta \mid \mathscr{F})$ összeg $\mathscr{F}$-mérhető, a feltételes várható érték majdnem mindenhol való egyértelműsége alapján

$$\mathbf{E}(\xi+\eta \mid \mathscr{F}) \overset{m.m.}{=} \mathbf{E}(\xi \mid \mathscr{F}) + \mathbf{E}(\eta \mid \mathscr{F}).$$

6. Mivel $-\xi, \xi \leq |\xi|$, ezért a 3. és 4. pontok alapján evidens.
7. Mivel $\mathscr{G} \subseteq \mathscr{F}$, ezért ha $G \in \mathscr{G}$, akkor egyúttal $G \in \mathscr{F}$ és így az $\mathbf{E}(\mathbf{E}(\xi \mid \mathscr{F}) \mid \mathscr{G})$ és az $\mathbf{E}(\xi \mid \mathscr{F})$ definíciója szerint

$$\int_G \mathbf{E}(\mathbf{E}(\xi \mid \mathscr{F}) \mid \mathscr{G})\,d\mathbf{P} = \int_G \mathbf{E}(\xi \mid \mathscr{F})\,d\mathbf{P} = \int_G \xi\,d\mathbf{P}.$$

De mivel az $\mathbf{E}(\mathbf{E}(\xi \mid \mathscr{F}) \mid \mathscr{G})$ $\mathscr{G}$-mérhető, és a feltételes várható érték majdnem mindenhol egyértelmű, ezért

$$\mathbf{E}(\mathbf{E}(\xi \mid \mathscr{F}) \mid \mathscr{G}) \overset{m.m.}{=} \mathbf{E}(\xi \mid \mathscr{G}).$$

8. A 2. alapján evidens, ugyanis a $\mathscr{G} \subseteq \mathscr{F}$ feltétel miatt az $\mathbf{E}(\xi \mid \mathscr{G})$ $\mathscr{F}$-mérhető.

9. A feltételes várható érték majdnem mindenhol őrzi a rendezést, ezért az $\mathbf{E}(\xi_n \mid \mathscr{F})$ sorozat majdnem mindenhol monoton nő, ezért majdnem minden kimenetre létezik a

$$\zeta \overset{\circ}{=} \lim_{n\to\infty} \mathbf{E}(\xi_n \mid \mathscr{F})$$

határérték. Tetszőleges $F \in \mathscr{F}$ halmazra definíció szerint

$$\int_F \mathbf{E}(\xi_n \mid \mathscr{F})\,d\mathbf{P} = \int_F \xi_n\,d\mathbf{P}.$$

A klasszikus monoton konvergencia tétel alapján

$$\begin{aligned}\int_F \zeta d\mathbf{P} &= \int_F \lim_{n\to\infty}\mathbf{E}(\xi_n \mid \mathscr{F})\,d\mathbf{P} = \lim_{n\to\infty}\int_F \mathbf{E}(\xi_n \mid \mathscr{F})\,d\mathbf{P} = \\ &= \lim_{n\to\infty}\int_F \xi_n d\mathbf{P} = \int_F \lim_{n\to\infty}\xi_n d\mathbf{P} = \int_F \xi d\mathbf{P},\end{aligned}$$

vagyis

$$\lim_{n\to\infty}\mathbf{E}(\xi_n \mid \mathscr{F}) \overset{\circ}{=} \zeta \overset{m.m.}{=} \mathbf{E}(\xi \mid \mathscr{F}).$$

10. Szemben a monoton konvergencia tétel igazolásával, nem evidens, hogy a (ξ_n) konvergenciájából következik az $(\mathbf{E}(\xi_n \mid \mathscr{F}))$ konvergenciája, ezért a bizonyítás valamelyest nehézkesebb. Ha

$$\zeta_n \overset{\circ}{=} \sup_{m\geq n}|\xi_m - \xi|,$$

akkor a $\xi_n \to \xi$ miatt

$$\zeta_n \searrow 0,$$

és mivel a feltételes várható érték tartja a rendezést, ezért majdnem mindenhol létezik a

$$\psi \overset{\circ}{=} \lim_{n\to\infty}\mathbf{E}(\zeta_n \mid \mathscr{F}) \overset{m.m.}{\geq} 0$$

határérték. Mivel

$$0 \overset{m.m.}{\leq} \zeta_n \overset{m.m.}{\leq} 2\eta,$$

ezért

$$0 \overset{m.m.}{\leq} \mathbf{E}(\zeta_n \mid \mathscr{F}) \overset{m.m.}{\leq} \mathbf{E}(2\eta \mid \mathscr{F}).$$

A feltétel miatt a 2η és az $\mathbf{E}(2\eta \mid \mathscr{F})$ integrálható, így a klasszikus majorált konvergencia tétel kétszeri alkalmazásával

$$\begin{aligned}\int_\Omega \psi d\mathbf{P} &\overset{\circ}{=} \int_\Omega \lim_{n\to\infty}\mathbf{E}(\zeta_n \mid \mathscr{F})\,d\mathbf{P} = \lim_{n\to\infty}\int_\Omega \mathbf{E}(\zeta_n \mid \mathscr{F})\,d\mathbf{P} \overset{\circ}{=} \\ &\overset{\circ}{=} \lim_{n\to\infty}\int_\Omega \zeta_n d\mathbf{P} = \int_\Omega \lim_{n\to\infty}\zeta_n d\mathbf{P} = 0,\end{aligned}$$

amiből $\psi \overset{m.m.}{=} 0$. Felhasználva, hogy a feltételes várható érték rendezéstartó, és mivel integrálható függvényekről van szó, ezért additív

$$\begin{aligned}0 &\overset{m.m.}{\leq} \limsup_{n\to\infty}|\mathbf{E}(\xi_n \mid \mathscr{F}) - \mathbf{E}(\xi \mid \mathscr{F})| \overset{m.m.}{=} \limsup_{n\to\infty}|\mathbf{E}(\xi_n - \xi \mid \mathscr{F})| \overset{m.m.}{\leq} \\ &\overset{m.m.}{\leq} \limsup_{n\to\infty}\mathbf{E}(|\xi_n - \xi| \mid \mathscr{F}) \overset{m.m.}{\leq} \limsup_{n\to\infty}\mathbf{E}(\zeta_n \mid \mathscr{F}) \overset{m.m.}{=} 0.\end{aligned}$$

11. Első lépésben $\eta \stackrel{\circ}{=} \chi_G$. Az η $\mathscr{F}$-mérhető, ezért $G \in \mathscr{F}$. Ha $F \in \mathscr{F}$ tetszőleges, akkor felhasználva, hogy a ξ-nek létezik feltételes várható értéke, és hogy $G \cap F \in \mathscr{F}$

$$\begin{aligned}\int_F \eta\xi d\mathbf{P} &\stackrel{\circ}{=} \int_F \xi\chi_G d\mathbf{P} \stackrel{\circ}{=} \int_{G\cap F} \xi d\mathbf{P} \stackrel{\circ}{=} \int_{G\cap F} \mathbf{E}(\xi \mid \mathscr{F})\, d\mathbf{P} \stackrel{\circ}{=} \\ &\stackrel{\circ}{=} \int_F \chi_G \mathbf{E}(\xi \mid \mathscr{F})\, d\mathbf{P} \stackrel{\circ}{=} \int_F \eta\mathbf{E}(\xi \mid \mathscr{F})\, d\mathbf{P},\end{aligned}$$

amiből

$$\mathbf{E}(\eta\xi \mid \mathscr{F}) = \eta\mathbf{E}(\xi \mid \mathscr{F}).$$

Legyen $\xi \geq 0$. A feltételes várható érték linearitása szerint az egyenlőség érvényben marad ha az $\eta \geq 0$ lépcsős függvény. Ha $0 \leq \eta_n \nearrow \eta$, akkor $0 \leq \xi\eta_n \nearrow \xi\eta$, és a monoton konvergencia tétel alapján

$$\begin{aligned}\mathbf{E}(\eta\xi \mid \mathscr{F}) &= \mathbf{E}\left(\lim_{n\to\infty} \eta_n\xi \mid \mathscr{F}\right) \stackrel{m.m.}{=} \lim_{n\to\infty} \mathbf{E}(\eta_n\xi \mid \mathscr{F}) \stackrel{m.m.}{=} \\ &= \lim_{n\to\infty} \eta_n \mathbf{E}(\xi \mid \mathscr{F}) = \eta\mathbf{E}(\xi \mid \mathscr{F}).\end{aligned}$$

Feltettük, hogy az $\mathbf{E}(\xi)$ várható érték létezik, ezért a B.4.2. állítás miatt

$$\mathbf{E}(\xi \mid \mathscr{F}) = \mathbf{E}\left(\xi^+ \mid \mathscr{F}\right) - \mathbf{E}\left(\xi^- \mid \mathscr{F}\right). \tag{B.4.2}$$

Legyen $\eta \geq 0$ és ξ tetszőleges. A feltétel szerint az η valószínűségi változó, tehát definíció szerint mindig véges, ezért a (B.4.2) egyenlőség beszorozható η-val:

$$\begin{aligned}\eta\mathbf{E}(\xi \mid \mathscr{F}) &= \eta\mathbf{E}\left(\xi^+ \mid \mathscr{F}\right) - \eta\mathbf{E}\left(\xi^- \mid \mathscr{F}\right) \stackrel{m.m.}{=} \\ &= \mathbf{E}\left(\eta\xi^+ \mid \mathscr{F}\right) - \mathbf{E}\left(\eta\xi^- \mid \mathscr{F}\right) = \\ &= \mathbf{E}\left((\eta\xi)^+ \mid \mathscr{F}\right) - \mathbf{E}\left((\eta\xi)^- \mid \mathscr{F}\right) \stackrel{m.m.}{=} \mathbf{E}(\eta\xi \mid \mathscr{F}),\end{aligned}$$

ahol az utolsó egyenlőségnél felhasználtuk, hogy az $\eta\xi$-nek van várható értéke. Legyen végül η tetszőleges előjelü. A feltétel szerint létezik az $\mathbf{E}(\eta\xi)$ várható érték.

$$\mathbf{E}(\eta\xi \mid \mathscr{F}) = \mathbf{E}\left(\eta^+\xi^+ + \eta^-\xi^- - \eta^+\xi^- - \eta^-\xi^+ \mid \mathscr{F}\right).$$

Ahhoz, hogy a várható értéket szét lehessen szedni elegendő, ha a négy összeadandó pozitív részének integrálja véges. Ez azonban teljesül, hiszen például

$$\mathbf{E}\left(\eta^+\xi^+\right) \leq \mathbf{E}\left(\eta^+\xi^+ + \eta^-\xi^-\right) = \mathbf{E}\left((\xi\eta)^+\right) < \infty,$$

illetve

$$\mathbf{E}\left(\left(-\eta^+\xi^-\right)^+\right) = 0 < \infty.$$

Ismételten felhasználva, hogy az η véges, illetve, hogy érvényes a (B.4.2) egyenlőség

$$\begin{aligned}\mathbf{E}(\eta\xi \mid \mathscr{F}) &= \sum \pm\mathbf{E}\left(\eta^\pm\xi^\pm \mid \mathscr{F}\right) = \sum \pm\eta^\pm\mathbf{E}\left(\xi^\pm \mid \mathscr{F}\right) = \\ &= \eta^+\mathbf{E}(\xi \mid \mathscr{F}) - \eta^-\mathbf{E}(\xi \mid \mathscr{F}) = \eta\mathbf{E}(\xi \mid \mathscr{F}).\end{aligned}$$

□

C.

KORLÁTLAN OSZTHATÓSÁG

Ebben a függelékben egy az anyaghoz szorosan kapcsolódó, bár azon erősen túlmutató kérdést vizsgálunk meg: A fejezet legfőbb állítása a következő:

C.0.1. Tétel (Thorin). *A lognormális, a másodfajú béta és a Student t_n eloszlások korlátlanul oszthatóak.*

Miként látni fogjuk, a tétel igazolása igen impozáns és alapos matematikai ismereteket tételez fel. Természetesen ez a függelék nem része az elemi valószínűségszámításnak, de jól mutatja, hogy a haladottabb valószínűségszámítás mennyire nehéz és szerteágazó matematikai eszközökre épül. A tárgyalás során nem a megközelítés nehézségét, hanem a szükséges háttérismeretek kiterjedtségét hangsúlyozom. A legfontosabb felhasznált terület a komplex változós analízis, pontosabban a komplex Laplace-transzformáció. Célom természetesen nem az olvasó elriasztása, hanem annak bemutatása, hogy a haladó valószínűségszámítás megértéséhez, komoly matematikai előkészületek szükségesek. Az elemi valószínűségszámítás keretében valamely eloszlás korlátlanul való oszthatóságát a legegyszerűbben úgy tudjuk igazolni, ha közvetlenül megmutatjuk, hogy bizonyos eloszlástípusok összege is a megadott típusba tartozik. Ezt követtük a gamma és Poisson-eloszlások esetén. Egy másik lehetőség, ha tekintjük az eloszlás karakterisztikus függvényét és annak vesszük az n-edik gyökét, amelyről belátjuk, hogy szintén valamilyen eloszlás karakterisztikus függvénye.

C.0.2. Példa. Vizsgáljuk meg a Cauchy-eloszlás karakterisztikus függvényét.

A karakterisztikus függvény képlete alapján

$$\begin{aligned}\varphi(t) &\overset{\circ}{=} \frac{1}{\pi}\int_{\mathbb{R}} \frac{1}{1+x^2}(\cos tx + i\sin tx)\,dx = \\ &= \frac{1}{\pi}\int_{\mathbb{R}} \frac{1}{1+x^2}\cos tx dx = \\ &= \frac{2}{\pi}\int_0^{\infty} \frac{1}{1+x^2}\cos tx dx.\end{aligned}$$

Mivel

$$\varphi(t) = \varphi(-t)$$

ezért elég a $t > 0$ esettel foglalkozni. Szemben a normális eloszlás esetével most nem tudunk t szerint az integrál mögött deriválni, mert a formális deriváltnak nincsen integrálható majoránsa. Vezessük be a

$$\varphi(t,\alpha) \overset{\circ}{=} \frac{2}{\pi}\int_0^{\infty} \frac{\exp(-\alpha x)}{1+x^2}\cos tx dx$$

paraméteres integrált. Mivel a $\cos tx/\left(1+x^2\right)$ integrálható, ezért

$$\lim_{\alpha\searrow 0} \varphi(t,\alpha) = \varphi(t).$$

Ha $\alpha > 0$, akkor t szerint az integrál alatt deriválhatunk

$$\varphi'(t,\alpha) = -\frac{2}{\pi}\int_0^\infty x\frac{\exp(-\alpha x)}{1+x^2}\sin txdx.$$

Újra deriválva, elemi átalakításokkal

$$\begin{aligned}\varphi''(t,\alpha) &= -\frac{2}{\pi}\int_0^\infty x^2\frac{\exp(-\alpha x)}{1+x^2}\cos txdx = \\ &= -\frac{2}{\pi}\int_0^\infty \left(x^2+1-1\right)\frac{\exp(-\alpha x)}{1+x^2}\cos txdx = \\ &= -\frac{2}{\pi}\int_0^\infty \exp(-\alpha x)\cos txdx + \varphi(t,\alpha).\end{aligned}$$

Az exponeciális eloszlás karakterisztikus függvényének képlete alapján

$$\begin{aligned}\varphi''(t,\alpha) &= -\frac{1}{\alpha}\frac{2}{\pi}\int_0^\infty \alpha\exp(-\alpha x)\cos txdx + \varphi(t,\alpha) = \\ &= -\frac{1}{\alpha}\frac{2}{\pi}\int_0^\infty \alpha\exp(-\alpha x)\cos txdx + \varphi(t,\alpha) = \\ &= -\frac{2}{\pi}\frac{\alpha}{\alpha^2+t^2} + \varphi(t,\alpha).\end{aligned}$$

Összefoglalva:

$$\varphi''(t,\alpha) - \varphi(t,\alpha) = -\frac{2}{\pi}\frac{\alpha}{\alpha^2+t^2}. \tag{C.0.1}$$

Természetesen a fenti (C.0.1) egyenletet közvetlenül is megoldhatjuk az állandók variálásának módszerével. Majd $\alpha \to 0$ adja a φ függvényt. Talán valamivel egyszerűbb, ha a két oldalt egymás után kétszer integráljuk:

$$\int_0^s \varphi''(r,\alpha)dr - \int_0^s \varphi(r,\alpha)dr = \int_0^s -\frac{2}{\pi}\frac{\alpha}{\alpha^2+r^2}dr.$$

Az első integrál

$$\int_0^s \varphi''(r,\alpha)dr = \varphi'(s,\alpha) - \varphi'(0,\alpha).$$

De

$$\varphi'(0,\alpha) = -\frac{2}{\pi}\int_0^\infty x\frac{\exp(-\alpha x)}{1+x^2}\sin 0xdx = 0,$$

így újra integrálva

$$\int_0^t\int_0^s \varphi''(r,\alpha)drds = \int_0^t \varphi'(s,\alpha)ds.$$

Ismételten az integrálást elvégezve:

$$\int_0^t \varphi'(s,\alpha)ds = \varphi(t,\alpha) - \varphi(0,\alpha).$$

Tehát

$$\begin{aligned}\varphi(t,\alpha)-\varphi(0,\alpha)-\int_0^t\int_0^s\varphi(r,\alpha)\,drds&=\\=\int_0^t\int_0^s-\frac{2}{\pi}\frac{\alpha}{\alpha^2+r^2}drds&=\\=-\frac{2}{\pi}\int_0^t\left[\arctan\left(\frac{r}{a}\right)\right]_0^s ds&=\\=-\frac{2}{\pi}\int_0^t\arctan\left(\frac{s}{a}\right)ds.\end{aligned}$$

Ha $\alpha\to 0$, akkor $\arctan(s/a)\to\pi/2$, tehát

$$-\frac{2}{\pi}\int_0^t\arctan\left(\frac{s}{a}\right)ds\to -t,$$

tehát

$$\varphi(t,0)-\varphi(0,0)-\int_0^t\int_0^s\varphi(r,0)\,drds=-t.$$

Kétszer deriválva $\varphi''-\varphi=0$, ami már megoldható ugyanis egy másodrendű konstans együtthatós egyenlet. Az egyenlet két karakterisztikus gyöke $\lambda=\pm1$, vagyis az általános megoldás $c_1\exp(t)+c_2\exp(-t)$. Mivel $\varphi(0)=1$ és $|\varphi|\le 1$ ezért

$$\varphi(t)=\exp(-t),\quad t\ge 0.$$

Vagyis

$$\varphi(t)=\exp(-|t|).$$

Ennek n-edik gyöke

$$\varphi_n(t)=\sqrt[n]{\varphi(t)}=\exp\left(-\frac{|t|}{n}\right),$$

amely nyilvánvalóan szintén egy karakterisztikus függvény.

□

A karakterisztikus függvények azonban általában komplex értékű függvények[1], így a gyökvonás nem feltétlenül triviális művelet, ezért általában egyszerűbb, ha valamilyen valós transzformálttal dolgozunk. Ezt nem negatív változók, mint például a lognormális, vagy a másodfajú béta eloszlás esetén mindig megtehetjük. A gond természetesen az, hogy az eloszlások Laplace-transzformáltja közvetlenül nem mindig adható

[1] A Cauchy-eloszlás szimmetrikus, így a karakterisztikus függvénye valós. De sem a lognormális, sem a másodfajú béta eloszlás karakterisztikus függvénye nem lesz valós.

meg egyszerű képlettel. Nem beszélve arról, hogy a gyökvonás eredményeként kapott függvényről sem lehet esetleg könnyen igazolni, hogy az egy eloszlás Laplace-transzformáltja. Ez a módszer például a gamma eloszlásra működik, de például a lognormális eloszlás esetén nem tűnik járhatónak. Végezetül érdemes még egy megjegyzést tenni. Az exponenciális eloszlás példája mutatja, hogy valamely korlátlanul osztható eloszlás „feltördelése" nem feltétlenül az eredetivel azonos típusú eloszlást eredményez. Vagyis a lognormális eloszlás korlátlanul való oszthatósága nem jelenti azt, hogy az eloszlást lognormálisokra lehet felosztani. A korlátlanul való oszthatóság azért fontos, mert ezek azok az eloszlások, amelyek Lévy-folyamatok eloszlásaként előállnak. A lognormális eloszlás esetén ez azt jelent, hogy van olyan X Lévy-folyamat, amelyre $X(1)$ eloszlása lognormális. De ebből nem következik az, hogy a $t = 1$ ponton kívül bármelyik t pontban az $X(t)$ eloszlása szintén lognormális lenne.

