

A sztochasztika alapjai

Kevei Péter

2021. február 9.

Tartalomjegyzék

1. Bevezetés	4
1.1. Alapfogalmak	4
1.2. A valószínűségi mérték	5
1.3. Klasszikus valószínűségi mező	8
2. Néhány klasszikus probléma	8
2.1. de Mére paradoxona	8
2.2. A párosítási probléma	9
2.3. Az igazságos osztozkodás problémája	10
3. Geometriai valószínűségi mező	11
3.1. Bertrand paradoxon (1888)	12
3.2. Buffon-féle tűprobléma (1777)	12
4. Feltételes valószínűség	14
5. Függetlenség	18
5.1. Craps játék	20
6. Véletlen változók	21
6.1. Diszkrét véletlen változók	25
6.2. Folytonos véletlen változók	25
6.3. Véletlen vektorváltozók	26
6.4. Véletlen változók függetlensége	27
6.5. Függetlenség és geometriai valószínűség	29
7. Várható érték	29
7.1. Várható érték tulajdonságai	30
7.2. Szórás, kovariancia, korreláció	33
8. Nevezetes eloszlások	37
8.1. Bernoulli-eloszlás	37
8.2. Binomiális eloszlás	37
8.3. Poisson-eloszlás	38
8.4. Geometriai eloszlás	39
8.5. Egyenletes eloszlás	40
8.6. Exponenciális eloszlás	41
8.7. Normális eloszlás	42

9. Véletlen változók konvergenciája	44
9.1. Markov és Csebisev egyenlőtlenségei	44
9.2. Nagy számok gyenge törvénye	46
9.3. Centrális határeloszlás-tétel	47
10. Statisztikai alapfogalmak	48
10.1. Alapstatisztikák	49
11. Pontbecslések	51
11.1. Német tank probléma	51
11.2. Torzítatlanság	51
11.3. Hatásosság	52
11.4. Konzisztencia	53
12. Becslési módszerek	56
12.1. Maximum likelihood módszer	56
12.2. Momentumok módszere	60
13. Konfidenciintervallumok	62
13.1. Konfidenciintervallum normális eloszlás várható értékére ismeret szórás esetén	62
13.2. Konfidenciintervallum normális eloszlás várható értékére ismeretlen szórás esetén	63
13.3. Konfidenciintervallum normális eloszlások várható értékének különbségére	65
14. Hipotézisvizsgálat	67
14.1. u-próba	67
14.2. t-próba	69
14.3. Kétmintás u- és t-próba	70

1. Bevezetés

1.1. Alapfogalmak

Véletlen (valószínűségi) kísérlet: lényegében azonos körülmények között tetszőlegesen sokszor megismételhető megfigyelés, melynek többféle kimenetele lehet, és a figyelembe vett körülmények nem határozzák meg egyértelműen a kimenetelt.

A véletlen kísérlet lehetséges kimeneteleinek halmaza az **eseménytér**, jele Ω .

Az **esemény** olyan a kísérlettel kapcsolatban tett állítás, melynek igaz vagy hamis volta eldönthető a kísérlet lefolytatása után. Az **események halmaza** az Ω részhalmazainak egy olyan rendszere, mely σ -algebra. Az (Ω, \mathcal{A}) párt mérhetőségi térnek nevezzük.

Egy $\mathcal{A} \subset 2^\Omega$ halmazrendszert akkor nevezünk **σ -algebrának**, ha

- $\emptyset \in \mathcal{A}$;
- valahányszor $A \in \mathcal{A}$, mindannyiszor $A^c = \Omega \setminus A \in \mathcal{A}$ (azaz a halmazrendszer zárt a komplementerképzésre);
- valahányszor $A_1, A_2, \dots \in \mathcal{A}$, mindannyiszor $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$ (azaz a halmazrendszer zárt a megszámlálható unióképzésre).

Megjegyzés. • Vegyük észre, hogy a $\{\emptyset, \Omega\}$ halmazrendszer σ -algebra. Ez a triviális σ -algebra.

- A 2^Ω halmazrendszer, az Ω hatványhalmaza, azaz az összes részhalmazának halmaza is σ -algebra. Abban az esetben, amikor az Ω alaphalmaz véges, akkor az események halmaza mindig a hatványhalmaz.

Események jelölése: A, B, A_1, \dots

- $|A| = 1 \Leftrightarrow A = \{\omega\}$, $\omega \in \Omega$, elemi esemény
- \emptyset a lehetetlen esemény
- Ω a biztos esemény
- A^c az ellentett esemény
- $A \cap B$ mindkét esemény bekövetkezik (A és B)
- $A \cup B$ a két esemény közül legalább az egyik bekövetkezik
- $A \cap B = \emptyset$ a két esemény kizárja egymást
- $A - B$ a bekövetkezik de B nem
- $A \subset B$ az A esemény maga után vonja B -t

1.1. *Példa.* Háromszor földobunk egy pénzérmét. Ekkor az eseménytér

$$\Omega = \{(F, F, F), (F, F, I), (F, I, F), (F, I, I), \\ (I, F, F), (I, F, I), (I, I, F), (I, I, I)\},$$

azaz $|\Omega| = 2^3 = 8$ darab elemi esemény van, és $|2^\Omega| = 2^8 = 256$ az összes esemény száma.