C.1. A komplex Laplace-transzformált

Miként említettük a fejezetben használt fő eszköz a komplex Laplace-transzformáció. Mielőtt ennek tárgyalását elkezdenénk, röviden emlékeztetek a komplex változós függvények néhány elemi tulajdonságára. Jelölje $\mathbb{C}$ a komplex számok halmazát. A $\mathbb{C}$ valamely részhalmazát a $\mathbb{C}$ egy részhalmazába képező f függvényeket komplex függvényeknek nevezzük. Az f értelmezési tartományának geometriai, topológiai tulajdonságai igen fontos szerepet játszanak, ezért azokat célszerű már a tárgyalás elején explicite rögzíteni. Mivel deriválható függvényekről fogunk beszélni, ezért az értelmezési tartományról mindig feltesszük, hogy nyílt. Ugyanakkor a valós esetben a függvények általában intervallumokon vannak értelmezve, amelyeknek a komplex síkban való megfelelőjét nem is olyan egyszerű megmondani. Mivel nem célunk túl általánosan fogalmazni ezért minden további említés nélkül feltesszük, hogy az értelmezési tartomány mindig csillagszerű. Ez alatt azt értjük, hogy van egy olyan pont, amiből minden másik pont is látható, vagyis amely a halmaz minden másik pontjával összeköthető egy a halmazban haladó egyenessel. A legegyszerűbb komplex változós függvények a polinomok, illetve a hatványsorok. A hatványsorok közül is kiemelkedik az $\exp(z) \mathrel{\hat{=}} \Sigma_{n=0}^{\infty} z^n/n!$ komplex exponenciális függvény. Miként az analízisből ismert, ha $z = x + iy$, akkor az Euler-képlet szerint

$$\exp(z) = e^x (\cos y + i \sin y)$$

és az $\exp(z)$ függvény értelmezési tartománya a teljes komplex sík az értékkészlete pedig a $\mathbb{C} \setminus \{0\}$ halmaz. Nem túl meglepő módon szükségünk lesz a exponenciális függvény inverzére, a $\log z$ módon jelölt komplex logaritmusfüggvényre. Mivel az értelmezési tartományt csillagszerűnek akarjuk, és mivel a $z = 0$ nem lehet a $\log(z)$ függvény értelmezési tartományában, ezért a $\log z$ értelmezési tartományából ki kell zárni a pozitív számokból nem látható pontokat, vagyis a valós tengely negatív részét, természetesen beleértve a nulla pontot is. Érdemes megjegyezni, hogy a sík felvágása miatt a pontokhoz tartozó φ szögeket a $-\pi < \varphi < \pi$ intervallumba korlátozzuk. Vagyis a teljes

$\mathbb{C}$ mellett értelmezési tartományként legtöbbször a $\mathbb{C}\setminus(-\infty,0]$ halmazzal fogunk találkozni. Könnyen látható, hogy ha a z pontot (ρ,φ) módon polárkoordinátákkal írjuk fel, akkor $\log(z)=\ln\rho+i\varphi$, ugyanis ezzel

$$\begin{aligned}\exp(\log z) &= e^{\ln\rho}(\cos\varphi+i\sin\varphi)= \\ &= \rho(\cos\varphi+i\sin\varphi)=z.\end{aligned}$$

Vegyük észre, hogy a sík felvágása miatt a most definiált logaritmusfüggvény folytonos, de a teljes síkra nem terjeszthető ki folytonos függvényként. A logaritmusfüggvény segítségével egy sor a valós analízisben ismert függvény komplex kiterjesztését adhatjuk meg. Ezek közé tartoznak például a $z^{\alpha}\stackrel{\circ}{=}\exp(\alpha\log z)$ alakú hatványfüggvények. Polárkoordinátákban felírva

$$\begin{aligned}z^{\alpha} &= \exp(\alpha\log z)=\exp(\alpha(\ln\rho+i\varphi))= \\ &= e^{\alpha\ln\rho}(\cos\alpha\varphi+i\sin\alpha\varphi)= \\ &= \rho^{\alpha}(\cos\alpha\varphi+i\sin\alpha\varphi).\end{aligned}$$

Komplex függvények deriválása

A komplex függvények folytonosságának definíciója megegyezik az $\mathbb{R}^2$-ből $\mathbb{R}^2$-be képező függvények folytonosságának definíciójával. A deriválhatóság definíciója azonban jelentős mértékben eltér.

C.1.1. Definíció. Egy f komplex változós függvényt egy z pontban deriválhatónak mondunk, ha egyrészt deriválható a z pontban mint a síkot a síkba képező függvény, másrészt a deriváltmátrix $\begin{pmatrix} a & -b \\ b & a\end{pmatrix}$ alakú. A mátrixszorzás szabályai alapján könnyen látható, hogy az ilyen alakú mátrixokkal való szorzás azonosítható az $f'(z)$ módon jelölt $a+bi$ komplex számmal való szorzással. Ennek következtében a komplex derivált éppen a

$$\lim_{h\to 0}\frac{f(z+h)-f(z)}{h}=f'(z) \tag{C.1.1}$$

határértékkel egyezik meg.

A komplex analízis kiinduló észrevétele az, hogy derivált (C.1.1) definíciója miatt a valós analízisben tanult összes deriválási szabály változtatás nélkül formálisan érvényben marad.

C.1.2. Példa. Számoljuk ki a logaritmusfüggvény deriváltját.

Miként ismert, ha az invertálható, valós-valós f függvény differenciálható az értelmezési tartománya egy u pontjában és f^{-1} folytonos az $f(u)$-ban, továbbá $f'(u) \neq 0$, akkor f^{-1} differenciálható $f(u)$-ban és

$$\left(f^{-1}\right)'(f(u)) = \frac{1}{f'(u)}.$$

Könnyen ellenőrizhető, hogy ennek igazolása csak a derivált definíciójára épül és a komplex esetre szó szerint átvihető. Ebből következően, felhasználva, hogy a komplex exponenciális függvény deriváltja önmaga, ha $z = \exp(u)$, akkor

$$\frac{d}{dz}\log z = \frac{1}{\exp(u)} = \frac{1}{\exp(\log z)} = \frac{1}{z}.$$

□

C.1.3. Példa. Számoljuk ki a z^{α} komplex hatványfüggvények deriváltját.

Legyen $\alpha \neq 0$. A $z^{\alpha} \mathrel{\circeq} \exp(\alpha \log z)$ definíció és az összetett függvény deriválási szabálya alapján ha $z \notin \mathbb{C} \setminus (-\infty, 0]$

$$\begin{aligned}
\frac{d}{dz} z^{\alpha} &\circeq \frac{d}{dz}\exp(\alpha \log z) = \exp(\alpha \log z)\,\alpha \frac{d}{dz}\log z = \\
&= \alpha \exp(\alpha \log z)\frac{1}{z} = \alpha \exp(\alpha \log z)\exp(-\log z) = \\
&= \alpha \exp((\alpha - 1)\log z) = \alpha z^{\alpha - 1}.
\end{aligned}$$

Érdemes hangsúlyozni, hogy a z^{α} nincs a $z = 0$ pontban definiálva. Ez nem meglepő, ugyanis például a valós $y = \sqrt{x}$ az $x = 0$ pontban nem deriválható, így a komplex definíció bizonyos értelemben szerencsésebb.

□

Komplex változós függvények integrálása

A deriválással szemben a komplex változós integrálás lényegesen eltér a valós esettől. Az első szembetűnő probléma, hogy két komplex szám, a és b esetén önmagában értelmetlen az $\int_a^b f(z)\,dz$ szimbólum. Ehelyett csak görbe menti integrálokról beszélhetünk. Mivel nem célunk egy teljes komplex változós analízist felépíteni, ezért csak a következő egyszerűsített definícióval élünk:

C.1.4. Definíció. Legyen $\gamma(t)$ valamely $[a,b]$ valós intervallumot a komplex számokba képező folytonos és szakaszonként folytonosan deriválható görbe. Jelölje ekkor

$$\int_{\gamma} f(z)\,dz \circeq \int_a^b f(\gamma(t))\,\gamma'(t)\,dt.$$

Bár az integrálás fenti definíciója számunkra teljesen kielégítő mégis érdemes röviden a tartalmát megvilágítani. Legyen f egy folytonos komplex függvény és γ egy az $[a,b]$ valós intervallumot a komplex számokba képező görbe. Ha tekintjük a szakasz egy felosztását, akkor képezhetjük a

$$\sum_k f(\gamma(\tau_k))(\gamma(t_k)-\gamma(t_{k-1})) \stackrel{\circ}{=} \sum_k f_k \Delta\gamma_k$$

alakú közelítő összegeket. A komplex számok szorzási szabálya miatt ennek valós része

$$\sum_k \operatorname{Re} f_k \operatorname{Re}\Delta\gamma_k - \sum_k \operatorname{Im} f_k \operatorname{Im}\Delta\gamma_k,$$

imaginárius része pedig

$$\sum_k \operatorname{Re} f_k \operatorname{Im}\Delta\gamma_k + \sum_k \operatorname{Im} f_k \operatorname{Re}\Delta\gamma_k.$$

Ez négy darab Stieltjes-integrált definiál. Ennek megfelelően a közelítő összegek határértékeként definiált integrál valós része

$$\int_a^b \operatorname{Re} f(\gamma(t))\, d\operatorname{Re}\gamma(t) - \int_a^b \operatorname{Im} f(\gamma(t))\, d\operatorname{Im}\gamma(t),$$

imaginárius része

$$\int_a^b \operatorname{Im} f(\gamma(t))\, d\operatorname{Re}\gamma(t) + \int_a^b \operatorname{Re} f(\gamma(t))\, d\operatorname{Im}\gamma(t).$$

Hangsúlyozni kell, hogy mind a négy integrál egyszerű Stieltjes-integrál[2]. Az egyetlen eltérés, hogy most nem tudjuk, hogy a $\operatorname{Re}\gamma$ és az $\operatorname{Im}\gamma$ súlyfüggvények monoton növekedőek-e. Ennek következtében nem tudjuk, hogy az integrálok léteznek. Ezt feloldandó, tegyük fel, hogy a γ görbe folytonos, és véges sok ponttól eltekintve folytonosan deriválható. Ilyenkor a Newton–Leibniz-szabály alapján γ a deriváltjának integráljaként áll elő. Ha a derivált pozitív és negatív részének külön vesszük az integrálját, akkor γ két koordinátája két növekedő függvény különbsége, így ha f folytonos, akkor a négy Stieltjes-integrál létezik. Ilyenkor a sűrűségfüggvényre vonatkozó formula alapján az integrál

$$\int_a^b \operatorname{Re} f(\gamma(t)) \operatorname{Re}\gamma'(t)\, dt - \int_a^b \operatorname{Im} f(\gamma(t)) \operatorname{Im}\gamma'(t)\, dt +$$
$$i\int_a^b \operatorname{Im} f(\gamma(t)) \operatorname{Re}\gamma'(t)\, dt + i\int_a^b \operatorname{Re} f(\gamma(t)) \operatorname{Im}\gamma'(t)\, dt.$$

[2] Az általános esetben elég azt feltenni, hogy az $\Sigma_k |\Delta\gamma(t_k)|$ összegek egy a felbontástól független korlát alatt maradnak. Belátható, hogy ez ekvivalens avval, hogy a γ görbének van véges ívhossza.

Ez a komplex számok szorzási szabálya alapján éppen azt jelenti, hogy

$$\int_a^b f(\gamma(t))\gamma'(t)\,dt \stackrel{\circ}{=} \int_\gamma f(z)\,dz.$$

A valós esetben megszokott és az integráltól elvárt számos tulajdonság érvényben marad:

1. Az integrál lineáris, vagyis
$$\int_\gamma a\cdot f_1(z)+b\cdot f_2(z)\,dz = a\int_\gamma f_1(z)\,dz + b\int_\gamma f_2(z)\,dz.$$
2. Az integrál az integrációs tartomány szerint additív, vagyis ha a γ_1 végpontja megegyezik a γ_2 kezdőpontjával, akkor a $\gamma_1\cup\gamma_2$ értelemszerű definíciójával
$$\int_{\gamma_1\cup\gamma_2} f(z)\,dz = \int_{\gamma_1} f(z)\,dz + \int_{\gamma_2} f(z)\,dz.$$
3. Ha γ_1 egy $[a,b]$ szakaszon értelmezett görbe és $\gamma_2(t)=\gamma_1(b-t)$, $t\in[a,b]$ éppen a γ_1 értékkészletének ellenkező irányban való befutása, akkor
$$\int_{\gamma_1} f(z)\,dz + \int_{\gamma_2} f(z)\,dz = 0.$$
4. Komplex számok nem rendezettek, így a nem negativitás, illetve a monotonitás szerepét a következő becslés veszi át:
$$\begin{aligned}\left|\int_\gamma f(z)\,dz\right| &\stackrel{\circ}{=} \left|\int_a^b f(\gamma(t))\gamma'(t)\,dt\right| \le \\ &\le \int_a^b \left|f(\gamma(t))\gamma'(t)\right| dt \le \\ &\le \sup_{a\le t\le b}|f(\gamma(t))| \int_a^b \left|\gamma'(t)\right| dt.\end{aligned}$$

   Egyedül az első egyenlőtlenséget érdemes indokolni, de az könnyen látható, hogy minden vektor értékű integrál esetén az egyenlőtlenség igaz, ugyanis az integrál a közelítő összegek határértéke és az egyenlőtlenség pedig a közelítő összegekre alkalmazott háromszög egyenlőtlenség következménye. Ugyanakkor
$$\int_a^b \left|\gamma'(t)\right| dt = \int_a^b \sqrt{(\operatorname{Re}\gamma'(t))^2+(\operatorname{Im}\gamma'(t))^2}\,dt$$
   éppen a γ görbe ívhossza.
$$\int_a^b \left|f(\gamma(t))\gamma'(t)\right| dt = \int_a^b |f(\gamma(t))|\,dV(t),$$

ahol $V(t) \triangleq \int_a^t |\gamma'(s)|\, ds$, amely éppen a γ által befutott görbe hossza az $[a,t]$ szakaszon. Az egyenlőtlenséget szokás

$$\left|\int_\gamma f(z)\, dz\right| \leq \int_\gamma |f(z)|\, d|z|$$

módon is írni, ahol a $d|z|$ az ívhossz szerinti integrálást jelöli, amely egy közönséges Stieltjes-integrál.

A komplex változós analízis legfontosabb tétele a következő:

C.1.5. Tétel (Cauchy). *Ha az f függvény értelmezési tartománya egy csillagszerű*[3] *nyílt halmaz és f az értelmezési tartományának minden pontjában deriválható, akkor az f minden zárt görbén vett integrálja nulla. Értelemszerűen egy $[a,b]$ szakaszon értelmezett γ görbét zárt görbének mondunk, ha $\gamma(a) = \gamma(b)$.*

Bizonyítás: Csak a bizonyítás alapgondolatát vázoljuk, a részletes bizonyítás a komplex változós függvényekkel foglakozó tankönyvek mindegyikében megtalálható. A feltételezett csillagszerűség miatt az f értelmezési tartománynak van egy olyan z_0 pontja, amelyből a γ görbe minden pontja látható, vagyis amely a görbe minden pontjával összeköthető. Ha a γ görbén kijelölünk véges sok pontot és ezeket összekötjük a z_0 ponttal, akkor az így kapott háromszögeken, miként alább látni fogjuk, az integrálok értéke nulla. Ugyanakkor a háromszögeken körbejárva minden a z_0-ból kiinduló oldalon kétszer kell végigmenni és mindig másik irányban, így a görbén levő zárt törtvonalon az integrál értéke szintén nulla. Mivel azonban a görbe törtvonalakkal tetszőlegesen megközelíthető, ezért egyszerű becsléssel látható, hogy a görbén vett integrál is nulla kell hogy legyen. Elég tehát a bizonyítást háromszögekre elvégezni. Indirekt módon tegyük fel, hogy egy Δ_0 háromszögön az integrál I értéke nem nulla. Legyen a háromszög kerülete L. Az oldalfelezők mentén a háromszöget négy egybevágó háromszögre bontjuk. Az ezeken vett négy integrál közül az egyik abszolút értéke nagyobb, vagy egyenlő mint $|I|/4$, ugyanis ellenkező esetben a négy integrál összegének abszolút értéke kisebb lenne mint $|I|$, ami ellentmondás, ugyanis a négy integrál összege, az extra oldalak kétszer való bejárása miatt, éppen I. Jelölje ezt a háromszöget Δ_1. Érdemes hangsúlyozni, hogy a Δ_1 kerülete $L/2$. Az eljárást folytatva egy olyan a (Δ_n) sorozatot kapunk, amelyre egyrészt $\Delta_{n+1} \subset \Delta_n$ és minden Δ_n kerülete $L/2^n$ és a rajta vett integrál abszolút értéke nagyobb, vagy egyenlő mint $|I|/4^n$. A háromszögek egy z^* pontra húzódnak össze. A z^* pont eleme az f értelmezési tartományának, így ott az f deriválható. A derivált definíciója miatt

$$f(z) = f(z^*) + f'(z^*)(z - z^*) + o(|z - z^*|).$$

[3] Az állítás jóval általánosabb feltételek között is érvényes, de ennek tárgyalása szükségszerűen messze vezetne. Érdemes hangsúlyozni, hogy a bizonyítás alapgondolata igen egyszerű, de a sík geometriai tulajdonságai néhány meglepő nehézséget támasztanak az általános eset bizonyítása során.

Számoljuk ki az $\int_{\Delta_n} f(z)\,dz$ integrált. Az integrál additivitása miatt elég az integrálokat a fenti egyenlőségben külön venni. Először is:

$$\int_{\Delta_n} f(z^*)\,dz = 0$$

ugyanis egy zárt görbéről van szó, amelyen minden konstans integrálja nulla. Ugyancsak az integrál linearitása miatt, illetve mivel a konstansok integrálja nulla

$$\begin{aligned}\int_{\Delta_n} f'(z^*)(z-z^*)\,dz &= f'(z^*)\int_{\Delta_n}(z-z^*)\,dz = \\ &= f'(z^*)\int_{\Delta_n} zdz.\end{aligned}$$

Tetszőleges zárt C görbe esetén $\int_C zdz = 0$, ugyanis az integrál értéke nem függ a közbülső pont választási módjától, ezért

$$\begin{aligned}\int_C zdz &= \lim_{n\to\infty}\sum_k z_k^{(n)}\left(z_k^{(n)} - z_{k-1}^{(n)}\right) = \\ &= \lim_{n\to\infty}\sum_k z_{k-1}^{(n)}\left(z_k^{(n)} - z_{k-1}^{(n)}\right),\end{aligned}$$

amiből, felhasználva, hogy a C zárt görbe

$$\begin{aligned}2\int_C zdz &= \lim_{n\to\infty}\sum_k \left(z_k^{(n)} + z_{k-1}^{(n)}\right)\left(z_k^{(n)} - z_{k-1}^{(n)}\right) = \\ &= \lim_{n\to\infty}\sum_k \left(\left(z_k^{(n)}\right)^2 - \left(z_{k-1}^{(n)}\right)^2\right) = \\ &= \lim_{n\to\infty} 0 = 0.\end{aligned}$$

Végezetül tetszőleges $\varepsilon > 0$ esetén ha n elég nagy, vagyis a Δ_n elég kicsi

$$\begin{aligned}\frac{I}{4^n} &\leq \left|\int_{\Delta_n} o(|z-z^*|)\,dz\right| \leq \int_{\Delta_n} |o(|z-z^*|)|\,d|z| \leq \\ &\leq \int_{\Delta_n} \varepsilon\,|z-z^*|\,d|z| \leq \int_{\Delta_n} \varepsilon\frac{L}{2^n}\,d|z| \leq \varepsilon\frac{L^2}{4^n},\end{aligned}$$

amely tetszőleges ε esetén nem teljesülhet.

□

C.1.6. Példa. Számoljuk ki az $\int_\gamma zdz = \int_a^b \gamma(t)\,\gamma'(t)\,dt$ integrált.

A bizonyítás során az integrál értékét már kiszámoltuk, de érdemes közvetlenül a definícióból is kiszámolni. A valós és az imaginárius részt szétválasztva

$$\int_a^b \operatorname{Re}\gamma(t)\operatorname{Re}\gamma'(t) - \operatorname{Im}\gamma(t)\operatorname{Im}\gamma'(t)\,dt,$$

$$\int_a^b \operatorname{Re}\gamma(t)\operatorname{Im}\gamma'(t) + \operatorname{Im}\gamma(t)\operatorname{Re}\gamma'(t)\,dt.$$

Az első integrálban az integrál additivitását használva elég külön kiszámolni az integrálokat.

$$\int_a^b \operatorname{Re}\gamma(t)\operatorname{Re}\gamma'(t)\,dt = \left[(\operatorname{Re}\gamma(t))^2\right]_a^b = 0$$

ugyanis a γ egy zárt görbe. Hasonlóan

$$\int_a^b \operatorname{Im}\gamma(t)\operatorname{Im}\gamma'(t)\,dt = \left[(\operatorname{Im}\gamma(t))^2\right]_a^b = 0.$$

Az imaginárius részre vonatkozó integrál a szorzat deriválási szabálya alapján

$$\int_a^b (\operatorname{Re}\gamma(t)\operatorname{Im}\gamma(t))'\,dt = [\operatorname{Re}\gamma(t)\operatorname{Im}(t)]_a^b = 0.$$

□

C.1.7. Példa. Legyen γ az origó középpontú r sugarú kör, és számoljuk ki a $\int_\gamma 1/z dz$ integrált.