Legyen $A_i = \{\text{az } i\text{-edik dobás fej}\}$, $i = 1, 2, 3$. Ekkor

$$A_1 = \{(F, F, F), (F, F, I), (F, I, F), (F, I, I)\}.$$

$$B = \{\text{csak az 1. fej}\} = \{(F, I, I)\} = A_1 \cap A_2^c \cap A_3^c$$

$$C = \{\text{egyik sem fej}\} = \{(I, I, I)\} = A_1^c \cap A_2^c \cap A_3^c$$

1.2. *Példa.* Véletlen sorrendben leírjuk a MATEMATIKA szó betűit.

1. *megoldás:* Az azonos betűket nem különböztetjük meg.

$$\Omega = \{\text{AAAEIKMMTT}, \text{AAAEIKMTMT}, \dots, \text{TTMMKIEAAA}\}$$

$$|\Omega| = \frac{10!}{3!2!2!}$$

$A = \{\text{MATEMATIKA szót kapjuk}\} = \{\text{MATEMATIKA}\}$, azaz A elemi esemény.

2. *megoldás:* Az azonos betűket megkülönböztetjük.

$$\Omega = \{A_1A_2A_3EIKM_1M_2T_1T_2, A_1A_2A_3EIKM_1M_2T_2T_1, \dots, \\ T_2T_1M_2M_1KIEA_3A_2A_1\},$$

$|\Omega| = 10!$, és ha $A = \{\text{MATEMATIKA szót kapjuk}\}$, $|A| = 3!2!2!$.

1.2. A valószínűségi mérték

1.3. **Definíció.** Egy $\mathbf{P} : \mathcal{A} \rightarrow [0, 1]$ halmazfüggvény *valószínűségi mérték* az (Ω, \mathcal{A}) mérhetőségi téren, ha

- $\mathbf{P}(\Omega) = 1$;
- ha az $A_1, A_2, \dots \in \mathcal{A}$ halmazok (páronként) diszjunktak, akkor

$$\mathbf{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbf{P}(A_i),$$

azaz a halmazfüggvény σ -additív.

A fenti tulajdonságokkal rendelkező $(\Omega, \mathcal{A}, \mathbf{P})$ hármast **valószínűségi mezőnek** nevezzük.

1.4. Állítás (A valószínűség tulajdonságai). Legyen $(\Omega, \mathcal{A}, \mathbf{P})$ egy valószínűségi mező, $A, B, A_1, A_2, \dots \in \mathcal{A}$ események.

(i) Ha $A_i \cap A_j = \emptyset$, minden $i \neq j$ párra, akkor

$$\mathbf{P}(A_1 \cup \dots \cup A_n) = \mathbf{P}(A_1) + \dots + \mathbf{P}(A_n).$$

(ii) $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$.

(iii) $A \subset B \Rightarrow \mathbf{P}(B - A) = \mathbf{P}(B) - \mathbf{P}(A)$, és $\mathbf{P}(A) \leq \mathbf{P}(B)$.

(iv) $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.

(v) Szita formula:

$$\mathbf{P}(A_1 \cup \dots \cup A_n) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbf{P}(A_{i_1} \cap \dots \cap A_{i_k}).$$

(vi) $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$.

(vii) $\mathbf{P}(A_1 \cup \dots \cup A_n) \leq \mathbf{P}(A_1) + \dots + \mathbf{P}(A_n)$.

(viii) Ha A_n monoton növekvő halmazzorozat, azaz $A_1 \subset A_2 \subset \dots$, akkor $\lim_{n \rightarrow \infty} \mathbf{P}(A_n) = \mathbf{P}(\cup_{i=1}^{\infty} A_i)$.

(ix) Ha A_n monoton csökkenő halmazzorozat, azaz $A_1 \supset A_2 \supset \dots$, akkor $\lim_{n \rightarrow \infty} \mathbf{P}(A_n) = \mathbf{P}(\cap_{k=1}^{\infty} A_k)$.

(x) $\mathbf{P}(\cup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} \mathbf{P}(A_k)$ (megszámlálható szubadditivitás).

Bizonyítás. (i) Legyen $A_{n+1} = A_{n+2} = \dots = \emptyset$.

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

(iii) $B = A \cup (B - A)$,

$$\mathbf{P}(B) = \mathbf{P}(A) + \mathbf{P}(B - A) \Rightarrow \mathbf{P}(B - A) = \mathbf{P}(B) - \mathbf{P}(A) \geq 0.$$

(iv) $\mathbf{P}(A \cup B) = \mathbf{P}(A \cup (B - (A \cap B))) = \mathbf{P}(A) + \mathbf{P}(B - (A \cap B)) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.