A $[0,2\pi]$ szakaszon értelmezett $\gamma(t) \triangleq r\exp(it)$ függvény éppen az r sugarú kört adja meg, így az integrál definíciója alapján, felhasználva, hogy a $t \mapsto \exp(it)$ függvény deriváltja $i\exp(it)$,

$$\int_\gamma \frac{1}{z}dz \triangleq \int_0^{2\pi} \frac{1}{\exp(it)} i\exp(it)\,dt = i\int_0^{2\pi} dt = 2\pi i.$$

Az integrál értéke azért nem nulla, mert az $f(z) = 1/z$ függvény értelmezési tartománya nem csillagszerű.

□

C.1.8. Példa. Legyen γ az origó középpontú r sugarú kör és legyen $n \neq 1$. Számoljuk ki a $\int_\gamma 1/z^n dz$ integrált.

Ha $n \neq 1$, akkor

$$\begin{aligned}
\int_\gamma \frac{1}{z^n} dz &\overset{\circ}{=} \int_0^{2\pi} \frac{1}{\exp(int)} i \exp(it)\, dt = \\
&= i \int_0^{2\pi} \exp(it(1-n))\, dt = \\
&= i \int_0^{2\pi} \cos((1-n)t) + i \sin((1-n)t)\, dt = \\
&= 0.
\end{aligned}$$

□

C.1.9. Példa. Számoljuk ki a standard normális eloszlás karakterisztikus függvényét.

Bár a függvényt közvetlenül, komplex analízis nélkül már meghatároztuk, mivel gyakran találkozunk az itt tárgyalt megoldással, így a részletes bemutatás nem feltétlenül érdektelen. A karakterisztikus függvény definíciója alapján

$$\begin{aligned}
\varphi(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(it) \exp\left(-\frac{x^2}{2}\right) dx = \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \int_{-\infty}^{\infty} \exp\left(-\frac{(x-it)^2}{2}\right) dt.
\end{aligned}$$

Igen csábító a gondolat, hogy $u = x - it$ helyettesítést végezzünk, de ezt nem tehetjük meg. A megoldást a Cauchy-tétel szolgáltatja. Tekintsük a $(-R,0),(R,0)$, $(R,it),(-R,it)$ pontok által meghatározott téglalap határát képező görbét. Ezen az $\exp\left(-z^2/2\right)$ integrálja nulla. A valós egyenesen levő $(-R,0),(R,0)$ szakaszon vett integrál éppen a

$$\int_{-R}^{R} \exp\left(-\frac{x^2}{2}\right) dx$$

valós integrál. Az $(R,it),(-R,it)$ ív éppen

$$x \mapsto -x + it, \quad -R \leq x \leq R,$$

így az integrál

$$-\int_{-R}^{R} \exp\left(-\frac{(-x+it)^2}{2}\right) dx = -\int_{-R}^{R} \exp\left(-\frac{(x-it)^2}{2}\right) dx.$$

Számoljuk ki a képzetes tengellyel párhuzamos két integrált. Az $(R,0)$ (R,it) közötti integrál

$$
\begin{gathered}
it\int_0^1 \exp\left(-\frac{(R+u\cdot it)^2}{2}\right)du = \\
it\int_0^1 \exp\left(-\frac{R^2-u^2\cdot t^2+2Ruit}{2}\right)du = \\
= it\exp\left(-\frac{R^2}{2}\right)\int_0^1 \exp\left(\frac{u^2\cdot t^2}{2}\right)\exp\left(-uRit\right)du.
\end{gathered}
$$

A $(-R,0)$ és a $(-R,it)$ közötti integrál értékét úgy kapjuk, hogy az előző képletben az R helyébe betesszük a $-R$ értéket. Vegyük észre, hogy a $(-R,it)$ és a $(-R,0)$ közötti integrál éppen ennek az ellentettje, így a két az imaginárius tengellyel párhuzamos integrál összege nem nulla, de ha $R\nearrow\infty$, akkor az összeg nullához tart. Következésképpen

$$
\begin{aligned}
\int_{\infty}^{\infty} \exp\left(-\frac{(x-it)^2}{2}\right)dx &= \int_{\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right)dx \\
&= \sqrt{2\pi}.
\end{aligned}
$$

Ezt visszaírva $\varphi(t)=\exp\left(-t^2/2\right)$.

□

C.1.10. Példa. Számoljuk ki a Cauchy-eloszlás karakterisztikus függvényét.

Ismét a példa önmaga miatt érdekes. A kiszámolandó integrál

$$
\varphi(t)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\exp(itx)}{1+x^2}dx.
$$

Mivel az eloszlás szimmetrikus, ezért feltehető, hogy $t>0$. Legyen C_R a R sugarú, origó középpontú és a felső félsíkban haladó félkörív. Az előző példához hasonlóan tekintsük a $(-R,0),(R,0)$ és C_R ívek által megadott γ_R zárt görbét. Sajnos most a γ_R görbe mentén vett integrál nem nulla, mert a $z=i$ pontban az

$$
f(z)=\frac{1}{\pi}\frac{\exp(itz)}{1+z^2}
$$

komplex változós függvény nevezője nulla. Ezt kikerülendő, tekintsük a $z=i$ pont körüli r sugarú c_r kört. A képzetes tengely mentén a síkot két felé bontva és a c_r görbe mentén a $z=i$ pontot jobbról, illetve balról kikerülve könnyen látható, hogy

$$
\int_{\gamma_R} f(z)\,dz=\int_{c_r} f(z)\,dz.
$$

Ahhoz, hogy a $\varphi(t)$ értékét ki tudjuk számolni először meg kell mutatni, hogy

$$\lim_{R\to\infty}\int_{C_R} f(z)\,dz = \varphi(t).$$

Ehhez, miként az előző példában is meg kell mutatni, hogy a felső köríven az integrál nullához tart. Mivel a C_R félkör éppen az

$$R\exp(iu), \quad 0\le u\le\pi$$

görbe, ezért az integrál definíciója alapján

$$\begin{aligned}&\int_{C_R} f(z)\,dz = \\ &= \frac{1}{\pi}\int_0^{\pi}\frac{\exp(itR\exp(iu))}{1+R^2\exp(2iu)}Ri\exp(iu)\,du = \\ &= \frac{i}{\pi}\int_0^{\pi}\frac{\exp(itR(\cos u+i\sin u))}{1+R^2\exp(2iu)}R\exp(iu)\,du = \\ &= \frac{i}{\pi}\int_0^{\pi}\frac{\exp(-tR\sin u)\exp(itR\cos u)}{1+R^2\exp(2iu)}R\exp(iu)\,du.\end{aligned}$$

Mivel $t>0$, és az integrációs tartományban a végpontoktól eltekintve $\sin u>0$, ezért $R\to\infty$ esetén az integrandus nullához tart. Mivel ha $|z|$ elég nagy, akkor

$$\left|\frac{z}{1+z^2}\right| = \frac{|z|}{\left|1+z^2\right|} \le \frac{|z|}{\left|z^2\right|-1} \le 2\frac{|z|}{\left|z^2\right|} = \frac{2}{|z|},$$

így az integrandus korlátos, tehát a határérték és az integrálás felcserélhető, vagyis a külső íven az integrál értéke valóban nullához tart. Ebből következően minden $0<r<1$ esetén

$$\varphi(t) = \int_{C_r} f(z)\,dz.$$

Az integrál kiszámításának problémáját érdemes egy kicsit általánosabban nézni. Tegyük fel, hogy $f(z)=g(z)/h(z)$ alakú, ahol $h(z_0)=0$, $h'(z_0)\neq 0$ és $g(z_0)\neq 0$. Ilyenkor

$$\begin{aligned}\varphi(t) &= \lim_{r\searrow 0}\int_{C_r}\frac{g(z)}{h(z)}dz = \lim_{r\searrow 0}\int_{C_r}\frac{g(z)}{h(z)-h(z_0)}dz = \\ &= \lim_{r\searrow 0} i\int_0^{2\pi}\frac{g(z_0+r\exp(iu))}{h(z_0+r\exp(iu))-h(z_0)}r\exp(z_0+iu)\,du = \\ &= i\int_0^{2\pi}\frac{g(z_0)}{h'(z_0)}du = 2\pi i\frac{g(z_0)}{h'(z_0)},\end{aligned}$$

ahol kihasználtuk, hogy az integrálás és a határérték felcserélhető. Ez utóbbi azért tehető meg, mert az integrandus korlátos. A konkrét esetben

$$\varphi(t) = \frac{2\pi i}{\pi}\frac{\exp(iti)}{2i} = \exp(-t).$$

Ebből, felhasználva, hogy az eloszlás szimmetrikus, $\varphi(t) = \exp(-|t|)$.

$\square$

A deriválható komplex változós függvények egy további meglepő, a Cauchy-tételből egyszerűen belátható, de alapvetően fontos tulajdonsága a következő:

C.1.11. Tétel. *Ha az f komplex változós függvény egy nyílt halmaz minden pontjában komplex értelemben deriválható, akkor ott hatványsorba is fejthető, pontosabban ha egy $K(z_0, r)$ zárt körlap része az f értelmezési tartományának, akkor ott az f*

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^n, \quad |z - z_0| < r$$

módon reprezentálható. Ha egy függvény az értelmezési tartományának minden pontjának valamely környezetében hatványsorba fejthető, akkor a függvényt analitikusnak mondjuk. Így minden deriválható komplex változós függvény analitikus.

Bizonyítás: Legyen $|z - z_0| \leq r' < r$ és tekintsük a

$$g(u) \stackrel{\circ}{=} \frac{f(u) - f(z)}{u - z}$$

függvényt. A $g(u)$ függvény a $K(z_0, r)$ minden $u \neq z$ pontjában deriválható, és a z pontban g folytonos. Legyen γ a z_0 körüli r sugarú kör. A Cauchy-tétel közvetlenül nem alkalmazható, ugyanis g nem deriválható, de miként már láttuk, ha C egy a z körüli elegendően kicsi δ sugarú kör, akkor a külső körlapot kétfelé vágva, és a belső kört kétfelől körbejárva azonnal látható, hogy

$$\int_{\gamma} g(u)\,du = \int_{C} g(u)\,du.$$

Mivel a belső kör sugara tetszőlegesen kicsi lehet és a g a z pontban folytonossá tehető, ezért a belső körön az integrál tetszőlegesen kicsi lehet, így a külső körön vett integrál értéke nulla. Ebből a fenti C.1.7. példa alapján

$$\int_{\gamma} \frac{f(u)}{u - z} du = \int_{\gamma} \frac{f(z)}{u - z} du = f(z) \int_{\gamma} \frac{1}{u - z} du = f(z)\, 2\pi i. \tag{C.1.2}$$

Vagyis, felhasználva, hogy $|z-z_0| \leq r' < r = |u-z_0|$, így az alábbi geometriai sor egyenletesen konvergens,

$$\begin{aligned}
f(z) &= \frac{1}{2\pi i}\int_\gamma \frac{f(u)}{u-z}du = \\
&= \frac{1}{2\pi i}\int_\gamma \frac{f(u)}{u-z_0-(z-z_0)}du = \\
&= \frac{1}{2\pi i}\int_\gamma \frac{1}{u-z_0}\frac{f(u)}{1-\frac{z-z_0}{u-z_0}}du = \\
&= \frac{1}{2\pi i}\int_\gamma \frac{1}{u-z_0}f(u)\sum_{k=0}^{\infty}\left(\frac{z-z_0}{u-z_0}\right)^k du = \\
&= \sum_{k=0}^{\infty}\frac{1}{2\pi i}\int_\gamma \frac{1}{u-z_0}f(u)\left(\frac{z-z_0}{u-z_0}\right)^k du = \\
&= \sum_{k=0}^{\infty}\frac{(z-z_0)^k}{2\pi i}\int_\gamma \frac{f(u)}{(u-z_0)^k}du \overset{\circ}{=} \sum_{k=0}^{\infty} a_k (z-z_0)^k,
\end{aligned}$$

ahol

$$a_k \overset{\circ}{=} \frac{1}{2\pi i}\int_\gamma \frac{f(u)}{(u-z_0)^{k+1}}du.$$

Mivel az a_k értéke nem függ az r'-től ezért az előállítás minden $|z-z_0| < r$ esetén teljesül.

□

C.1.12. Következmény (Cauchy-formula)**.** *A fenti tétel feltételeinek teljesülése esetén*

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(u)}{u-z}du.$$

Valós függvények komplex kiterjesztései

Mivel a deriválható komplex változós függvények sokkal gazdagabb matematikai tulajdonságokkal rendelkező osztály, mint a deriválható valós függvények, ezért alapvető kérdése a matematikai analízisnek, és így a valószínűségszámításnak is, hogy mely valós függvények terjeszthetők ki a komplex síkra deriválható komplex változós függvényként.

1. Mivel a deriválható komplex függvények hatványsorba fejthetők, ezért akárhányszor deriválhatóak sőt a hatványsorba fejtésben az együtthatók a függvény deriváltjai segítségével számíthatók, így a számegyenesről való kiterjesztés egyértelmű. Ezen elnagyolt gondolatmenet pontosításához, elég megmutatni, hogy ha G egy összefüggő, nyílt halmaz, amely tartalmazza a számegyenes egy intervallumát és f egy olyan deriválható

komplex változós függvény, amely ezen az intervallumon nulla, akkor az f a teljes G halmazon is nulla. Ha $Z \subseteq G$ azon pontok halmaza, ahol $f(z) = 0$, akkor tetszőleges $z \in Z$ esetén a hatványsorba való fejtés miatt két lehetőség van: Ha a kifejtésben az összes együttható nulla, akkor az f a z egy környezetében is nulla, ha pedig van nem nulla együttható, akkor az f függvény $f(u) = (u-z)^m g(u)$ alakba írható, ahol $g(z) \neq 0$ és g folytonos a z pontban. A g folytonossága miatt a g a z egy környezetében sem nulla, és az $(u-z)^m$ is csak az $u = z$ pontban lehet nulla, ezért z a Z egy izolált pontja. Jelölje H a Z torlódási pontjainak halmazát. Az f folytonossága miatt $H \subseteq Z$. Világos, hogy a H a G halmazon belül zárt halmaz, de az elmondottak alapján a H egyúttal nyílt halmaz is. Mivel G összefüggő, ez csak úgy lehet, ha vagy $H = \emptyset$, vagy $H = G$. De mivel a H tartalmazza a számegyenes egy intervallumát ezért $H = G$.

2. A valós számegyenes értelmezett valós függvények deriválható komplex változós kiterjesztéseinek van egy további szemléletes tulajdonsága, amit szokás tükrözési elvnek is mondani. Eszerint az elv szerint egy a számegyenesre szimmetrikus, nyílt és összefüggő tartományban a deriválható f kiterjesztésre igaz az $f(z) = \overline{f}(\overline{z})$ azonosság. Valóban, ha $z = x + iy$ és

$$f(z) = u(x,y) + iv(x,y),$$

akkor a

$$g(z) \stackrel{\circ}{=} \overline{f}(\overline{z}) = u(x,-y) - iv(x,-y)$$

deriváltmátrixa

$$\begin{pmatrix} \partial u/\partial x(x,-y) & -\partial u/\partial y(x,-y) \\ -\partial v/\partial x(x,-y) & \partial v/\partial x(x,-y) \end{pmatrix}.$$

Mivel az f deriválható, ezért

$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}, \quad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$$

következésképpen a g deriváltmátrixa $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ alakú, vagyis a g is deriválható[4] komplex értelemben. De mivel a valós egyenes mentén $g = f$, és mivel a deriválható kiterjesztés egyértelmű, ezért mindenhol $f = g$.

A Laplace-transzformáció komplex kiterjesztése

Ebben az alpontban az

$$L(s) \stackrel{\circ}{=} \int_0^\infty \exp(-sx)\, dK(x) \tag{C.1.3}$$

alakú valós függvények kiterjesztését tárgyaljuk, ahol K egy monoton súlyfüggvény. Érdemes hangsúlyozni, hogy a K-nak nem kell feltétlenül korlátosnak lennie. Nyilvánvaló módon ha valamely s pontra $L(s)$ véges, akkor minden $u > s$ pontra is véges, és ha

[4]Az $f(z)$-t az $\overline{f}(\overline{z})$ képező transzformációk lineárisak, így ha az $f(z)$ deriválható mint a sík egy önmagára való leképezése, akkor az $\overline{f}(\overline{z})$ is az.

az $L(s)$ végtelen, akkor minden $u < s$ pontra is végtelen. Ennek következtében van egy olyan $0 \leq s_0 \leq \infty$ érték, amelynél nagyobb $s > s_0$ pontokra $L(s) < \infty$ és a nálánál kisebb $s < s_0$ pontokban $L(s) = \infty$. Az s_0 pontról általában nem tudunk semmit. Az integráljel alatt deriválva könnyen látható, hogy minden $s > s_0$ esetben az $L(s)$ végtelen sokszor deriválható és

$$L^{(n)}(s) = \int_0^\infty x^n \exp(-sx)\, dK(x).$$

Az s_0 pontot a fenti (C.1.3) konvergenciasugarának mondjuk.

C.1.13. Állítás. *Ha s_0 a fenti (C.1.3) konvergenciasugara, akkor az $L(s)$ függvény komplex deriválható függvényként kiterjeszthető a* $\operatorname{Re} z > s_0$ *félsíkra.*

Bizonyítás: Az állítás igazolása majdnem nyilvánvaló. Legyen $z = s + iu$ és tegyük fel, hogy $s > s_0$. A komplex exponenciális függvényre vonatkozó Euler-képlet alapján

$$\begin{aligned}
L(z) &\overset{\circ}{=} \int_0^\infty \exp(-zx)\, dK(x) = \\
&= \int_0^\infty \exp(-(s+iu)x)\, dK(x) = \\
&= \int_0^\infty \exp(-sx)(\cos ux - i \sin ux)\, dK(x) = \\
&= \int_0^\infty \exp(-sx)\cos ux dK(x) - i\int_0^\infty \exp(-sx)\sin ux dK(x),
\end{aligned}$$

amely két integrál nyilvánvalóan konvergens, ugyanis például

$$|\exp(-sx)\cos ux| \leq \exp(-sx).$$

Lássuk most a kiterjesztett $L(z)$ függvény komplex értelemben való deriválhatóságát. Számoljuk ki a parciális deriváltakból álló Jacobi-mátrixot. Könnyen látható, hogy ha $s > s_0$, akkor a két integrálban az integráljel alatt deriválhatunk. Az egyszerűbb jelölés céljából az integrálokat „kiemelve”:

$$J(s,u) = \int_0^\infty e^{-sx} x \begin{pmatrix} -\cos ux & -\sin ux \\ +\sin ux & -\cos ux \end{pmatrix} dK(x).$$

Világos, hogy a $J(s,u)$ mátrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ alakú. Ugyancsak világos, hogy a $J(s,u)$ folytonos, vagyis a függvény mint a síkot a síkba képező leképezés deriválható, így a komplex értelemben való deriválhatóság teljesül, amivel az állítást igazoltuk[5].

□

[5]Egy másik kézenfekvő bizonyítás a következő: Az majorált konvergencia tétel miatt az $L(z)$ folytonos. Tetszőleges γ zárt görbe esetén az $\int_\gamma L(z)\, dz$ integrálban az $|\exp(-zx)| \leq \exp(-sx)$ miatt a két integrál felcserélhető és a Cauchy-tétel miatt $\int_\gamma \exp(-zx)\, dz = 0$. Így $\int_\gamma L(z)\, dz = 0$. Így a Cauchy-tétel megfordításának is tekinthető úgynevezett Morera-tétel miatt az L deriválható.

A $\operatorname{Re} z > s_0$ a maximális olyan félsík, amelyre a transzformáció deriválható módon kiterjeszthető:

C.1.14. Állítás. *Ha az L komplex Laplace-transzformált deriválható módon kiterjeszthető egy $\operatorname{Re} z > b$ félsíkra, akkor ott*

$$L(z) = \int_0^\infty \exp(-zx)\, dK(x),$$

következésképpen $b \geq s_0$.

Bizonyítás: Indirekt tegyük fel, hogy van egy olyan $g(z)$ deriválható függvény amely $\operatorname{Re} z > b$ félsíkra való kiterjesztése a transzformáltnak és $b < s_0$. Ekkor g egy alkalmas $s_0 > a$ pontban sorba fejthető úgy, hogy a konvergenciasugár nagyobb lesz, mint $s_0 - a$. Ez azt jelenti, hogy a

$$\sum_{k=0}^\infty \frac{L^{(k)}(a)}{k!}(s-a)^k$$

Taylor-sor egy $s < s_0$ pontban konvergens. A derivált képletét beírva és a Fubini-tétellel az összegzést és az integrálást felcserélve

$$\begin{aligned}
\sum_{k=0}^\infty & \frac{(-1)^k \int_0^\infty x^k \exp(-ax)\, dK(x)}{k!}(-1)^k (a-s)^k = \\
&= \sum_{k=0}^\infty \frac{\int_0^\infty x^k \exp(-ax)\, dK(x)}{k!}(a-s)^k = \\
&= \int_0^\infty \exp(-ax) \sum_{k=0}^\infty \frac{x^k (a-s)^k}{k!} dK(x) = \\
&= \int_0^\infty \exp(-ax) \exp(x(a-s))\, dK(x) = \\
&= \int_0^\infty \exp(-sx)\, dK(x)
\end{aligned}$$

véges, vagyis a Laplace-transzformált az $s < s_0$ pontban is konvergens, miközben a feltétel szerint a konvergenciatartomány határa az s_0 pont. Így tehát $b \geq s_0$. Következésképpen L az előírt alakú.