(v) Teljes indukcióval. $n = 1, 2$ -re igaz. Tegyük fel, hogy n -ig igaz. A $B_i = A_i$, $i \leq n - 1$, $B_n =$

$A_n \cup A_{n+1}$ jelöléssel

$$\begin{aligned}
\mathbf{P}(A_1 \cup \dots \cup A_n \cup A_{n+1}) &= \mathbf{P}(B_1 \cup \dots \cup B_n) \\
&= \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbf{P}(B_{i_1} \cap \dots \cap B_{i_k}) \\
&= \sum_{k=1}^n (-1)^{k+1} \left(\sum_{i_k < n} \mathbf{P}(A_{i_1} \cap \dots \cap A_{i_k}) \right. \\
&\quad \left. + \sum_{i_{k-1} \leq n-1} \mathbf{P}(A_{i_1} \cap \dots \cap A_{i_{k-1}} \cap (A_n \cup A_{n+1})) \right) \\
&= \sum_{k=1}^n (-1)^{k+1} \left(\sum_{i_k < n} \mathbf{P}(A_{i_1} \cap \dots \cap A_{i_k}) \right. \\
&\quad \left. + \sum_{i_{k-1} \leq n-1} \left[\mathbf{P}(A_{i_1} \cap \dots \cap A_{i_{k-1}} \cap A_n) + \mathbf{P}(A_{i_1} \cap \dots \cap A_{i_{k-1}} \cap A_{n+1}) \right. \right. \\
&\quad \left. \left. - \mathbf{P}(A_{i_1} \cap \dots \cap A_{i_{k-1}} \cap A_n \cap A_{n+1}) \right] \right) \\
&= \sum_{k=1}^{n+1} (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbf{P}(A_{i_1} \cap \dots \cap A_{i_k})
\end{aligned}$$

- (vi) $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) \leq \mathbf{P}(A) + \mathbf{P}(B)$.
(vii) Teljes indukcióval.
(viii) Vezessük be a $B_1 = A_1$, $B_n = A_n \setminus A_{n-1}$, $n \geq 2$ jelölést. Ekkor a B_n halmazok diszjunktak, $\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n$, és $\cup_{k=1}^n B_k = \cup_{k=1}^n A_k = A_n$, $n \geq 1$. Így

$$\begin{aligned}
\mathbf{P}(\cup_{k=1}^{\infty} A_k) &= \mathbf{P}(\cup_{k=1}^{\infty} B_k) = \sum_{k=1}^{\infty} \mathbf{P}(B_k) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mathbf{P}(B_k) \\
&= \lim_{n \rightarrow \infty} \mathbf{P}(\cup_{k=1}^n B_k) = \lim_{n \rightarrow \infty} \mathbf{P}(A_n),
\end{aligned}$$

amint állítottuk.

- (ix) Mivel A_n monoton csökkenő, A_n^c monoton növekvő halmzsorozat, ezért használhatjuk az előző pont állítását. Ezért

$$\begin{aligned}
\mathbf{P}(\cap_{k=1}^{\infty} A_k) &= 1 - \mathbf{P}((\cap_{k=1}^{\infty} A_k)^c) = 1 - \mathbf{P}(\cup_{k=1}^{\infty} A_k^c) \\
&= 1 - \lim_{n \rightarrow \infty} \mathbf{P}(A_n^c) = 1 - \lim_{n \rightarrow \infty} (1 - \mathbf{P}(A_n)) = \lim_{n \rightarrow \infty} \mathbf{P}(A_n),
\end{aligned}$$

amivel az állítást igazoltuk.

- (x) Tetszőleges n természetes számra a véges szubadditivitás alapján

$$\mathbf{P}(\cup_{k=1}^n A_k) \leq \sum_{k=1}^n \mathbf{P}(A_k) \leq \sum_{k=1}^{\infty} \mathbf{P}(A_k).$$

Mivel a $B_n = \sum_{k=1}^n A_k$ halmazosorozat monoton növekvő, így

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\sum_{k=1}^n A_k \right) = \mathbf{P} \left(\cup_{k=1}^{\infty} A_k \right),$$

ezért az egyenlőtlenségből határátmenettel kapjuk az állítást. □

Az $(\Omega, \mathcal{A}, \mathbf{P})$ hármast **valószínűségi mezőnek** nevezzük.

1.3. Klasszikus valószínűségi mező

Az $(\Omega, 2^\Omega, \mathbf{P})$ valószínűségi mező **klasszikus**, ha minden kimenetel egyformán valószínű, azaz $\mathbf{P}(\{\omega\}) = c$ minden $\omega \in \Omega$ esetén. Ekkor persze szükségképpen $c = 1/|\Omega|$. Tetszőleges A eseményre $\mathbf{P}(A) = \frac{|A|}{|\Omega|} = \frac{\text{kedvező}}{\text{összes}}$.