□

Tegyük most fel, hogy a K egy nem negatív valószínűségi változó eloszlásfüggvénye. Tegyük fel továbbá, hogy az eloszlásnak van sűrűségfüggvénye, amelyet jelöljön f. Jelölje továbbá L a transzformációt. Ekkor az

$$L(s) \overset{\circ}{=} \int_0^\infty \exp(-sx) f(x)\, dx$$

deriválható módon kiterjeszthető a $\operatorname{Re} z > 0$ félsíkra. Ez azonban nem jelenti azt, hogy nincs egy nagyobb halmaz, amire az L esetleg kiterjeszthető deriválható módon. A további kiterjeszthetőséghez szükségünk lesz a következő feltételekre: Az f sűrűségfüggvényre a következő megkötéseket tesszük:

1. Az f sűrűségfüggvény komplex deriválható módon kiterjeszthető a $\operatorname{Re} z > 0$ félsíkra.

2. Jelölje $\arg z$ a z komplex szám polárfelbontásához tartozó $(-\pi, \pi)$-beli szöget, és legyen
$$S(\varepsilon) \circeq \{z \mid |\arg z| \leq \pi/2 - \varepsilon\}.$$
Tetszőleges $\varepsilon > 0$ esetén az $S(\varepsilon)$ halmazon van olyan $\alpha > 0$, hogy
$$f(z) = z^{\alpha-1} h(z)$$
ahol a $h(z)$ korlátos.

C.1.15. Állítás. *A megadott feltételek teljesülése esetén az $L(z)$ deriválható módon kiterjeszthető a negatív valós számok mentén felvágott komplex síkra, vagyis a $\mathbb{C} \setminus (-\infty, 0]$ halmazra.*

Bizonyítás: Legyen
$$L(z) \circeq \int_0^\infty \exp(-z \cdot x) f(x)\, dx, \quad \operatorname{Re} z > 0$$
a komplex transzformált. Az L értelmezve van és deriválható a $\operatorname{Re} z > 0$ félsíkban. Megmutatjuk, hogy az L deriválható módon kiterjeszthető az $\operatorname{Im} z > 0$ felső félsíkra. Az $\operatorname{Im} z < 0$ eset indoklása analóg módon végezhető el. A bizonyítás alapgondolata viszonylag szemléletes. Az f komplex, így definiálható a némiképpen homályosan és félrevezető módon jelölt
$$L(s) \circeq \int_0^\infty \exp(-s \cdot z) f(z)\, dz \tag{C.1.4}$$
komplex integrál. A lényeges gondolat, hogy a z integrációs változó komplex. A jelölés homályossága abból ered, hogy az $\int_a^b g(z)\, dz$ jelölés csak akkor értelmes, ha az a és a b pontokat összekötő minden görbén az integrál értéke azonos. Ennek megfelelően az $\int_0^\infty g(z)\, dz$ komplex integrál csak akkor értelmes, ha minden a nulla pontból a végtelenben tartó görbe mentén az integrál értéke azonos[6]. A bizonyítás lényege, hogy az $S(\varepsilon)$-ban haladó görbék esetén ez teljesül. Tekintsünk két a 0 és a ∞ között az $S(\varepsilon)$-ban haladó görbét. A deriválható függvények zárt görbéken vett vonalintegrálja nulla. Legyen $r < R$ és tekintsük az r és a R sugarú körökön, illetve a görbéken vett integrálokat. Elég megmutatni, hogy ha $r \searrow 0$, illetve, ha $R \nearrow \infty$, akkor a köríveken vett integrálok

[6] Bár ennek itt nincs jelentősége, a komplex sík felfogható, mint egy a komplex síkra állított gömb pontjainak az északi pólusból való levetítése. Ilyenkor az északi pólusnak nincs a síkban képe, de tekinthető a ∞ pontnak. Ebben az ábrázolásban az déli pólusból az északi pólus felé haladó görbék mentén kell az összes integrálnak azonosnak lenni.

nullához tartanak. A belső köríven vett integrál, ha $-\pi/2+\varepsilon \leq \alpha < \beta \leq \pi/2-\varepsilon$ a két görbe metszéspontjához tartozó szög, akkor

$$i\int_{\alpha}^{\beta} \exp(-sr\exp(it)) f(r\exp(it)) r\exp(it)\,dt.$$

A második feltétel miatt

$$zf(z) = zz^{\alpha-1}h(z) = z^{\alpha}h(z).$$

A $h(z)$ egyenletesen korlátos, és ha $r \searrow 0$, akkor

$$z^{\alpha} = r^{\alpha}(\cos\alpha\varphi + i\sin\alpha\varphi)$$

egyenletesen nullához tart, így a belső körön az integrál nullához tart. A külső köríven vett integrálra ha $R \nearrow \infty$, akkor

$$\begin{aligned}
&\left|\int_{\alpha}^{\beta} \exp(-sR\exp(it)) f(R\exp(it)) iR\exp(it)\,dt\right| \leq \\
&\leq \int_{-\pi/2+\varepsilon}^{\pi/2-\varepsilon} R\cdot|\exp(-sR\exp(it))|\,|f(R\exp(it))|\,dt = \\
&= R\int_{-\pi/2+\varepsilon}^{\pi/2-\varepsilon} \exp(-R(s\cos t)) R^{\alpha-1}\,|h(R\exp(it))|\,dt \leq \\
&\leq KR^{\alpha}\int_{-\pi/2+\varepsilon}^{\pi/2-\varepsilon} \exp\left(-R\left(s\cos\left(\frac{\pi}{2}-\varepsilon\right)\right)\right) dt \circeq \\
&\circeq C\cdot R^{\alpha}\exp(-c\cdot R) \to 0.
\end{aligned}$$

Ebből következően az integrált bármely görbén vehetjük, vagyis a fenti (C.1.4) kifejezés értelmes. Ha most valós transzformált komplex kiterjesztéshez tartozó korábbi trükköt egy φ szöghöz tartozó egyenes mentén akarjuk alkalmazni, akkor az

$$L(s) = \exp(i\varphi)\int_{0}^{\infty} \exp(-s\exp(i\varphi)\cdot t) f(\exp(i\varphi)\cdot t)\,dt \tag{C.1.5}$$

integrált kell vizsgálni. Ugyanakkor a valós tengely mentén vett integrál abszolút konvergens, de az általános egyenes mentén vett integrál esetleg csak improprius értelemben létezik. Az integrál nyilvánvalóan kétszeresen improprius. A végtelen oldalon az exponenciális függvény biztosítja az abszolút érték integráljának konvergenciáját, a nulla oldalon pedig a $O\left(|z|^{\alpha-1}\right)$ feltétel garantálja az abszolút konvergenciát. Az integrál deriválható kiterjesztését olyan $z = x+iy$ komplex számokra tudjuk elvégezni, amelyekre a z-hez tartó exponenciális kifejezés abszolút értéke azonos a fenti (C.1.5) integrálban szereplő exponenciális függvény abszolút értékével, vagyis

$$s\cos\varphi = x\cos\varphi - y\sin\varphi.$$

Vagyis a z akkor szerepel a kiterjesztett transzformáció értelmezési tartományában, ha a $z-s = x-s+iy$ merőleges a $\exp(-i\varphi)$ egységvektorra. Legyen $0<\varphi<\pi/2$ és tekintsük a $\operatorname{Re} z>0$ félsíkban haladó

$$t \mapsto \exp(-i\varphi)t \mathrel{\hat{=}} \vartheta t$$

sugarat[7]. Az $u \mapsto \exp(-z\cdot u) f(u)$ deriválható komplex függvény ezen sugáron vett vonalintegrálja legyen

$$\widehat{L}(z) \mathrel{\hat{=}} \vartheta \int_0^\infty \exp(-z\vartheta\cdot t) f(\vartheta\cdot t)\, dt.$$

Ez az integrál minden olyan z esetén értelmes, amely az $\exp(i\varphi)$ vektorra merőleges egyenes felett van. Ezek éppen azok a vektorok, amelyeket a $\operatorname{Re} z>0$ félsík φ szöggel való elforgatásával kapunk. Az elmondottak miatt az $\widehat{L}$ deriválható és a kifejezésben szereplő integrál abszolút konvergens. Ha most $z=s$, ahol az s valós, vagyis az $\widehat{L}$ értékét a valós tengelyen vesszük, amely része az elforgatott síknak, akkor éppen a fenti (C.1.5) integrált kapjuk csak a φ helyébe a $-\varphi$ szöget kell írni. Mivel az integrál értéke $L(s)$, ezért az $\widehat{L}$ az L kiterjesztése.

□

C.1.16. Példa. Néhány nem negatív eloszlás Laplace-transzformáltja.

Az előző állítás feltételei igen gyengék és a legtöbb ismert eloszlás esetén teljesülnek.

1. Az exponenciális eloszlás Laplace-transzformáltja $L(s) = \lambda/(\lambda+s)$, amely a $s = -\lambda$ ponttól eltekintve minden komplex számra deriválható. Az $f(x) = \lambda\exp(-\lambda x)$ függvényre nyilván teljesülnek a tétel feltételei, ugyanis

$$|f(z)| = \lambda\exp(-\lambda\operatorname{Re} z).$$

2. A gamma eloszlás Laplace-transzformáltja

$$L(s) = \left(\frac{\lambda}{\lambda+s}\right)^a = \exp\left(a\ln\frac{\lambda}{\lambda+s}\right)$$

amely szintén kiterjeszthető a $\mathbb{C}\setminus(-\infty,\lambda]$ halmazra. Ismét könnyen látható, hogy az f sűrűségfüggvény teljesíti az állításban szereplő feltételeket. A definíciókból evidens, hogy az alább bevezetett általánosított gamma-konvolúciók Laplace-transzformáltja kiterjeszthető a negatív oldalon felvágott komplex síkra. Ilyenkor persze a sűrűségfüggvény képlete nem egyszerűen meghatározható.

[7] Ügyeljünk az előjelre!

3. A lognormális eloszlás Laplace-transzformációjának képlete közvetlenül nem határozható meg, ezért a tétel alapján tudjuk csak a kiterjeszthetőséget indokolni. Mivel a logaritmusfüggvény deriválható módon kiterjeszthető a $\mathbb{C}\setminus(-\infty,0]$ és így a $\operatorname{Re} z>0$ halmazra, ezért az

$$f(x)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right)$$

függvény is kiterjeszthető. Könnyen látható, hogy a kiterjesztett függvény korlátos.

4. A másodfajú béta eloszlás sűrűségfüggvénye

$$f(x)=\frac{1}{B(\alpha,\beta)}x^{\alpha-1}\frac{1}{(1+x)^{\alpha+\beta}},\quad x>0,$$

amely kifejezés nyilvánvalóan kiterjeszthető a $\mathbb{C}\setminus(-\infty,0]$ és így a $\operatorname{Re} z>0$ halmazra. Mivel $\alpha,\beta>0$ könnyen igazolható, hogy az állítás feltételei teljesülnek.

5. Legyen $r>0$ és az eloszlás sűrűségfüggvénye legyen

$$f(x)\circeq\frac{1}{\Gamma(r,\lambda)}\left(\frac{1}{x}\right)^{r+1}\exp\left(-\frac{\lambda}{x}\right),\quad x>0.$$

Könnyen látható, hogy ez éppen egy (r,λ) paraméterű gamma eloszlás reciprokának a sűrűségfüggvénye, amely ismét deriválható módon kiterjeszthető a $\mathbb{C}\setminus(-\infty,0]$ halmazra. A fenti állítás feltételei triviálisan teljesülnek ugyanis a kiterjesztett sűrűségfüggvény korlátos.

□

C.2. Teljesen monoton függvények

Folytassuk a probléma tárgyalását a nem negatív változókhoz tartozó Laplace-transzformáltak karakterizációjával. A továbbiakban, ha csak másképpen nem mondjuk, függvénytranszformált alatt mindig egy nem negatív változó Laplace-transzformáltját értjük.

C.2.1. Definíció. Egy (a,b) nyílt intervallumon értelmezett f függvényt teljesen monotonnak mondunk, ha az f végtelen sokszor deriválható és minden $n\geq 0$-ra

$$(-1)^n\frac{d^n}{dx^n}f(x)\geq 0,\quad a<x<b.$$

Ha mást nem mondunk, akkor az intervallumon a $(0,\infty)$ félegyenest értjük. Egy $[a,b]$ zárt intervallumon értelmezett f függvényt teljesen monotonnak mondunk ha az (a,b) nyílt intervallumon teljesen monoton és az $[a,b]$ zárt intervallumon folytonos. Ha másképpen nem mondjuk, akkor mindig a $[0,\infty)$ zárt intervallumra gondolunk. A teljesen monoton függvények nem negatívak és csökkennek, vagyis ilyenkor a definíció alapján minden teljesen monoton függvény a megadott zárt intervallumon korlátos.

C.2.2. Példa. Ha $\beta > 0$, akkor az $f(x) = \exp(-\beta x)$ függvény teljesen monoton a $[0,\infty)$ félegyenesen és az $f(x) = 1/x^{\beta}$ függvény teljesen monoton a $(0,\infty)$ félegyenesen.

C.2.3. Példa. Ha az f és a g függvény teljesen monoton, akkor az fg és az $f+g$ is teljesen monoton.

Az összeg igazolása triviális. A szorzatra

$$\begin{aligned}
(-1)^n (fg)^{(n)} &= (-1)^n \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)} = \\
&= \sum_{k=0}^{n} \binom{n}{k} (-1)^k f^{(k)} (-1)^{n-k} g^{(n-k)} \geq 0.
\end{aligned}$$

□

C.2.4. Példa. Ha az f teljesen monoton és a g' derivált szintén teljesen monoton, akkor az $f(g(x))$ összetett függvény is teljesen monoton.

Az összetett függvény deriválási szabálya alapján ha az f teljesen monoton és a g' derivált is teljesen monoton, akkor először is $f(g(x)) \geq 0$, továbbá

$$\frac{d}{dx} f(g(x)) = f'(g(x))\, g'(x) \leq 0,$$

ugyanis $g'(x) \geq 0$ a g' teljes monotonitása miatt. A második deriváltat számolva

$$\frac{d^2}{dx^2} f(g(x)) = f''(g(x)) \left(g'(x)\right)^2 + f'(g(x))\, g''(x).$$

Az első tag nem negatív, ugyanis $f'' \geq 0$. A második tag szintén nem negatív, ugyanis a benne szereplő tagok mindegyike nem pozitív. Indukcióval folytatva tegyük fel, hogy valamilyen n-re $(-1)^n \frac{d^n}{dx^n} f(g(x)) \geq 0$.

$$\begin{aligned}
\frac{d^{n+1}}{dx^{n+1}} f(g(x)) &= \frac{d^n}{dx^n} \left(f'(g(x))\, g'(x)\right) = \\
&= \sum_{k=0}^{n} \binom{n}{k} \left(f'(g(x))\right)^{(k)} \left(g'(x)\right)^{(n-k)}.
\end{aligned}$$

Ha az f teljesen monoton, akkor a $(-1) f'$ függvény is teljesen monoton. Ezért az indukciós feltétel miatt minden k indexre

$$(-1)^{k+1} \left(f'(g(x))\right)^{(k)} \geq 0.$$

Ugyanakkor a g' teljesen monoton, ezért

$$(-1)^{n-k} \left(g'(x)\right)^{(n-k)} \geq 0,$$

amiből

$$(-1)^{n+1}\left(f'(g(x))\right)^{(k)}\left(g'(x)\right)^{(n-k)} \geq 0,$$

tehát

$$(-1)^{n+1}\frac{d^{n+1}}{dx^{n+1}}f(g(x)) \geq 0.$$

Vagyis az $f(g(x))$ összetett függvény is teljesen monoton.

□

C.2.5. Tétel (Bernstein)**.** *A $[0,\infty)$ szakaszon értelmezett $L(s)$ függvény pontosan akkor egy a nem negatív számokra koncentrált F valószínűségeloszlás Laplace-transzformáltja, ha $L(0)=1$ és az $L(s)$ teljesen monoton a $[0,\infty)$-en.*

A bizonyítás vázlata: Bár a tételnek létezik elemi bizonyítása, mégis egy absztrakt, funkcionálanalízisre épülő bizonyítást választottam. Pontosabban egy funkcionálanalízisre épülő bizonyítás vázlatát ismertetem. Az elemi bizonyítások hátránya, hogy a gondolatmenet némiképpen szövevényes, az absztrakt bizonyítások hátránya ugyanakkor, hogy komoly előképzettséget kíván a modern matematika fogalomkörében. A bizonyítás első olvasásra kihagyható.

1. Legyen L egy $\xi \geq 0$ változó Laplace-transzformáltja, vagyis

$$L(s) \doteq \mathbf{E}(\exp(-s\xi)),$$

ahol $\xi \geq 0$. A $\xi \geq 0$ feltétel miatt az L értelmezve van minden $s \geq 0$ esetén. A paraméteres integrálok deriválási szabálya alapján ha $s > 0$, akkor $L^{(n)}(s) = (-1)^n \mathbf{E}(\xi^n \exp(-s\xi))$, amiből

$$\begin{aligned}(-1)^n L^{(n)}(s) &= (-1)^{2n}\mathbf{E}(\xi^n \exp(-s\xi)) = \\ &= \mathbf{E}(\xi^n \exp(-s\xi)) \geq 0,\end{aligned}$$

vagyis egy nem negatív változó Laplace-transzformáltja teljesen monoton függvény. Vegyük észre, hogy az $L^{(n)}(s)$ derivált minden $s > 0$ esetén létezik, de a deriváltak korlátosságáról nem tudunk semmit mondani, vagyis az $s = 0$ pontba a derivált nem feltétlenül terjeszthető ki véges módon.

2. A fordított irány igazolása jóval nehezebb. Ha L egy teljesen monoton függvény, akkor $L \in C^\infty((0,\infty))$, ahol a $C^\infty \doteq C^\infty((0,\infty))$ a $(0,\infty)$ szakaszon végtelenül sokszor deriválható függvények halmaza. A modern analízis szemlélete alapján a végtelen sokszor deriválható függvények halmazát, vagyis a C^∞ teret, egy lineáris térnek tekintjük, és ha M a teljesen monoton függvények halmaza, akkor azt kérdezzük, hogy az M milyen geometriai tulajdonságokkal rendelkező részhalmaza a C^∞ vektortérnek. Mivel a teljesen monoton függvények összege és nem negatív számmal való szorzata is teljesen monoton, ezért az M egy kúp a C^∞ térben. Minden a $(0,\infty)$ nyílt félegyenesen teljesen monoton függvény monoton csökken, így az $L(0+)$ határérték mindig létezik, természetesen ez a határérték esetleg végtelen is lehet. Jelölje K az olyan L teljesen monoton

függvények halmazát, amelyekre $L(0+) \leq 1$. A bizonyítás első lépése annak igazolása, hogy a K egy konvex és alkalmas topolológiában kompakt halmaz. A konvexitás igazolása triviális, a kompaktság azonban távolról sem magától érthetődő. Kezdjük talán avval, hogy ahhoz, hogy kompaktságról beszélhessünk, ahhoz konvergenciát és távolságot is definiálni kell. A C^{∞} térben a természetes konvergencia a kompakt részhalmazokon az összes derivált egyenletes konvergenciája. Vagyis ha (f_n) egy C^{∞} függvényekből álló sorozat, akkor definíció szerint $f_n \to f$, pontosan akkor, ha minden $[a,b] \subset (0,\infty)$ zárt intervallumon az összes deriváltra $f_n^{(k)} \to f^{(k)}$, ahol a deriváltak konvergenciája az $[a,b]$ szakaszon egyenletes. Mivel az $f_n \to f$ konvergencia az összes derivált véges szakaszokon való egyenletes konvergenciáját implikálja, ezért az M és a K a C^{∞} konvex és zárt részhalmazait alkotják. A kompaktság igazolásához meg kell mutatni, hogy minden K-ból vett sorozatnak van konvergens részsorozata. Ehhez elegendő megmutatni, hogy minden fix $[a,b]$ és minden fix k esetén az $\left(f_n^{(k)}\right)$ sorozatnak van egyenletesen konvergens részsorozata, ugyanis akkor a Cantor-féle diagonális eljárással képezhetünk egy a C^{∞} térben is konvergens további részsorozatot. Az Arzelà–Ascoli tétel alapján ehhez be kell mutatni, hogy az $\left(f_n^{(k)}\right)$ sorozat egyrészt egyenletesen korlátos, másrészt egyenlő mértékben egyenletesen folytonos. Ez azonban majdnem nyilvánvaló. Ha $k = 0$, akkor a K definíciója miatt ha $f \in K$, akkor $0 \leq f \leq 1$, ugyanis az első deriváltra vonatkozó $f' \leq 0$ feltétel miatt az f csökken és $f(0+) \leq 1$. Mivel a deriváltak felváltva nem negatívak és nem pozitívak, ezért a derivált függvények vagy csökkenek vagy nőnek. Amikor nőnek, akkor nem pozitívak, ha csökkennek, akkor nem negatívok. Ebből következően az egyenletes korlátossághoz elegendő megmutatni, hogy minden x pontban a deriváltak sorozata korlátos. Tegyük fel, hogy beláttuk, hogy az $\left(f_n^{(k)}(x)\right)$ sorozat minden $x > 0$ esetén korlátos. Tekintsük azt az esetet, amikor az $f^{(k+1)}$ nem negatív és csökken, a másik eset analóg módon igazolható. Mivel $x > 0$, ezért alkalmas $\varepsilon > 0$ számra a középértéktétel alapján

$$\frac{f_n^{(k)}(x) - f_n^{(k)}(x-\varepsilon)}{\varepsilon} = f_n^{(k+1)}(c) \geq f_n^{(k+1)}(x).$$

Mivel az $\left(f_n^{(k)}(y)\right)$ sorozat minden $y > 0$ pontban korlátos, ezért minden $x > 0$ pontban az $\left(f_n^{(k+1)}(x)\right)$ is korlátos. Az $[a,b]$ szakaszon való egyenlő mértékben való egyenletes folytonosság szintén a középértéktétel, illetve a már belátott egyenletes korlátosság következménye ugyanis

$$\left|f_n^{(k)}(x) - f_n^{(k)}(y)\right| \leq \sup_{c\in[a,b]} \left|f_n^{(k+1)}(c)\right| |x-y| \leq L_{k+1} |x-y|$$

minden $x, y \in [a,b]$ esetén, amiből az egyenlő mértékben való egyenletes folytonosság már evidens. Így tehát a K egy kompakt és konvex részhalmaza a $C^{\infty}((0,\infty))$ térnek.