1.5. *Példa. Születésnap probléma.* Mekkora a valószínűsége annak, hogy n ember között van két olyan, akiknek ugyanazon a napon van a születésnapjuk?

$$\begin{aligned} f(n) &= \mathbf{P}(n \text{ ember között van } 2, \text{ akiknek ugyanazon} \\ &\quad \text{a napon van a születésnapja}) \\ &= 1 - \mathbf{P}(\text{mindenkinek különböző napon van a születésnapja}) \\ &= 1 - \frac{365 \cdot 364 \cdot \dots \cdot (365 - n + 1)}{365^n}. \end{aligned}$$

$$f(22) \approx 0,4757 < 1/2 < 0,5073 \approx f(23).$$

2. Néhány klasszikus probléma

2.1. de Méré paradoxona

1654: Pascal és Fermat levelezése de Méré lovag feladatairól, majd a „véletlen matematikájának” megalapozásáról.

de Méré lovag paradoxona: Miért nem ugyanakkora valószínűségű a következő két esemény:

- 1 kockával 4-szer dobva legalább egy hatost dobunk;
- 2 kockával 24-szer dobva legalább egy dupla hatost dobunk.

Legyen A az az esemény, hogy 1 kockával 4-szer dobva legalább egyszer dobunk 6-ost. Ekkor $\Omega = \{(1, 1, 1, 1), \dots, (6, 6, 6, 6)\}$, azaz $|\Omega| = 6^4$. Mivel minden kimenetel egyformán valószínű, a valószínűségi mező klasszikus. A^c az az esemény, hogy nem dobunk 6-ost, így $|A^c| = 5^4$. Ezért

$$\mathbf{P}(A) = 1 - \mathbf{P}(A^c) = 1 - \left(\frac{5}{6}\right)^4 \approx 0,5177.$$

Vizsgáljuk most azt a kísérletet, hogy 2 kockával dobunk 24-szer, és legyen B az az esemény, hogy dobunk dupla 6-ost. Ekkor $|\Omega| = 36^{24}$, és $|B^c| = 35^{24}$, ezért

$$\mathbf{P}(B) = 1 - \mathbf{P}(B^c) = 1 - \left(\frac{35}{36}\right)^{24} \approx 0,4914.$$

A rossz(!) intuíció az, hogy ha 4-szer megyünk neki egy $1/6$ valószínűségű eseménynek, akkor a siker valószínűsége ugyanannyi, mint ha 24-szer megyünk neki egy $1/36$ -od valószínűségűnek, hiszen $4/6 = 24/36$.

2.2. A párosítási probléma

Veszünk n darab kártyát 1-től n -ig megszámozva. Összekeverjük, és véletlen sorrendben lerakjuk őket egy sorba. A k -adik helyen párosítás történik, ha a k -adik helyre a k sorszámú kártya kerül. (Tehát véletlen permutációk fixpontjait tekintjük.)

Arra keressük a választ, hogy mennyi a valószínűsége, hogy nem történik párosítás. Jelölje p_n ezt a valószínűséget.

Jelölje A_k azt az eseményt, hogy a k -adik helyen párosítás történik, $k = 1, 2, \dots, n$. Ekkor az az esemény, hogy legalább egy párosítás történik éppen $A_1 \cup A_2 \cup \dots \cup A_n$. Ennek a valószínűségét a szita formulával határozhatjuk meg. Eszerint

$$\begin{aligned} \mathbf{P}(A_1 \cup \dots \cup A_n) &= \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbf{P}(A_{i_1} \cap \dots \cap A_{i_k}) \\ &= \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \frac{(n-k)!}{n!} \\ &= \sum_{k=1}^n (-1)^{k+1} \binom{n}{k} \frac{(n-k)!}{n!} \\ &= \sum_{k=1}^n \frac{(-1)^{k+1}}{k!}. \end{aligned}$$

Ezek szerint

$$\begin{aligned} p_n &= \mathbf{P}(\text{nincs párosítás}) = \mathbf{P}(A_1^c \cap \dots \cap A_n^c) \\ &= 1 - \mathbf{P}(A_1 \cup \dots \cup A_n) = 1 - \left(-\sum_{k=1}^n \frac{(-1)^k}{k!}\right) \\ &= \sum_{k=0}^n \frac{(-1)^k}{k!}. \end{aligned}$$

Analízisből tudjuk, hogy tetszőleges x valós számra

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots,$$

ahonnan látjuk, hogy

$$\lim_{n \rightarrow \infty} p_n = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = e^{-1} \approx 0,368.$$

Ezek után határozzuk meg azt a valószínűséget, hogy pontosan k darab párosítás történik. Vezessük be a

$$p_{n,k} = \mathbf{P}(n \text{ kártya van, és pontosan } k \text{ párosítás történik}), \quad k = 0, 1, \dots, n.$$

Nyilván $p_n = p_{n,0}$. Jelölje $N_{n,k}$ azon kimenetek számát, amikor pontosan k párosítás történik n kártyával. Ezekkel a jelölésekkel $p_m = N_{m,0}/m!$ minden m természetes szám esetén. Könnyen meggondolható, hogy

$$\begin{aligned} p_{n,k} &= \frac{N_{n,k}}{n!} = \frac{\binom{n}{k} N_{n-k,0}}{n!} = \frac{\binom{n}{k} (n-k)! p_{n-k}}{n!} \\ &= \frac{p_{n-k}}{k!} = \frac{1}{k!} \sum_{j=0}^{n-k} \frac{(-1)^j}{j!}. \end{aligned}$$

Az utóbbi alakból rögtön látjuk, hogy

$$\lim_{n \rightarrow \infty} p_{n,k} = \frac{1}{k!} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} = \frac{e^{-1}}{k!}.$$

Megjegyzés. Valójában azt bizonyítottuk be, hogy egy véletlen permutáció fix-pontjainak száma nagy n esetén közelítőleg Poisson-eloszlású, pontosabban a fix-pontok száma eloszlásban konvergál egy 1-paraméterű Poisson-eloszlású véletlen változóhoz. De erről majd később.