3. A bizonyítás nagyágyúja a funkcionálanalízis ünnepelt Krein–Milman tétele, amely szerint minden konvex kompakt halmaz az extremális pontjai konvex burkának a lezártja[8]. A jelen bizonyítás „lényege” a K extremális pontjainak meghatározása. Legyen tehát f a K egy extremális pontja. Minden $y>0$ esetén tekintsük az $u(x) = f(x+y) - f(x)f(y)$ függvényt. Alább megmutatjuk, hogy $f \pm u \in K$. Mivel az f extremális pont, ez csak úgy lehetséges, ha $u=0$, vagyis minden $x,y>0$ esetén $f(x+y) = f(x)f(y)$. Mivel az f folytonos és $0 \leq f \leq 1$, ezért $f(x) = \exp(-\beta x)$ alakú, ahol $\beta \in [0,\infty]$. Lássuk be tehát, hogy $f \pm u \in K$. Legyen $\alpha \triangleq f(y)$. A feltételek miatt $0 \leq \alpha \leq 1$. Vegyük észre,

$$
\begin{gathered}
(f+u)(0+) = (1-\alpha) f(0+) + \alpha \leq 1, \\
(f-u)(0+) = f(0+) - \alpha(1-f(0+)) \leq f(0+) \leq 1, \\
(-1)^n (f+u)^{(n)}(x) = (-1)^n f^{(n)}(x) + (-1)^n u^{(n)}(x) = \\
= (-1)^n f^{(n)}(x) + (-1)^n \left(f^{(n)}(x+y) - f(y) f^{(n)}(x)\right) = \\
= (-1)^n f^{(n)}(x)(1-\alpha) + (-1)^n f^{(n)}(x+y) \geq 0, \\
(-1)^n (f-u)^{(n)}(x) = (-1)^n f^{(n)}(x) - (-1)^n u^{(n)}(x) = \\
= (-1)^n f^{(n)}(x) - (-1)^n \left(f^{(n)}(x+y) - f(y) f^{(n)}(x)\right) = \\
= (-1)^n \left(f^{(n)}(x) - f^{(n)}(x+y)\right) + (-1)^n \alpha f^{(n)}(x) \geq 0.
\end{gathered}
$$

Az utolsó lépésben kihasználtuk, hogy egyrészt $(-1)^n f^{(n)}(x) \geq 0$, másrészt, hogy a $h \triangleq (-1)^n f^{(n)}$ függvény deriváltja nem pozitív, ezért a h függvény nem nő, így mivel $y \geq 0$ ezért az első tag is nem negatív. A Krein–Milman tétel szerint tehát mind f teljesen monoton függvény esetén ha $x>0$. akkor

$$
f(x) = \lim_{n\to\infty} \sum_i \alpha_i^{(n)} \exp\left(-\beta_i^{(n)} x\right) = \lim_{n\to\infty} \int_0^\infty \exp(-\beta x)\, dF_n(\beta),
$$

ahol F_n az $\left(\alpha_i^{(n)}\right)$ diszkrét eloszlás eloszlásfüggvénye. A konvergencia x szerint a kompakt halmazokon egyenletes. Mivel az extremális pontok között a $\beta = \infty$ is lehetséges, vagyis az $f=0$ is egy lehetséges extremális pont, ezért csak a $\Sigma_i \alpha_i^{(n)} \leq 1$ egyenlőtlenség teljesül, vagyis $0 \leq F_n \leq 1$.

4. Következő lépésként ki kell használni a Helly-féle kiválasztási tételt[9]. A monoton növekedő függvényekből álló (F_n) sorozat egyenletesen korlátos, így az említett kiválasztási tétel alapján van olyan F monoton növekedő, balról folytonos függvény, hogy egy alkalmas részsorozatra $F_{n_k}(x) \to F(x)$ az F minden x folytonossági pontjában.

[8] Az állítás „csak“ lokálisan konvex Hausdorff topologikus vektortérben érvényes. Véges dimenzió esetén a lezárás elhagyható. A C^∞ tér metrizálható és lokálisan konvex, így az állítás használható.

[9] V.ö.: B.3.13. Állítás, 256. oldal.

Hangsúlyozni kell, hogy az F bár monoton nő és balról folytonos általában nem feltétlenül lesz eloszlásfüggvény, ugyanis nem feltétlenül érvényes az $F(\infty) = 1$ egyenlőség. Az egyszerűbb jelölés céljából tegyük fel, hogy a részsorozat megegyezik az eredeti sorozattal. Tegyük fel, hogy a b végpont folytonossági pont. Könnyen látható, hogy tetszőleges g folytonos függvény esetén[10]

$$\lim_{n\to\infty} \int_0^b g(\beta)\, dF_n(\beta) = \int_0^b g(\beta)\, dF(\beta).$$

Mivel az $\exp(-\beta x)$ függvény elég nagy β-ra tetszőlegesen kicsivé tehető, ezért azonnal látható, hogy

$$f(x) = \lim_{n\to\infty} \int_0^\infty \exp(-\beta x)\, dF_n(\beta) = \int_0^\infty \exp(-\beta x)\, dF(\beta).$$

5. Ha $\lim_{x\searrow 0} f(x) = 1$, akkor a monoton konvergencia tétel miatt $1 = f(0+) = \int_0^\infty 1 dF$, vagyis ilyenkor az F egy valószínűségi mérték és az f egy valószínűségi mérték Laplace-transzformáltja.

□

C.2.6. Példa. Ha A egy monoton növekedő súlyfüggvény, akkor az

$$L(s) \stackrel{\circ}{=} \exp\left(\int_{0+}^\infty \ln\left(\frac{\lambda}{\lambda+s}\right) dA(\lambda)\right), \quad s \geq 0$$

egy nem negatív, korlátlanul osztható eloszlás Laplace-transzformáltja.

Az A súlyfüggvényre tett feltétel természetesen úgy értendő, hogy az integrál minden $s \geq 0$ esetén létezik és véges. Ha $s = 0$, akkor $L(s) = 1$. Ha $s \searrow 0$, akkor $\lambda/(\lambda+s) \nearrow 1$ és a logaritmust tartalmazó kifejezés monoton növekedőleg nullához tart. Meg kell mutatni, hogy az integrál is nullához tart. A logaritmust tartalmazó integrandus mindig nem pozitív és az s függvényében nő, ezért bármely $s > 0$ értékhez tartozó integrandus tekinthető integrálható majoránsnak, vagyis az integrál és a határérték felcserélhető, tehát $L(0+) = 1$. Meg kell mutatni, hogy az L végtelen sokszor deriválható. Ha

$$r(s) \stackrel{\circ}{=} \int_{0+}^\infty \frac{1}{\lambda+s} dA(\lambda),$$

akkor a Fubini-tétel alapján az integrálokat felcserélve minden $s > 0$ esetén

$$\begin{aligned}
\int_0^s r(u)\, du &= \int_0^s \int_{0+}^\infty \frac{1}{\lambda+u} dA(\lambda)\, du = \\
&= \int_{0+}^\infty \int_0^s \frac{1}{\lambda+u} du dA(\lambda) = \\
&= \int_{0+}^\infty [\ln(\lambda+u)]_0^s\, dA(\lambda) = \\
&= \int_{0+}^\infty \ln\frac{\lambda+s}{\lambda} dA(\lambda).
\end{aligned}$$

[10] V.ö.: B.3.7. Lemma, 251. oldal.

Mivel ez a kifejezés az A-ra tett feltétel miatt véges, ezért az $r(u)$ is majdnem minden u esetén véges. Következésképpen a

$$\lim_{u\to s}\int_{0+}^{\infty}\frac{1}{\lambda+u}dA(\lambda)$$

határértékben mindig van integrálható majoráns, így r folytonos, tehát

$$\frac{d}{ds}\int_{0+}^{\infty}\ln\frac{\lambda+s}{\lambda}dA(\lambda)=r(s)=\int_{0+}^{\infty}\frac{1}{\lambda+s}dA(\lambda).$$

Az $1/(\lambda+s)$ függvény minden $\lambda>0$ esetén teljesen monoton, és a Fubini-tétellel a már látott módon igazolható, hogy az integrálás és a deriválás mindig felcserélhető. Mivel

$$L(s)=\exp\left(\int_{0+}^{\infty}\ln\left(\frac{\lambda}{\lambda+s}\right)dA(\lambda)\right)=\exp\left(-\int_{0+}^{\infty}\ln\left(\frac{\lambda+s}{\lambda}\right)dA(\lambda)\right),$$

így az L egy olyan összetett függvény ahol a külső függvény teljesen monoton és a belső függvény deriváltja szintén teljesen monoton. Ezért az L is teljesen monoton, így az L egy $\xi\geq 0$ eloszlás transzformáltja. Világos, hogy minden n esetén az $\sqrt[n]{L(s)}$ szintén a megadott alakú, ezért az eloszlás korlátlanul osztható.

□

C.3. Hiperbolikusan teljes monotonitás

C.3.1. Definíció. A nem negatív számokon értelmezett f függvényt hiperbolikusan teljesen monotonnak mondjuk, ha minden $u>0$ esetén az $f(uv)\cdot f(u/v)$ teljesen monoton a $w\circeq v+1/v$ változó szerint. Vagyis az $f(uv)\cdot f(u/v)$ felírható mint a $w\circeq v+1/v$ függvénye és az így kapott $g(w)$ függvény teljesen monoton. Világos, hogy $w\geq 2$, így a teljes monotonitás a $[2,\infty)$ intervallumon értendő. Nyilván $v^2-wv+1=0$, amiből

$$v=\frac{w\pm\sqrt{w^2-4}}{2}.$$

Ennek egyik ága a v a másik ága az $1/v$ értéket szolgáltatja, ugyanis

$$\frac{w+\sqrt{w^2-4}}{2}\frac{w-\sqrt{w^2-4}}{2}=\frac{w^2-\left(w^2-4\right)}{4}=1.$$

Ennek megfelelően feltehető, hogy $v\geq 1$. Egy eloszlást hiperbolikusan teljesen monotonnak mondunk, ha a sűrűségfüggvénye hiperbolikusan teljesen monoton.

C.3.2. Példa. A másodfajú béta eloszlás hiperbolikusan teljesen monoton.

Általános paramétereket használva az f függvény egy másodfajú béta eloszlás sűrűségfüggvénye, ha alkalmas konstansokkal

$$f(x) = Cx^{\delta} \frac{1}{(1+x)^{\gamma}}, \quad x > 0$$

alakú. A hiperbolikus teljesen monotonitás igazolásához elegendő annyit feltenni, hogy $\gamma > 0$.

$$\begin{aligned} f(uv) f\left(\frac{u}{v}\right) &= C^2 u^{2\delta} \frac{1}{(1+uv)^{\gamma}} \frac{1}{(1+u/v)^{\gamma}} = \\ &= C^2 u^{2\delta} \frac{1}{\left(1+u(v+1/v)+u^2\right)^{\gamma}}. \end{aligned}$$

Elegendő belátni, hogy ha $b > 0$, akkor a

$$h(w) \mathrel{\hat{=}} \frac{1}{(a+bw)^{\gamma}}$$

végtelen sokszor deriválható függvény teljesen monoton, vagyis a deriváltjai alternálnak. Ez utóbbi azonban evidens.

□

C.3.3. Példa. A lognormális eloszlás hiperbolikusan teljesen monoton.

A lognormális eloszlás sűrűségfüggvénye alapján

$$\begin{gathered} f(uv) f\left(\frac{u}{v}\right) = \\ = \frac{1}{u^2\sigma^2 2\pi} \exp\left(-\frac{(\ln(uv)-\mu)^2 + (\ln(u/v)-\mu)^2}{2\sigma^2}\right). \end{gathered}$$

A kitevőben a négyzetre emeléseket elvégezve

$$(\ln(uv))^2 + \left(\ln\left(\frac{u}{v}\right)\right)^2 - 2\mu\left(\ln uv + \ln\frac{u}{v}\right) + 2\mu^2.$$

Mivel $\log uv + \log u/v = 2\log u$, ezért a kitevőben a v-től független tagokat összevonva elegendő megvizsgálni a

$$
\begin{aligned}
h(w) &\overset{\circ}{=} C\exp\left(-\frac{(\ln(uv))^2 + (\ln(u/v))^2}{2\sigma^2}\right) = \\
&= C\exp\left(-\frac{(\ln u + \ln v)^2 + (\ln u - \ln v)^2}{2\sigma^2}\right) = \\
&= C\exp\left(-\frac{(\ln(u))^2 + (\ln(v))^2}{\sigma^2}\right) \overset{\circ}{=} \\
&\overset{\circ}{=} K\exp\left(-\frac{(\ln(v))^{\,2}}{\sigma^2}\right)
\end{aligned}
$$

kifejezés teljes monotonitását. Mivel ha $\alpha > 0$, akkor az $\exp(-\alpha x)$ függvény teljesen monoton, ezért elég belátni, hogy a

$$\frac{d}{dw}(\ln v)^2 = 2\frac{\ln v}{v}\frac{dv}{dw}$$

függvény mint w függvénye teljesen monoton. $w = v + 1/v$, így

$$\frac{dw}{dv} = 1 - \frac{1}{v^2}.$$

Ha $v \neq 1$, vagyis ha $w > 2$, akkor az inverz függvény deriválási szabálya alapján

$$
\begin{aligned}
\frac{d}{dw}(\ln v)^2 &= 2\frac{\ln v}{v}\frac{1}{1 - 1/v^2} = \\
&= \frac{\ln v - \ln 1/v}{v - 1/v}.
\end{aligned}
$$

Ugyanakkor

$$
\begin{aligned}
\int_0^\infty \frac{1}{x+a} - \frac{1}{x+b}dx &= [\ln(x+a) - \ln(x+b)]_0^\infty = \\
&= \left[\ln\frac{x+a}{x+b}\right]_0^\infty = \ln 1 - \ln\frac{a}{b} = \\
&= \ln b - \ln a.
\end{aligned}
$$

Ebből

$$\begin{aligned}
\frac{d}{dw}(\ln v)^2 &= \frac{1}{v-1/v}\int_0^\infty \frac{1}{x+1/v}-\frac{1}{x+v}dx = \\
&= \frac{1}{v-1/v}\int_0^\infty \frac{(x+v)-(x+1/v)}{(x+1/v)(x+v)}dx = \\
&= \int_0^\infty \frac{1}{x^2+x(v+1/v)+1}dx = \\
&= \int_0^\infty \frac{1}{x^2+xw+1}dx.
\end{aligned}$$

Mivel az $1/(a+bw)$ függvény teljesen monoton, ezért az integrál mögött deriválva egyszerűen látható, hogy a $h(w)$ függvény is teljesen monoton.

□

Vegyük észre, hogy az előző számításokban a w függvényeként előálló képletek nem csak a $w \geq 2$ halmazon voltak teljesen monotonok, hanem a teljes $w \geq 0$ félegyenesen is. Ez nem véletlen.

C.3.4. Állítás. *Ha az f hiperbolikusan teljesen monoton függvény deriválható módon kiterjeszthető a* $\operatorname{Re} z > 0$ *komplex félsíkra*[11]*, akkor minden $u > 0$ esetén a $g(w) \stackrel{\circ}{=} f(uv)f(u/v)$ függvény is deriválhatóan kiterjeszthető a* $\operatorname{Re} w > 0$ *félsíkra, és ott alkalmas K növekedő súlyfüggvénnyel*

$$g(w) = \int_0^\infty \exp(-\beta w)\,dK(\beta), \quad \operatorname{Re} w > 0.$$

Bizonyítás: A hiperbolikus teljes monotonitás definíciója miatt az

$$f(uv)f\left(\frac{u}{v}\right) \stackrel{\circ}{=} g(w)$$

függvény minden $u > 0$ esetén teljesen monoton, ha $w \geq 2$. Bernstein tétele alapján $z \stackrel{\circ}{=} w-2$ helyettesítéssel

$$\begin{aligned}
g(w) &= g(z+2) \stackrel{\circ}{=} u(z) = \int_0^\infty \exp(-\beta z)\,dF(\beta) = \\
&= \int_0^\infty \exp(-\beta(w-2))\,dF(\beta) = \\
&= \int_0^\infty \exp(-\beta w)\exp(2\beta)\,dF(\beta) \stackrel{\circ}{=} \\
&\stackrel{\circ}{=} \int_0^\infty \exp(-\beta w)\,dK(\beta).
\end{aligned}$$

alakba írható, hacsak $w \geq 2$. A Bernstein-tétel igazolásakor kihasználtuk, hogy a reprezentálandó függvény a $(0,\infty)$ halmazon korlátos, következésképpen a nullában vett

[11] Mivel a $\log z$ komplex logaritmusfüggvény a $\operatorname{Re} z > 0$ félsíkban deriválható, ezért ez biztosan teljesül a lognormális eloszlásra. Mivel a $z^\alpha \stackrel{\circ}{=} \exp(\alpha \log z)$ hatványfüggvényeket a logaritmusfüggvény segítségével származtatjuk ez a tulajdonság szintén érvényes a másodfajú béta eloszlásra is.

határérték is véges, így a súlyfüggvény korlátos. Ugyanakkor az $\exp(2\beta)$ függvény nem korlátos, következésképpen a K súlyfüggvény bár minden pontban véges, nem feltétlenül korlátos. Tegyük fel, hogy az f függvény a $\operatorname{Re} z > 0$ félsíkban deriválható. Az $f(uv)\, f(u/v)$ kifejezés minden u pozitív valós szám esetén a $\operatorname{Re} v > 0$ halmazon szintén deriválható. Tekintsük a

$$v \mapsto \frac{v-1}{v+1} \overset{\circ}{=} z$$

leképezést, amely a $\operatorname{Re} v > 0$ nyílt félsíkot a $|z| < 1$ komplex nyílt egységkörre képezi. Valóban

$$\begin{aligned} \left|\frac{v-1}{v+1}\right|^2 &= \frac{v-1}{v+1}\overline{\left(\frac{v-1}{v+1}\right)} = \\ &= \frac{|v-1|^2}{|v+1|^2} = \\ &= \frac{(\operatorname{Re}(v)-1)^2 + (\operatorname{Im}(v))^2}{(\operatorname{Re}(v)+1)^2 + (\operatorname{Im}(v))^2} < 1 \end{aligned}$$

A leképezés ráképezés, ugyanis ha $|z| < 1$, akkor az inverz leképezés

$$\begin{aligned} v &\overset{\circ}{=} \frac{1+z}{1-z} = \frac{(1+z)\overline{(1-z)}}{|1-z|^2} = \\ &= \frac{(1+z)(1-\overline{z})}{|1-z|^2} = \frac{1-\overline{z}+z-|z|^2}{|1-z|^2} = \\ &= \frac{1-|z|^2+2i\operatorname{Im} z}{|1-z|^2}, \end{aligned}$$

amely valós része mindig pozitív. Tekintsük a $f(uv)\, f(u/v)$ függvényt a z függvényeként, vagyis végezzünk

$$v = \frac{1+z}{1-z}, \quad |z| < 1$$

helyettesítést. Jelölje a $|z| < 1$ halmazon értelmezett így kapott függvényt $h(z)$. Ha a z helyébe a $-z$ értéket tesszük, akkor a v helyébe éppen az

$$\frac{1-z}{1+z} = \frac{1}{v}$$

érték kerül, következésképpen $h(z) = h(-z)$. A h függvény deriválható, így az egységkörön belül hatványsorba fejthető. A $h(z) = h(-z)$ reláció miatt a h sorfejtésében csak páros kitevők szerepelnek. Vagyis $h(z) = \Sigma_k a_{2k} z^{2k}$. Ennek következtében

$$f(uv)\, f\left(\frac{u}{v}\right) = h\left(\frac{v-1}{v+1}\right) = \sum_k a_{2k}\left(\frac{v-1}{v+1}\right)^{2k}.$$

Viszont

$$\left(\frac{v-1}{v+1}\right)^2=\frac{1+v^2-2v}{1+v^2+2v}=\frac{1/v+v-2}{1/v+v+2}=\frac{w-2}{w+2}.$$

Ez utóbbi kifejezés deriválható a $\operatorname{Re} w>-2$ félsíkon. Ha $\operatorname{Re} w>0$, akkor

$$\left|\frac{w-2}{w+2}\right|<1,$$

így az

$$f(uv)\,f\left(\frac{u}{v}\right)=h\left(\frac{w-2}{w+2}\right)=g(w)$$

függvény kiterjeszthető a $\operatorname{Re} w>0$ félsíkra. A $g(w)$ a fenti K súlyfüggvény Laplace-transzformáltja a $\operatorname{Re} w>2$ félsíkban. Miként láttuk, egy növekedő súlyfüggvény Laplace-transzformáltjának értelmezési tartománya határán szükségszerűen szingularitás van, vagyis azon túl a függvény nem terjeszthető ki analitikus módon, miközben a $g(w)$ minden $\operatorname{Re} w>0$ esetén analitikus, ezért a

$$g(w)=\int_0^\infty \exp(-\beta w)\,dK(\beta)$$

előállítás minden $\operatorname{Re} w>0$ esetén is érvényes.