2.3. Az igazságos osztozkodás problémája

Két játékos, Anna és Balázs játszanak. Mindketten $1/2 - 1/2$ valószínűséggel nyernek egy-egy játékot. Az előre befizetett tétet az kapja, aki előbb nyer 10 játékot. A játék azonban 8-7 -es állásnál abbamaradt. Hogyan osszák el igazságosan a befizetett tétet?

Kicsit általánosabban a következő feladatot vizsgáljuk. Két játékos, Anna és Balázs játszanak. Annának a pont hiányzik a győzelemhez, Balásznak pedig b . Egy-egy játékot Anna $p \in (0, 1)$ valószínűséggel nyer meg, Balázs pedig $1 - p$ valószínűséggel. Hogyan osszák el a tétet?

Már 15. században írnak a problémáról, de Méré előtt (Luca Pacioli (1494), Tartaglia (~1550)). Speciális esetekben meg is tudják oldani a feladatot, azonban a teljes megoldást Pascal és Fermat adják meg. A következőkben Fermat megoldását mutatjuk be.

Vegyük észre, hogy a játéksorozat $a + b - 1$ játékkal biztos véget ér. Anna pontosan akkor nyer, ha az $a + b - 1$ játékból legalább a játékot nyer meg. Így Anna nyerésének valószínűségét $P_A(a, b, p)$ -vel jelölve azt kapjuk, hogy

$$\begin{aligned} P_A(a, b, p) &= \mathbf{P}(a + b - 1 \text{ játékból Anna legalább } a \text{ pontot szerez}) \\ &= \sum_{k=a}^{a+b-1} \mathbf{P}(a + b - 1 \text{ játékból Anna pontosan } k \text{ pontot szerez}) \\ &= \sum_{k=a}^{a+b-1} \binom{a+b-1}{k} p^k (1-p)^{a+b-1-k} \end{aligned}$$

Mivel pontosan az egyikük nyer $P_A + P_B = 1$, a binomiális tétel szerint pedig

$$\sum_{k=0}^{a+b-1} \binom{a+b-1}{k} p^k (1-p)^{a+b-1-k} = (p + (1-p))^{a+b-1} = 1,$$

ezért Balázs nyerésének valószínűsége

$$P_B(b, a, 1-p) = \sum_{k=0}^{a-1} \binom{a+b-1}{k} p^k (1-p)^{a+b-1-k}.$$

A $p = 1/2$ esetben ez a következőt adja:

$$\begin{aligned} P_A\left(a, b, \frac{1}{2}\right) &= \frac{1}{2^{a+b-1}} \sum_{k=a}^{a+b-1} \binom{a+b-1}{k}, \\ P_B\left(b, a, \frac{1}{2}\right) &= \frac{1}{2^{a+b-1}} \sum_{k=0}^{a-1} \binom{a+b-1}{k}. \end{aligned}$$

Ezek szerint a tétet

$$\frac{P_A\left(a, b, \frac{1}{2}\right)}{P_B\left(b, a, \frac{1}{2}\right)} = \frac{\sum_{k=a}^{a+b-1} \binom{a+b-1}{k}}{\sum_{k=0}^{a-1} \binom{a+b-1}{k}}$$

arányban kell elosztani. Ezt úgy tudjuk egyszerűen kiszámolni, hogy a Pascal-háromszög $(a + b - 1)$ -edik sorában összeadjuk az elemeket az a -adiktól az $(a + b - 1)$ -edikig, majd az eredményt elosztjuk a 0 -adiktól az $(a - 1)$ -edik elemig vett összeggel.

A kiinduló példánkban Annának 2 pont Balázsnak 3 pont hiányzott a győzelemhez. Tehát a tétet

$$\frac{P_A(2, 3, 1/2)}{P_B(3, 2, 1/2)} = \frac{6 + 4 + 1}{1 + 4} = \frac{11}{5}$$

arányban kell elosztani.

3. Geometriai valószínűségi mező

Akkor beszélünk geometriai valószínűségi mezőről, ha a kísérlettel kapcsolatos események egy geometriai alakzat részhalmazainak feleltethetők meg. Ekkor a lehetséges kimenetek halmaza $\Omega = H \subset \mathbb{R}^n$, aminek a mértéke (hossza, területe, térfogata) pozitív és véges. Ekkor egy $A \subset H$ esemény valószínűsége arányos a halmaz mértékével, azaz

$$\mathbf{P}(A) = \frac{\lambda(A)}{\lambda(H)},$$

ahol λ az n -dimenziós Lebesgue-mérték (hossz, terület, térfogat).