□

Nem kétséges, hogy a hiperbolikusan teljesen monoton eloszlás fogalma nem túlzottan kézenfekvő vagy szemléletes. Ugyanakkor a függelék fő állítása a következő:

C.3.5. Tétel (Thorin)**.** *Ha valamely eloszlás hiperbolikusan teljesen monoton*[12]*, akkor az eloszlás korlátlanul osztható.*

C.4. Általánosított gamma-konvolúciók

Az általánosított gamma-konvolúciók elmélete három észrevételre épül:

1. A gamma eloszlás korlátlanul osztható. Ha a ξ eloszlás (a,λ) paraméterű gamma eloszlás, akkor a ξ Laplace-transzformáltja

$$L(s)=\left(\frac{\lambda}{\lambda+s}\right)^a.$$

[12] A bizonyításhoz a sűrűségfüggvényre vonatkozóan néhány további, korábban már bevezetett technikai feltételre szükségünk lesz. Ezek a feltételek azonban könnyen ellenőrizhetők és általában teljesülő enyhe megkötések.

2. Független, korlátlanul osztható eloszlások összege szintén korlátlanul osztható. Mivel független eloszlások összegének transzformáltjai[13] a transzformáltak szorzata, ezért ha valamely $\xi \geq 0$ eloszlás transzformáltja

$$L(s) = \prod_{n=1}^{N} \left(\frac{\lambda_n}{\lambda_n + s} \right)^{a_n}$$

alakú, akkor a ξ korlátlanul osztható. Értelemszerűen a képletben szereplő (a_n, λ_n) paraméterek pozitívak. A képletet

$$\begin{aligned} L(s) &= \exp\left(\sum_{n=1}^{N} a_n \ln\left(\frac{\lambda_n}{\lambda_n + s} \right) \right) = \\ &= \exp\left(\int_{0+}^{\infty} \ln\left(\frac{\lambda}{\lambda + s} \right) dA_N(\lambda) \right) \end{aligned}$$

módon is írhatjuk, ahol A_N a λ_n pontokhoz a_n súlyt rendelő monoton növekedő súlyfüggvény. Mivel a konstansok szintén korlátlanul oszthatók, ezért az

$$L(s) = \exp\left(-sc + \int_{0+}^{\infty} \ln\left(\frac{\lambda}{\lambda + s} \right) dA_N(\lambda) \right)$$

alakú transzformáltak is korlátlanul oszthatóak.

3. Korlátlanul osztható eloszlások gyenge konvergenciában vett határértéke szintén korlátlanul osztható. Általánosított gamma-konvolúción a fenti véges konvolúciók gyenge konvergenciában való lezártját értjük. Nem negatív változók esetén a gyenge konvergenciából következik a Laplace-transzformációk konvergenciája[14], így a Laplace-transzformáció

$$\begin{aligned} L(s) &= \lim_{N \to \infty} \exp\left(-sc + \int_{0+}^{\infty} \ln\left(\frac{\lambda}{\lambda + s} \right) dA_N(\lambda) \right) = \\ &= \exp\left(-sc + \lim_{N \to \infty} \int_{0+}^{\infty} \ln\left(\frac{\lambda}{\lambda + s} \right) dA_N(\lambda) \right) \end{aligned}$$

alakú. Vegyük észre, hogy most nem tudjuk a Helly-féle kiválasztási tételt alkalmazni, mert az (A_N) sorozatról nem tudjuk, hogy egyenletesen korlátos volna. A problémák elkerülése céljából az alábbi definícióval élünk. Ugyancsak vegyük észre, hogy mivel a $\lambda = 0$ nem megengedett érték, ezért az integrálokat a $(0, \infty)$ halmazon kell venni, emiatt kell az alsó határra a $0+$ értéket írni.

[13] Generátorfüggvény, Laplace-transzformált, karakterisztikus függvény.

[14] Ugyanis az $x \mapsto \exp(-sx)$ függvények korlátosak az $x \geq 0$ halmazon feltéve, ha $s \geq 0$.

C.4.1. Definíció. A $\xi \geq 0$ változó eloszlását általánosított gamma-konvolúciónak mondjuk, ha a Laplace-transzformáltja

$$L(s) \triangleq \mathbf{E}(\exp(-s\xi)) = \exp\left(-sc + \int_{0+}^{\infty} \ln\left(\frac{\lambda}{\lambda+s}\right) dA(\lambda)\right), \quad s \geq 0$$

alakú.

Vegyük észre, hogy mivel az integrandus nem pozitív, ezért az integrál a közelítő összegek határértéke, így az $L(s)$ véges sok független gamma eloszlás konvolúciójából álló sorozat transzformáltjának határértéke. Korábban[15] a Berstein-tétel alkalmazásaként beláttuk, hogy az ilyen alakú $L(s)$ függvények mögött egyrészt egy eloszlás van, másrészt ez az eloszlás lényegében triviálisan korlátlanul osztható[16]. Belátható a fordított irány is, miszerint véges elemből álló gamma-konvolúciók gyenge konvergenciában vett határértékeinek Laplace-transzformáltjai a megadott képlet szerint írhatók fel. Ez utóbbira azonban nem lesz szükségünk. Az A súlyfüggvényt szokás a ξ-hez tartozó eloszlás Thorin-függvényének[17] mondani. Az A interpretációja igen kézenfekvő: azt mondja meg, hogy miként írható fel a ξ gamma eloszlások általánosított konvolúciójaként, másképpen A a ξ gamma eloszlások szerinti spektrálfüggvénye.

Ha $s > 0$, akkor az L, mint minden transzformált, végtelen sokszor deriválható. Az elméletben fontos szerepet játszik a

$$\rho(s) \triangleq \frac{d}{ds} \ln L(s) = \frac{L'(s)}{L(s)}$$

függvény. A ρ az egyetlen olyan függvény, amely segítségével

$$L(s) = \exp\left(\int_0^s \rho(u)\, du\right).$$

Ha $r(s) \triangleq \int_{0+}^{\infty} 1/(\lambda+s)\, dA(\lambda)$, akkor a Fubini-tétel alapján az integrálokat felcserélve minden $s > 0$ esetén

$$\begin{aligned}
\int_0^s r(u)\, du &= \int_0^s \int_{0+}^{\infty} \frac{1}{\lambda+u} dA(\lambda)\, du = \\
&= \int_{0+}^{\infty} \int_0^s \frac{1}{\lambda+u} du dA(\lambda) = \\
&= \int_{0+}^{\infty} [\ln(\lambda+u)]_0^s\, dA(\lambda) = \\
&= \int_{0+}^{\infty} \ln \frac{\lambda+s}{\lambda} dA(\lambda),
\end{aligned}$$

---

[15] V.ö.: C.2.6. példa, 299. oldal.

[16] Vagyis a nehéz rész, hogy van az $L(s)$ mögött egy eloszlás. Ehhez kell a Bernstein-tétel. Az n részre való oszthatósághoz elég az A/n súlyfüggvényt venni. Ugyancsak v.ö: B.3.17. állítás, 259. oldal.

[17] Olaf Thorin, az elmélet kidolgozója, tiszteletére.

amiből a logaritmusfüggvény alaptulajdonsága alapján

$$L(s) = \exp\left(-sc - \int_0^s r(u)\,du\right).$$

Az r függvény folytonos, így

$$\rho(s) = \frac{L'(s)}{L(s)} = -c - r(s) \overset{\circ}{=} a - \int_{0+}^{\infty} \frac{1}{\lambda + s} dA(\lambda). \qquad \text{(C.4.1)}$$

A számolásból egyúttal az is következik, hogy a ρ képletében szereplő integrál minden s esetén véges. Ebből következően az $L(s)$ is deriválható módon kiterjeszthető a $\mathbb{C}\setminus(-\infty,0]$ halmazra. Ha $z \in \mathbb{C}\setminus(-\infty,0]$, akkor a kiterjesztett ρ függvény

$$\begin{aligned} \rho(z) &= a - \int_{0+}^{\infty} \frac{1}{\lambda + z} dA(\lambda) = \\ &= a - \int_{0+}^{\infty} \frac{\lambda + \operatorname{Re} z - i \operatorname{Im} z}{(\lambda + \operatorname{Re} z)^2 + (\operatorname{Im} z)^2} dA(\lambda). \end{aligned}$$

Így ha $\operatorname{Im} z > 0$, akkor $\operatorname{Im} \rho(z) > 0$. A függelék legfőbb technikai eszköze, hogy a fordított irány is igaz. Vagyis igaz a következő:

C.4.2. Állítás. *Ha a pozitív valós számokon értelmezett valós ρ függvényre $\rho(s) \leq 0$, és ρ folytonosan kiterjeszthető deriválható komplex függvényként az $\operatorname{Im} z > 0$ félsíkra, és a kiterjesztett függvény imaginárius része nem negatív, akkor a ρ reprezentálható*

$$\rho(s) = a - \int_0^{\infty} \frac{1}{u + s} dA(u). \quad s > 0,$$

módon ahol A egy alkalmas monoton növekedő súlyfüggvény, és a egy valós konstans.

A következő alpont célja az állítás igazolása, amely első olvasásra elhagyható.

Néhány közismert komplex függvénytani tétel

Az előző alfejezet végén tett észrevétel indokolja a következő definíciót:

C.4.3. Definíció. Az $\operatorname{Im} z > 0$ nyílt felső komplex félsíkon értelmezett $f(z)$ deriválható függvényt Pick- vagy Nevanlinna-függvénynek mondjuk, ha itt $\operatorname{Im} f(z) \geq 0$. Vagyis az f függvény az $\operatorname{Im} z > 0$ felső nyílt félsík pontjait vagy ugyanezen félsíkba, vagy a valós egyenesre képezi. Érdemes hangsúlyozni, hogy a Pick-függvények értelmezési tartományához nem tartoznak feltétlenül a valós számok és a definíció alapján nem világos, hogy egy Pick-függvény folytonosan kiterjeszthető-e a valós számokra.

Ebben az alpontban az ilyen típusú függvények integrálreprezentációs tulajdonságát fogjuk megvizsgálni. Kezdjük egy közismert tétellel amely igazolása minden komplex analízis könyvben megtalálható:

C.4.4. Állítás (Poisson-formula). *Ha $f(z)$ egy olyan komplex függvény, amely a $|z|<1$ nyílt körlapon deriválható és a $|z|\le 1$ zárt körlapon folytonos, akkor*

$$f(z)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{Re}\frac{\exp(it)+z}{\exp(it)-z}\cdot f(\exp(it))\,dt,\qquad |z|<1.$$

Bizonyítás: Mielőtt a bizonyításra rátérnénk, érdemes annak tartalmát megvizsgálni. Az $\exp(it)$ mint t függvénye végigfutja az egységkört és így az $f(\exp(it))$ egy a körvonalon értelmezett függvény. Az állítás szerint a nyílt körlemez egy pontjában az f értéke a képletben szereplő magfüggvény szerinti integrálja a körvonalon felvett értékeknek. Következésképpen az f egyértelműen meghatározott a körvonalon felvett értékek által. Továbbá az integrál mögött deriválva könnyen látható, hogy minden a körvonalon folytonos függvény esetén az integrállal definiált f deriválható, így a körvonalon folytonos függvények és az állításban szereplő függvények egyértelműen megfeleltethetők egymásnak. A magfüggvény

$$\begin{aligned}\operatorname{Re}\frac{\exp(it)+z}{\exp(it)-z} &= \operatorname{Re}\frac{(\exp(it)+z)(\exp(-it)-\overline{z})}{|\exp(it)-z|^2}=\\ &= \operatorname{Re}\frac{1-|z|^2+ze^{-it}-e^{it}\overline{z}}{|e^{it}-z|^2}=\\ &= \operatorname{Re}\frac{1-|z|^2+ze^{-it}-\overline{e^{-it}z}}{|e^{it}-z|^2}=\\ &= \operatorname{Re}\frac{1-|z|^2+2\operatorname{Im}ze^{-it}}{|e^{it}-z|^2}=\\ &= \frac{1-|z|^2}{|\exp(it)-z|^2}.\end{aligned}$$

Ha most polárkoordinátákban $z=r\exp(i\varphi)$, akkor

$$\begin{gathered}|\exp(it)-z|^2=\\ =(\exp(it)-r\exp(i\varphi))(\exp(-it)-r\exp(-i\varphi))=\\ =1-r\exp(i(t-\varphi))-r\exp(i\varphi-t)+r^2=\\ =1-2r\cos(t-\varphi)+r^2.\end{gathered}$$

Vagyis a magfüggvény

$$\frac{1-r^2}{1-2r\cos(t-\varphi)+r^2}.$$

A bizonyítás során némi probléma származik abból, hogy a feltételek szerint az f csak az egységkör belsejében deriválható. Legyen $0<r<1$ és $f_r(z)\doteq f(rz)$. Az f_r a $|z|<$

$1/r$ nyílt körlapon deriválható. Miként láttuk[18], ha $|z|<1$, és γ az egységkör, akkor

$$\begin{aligned} f(rz) &\overset{\circ}{=} f_r(z) = \frac{1}{2\pi i}\int_\gamma \frac{f_r(u)}{u-z}du = \\ &= \frac{1}{2\pi i}\int_{-\pi}^{\pi} \frac{f(r\exp(it))}{\exp(it)-z} i\exp(it)\,dt. \end{aligned}$$

Ugyanakkor az f a feltételek szerint folytonos, így ha $r \nearrow 1$, akkor

$$f(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(\exp(it))}{\exp(it)-z}\exp(it)\,dt,$$

feltéve, hogy a határérték és az integrálás felcserélhető. Ez utóbbi azonban azért tehető meg, mert mivel $|z|<1$, ezért a z távolsága a körvonaltól pozitív. Legyen most $z^* = 1/\bar{z}$. Egyszerű számolással $|z^*|>1$. Az $f(z)/(z-z^*)$ függvény a körön belül deriválható, így ha γ a körvonal, akkor a már látott határértékes megfontolással a Cauchy-tétel alapján ha $z = r\exp(i\varphi)$

$$\begin{aligned} 0 &= \frac{1}{2\pi i}\int_\gamma \frac{f(u)}{u-z^*}du = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(\exp(it))}{\exp(it)-z^*}\exp(it)\,dt = \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(\exp(it))}{\exp(it)-\exp(i\varphi)/r}\exp(it)\,dt = \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{r\exp(it)}{r\exp(it)-\exp(i\varphi)} f(\exp(it))\,dt. \end{aligned}$$

A két integrált egymásból kivonva

$$\begin{aligned} f(z) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f\left(e^{it}\right)\left(\frac{\exp(it)}{\exp(it)-r\exp(i\varphi)} - \frac{r\exp(it)}{r\exp(it)-\exp(i\varphi)}\right)dt = \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f\left(e^{it}\right)\left(\frac{1}{1-r\exp(i(\varphi-t))} - \frac{r}{r-\exp(i\varphi-t)}\right)dt = \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f\left(e^{it}\right)\frac{1-r^2}{1-2r\cos(t-\varphi)+r^2}dt. \end{aligned}$$

□

C.4.5. Állítás (Cauchy–Schwarz-formula)**.** *Ha $f(z)$ egy olyan komplex függvény, amely a $|z|<1$ nyílt körlapon deriválható és a $|z|\le 1$ zárt körlapon folytonos, akkor a nyílt körlap minden pontjában*

$$\begin{aligned} f(z) &= i\operatorname{Im} f(0) + \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z}\operatorname{Re} f(\exp(it))\,dt. \\ f(z) &= \operatorname{Re} f(0) + \frac{i}{2\pi}\int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z}\operatorname{Im} f(\exp(it))\,dt. \end{aligned}$$

[18] V.ö.: (C.1.2) sor.

Bizonyítás: A két formula igen hasonló és egyik a másikból egyszerűen származtatható és mindkettő a Poisson-formula alternatív felírási módja. Jelölje $F(z)$ az első formulában szereplő integrált. Az integráljel mögöttt deriválva azonnal látható, hogy az F is deriválható a $|z| < 1$ tartományon. A Poisson-formula alapján

$$\begin{aligned}
\operatorname{Re} F(z) &= \frac{1}{2\pi} \operatorname{Re} \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} \operatorname{Re} f(\exp(it))\, dt = \\
&= \frac{1}{2\pi} \int_{0}^{2\pi} \operatorname{Re} \frac{\exp(it)+z}{\exp(it)-z} \operatorname{Re} f(\exp(it))\, dt = \\
&= \frac{1}{2\pi} \operatorname{Re} \int_{-\pi}^{\pi} \operatorname{Re} \frac{\exp(it)+z}{\exp(it)-z} f(\exp(it))\, dt = \\
&= \operatorname{Re} f(z).
\end{aligned}$$

A $g(z) \circeq F(z) - f(z)$, függvény valós része nulla, így a valós rész deriváltja is nulla. Így a derivált mátrix első sora nulla, de akkor a második sora is nulla, vagyis a derivált azonosan nulla. A $|g|^2$ függvényre mint kétváltozós valós értékű deriválható függvényre bármely

$$[0,z] = tz, \quad 0 \le t \le 1$$

egyváltozós szakaszon alkalmazva a valós középérték-tételt azonnal látható, hogy a g konstans. Következésképpen $F(z) - f(z) = F(0) - f(0)$. De ismételten a Poisson-formula miatt

$$\begin{aligned}
F(0) &\circeq \frac{1}{2\pi} \int_{\pi}^{\pi} \operatorname{Re} f(\exp(it))\, dt = \\
&= \frac{1}{2\pi} \operatorname{Re} \int_{-\pi}^{\pi} f(\exp(it))\, dt = \\
&= \frac{1}{2\pi} \operatorname{Re} \int_{-\pi}^{\pi} \operatorname{Re} \frac{\exp(it)+0}{\exp(it)-0} f(\exp(it))\, dt = \operatorname{Re} f(0).
\end{aligned}$$

Ebből

$$f(z) = F(z) + f(0) - F(0) = F(z) + i \operatorname{Im} f(0),$$

ami éppen az igazolandó első formula. Legyen most $g(z) \circeq i f(z)$. Erre az első formulát alkalmazva,

$$\begin{aligned}
f(z) &= \frac{g(z)}{i} = \frac{1}{i} \left(i \operatorname{Im} g(0) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} \operatorname{Re} g(\exp(it))\, dt \right) = \\
&= \operatorname{Im} g(0) + \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} \operatorname{Re} g(\exp(it))\, dt = \\
&= \operatorname{Re} f(0) - \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} \operatorname{Im} f(\exp(it))\, dt = \\
&= \operatorname{Re} f(0) + \frac{i}{2\pi} \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} \operatorname{Im} f(\exp(it))\, dt.
\end{aligned}$$

□

C.4.6. Állítás (Herglotz). *Legyen* $f(z)$ *a* $|z|<1$ *nyílt körlapon deriválható és tegyük fel, hogy* $\operatorname{Im} f(z) \geq 0$. *Ekkor alkalmas* H *korlátos, monoton növekedő súlyfüggvénnyel*

$$f(z) = \operatorname{Re} f(0) + i \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} dH(t).$$

Bizonyítás: Vegyük észre, hogy az f függvényről nem tettük, fel, hogy folytonos a zárt körlapon. Legyen $r<1$. Ekkor a $f_r(z) \stackrel{\circ}{=} f(rz)$ függvényre alkalmazva az előző állítást, minden $|z|<1$ esetén

$$\begin{aligned} f(rz) &= f_r(z) = \\ &= \operatorname{Re} f_r(0) + \frac{i}{2\pi} \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} \operatorname{Im} f_r(\exp(it))\, dt = \\ &= \operatorname{Re} f(0) + \frac{i}{2\pi} \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} \operatorname{Im} f(r\exp(it))\, dt \stackrel{\circ}{=} \\ &\stackrel{\circ}{=} \operatorname{Re} f(0) + i \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} dH_r(t), \end{aligned}$$

ahol az $\operatorname{Im} f(z) \geq 0$ feltétel miatt a H_r növekedő és

$$0 \leq H_r(x) \stackrel{\circ}{=} \frac{1}{2\pi} \int_{-\pi}^{x} \operatorname{Im} f(r\exp(it))\, dt \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \operatorname{Im} f(r\exp(it))\, dt = \operatorname{Im} f(0),$$

ugyanis a Cauchy–Schwarz-formula alapján a $z=0$ helyen

$$f(0) = \operatorname{Re} f(0) + \frac{i}{2\pi} \int_{-\pi}^{\pi} \operatorname{Im} f(r\exp(it))\, dt,$$

amiből

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \operatorname{Im} f(r\exp(it))\, dt = \frac{1}{i}(f(0) - \operatorname{Re} f(0)) = \operatorname{Im} f(0).$$

A H_r növekedő függvények családja egyenletesen korlátos, így a Helly-tétele[19] alapján van olyan $r_k \nearrow 1$ sorozat, és H súlyfüggvény, amelyre $H_{r_k} \to H$ a H minden folytonossági pontjában. Az

$$u(t) \stackrel{\circ}{=} \frac{\exp(it)+z}{\exp(it)-z}, \quad -\pi \leq t \leq \pi$$

függvény folytonos, a $-\pi$ és π pontokban a (H_r) súlyfüggvények konvergálnak, így

$$f(z) = \lim_{r \nearrow 1} f(rz) = \operatorname{Re} f(0) + i \int_{-\pi}^{\pi} \frac{\exp(it)+z}{\exp(it)-z} dH(t).$$

□

[19] V.ö.: B.3.13. Állítás, 256. oldal.