3.1. Bertrand paradoxon (1888)

Véletlenszerűen választunk egy húrt egy r sugarú körön. Mennyi a valószínűsége, hogy a húr hosszabb, mint a körbe írható szabályos háromszög oldala? Jelölje p ezt a valószínűséget.

1. *Megoldás.* A húr hosszát meghatározza a felezőpontjának (F) a kör középpontjától (O) vett távolsága. Ha $|OF| > r/2$, akkor a húr hosszabb, mint a szabályos háromszög oldala, különben rövidebb. Tehát a feladat megfeleltethető annak, hogy egy rögzített sugárról egyenletes eloszlás szerint választunk pontot. Ezért

$$p = \frac{r/2}{r} = \frac{1}{2}.$$

2. *Megoldás.* Ha az egyik végpontot rögzítjük, akkor a húr hosszát meghatározza, hogy hova esik a másik végpont. Legyen ϑ a húrnak és a rögzített ponthoz húzott érintőnek a szöge. A húr pontosan akkor hosszabb, mint a háromszög oldala, ha $\vartheta \in (\pi/3, 2\pi/3)$. Ennek a valószínűsége nyilván $1/3$, tehát

$$p = \frac{1}{3}.$$

3. *Megoldás.* A húr pontosan akkor lesz hosszabb a háromszög oldalánál, ha a középpontja belesik a háromszög beírt körébe. Ennek a valószínűsége $(r/2)^2\pi/(r^2\pi) = 1/4$, tehát

$$p = \frac{1}{4}.$$

A problémát természetesen az okozza, hogy a húr választását nem mondja meg a feladat. Azaz a véletlen nincs jól megadva.

3.2. Buffon-féle tűprobléma (1777)

Egy padló mintázata párhuzamos egyenesekből áll. A szomszédos egyenesek távolsága d . A padlóra ledobunk egy ℓ hosszú tűt, ahol $\ell \leq d$. Mekkora a valószínűsége, hogy a tű metszi valamelyik egyenest?

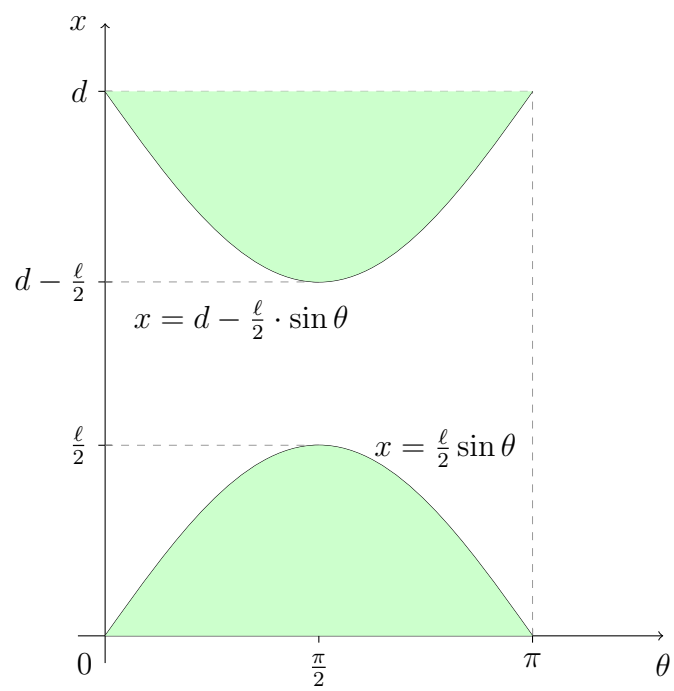
Jelölje x a tű középpontjának és a hozzá legközelebbi, tőle balra levő egyenesnek a távolságát! Legyen Θ a tű függőlegessel bezárt szöge. Vegyük észre, hogy a tű akkor metszi a baloldali egyenest, ha $0 \leq x \leq \frac{\ell}{2} \sin \Theta$, és akkor metszi a jobboldalit, ha $0 \leq d - x \leq \frac{\ell}{2} \sin \Theta$. Ha x egyenletes eloszlású $[0, d]$ -n és Θ egyenletes eloszlású $[0, \pi]$ -n, akkor a kísérlet megfeleltethető egy pont egyenletes eloszlás szerinti választásának a $[0, d] \times [0, \pi]$ téglalaphoz. A kedvező területrészt területe

$$2 \int_0^\pi \frac{\ell}{2} \sin \theta d\theta = 2\ell,$$

így a keresett valószínűség

$$\mathbf{P}(\text{van metszés}) = \frac{2\ell}{\pi d}.$$

Innen látjuk, hogy a π értékét meg lehet határozni empirikus módon.