C.4.7. Állítás. *Ha az előző tételben $f(z)$ a körvonal egy nyílt intervallumára folytonosan kiterjeszthető és ott az f valós, akkor az intervallum H szerinti mértéke nulla.*

Bizonyítás: Herglotz-tételben

$$H_r(x) \mathrel{\hat{=}} \frac{1}{2\pi}\int_{-\pi}^{x} \operatorname{Im} f(r\exp(it))\,dt.$$

Ha most valamely x_1 és x_2 pontok a H folytonossági pontjai és az általuk meghatározott íven az f folytonos valós függvényként terjeszthető ki, akkor

$$\begin{aligned}
2\pi(H(x_1)-H(x_2)) &= \lim_{r\nearrow 1}\int_{x_1}^{x_2} \operatorname{Im} f(r\exp(it))\,dt = \\
&= \int_{x_1}^{x_2} \lim_{r\nearrow 1} \operatorname{Im} f(r\exp(it))\,dt = \\
&= \int_{x_1}^{x_2} \operatorname{Im} f(\exp(it))\,dt = \int_{x_1}^{x_2} 0dt = 0,
\end{aligned}$$

ahol a folytonos kiterjeszthetőség miatt az integrandus egyenletesen korlátos, így a határérték és az integrálás sorrendje felcserélhető. Ebből az állítás igazolása már evidens. □

Végezetül legyen f egy Pick-függvény. Tekintsük a

$$v \mapsto \frac{i-v}{i+v} = -1 + \frac{2i}{i+v}$$

leképezést. Ha $\operatorname{Im} v \geq 0$, akkor

$$\left|\frac{i-v}{i+v}\right|^2 = \frac{i-v}{i+v}\frac{\overline{i-v}}{\overline{i+v}} = \frac{(\operatorname{Re} v)^2 + (\operatorname{Im} v - 1)^2}{(\operatorname{Re} v)^2 + (\operatorname{Im} v + 1)^2} \leq 1,$$

tehát a leképezés az $\operatorname{Im} v \geq 0$ halmazt az egységkörlapra képezi. A leképezés bijekció, ugyanis az inverze

$$\begin{aligned}
i\frac{1-v}{1+v} &= i\frac{(1-v)(1+\overline{v})}{|1+v|^2} = \\
&= i\frac{1-|v|^2+\overline{v}-v}{|1+v|^2} = \\
&= i\frac{1-|v|^2-2i\operatorname{Im} v}{|1+v|^2}
\end{aligned}$$

amely imaginárius része nem negatív valahányszor $|v| \leq 1$. A $|v| = 1$, $v \neq -1$ képe $2\operatorname{Im} v/|1+v|^2$ amely egy valós szám. A $v = -1$ ponthoz rendeljük hozzá $+\infty$ pontot. Tekintsük a

$$g(v) = f\left(i\frac{1-v}{1+v}\right), \quad |v| < 1$$

függvényt. Herglotz-tétele alapján ha $|v| < 1$, akkor $u = \tan t/2$ helyettesítéssel[20]

$$\begin{aligned}
f(z) &= g(v) = \operatorname{Re} g(0) + i\int_{-\pi}^{\pi} \frac{\exp(it)+v}{\exp(it)-v} dH(t) = \\
&= \alpha + i\mu\frac{-1+v}{-1-v} + \int_{-\pi+0}^{\pi-0} i\frac{\exp(it)+v}{\exp(it)-v} dH(t) = \\
&= \alpha + \mu z + \int_{-\pi+0}^{\pi-0} i\frac{\exp(it/2)+\exp(-it/2)v}{\exp(it/2)-\exp(-it/2)v} dH(t) = \\
&= \alpha + \mu z + \int_{-\pi+0}^{\pi-0} i\frac{\exp(it/2)+\exp(-it/2)\left(-1+\frac{2i}{i+z}\right)}{\exp(it/2)-\exp(-it/2)\left(-1+\frac{2i}{i+z}\right)} dH(t) = \\
&= \alpha + \mu z + \int_{-\pi+0}^{\pi-0} i\frac{\sin t/2+\exp(-it/2)\frac{1}{i+z}}{\frac{1}{i}\cos t/2-\exp(-it/2)\frac{1}{i+z}} dH(t) = \\
&= \alpha + \mu z + \int_{-\pi+0}^{\pi-0} \frac{(i+z)\sin t/2+\exp(-it/2)}{-(i+z)\cos t/2+i\exp(-it/2)} dH(t) \\
&= \alpha + \mu z + \int_{-\pi+0}^{\pi-0} \frac{(i+z)\tan t/2+1-i\tan t/2}{-(i+z)+i+\tan t/2} dH(t) = \\
&= \alpha + \mu z + \int_{-\pi+0}^{\pi-0} \frac{z\tan t/2+1}{\tan t/2-z} dH(t) \mathrel{\mathring{=}} \\
&\mathrel{\mathring{=}} \alpha + \mu z + \int_{-\infty}^{\infty} \frac{uz+1}{u-z} dP(u).
\end{aligned}$$

Vegyük észre, hogy $\mu \geq 0$ és P egy véges mérték. Ezzel igazoltuk a következőt:

C.4.8. Állítás. *Ha az $f(z)$ egy Pick-függvény, akkor az f minden $\operatorname{Im} z > 0$ esetén felírható*

$$f(z) = \alpha + \mu z + \int_{-\infty}^{\infty} \frac{1+uz}{u-z} dP(u)$$

módon, ahol $\mu \geq 0$ és P egy monoton növekedő korlátos súlyfüggvény. Ha az $f(z)$ Pick-függvény valós függvényként folytonosan kiterjeszthető a valós számok egy intervallumára, akkor az intervallum súlya a P szerint nulla.

Most már minden készen áll arra, hogy az előző alpont végén említett állítást igazoljuk:

C.4.9. Állítás. *Ha az $f(z)$ Pick-függvény folytonosan kiterjeszthető a pozitív valós számokra*[21]*, és ott a kiterjesztett $f(s)$ függvény valós és $f(s) \leq 0$, akkor*

$$f(s) = a - \int_0^{\infty} \frac{1}{u+s} dA(u). \quad s > 0, \tag{C.4.2}$$

ahol A egy alkalmas monoton növekedő súlyfüggvény, és a egy valós konstans.

[20] V.ö.: (4.2.3) sor, 54. oldal.

[21] A függvények eredendően az $\operatorname{Im} z > 0$ tartományon vannak értelmezve, amelynek a valós egyenes nem része.

Bizonyítás: Az előző karakterizációs tétel miatt minden $\mathrm{Im}\, z > 0$ esetén

$$f(z) = \alpha + \mu z + \int_{(-\infty,0]} \frac{1+uz}{u-z} dP(u).$$

Ha $\mathrm{Re}\, z > 0$, akkor mivel $u \leq 0$, ezért

$$|u-z| = \sqrt{(u - \mathrm{Re}(z))^2 + \mathrm{Im}(z)^2} \geq \mathrm{Re}(z),$$

így az integrálandó kifejezés korlátos és mivel a $P(u)$ is korlátos, így határértéket véve könnyen látható, hogy az integrálelőállítás a pozitív valós számokon is teljesül, vagyis

$$f(s) = \alpha + \mu s + \int_{(-\infty,0]} \frac{1+us}{u-s} dP(u), \quad s > 0.$$

Alább a Fatou-lemmát szeretnénk használni, de az integrandus nem lesz nem negatív, ezért két felé kell bontani. Mivel

$$\left| \frac{1}{u-s} \right| \leq \frac{1}{s}$$

ezért az

$$\int_{(-\infty,0]} \frac{1}{u-s} dP(u)$$

integrál véges és használható a majorált konvergencia tétele, így

$$\lim_{s\to\infty} \int_{(-\infty,0]} \frac{1}{u-s} dP(u) = 0.$$

Az $us/(u-s) \geq 0$ így

$$\begin{aligned} \liminf_{s\to\infty} \int_{(-\infty,0]} \frac{us}{u-s} dP(u) &\geq \int_{(-\infty,0]} \liminf_{s\to\infty} \frac{us}{u-s} dP(u) = \\ &= \int_{(-\infty,0]} -u dP(u). \end{aligned}$$

Mivel pedig $f(s) \leq 0$, ezért

$$\begin{aligned} 0 &\geq \liminf_{s\to\infty} \left(a + \mu s + \int_{(-\infty,0]} \frac{1+us}{u-s} dP(u) \right) \geq \\ &\geq a + \liminf_{s\to\infty} \mu s + \int_{(-\infty,0]} \liminf_{s\to\infty} \frac{us}{u-s} dP(u) \\ &\geq a + \liminf_{s\to\infty} \mu s - \int_{(-\infty,0]} u dP(u). \end{aligned}$$

Ez csak akkor teljesülhet, ha $\mu \leq 0$ és az $\int_{(-\infty,0]} udP(u)$ integrál véges. De mivel $\mu \geq 0$, ezért $\mu = 0$.

$$\frac{1+us}{u-s} = \frac{1+u^2-u(u-s)}{u-s} = \frac{1+u^2}{u-s} - u.$$

Az integrál értékét a konstansba beolvasztva,

$$\begin{aligned} f(s) &= a - \int_{(-\infty,0]} \frac{1+u^2}{-u+s} dP(u) = \\ &= a + \int_{(-\infty,0]} \frac{1}{-u+s} \left(1+u^2\right) d\left(P(\infty) - P(u)\right) = \\ &= a - \int_{[0,\infty)} \frac{1}{u+s} \left(1+u^2\right) d\left(P(\infty) - P(-u)\right) \stackrel{\circ}{=} \\ &\stackrel{\circ}{=} a - \int_0^{\infty} \frac{1}{u+s} dA(u). \end{aligned}$$

Ezzel az állítást és így az előző alpont végén kimondott állítás bizonyítását befejeztük.

$\square$

C.4.10. Példa. Az A függvény nulla pontban vett $\Delta A(0)$ ugrása lehet nem nulla.

A későbbiek szempontjából nem érdektelen megjegyezni, hogy az $f(z) = -1/z$ függvény kielégíti az állítás feltételeit. A hozzá tartozó A súlyfüggvény az $u = 0$ pontra egységnyi súlyt helyező ugrófüggvény, amelyre $\Delta A(0) = 1$.

$\square$

Általánosított gamma-konvolúciók komplex függvénytani karakterizációja

Térjünk vissza az általánosított gamma-konvolúciók vizsgálatára.

C.4.11. Állítás. *Ha egy $\xi \geq 0$ valószínűségi változó L Laplace-transzformáltja deriválható módon kiterjeszthető a $\mathbb{C} \setminus (-\infty, 0]$ halmazra*[22]*, és*

$$\operatorname{Im}\left(L'(z)\overline{L(z)}\right) \geq 0, \quad \textit{ha} \quad \operatorname{Im} z > 0,$$

akkor ξ egy általánosított gamma-konvolúció.

Bizonyítás: Első lépésként megmutatjuk, hogy a feltétel teljesülése esetén $L(z) \neq 0$ az $\operatorname{Im} z > 0$ halmazon. Ha mégis $L(z_0) = 0$ lenne, akkor mivel az L analitikus, ezért a z_0

[22]Miként korábban már láttuk általában egy $\xi \geq 0$ Laplace-transzformáltja az $s = 0$ pontba már nem terjeszthető ki deriválható módon.

pont körül sorba fejthető. Ezért alkalmas $k \geq 1$ és $a_k \neq 0$ számokkal, ahol a_k az első nem nulla együttható

$$\begin{aligned} L(z) &= a_k (z - z_0)^k + O\left(|z - z_0|^{k+1}\right). \\ L'(z) &= k a_k (z - z_0)^{k-1} + O\left(|z - z_0|^k\right). \end{aligned}$$

Ekkor

$$\operatorname{Im}\left(L'(z)\overline{L(z)}\right) = k|a_k|^2 |z - z_0|^{2k-2} \operatorname{Im}\overline{z - z_0} + O\left(|z - z_0|^{2k}\right).$$

Ha most $z - z_0 = iy$, ahol $y > 0$ és elég kicsi, akkor a most kapott kifejezés negatív, ami ellentmondás. Legyen

$$\rho(z) \stackrel{\circ}{=} \frac{L'(z)}{L(z)}, \quad \operatorname{Im} z > 0,$$

amely így értelmes. A feltétel szerint

$$\operatorname{Im} \rho(z) \stackrel{\circ}{=} \operatorname{Im} \frac{L'(z)}{L(z)} = \operatorname{Im} \frac{L'(z)\overline{L(z)}}{|L(z)|^2} \geq 0.$$

Ugyanakkor a pozitív valós számokon $\rho(s) \leq 0$ ugyanis a valós számokon

$$L(s) = \exp\left(\int_0^s \rho(u)\, du\right)$$

és az L monoton csökken, Ebből az előző alpont végén igazolt (C.4.2) előállítás miatt

$$\rho(s) = a - \int_0^\infty \frac{1}{u+s} dA(u), \quad s > 0.$$

Ebből az állítás már evidens ugyanis

$$\begin{aligned} L(s) &= \exp\left(\int_0^s \rho(u)\, du\right) = \\ &= \exp\left(sa - \int_0^s \int_0^\infty \frac{1}{\lambda + u} dA(\lambda)\, du\right) = \\ &= \exp\left(sa - \int_0^s \int_0^\infty \frac{1}{\lambda + u} dA(\lambda)\, du\right) = \\ &= \exp\left(sa - \int_0^s \int_0^\infty \frac{1}{u + \lambda} dA(\lambda)\, du\right) = \\ &= \exp\left(sa - \int_0^\infty \int_0^s \frac{1}{u + \lambda} du dA(\lambda)\right) = \\ &= \exp\left(sa + \int_0^\infty \ln \frac{\lambda}{\lambda + s} dA(\lambda)\right). \end{aligned}$$

Érdemes megjegyezni, hogy $\Delta A(0)=0$, ugyanis ellenkező esetben az integrál értéke minden $s>0$ esetén mínusz végtelen lenne:

$$\begin{aligned}\int_0^s\int_0^\infty \frac{1}{\lambda+u}dA(\lambda)\,du &= \int_0^s\int_{0+}^\infty \frac{1}{\lambda+u}dA(\lambda)+\frac{1}{u}\Delta A(0)\,du=\\ &= \int_0^s\int_{0+}^\infty \frac{1}{\lambda+u}dA(\lambda)+\Delta A(0)\int_0^s\frac{1}{u}du=\infty.\end{aligned}$$

□

C.4.12. Példa. $L(s)=\exp(-s^\alpha)$, $0<\alpha<1$ transzformációval rendelkező stabil eloszlás általánosított gamma-konvolúció.

Miként ismert, ha $0<\alpha<1$, akkor az $L(s)=\exp(-s^\alpha)$ transzformálttal megadott eloszlás korlátlanul osztható. A korlátlanul való oszthatóság evidens, ugyanis

$$\sqrt[n]{L(s)}=\exp\left(-\frac{s^\alpha}{n}\right)=\exp\left(-\left(\frac{s}{n^{1/\alpha}}\right)^\alpha\right),$$

amely mögötti változó csak konstansszorosa az $L(s)$-hez tartozó változónak, feltéve ha ilyen változó van, vagyis az L valóban egy eloszlás transzformáltja. Az L egy valószínűségi változó transzformáltja, ugyanis az $L(s)$ az $s\geq 0$ tartományon folytonos, $L(0)=1$ és teljesen monoton. Valóban az $\exp(-s)$ függvény teljesen monoton és az s^α kitevő deriváltja az $as^{\alpha-1}\overset{\circ}{=}\alpha/s^\gamma$ is teljesen monoton ugyanis $\gamma\overset{\circ}{=}1-\alpha>0$. Az

$$\frac{L'(s)}{L(s)}=-\alpha s^{\alpha-1}=-\alpha\exp((\alpha-1)\ln s)$$

függvény kiterjeszthető a $\mathbb{C}\backslash(-\infty,0]$ tartományra ugyanis a logaritmusfüggvény kiterjeszthető a megadott tartományra. A z komplex számot (ρ,φ) trigonometrikus alakba írva

$$\begin{aligned}\frac{L'(z)}{L(z)} &= -\alpha\exp((\alpha-1)(\ln\rho+i\varphi))=\\ &= -\alpha\rho^{\alpha-1}\exp((\alpha-1)i\varphi),\end{aligned}$$

amely imaginárius része

$$\operatorname{Im}\frac{L'(z)}{L(z)}=-\alpha\rho^{\alpha-1}\sin((\alpha-1)\varphi)>0$$

ugyanis ha $\operatorname{Im} z > 0$, akkor $0 < \varphi < \pi$. A másodfajú béta eloszlás sűrűségfüggvénye szerint

$$\begin{aligned}
\int_0^\infty \frac{1}{u+s} u^{\alpha-1} du &= \frac{1}{s} \int_0^\infty \frac{1}{u/s+1} u^{\alpha-1} du = \\
&= \frac{1}{s} \int_0^\infty \frac{1}{t+1} (st)^{\alpha-1} s dt = \\
&= s^{\alpha-1} \int_0^\infty \frac{1}{t+1} t^{\alpha-1} dt = \\
&= s^{\alpha-1} B(\alpha, 1-\alpha) = \\
&= s^{\alpha-1} \Gamma(1-\alpha) \Gamma(\alpha).
\end{aligned}$$

Emlékeztetünk, hogy egy általánosított gamma-konvolúció Thorin-függvénye az

$$L(s) = \exp\left(\int_{0+}^\infty \ln\left(\frac{\lambda}{\lambda+s}\right) dA(\lambda)\right), \quad s \geq 0$$

előállításban szereplő A súlyfüggvény. Miként láttuk ez éppen a

$$\rho(s) = a - \int_{0+}^\infty \frac{1}{u+s} dA(u), \quad s > 0$$

előállításban szereplő A súlyfüggvény. A fentiek alapján a jelen példában

$$\rho(s) = -\alpha s^{\alpha-1} = -\frac{\alpha}{\Gamma(1-\alpha)\Gamma(\alpha)} \int_0^\infty \frac{1}{u+s} u^{\alpha-1} du.$$

Ebből a példában szereplő stabil eloszlás Thorin-függvényének sűrűségfüggvénnye

$$\frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)} u^{\alpha-1}, \quad u > 0.$$

$\square$

A hiperbolikusan teljesen monoton eloszlások általánosított gamma-konvolúciók

A lognormális eloszlás vagy a másodfajú béta eloszlás korlátlanul való oszthatóságához elegendő tehát belátni, hogy ha f egy „jó tulajdonságú” hiperbolikusan teljesen monoton sűrűségfüggvény, akkor az $L(s)$ Laplace-transzformáció deriválható módon kiterjeszthető a $\mathbb{C} \setminus (-\infty, 0]$ halmazra, és a kiterjesztett függvényre

$$\operatorname{Im} L'(z) \overline{L(z)} \geq 0, \quad \operatorname{Im} z > 0.$$

1. Első lépésként biztosítani kell, hogy az L kiterjeszthető legyen a $\mathbb{C}\backslash(-\infty,0]$ halmazra. Ehhez természetesen szükségünk van a már bemutatott deriválhatósági feltevésekre[23]. Miként láttuk, a megadott feltételek teljesülése esetén az $L(z)$ kiterjeszthető a negatív valós tengely mentén felvágott komplex síkra, és ott

$$L(z)=\exp(-i\varphi)\int_0^\infty \exp(-z\exp(-i\varphi)t)\,f(\exp(-i\varphi)t)\,dt,$$

valamint az integrálok abszolút konvergensek és $0<\varphi<\pi/2$.