1. ábra. Kedvező terület a Buffon-féle problémánál

4. Feltételes valószínűség

Két szabályos dobókockával dobunk. Az első kockával hatost dobtunk, a második kocka elgurult. Mennyi a valószínűsége, hogy dupla hatost dobtunk? Jelölje A azt az eseményt, hogy az első kockával hatost dobtunk, B pedig azt, hogy mindkét kockával ugyanazt dobtuk. Tudjuk, hogy az A esemény bekövetkezett, azaz a kísérlet kimeneteléről van egy részinformációnk. Ekkor az eseménytér azon részhalmazán dolgozunk, ahol az adott A esemény bekövetkezett, azaz A -n. Ekkor a kedvező esetek száma $|A \cap B|$ és az összes esetek száma $|A|$. Tehát a keresett valószínűség $\mathbf{P}(A \cap B)/\mathbf{P}(B)$. Éppen ez a feltételes valószínűség definíciója.

Legyen $(\Omega, \mathcal{A}, \mathbf{P})$ egy valószínűségi mező, és ezen A, B események, és tegyük föl, hogy $\mathbf{P}(B) > 0$. Ekkor az A esemény B eseményre vonatkozó feltételes valószínűsége

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

Ha annyi információnk van a véletlen kísérletről, hogy a B esemény bekövetkezett, akkor az A esemény valószínűsége $\mathbf{P}(A|B)$.

4.1. Állítás. Rögzítsünk egy tetszőleges B eseményt, melyre $\mathbf{P}(B) > 0$. Ekkor $\mathbf{P}_B(A) = \mathbf{P}(A|B)$ valószínűségi mérték \mathcal{A} -n.

Bizonyítás. Világos, hogy tetszőleges A eseményre $\mathbf{P}_B(A) \geq 0$, és mivel $\mathbf{P}(A \cap B) \leq \mathbf{P}(B)$, így $\mathbf{P}_B(A) \leq 1$. Továbbá

$$\mathbf{P}_B(\Omega) = \frac{\mathbf{P}(\Omega \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(B)}{\mathbf{P}(B)} = 1.$$

Már csak az additivitás ellenőrzése maradt. Legyenek $A_1, A_2, \dots \in \mathcal{A}$ diszjunktak. Ekkor a definíció szerint és a \mathbf{P} valószínűségi mérték additivitása alapján

$$\begin{aligned} \mathbf{P}_B(\cup_{i=1}^{\infty} A_i) &= \mathbf{P}(\cup_{i=1}^{\infty} A_i | B) = \frac{\mathbf{P}((\cup_{i=1}^{\infty} A_i) \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(\cup_{i=1}^{\infty} (A_i \cap B))}{\mathbf{P}(B)} \\ &= \frac{\sum_{i=1}^{\infty} \mathbf{P}(A_i \cap B)}{\mathbf{P}(B)} = \sum_{i=1}^{\infty} \frac{\mathbf{P}(A_i \cap B)}{\mathbf{P}(B)} = \sum_{i=1}^{\infty} \mathbf{P}(A_i | B) \\ &= \sum_{i=1}^{\infty} \mathbf{P}_B(A_i), \end{aligned}$$

ami éppen a bizonyítandó egyenlőség. □

Ebből következik, hogy \mathbf{P}_B halmazfüggvényre is teljesülnek a valószínűségi mérték tulajdonságai, melyeket a későbbiekben említés nélkül fölhasználunk.

4.2. Tétel. *Szorzási szabály* Legyenek A_1, A_2, \dots, A_n tetszőleges olyan események, melyekre $\mathbf{P}(A_1 \cap \dots \cap A_{n-1}) > 0$. Ekkor

$$\mathbf{P}(A_1 \cap \dots \cap A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1 \cap A_2) \dots \mathbf{P}(A_n|A_1 \cap \dots \cap A_{n-1}).$$

A bizonyítás előtt megjegyezzük a következőket:

1. A formulában szereplő összes feltétel valószínűsége pozitív, azaz minden jóldefiniált. Ez a $\mathbf{P}(A_1 \cap \dots \cap A_{n-1}) > 0$ feltétel következménye.
2. Ha $\mathbf{P}(A_1 \cap \dots \cap A_n) > 0$ is teljesül, akkor $n!$ darab különböző ilyen szabály van.
3. A szabályt az $n = 2$ esetben használjuk legtöbbször. Ekkor

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

amennyiben A és B is pozitív valószínűségű esemény.

Bizonyítás. A feltételes valószínűség definíciója szerint

$$\begin{aligned} & \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1 \cap A_2) \dots \mathbf{P}(A_n|A_1 \cap \dots \cap A_{n-1}) \\ &= \mathbf{P}(A_1) \frac{\mathbf{P}(A_1 \cap A_2)}{\mathbf{P}(A_1)} \frac{\mathbf{P}(A_1 \cap A_2 \cap A_3)}{\mathbf{P}(A_1 \cap A_2)} \dots \frac{\mathbf{P}(A_1 \cap \dots \cap A_{n-1} \cap A_n)}{\mathbf{P}(A_1 \cap \dots \cap A_{n-1})} \\ &= \mathbf{P}(A_1 \cap \dots \cap A_n), \end{aligned}$$

amint állítottuk. □

4.3. *Példa.* Egy dobozban 12 kék és 3 fehér golyó van. Visszatevés nélkül húzunk két golyót egymás után. Mi annak a valószínűsége, hogy mindkét golyó kék?