2. Tekintsük a $L'(z)\overline{L(z)}$ kifejezést, ahol $\operatorname{Im} z>0$, rögzített komplex szám. A jobb olvashatóság kedvéért legyen $\vartheta \overset{\circ}{=} \exp(-i\varphi)$, illetve

$$h(x)\overset{\circ}{=}\exp(-z\vartheta x)\,f(\vartheta x).$$

Ekkor

$$\begin{aligned}\overline{L(z)} &= \overline{\vartheta}\int_0^\infty \overline{h(t)}dt=\frac{1}{\vartheta}\int_0^\infty \overline{h(t)}dt,\\ \frac{d}{dz}L(z) &= -\vartheta^2\int_0^\infty sh(s)\,ds.\end{aligned}$$

Ebből[24]

$$\begin{aligned}-L'(z)\overline{L(z)}=\vartheta\int_0^\infty \overline{h(t)}dt\int_0^\infty sh(s)\,ds=\\ =\vartheta\int_0^\infty\int_0^\infty sh(s)\overline{h(t)}dtds.\end{aligned}$$

Az $L(z)$ függvényt megadó integrál abszolút konvergens, ezért az $L'(z)$ és az $\overline{L(z)}$ értékeket megadó integrálok is abszolút konvergensek, így a kettős integrál is abszolút konvergens, következésképpen használható az integráltranszformációs tétel. Végezzünk $s=uv$, $t=u/v$ transzformációt[25]:

$$\begin{aligned}-L'(z)\cdot\overline{L(z)}=\vartheta\int_0^\infty\int_0^\infty uvh(uv)\overline{h\left(\frac{u}{v}\right)}\frac{2u}{v}dtds=\\ =2\vartheta\int_0^\infty\int_0^\infty u^2h(uv)\overline{h\left(\frac{u}{v}\right)}dvdu.\end{aligned}$$

Vegyük észre, hogy a Jacobi-determináns

$$\left|\det\begin{pmatrix} v & u\\ 1/v & -u/v^2\end{pmatrix}\right|=\frac{2u}{v},$$

[23] V.ö.: C.1.15. Állítás, 291. oldal.

[24] Vegyük észre az előjelet!

[25] Valójában az integrandus komplex értékű, ezért a valós és az imaginárius részben kell az integráltranszformációt elvégezni.

illetve, hogy a transzformáció a nem negatív számpárok halmazát önmagára képező bijekció. A h képletében szereplő két tagot kiírva, illetve f-re kihasználva a tükrözési elvet

$$\begin{aligned}\overline{h\left(\frac{u}{v}\right)} &= \overline{\exp\left(-z\vartheta\frac{u}{v}\right)f\left(\vartheta\frac{u}{v}\right)} = \exp\left(-\overline{z\vartheta\frac{u}{v}}\right)f\left(\overline{\vartheta\frac{u}{v}}\right) = \\ &= \exp\left(-\bar{z}\frac{u}{v\vartheta}\right)f\left(\frac{u}{v\vartheta}\right).\end{aligned}$$

Mivel $0<\varphi<\pi/2$, ezért a $\operatorname{Re}\vartheta$ és a $\operatorname{Re}1/\vartheta$ pozitív. Az f hiperbolikusan teljesen monoton, és a feltételek szerint kiterjeszthető a $\operatorname{Re}z>0$ félsíkra következésképpen, miként korábban láttuk

$$\begin{gathered}f\left(\frac{u}{\vartheta v}\right)\cdot f\left(u\cdot\vartheta v\right) = \\ = \int_0^\infty \exp\left(-t\left(\frac{u}{\vartheta v}+\vartheta uv\right)\right)dK_u\left(t\right)\end{gathered}$$

módon írható. Ezt beírva az alábbi hármas integrált kapjuk:

$$2\vartheta\int_0^\infty u^2\int_0^\infty\int_0^\infty \exp\left(-\overline{(z+t)}\cdot\frac{u}{v\vartheta}-(z+t)\cdot u\vartheta v\right)dK_u\left(t\right)dvdu.$$

Ismét az áttekinthető jelölés céljából legyen

$$w \stackrel{\circ}{=} w\left(t,u\right) \stackrel{\circ}{=} \frac{u}{\vartheta}\overline{(z+t)}.$$

Az integrál

$$2\vartheta\int_0^\infty u^2\int_0^\infty\int_0^\infty \exp\left(-\left(\frac{w}{v}+\overline{w}\cdot v\right)\right)dK_u\left(t\right)dvdu.$$

Megmutatjuk, hogy a belső két integrál felcserélhető. Fubini-tétele értelmében ehhez igazolni kell, hogy a hármas integrál abszolút konvergens. Ezt lényegében az előző számítások megfordításával igazolhatjuk:

$$\left|\exp\left(-\left(\frac{w}{v}+\overline{w}v\right)\right)\right| = \exp\left(-\operatorname{Re}\left(\frac{w}{v}+\overline{w}v\right)\right).$$

Mivel $\operatorname{Re}z\vartheta>0$ és $\operatorname{Re}\vartheta>0$ ezért a kitevőben szereplő valós rész alkalmas $a,b>0$ konstansokkal

$$(at+b)\left(\frac{u}{v}+uv\right)$$

alakba írható. Ebből következően a belső integrál

$$\begin{gathered}\int_0^\infty \exp\left(-(at+b)\left(\frac{u}{v}+uv\right)\right)dK_u\left(t\right) = \\ = \exp\left(-b\left(\frac{u}{v}+uv\right)\right)\int_0^\infty \exp\left(-t\left(\frac{au}{v}+auv\right)\right)dK_u\left(t\right) \\ = \exp\left(-b\left(\frac{u}{v}+uv\right)\right)f\left(\frac{au}{v}\right)\cdot f\left(au\cdot v\right).\end{gathered}$$

Ezt u és v szerint integrálva, majd kihasználva, hogy az integrandus nem negatív, így a helyettesítést „visszacsinálva”

$$\begin{aligned}
&\int_0^\infty 2u^2 \int_0^\infty \exp\left(-b\left(\frac{u}{v}+uv\right)\right)\times \\
&\quad\times f\left(\frac{au}{v}\right)\cdot f(au\cdot v)\,dudv = \\
&= \int_0^\infty \int_0^\infty x\exp(-b(x+y))\,f(ax)\cdot f(ay)\,dxdy = \\
&= \int_0^\infty x\exp(-bx)\,f(ax)\,dx \int_0^\infty \exp(-by)\,f(ay)\,dy,
\end{aligned}$$

amely két integrál véges. A belső két integrált felcserélése után tekintsük az imaginárius rész előjelét meghatározó

$$\vartheta \int_0^\infty \exp\left(-\left(\frac{w}{v}+\overline{w}\cdot v\right)\right)dv$$

kifejezést. A valós számok mentén való v szerinti integrálás helyett térjünk át a $h \mapsto wh$ sugáron vett integrálra. Ezt, kihasználva, hogy $\operatorname{Re} w > 0$, és így a külső körön való integrál nullához tart a korábban már látott módon megtehetjük. Az integrál értéke tehát

$$\vartheta w \int_0^\infty \exp\left(-\left(\frac{1}{h}+|w|^2 h\right)\right)dh.$$

A képletben szereplő valós integrál mindig pozitív, következésképpen a hármas integrál imaginárius részének előjelét az

$$\operatorname{Im}\vartheta w = \operatorname{Im}\left(u\cdot\overline{(z+t)}\right) = -u\operatorname{Im} z$$

kifejezés adja meg. Vagyis ha $\operatorname{Im} z > 0$, akkor a hármas integrál imaginárius része nem lehet pozitív, így a számítás elején használt negatív előjel miatt

$$\operatorname{Im}\left(L'(z)\overline{L(z)}\right) \geq 0, \quad \text{ha } \operatorname{Im} z > 0.$$

Ezzel Thorin tételének bizonyítását befejeztük.

□

C.5. Általánosított szimmetrikus gamma-konvolúciók

Térjünk rá a Student-eloszlásra. A t_n sűrűségfüggvénye

$$f(x) = \frac{1}{B(1/2, n/2)}\frac{1}{\left(1+x^2\right)^{(n+1)/2}}.$$

Valamivel általánosabban: ha $r > 0$, akkor t_r eloszláson szokás az

$$f(x) = \frac{1}{B(1/2, r/2)} \frac{1}{\left(1+x^2\right)^{(r+1)/2}}$$

sűrűségfüggvénnyel adott eloszlást érteni. A t_r eloszlás a teljes számegyenesen van értelmezve, így a karakterisztikus függvényét kell vizsgálni. Ha ξ és η független azonos paraméterű gamma eloszlású változók, akkor a $\xi - \eta$ eloszlását szimmetrikus gamma-konvolúciónak mondjuk. Az (a, μ) paraméterű gamma eloszlás Laplace-transzformáltjában $s = -it$ helyettesítést végezve, a változó karakterisztikus függvénye

$$\varphi(t) = \left(\frac{\mu}{\mu - it}\right)^a.$$

Ebből a szimmetrikus gamma-konvolúciók karakterisztikus függvénye

$$\left(\frac{\mu}{\mu - it}\right)^a \left(\frac{\mu}{\mu + it}\right)^a = \left(\frac{\mu^2}{\mu^2 + t^2}\right)^a.$$

Az egyszerűbb jelölés céljából a μ^2 helyett csak λ-t fogunk írni. Mivel független korlátlanul osztható eloszlások összege is korlátlanul osztható, ezért a szimmetrikus gamma-konvolúciók is korlátlanul oszthatóak. Miként az általánosított gamma-konvolúciók esetén is a

$$\begin{aligned} \prod_{i=1}^{N} \left(\frac{\lambda_i}{\lambda_i + t^2}\right)^{a_i} &= \exp\left(\sum_{i=1}^{N} a_i \log \frac{\lambda_i}{\lambda_i + t^2}\right) = \\ &= \exp\left(\int_{0+}^{\infty} \log \frac{\lambda}{\lambda + t^2} dA_N(\lambda)\right) \end{aligned}$$

alakú karakterisztikus függvénnyel rendelkező változók korlátlanul oszthatóak. Ebből, felhasználva, hogy korlátlanul osztható eloszlások gyenge konvergenciában vett határértéke szintén korlátlanul osztható minden

$$\varphi(t) = \exp\left(\int_{0+}^{\infty} \log \frac{\lambda}{\lambda + t^2} dA(\lambda)\right)$$

alakú karakterisztikus függvénnyel rendelkező eloszlás is korlátlanul osztható. Ez indokolja a következő definíciót:

C.5.1. Definíció. Egy eloszlást általánosított szimmetrikus gamma-konvolúciónak mondunk, ha a karakterisztikus függvénye felírható $\varphi(t) = L\left(t^2\right)$ módon, ahol L egy általánosított gamma-konvolúció Laplace-transzformáltja.

C.5.2. Állítás. *Az általánosított szimmetrikus gamma-konvolúciók korlátlanul oszthatók.*

Bizonyítás: Az állítás az elmondottak alapján evidens, de érdemes más oldalról is megvilágítani. Az egyszerűség kedvéért a $\varphi(t) = L\left(t^2/2\right)$ alakú kifejezéseket vizsgáljuk. Ha az $L(s) = \mathbf{E}(\exp(-s\xi))$ egy Laplace-transzformáció, akkor az $L(s/2)$ éppen a $\xi/2$ transzformáltja, így az általánosságot nem csorbítjuk. Legyen F az L-hez tartozó eloszlásfüggvény. Ezzel

$$\begin{aligned}
\varphi(t) &= L\left(\frac{t^2}{2}\right) = \int_0^\infty \exp\left(-u\frac{t^2}{2}\right) dF(u) = \int_0^\infty \exp\left(-\left(\sqrt{u}\right)^2 \frac{t^2}{2}\right) dF(u) = \\
&= \int_0^\infty \varphi_{N(0,\sqrt{u})}(t)\, dF(u) = \\
&= \int_0^\infty \frac{1}{\sqrt{u}\sqrt{2\pi}} \int_{-\infty}^\infty \exp(itx)\exp\left(-\frac{x^2}{2u}\right) dx dF(u) = \\
&= \int_{-\infty}^\infty \exp(itx) \int_0^\infty \frac{1}{\sqrt{2\pi u}} \exp\left(-\frac{x^2}{2u}\right) dF(u)\, dx = \\
&\stackrel{\circ}{=} \int_{-\infty}^\infty \exp(itx)\, g(x)\, dx,
\end{aligned}$$

ahol a Fubini-tétel ismételt alkalmazásával egyszerűen látható, hogy a belső

$$g(x) \stackrel{\circ}{=} \int_0^\infty \frac{1}{\sqrt{2\pi u}} \exp\left(-\frac{x^2}{2u}\right) dF(u)$$

integrál egy sűrűségfüggvény. Vagyis a φ valóban egy karakterisztikus függvény. Továbbá, ha az L egy korlátlanul osztható eloszlás Laplace-transzformáltja, akkor az $L\left(t^2/2\right)$, illetve az $L\left(t^2\right)$ egy korlátlanul osztható eloszlás karakterisztikus függvénye, ugyanis ilyenkor a $\varphi_n(t) = \sqrt[n]{L}\left(t^2/2\right)$, illetve az $\sqrt[n]{L}\left(t^2\right)$ szintén karakterisztikus függvény. A bizonyításban kulcsszerepet játszó $g(x)$ kifejezés egy kevert normális eloszlás sűrűségfüggvénye, ahol a keverés paramétere a variancia.

□

A számolást fordított irányban elvégezve beláthatjuk a következőt:

C.5.3. Következmény. *Egy eloszlás φ karakterisztikus függvénye pontosan akkor $\varphi(t) = L\left(t^2/2\right)$ alakú, ha az eloszlás egy alkalmas F eloszlás szerint kevert normális eloszlás. Ilyenkor a keverendő normális eloszlások várható értéke nulla és a keverés az $u = \sigma^2$ variancia szerint történik.*

C.5.4. Példa. Szimmetrikus stabil eloszlások.

Ha $0 < \alpha \leq 1$, és $c > 0$, akkor az $s \geq 0$ félegyenesen az $L(s) = \exp(-c \cdot s^\alpha)$ egy teljesen monoton függvény, tehát egy nem negatív eloszlás Laplace-transzformáltja. Az elmondottak alapján a

$$\varphi(t) = \exp\left(-c\left(t^2\right)^\alpha\right) \stackrel{\circ}{=} \exp\left(-c\,|t|^\gamma\right)$$

ahol $0 < \gamma = 2\alpha \leq 2$ egy eloszlás karakterisztikus függvénye. Könnyen látható, hogy az eloszlás korlátlanul osztható és az osztáskor kapott komponensek azonos típusba tartoznak, ugyanis ha

$$\varphi^{(n)}(t) \doteq \exp\left(-\frac{c}{n}|t|^{\gamma}\right),$$

akkor $\left(\varphi^{(n)}(t)\right)^n = \varphi(t)$. A φ-hez tartozó eloszlást szimmetrikus stabil eloszlásnak szokás nevezni. Érdemes hangsúlyozni, hogy a $\gamma = 1$ és $c = 1$ eset éppen a Cauchy-eloszlást adja, és a $\gamma = 2$ és $c = 1/2$ pedig a standard normális eloszlást. A többi γ esetén a sűrűségfüggvények nem fejezhetők ki egyszerű képletekkel.

□

Ha most a keverő F egy általánosított gamma-konvolúció eloszlásfüggvénye, akkor a g egy általánosított szimmetrikus gamma-konvolúció sűrűségfüggvénye. Legyen $r > 0$ és az F sűrűségfüggvénye legyen

$$f(x) \doteq \frac{1}{\Gamma(r)}\left(\frac{1}{x}\right)^{r+1}\exp\left(-\frac{1}{x}\right), \quad x > 0.$$

Az f egy sűrűségfüggvény[26] ugyanis nem negatív és

$$\begin{aligned}&\frac{1}{\Gamma(r)}\int_0^{\infty}\left(\frac{1}{x}\right)^{r+1}\exp\left(-\frac{1}{x}\right)dx = \\ &= \frac{1}{\Gamma(r)}\int_0^{\infty}u^{r+1}\exp(-u)\frac{1}{u^2}dt = 1.\end{aligned}$$

Tekintsük az $u = \sigma^2$ paraméter szerinti keverést:

$$\begin{aligned}g(x) &= \frac{1}{\sqrt{2\pi}\Gamma(r)}\int_0^{\infty}\frac{1}{\sqrt{u}}\left(\frac{1}{u}\right)^{r+1}\exp\left(-\frac{1}{u}\right)\exp\left(-\frac{x^2}{2u}\right)du = \\ &= \frac{1}{\sqrt{2}\Gamma(1/2)\Gamma(r)}\int_0^{\infty}t^{1/2}t^{r+1}\exp(-t)\exp\left(-t\frac{x^2}{2}\right)\frac{dt}{t^2} = \\ &= \frac{1}{\sqrt{2}\Gamma(1/2)\Gamma(r)}\int_0^{\infty}t^{r-1/2}\exp\left(-t\left(1+\frac{x^2}{2}\right)\right)dt = \\ &= \frac{\Gamma(r+1/2)}{\sqrt{2}\Gamma(1/2)\Gamma(r)}\left(\frac{1}{1+x^2/2}\right)^{r+1/2} = \\ &= \frac{1}{\sqrt{2}B(1/2,r)}\left(\frac{1}{1+\left(x/\sqrt{2}\right)^2}\right)^{r+1/2},\end{aligned}$$

[26] Könnyen látható, hogy az f éppen egy $(r,1)$ paraméterű gamma eloszlású változó reciprokának sűrűségfüggvénye. Ennek megfelelően a t_r lényegében gamma eloszlás reciproka szerint kevert normális eloszlású változó. A gamma eloszlás reciprokát szokás inverz gamma eloszlásnak is mondani.

amely a $\sqrt{2}t_r$ alakú változó sűrűségfüggvénye. Az f keverő sűrűségfüggvény azonban hiperbolikusan teljesen monoton, ugyanis

$$f(uv)f(u/v) =$$
$$\left(\frac{1}{\Gamma(r)}\right)^2\left(\frac{1}{uv}\right)^{r+1}\left(\frac{1}{u/v}\right)^{r+1}\exp\left(-\frac{1}{uv}\right)\exp\left(-\frac{1}{u/v}\right) =$$
$$= C(u)\exp\left(-\frac{1}{u}\left(v+\frac{1}{v}\right)\right) = C(u)\exp\left(-\frac{1}{u}w\right).$$

A t_r eloszlás tehát egy általánosított szimmetrikus gamma-konvolúció, így korlátlanul osztható.

□

C.6. Felhasznált irodalom

A könyvben leírtak a jelen fejezettől eltekintve közismertek, így a felhasznált irodalom bemutatásától eltekintek. Természetesen nagyban támaszkodtam az általam ismert irodalomra és igen gyakran szó szerint átvettem bizonyos példákat és feladatokat. A legtöbbet az egyetemi oktatásban korábban használt, Bárfai Pál által írt feladatgyűjteményből vettem át, de az évek során számtalan más forrásból is merítettem. Mivel manapság a plágium komoly vádnak számít, ezért nagy megtiszteltetésnek tekintem annak kijelentését, hogy minden ami a jegyzetben le van írva valaki mástól származik, és bár a forrásra általában nem emlékszem, valakitől tanultam és nem én magam találtam ki őket. Mint tudjuk minden kutató olyan törpe, aki óriásokon áll. Természetesen az irodalomjegyzék elsődleges célja nem az elsőség tisztázása, hanem az olvasó további orientálása. Amennyiben az olvasó a valószínűségszámítás területén további tanulmányokat szeretne folytatni, úgy a honlapomon, *medvegyev.hu*, tud tájékozódni. A jelen fejezetben leírtak azonban nem részei a standard valószínűségszámítási tananyagnak, így célszerű az idevágó irodalom rövid bemutatása. A fejezet megírásakor két hosszabb monográfiára támaszkodtam:

1. Lennart Bondesson: *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*, Lecture Notes in Statistics 76, Springer-Verlag, New York, 1992.

2. Fred W. Steutel és Klass van Harn: *Infinite Divisibility of Probability Distributions on the Real Line*, Pure and Applied Mathematics, A Program of Monographs, Textbooks, and Lecture Notes 259, Marcel Dekker, Inc., New York, Basel, 2004.

A terület legfontosabb eredményei G. Olof Thorin svéd matematikustól származnak. Idevágó legfontosabb dolgozatai:

Thorin O.: *On the infinite divisibility of the lognormal distribution*. Scand. Actuar. J., 1977 , 121–148. oldalak.

Thorin O.: *On the infinite divisibility of the Pareto distribution.* Scand. Actuar. J., 1977, 31-40. oldalak.

Thorin O.: *An extension of the notion of a generalized G-convolution.* Scand. Actuar. J. , 1978, 141–149. oldalak.

TÁRGYMUTATÓ

#05

Bevezetés a valószínűségszámításba

Miért olvassam el?

A könyv a valószínűségszámítás alapfogalmait foglalja össze. A könyv első fele a legfontosabb alapismeretek tartalmazza, a második rész, amely a könyvben függelékként jelenik meg a korlátlanul osztható eloszlások és az általánosított gamma konvolúciók elméletét tárgyalja.

Az első rész megértéséhez elegendő az egyetemi bevezető analízis ismerete, a második rész megértéséhez azonban a matematikai analízis haladottabb fejezeteit is ismerni kell.