Jelölje K_i az az eseményt, hogy az i -ediknek kihúzott golyó kék. Ekkor a szorzási szabály szerint

$$\mathbf{P}(K_1 \cap K_2) = \mathbf{P}(K_1)\mathbf{P}(K_2|K_1) = \frac{12}{15} \cdot \frac{11}{14}.$$

Ugyanezt kapjuk a klasszikus valószínűségi mezőre vonatkozó kedvező/összes formulával is, mely szerint

$$\mathbf{P}(K_1 \cap K_2) = \frac{\binom{12}{2}}{\binom{15}{2}}.$$

A B_1, B_2, \dots események **teljes eseményrendszert** alkotnak, ha

- minden $i \neq j$ párra $B_i \cap B_j = \emptyset$;
- $\cup_{n=1}^{\infty} B_n = \Omega$.

4.4. Tétel. *Teljes valószínűség tétele* Legyen B_1, B_2, \dots teljes eseményrendszer, melyre $\mathbf{P}(B_n) > 0$ minden n -re. Ekkor tetszőleges $A \in \mathcal{A}$ esemény esetén

$$\mathbf{P}(A) = \sum_{n=1}^{\infty} \mathbf{P}(A|B_n)\mathbf{P}(B_n).$$

Bizonyítás. A feltételes valószínűség definíciója és a valószínűség additivitása alapján

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbf{P}(A|B_n)\mathbf{P}(B_n) &= \sum_{n=1}^{\infty} \frac{\mathbf{P}(A \cap B_n)}{\mathbf{P}(B_n)} \cdot \mathbf{P}(B_n) = \sum_{n=1}^{\infty} \mathbf{P}(A \cap B_n) \\ &= \mathbf{P}(\cup_{n=1}^{\infty} (A \cap B_n)) = \mathbf{P}(A \cap (\cup_{n=1}^{\infty} B_n)) \\ &= \mathbf{P}(A \cap \Omega) = \mathbf{P}(A). \end{aligned}$$

□

4.5. Tétel. *Bayes-formula* Legyenek A és B olyan események, hogy $\mathbf{P}(A) > 0$, $\mathbf{P}(B) > 0$. Ekkor

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(A|B)\mathbf{P}(B)}{\mathbf{P}(A)}.$$

Bizonyítás. A definíció szerint

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

□

4.6. Tétel. *Bayes-tétel* Legyen B_1, B_2, \dots teljes eseményrendszer, melyre $\mathbf{P}(B_n) > 0$ minden n -re. Ekkor tetszőleges pozitív valószínűségű $A \in \mathcal{A}$ esemény esetén, tetszőleges k -ra

$$\mathbf{P}(B_k|A) = \frac{\mathbf{P}(A|B_k)\mathbf{P}(B_k)}{\sum_{n=1}^{\infty} \mathbf{P}(A|B_n)\mathbf{P}(B_n)}.$$

Bizonyítás. Előbb a teljes valószínűség tételét, majd a Bayes-formulát használva

$$\frac{\mathbf{P}(A|B_k)\mathbf{P}(B_k)}{\sum_{n=1}^{\infty} \mathbf{P}(A|B_n)\mathbf{P}(B_n)} = \frac{\mathbf{P}(A|B_k)\mathbf{P}(B_k)}{\mathbf{P}(A)} = \mathbf{P}(B_k|A).$$

□

4.7. *Példa. Doppingteszt.* Kifejlesztének egy új doppingtesztet, mely a doppingolók 99%-ánál pozitív eredményt ad, azonban a nem doppingoló sportolók 1%-nál is tévesen pozitív eredményt ad. Tegyük föl, hogy a sportolók 1%-a doppingol. Mennyi annak a valószínűsége, hogy egy véletlenül kiválasztott sportoló

(a) doppingtesztje pozitív?

(b) doppingolt, ha tudjuk, hogy a doppingtesztje pozitív?

Jelölje T azt az eseményt, hogy a teszt eredménye pozitív, és D azt az eseményt, hogy a sportoló doppingolt. Ekkor a feladat (a) része a $\mathbf{P}(T)$, a (b) része a $\mathbf{P}(D|T)$ valószínűséget kérdezi. A teljes valószínűség tételét alkalmazva a D, D^c eseményrendszerre kapjuk

$$\begin{aligned}\mathbf{P}(T) &= \mathbf{P}(D)\mathbf{P}(T|D) + \mathbf{P}(D^c)\mathbf{P}(T|D^c) \\ &= 0,01 \cdot 0,99 + 0,99 \cdot 0,01 = 0,0198.\end{aligned}$$

A Bayes-formula szerint

$$\mathbf{P}(D|T) = \frac{\mathbf{P}(T|D)\mathbf{P}(D)}{\mathbf{P}(T)} = \frac{0,99 \cdot 0,01}{0,0198} = \frac{1}{2}.$$

A feladat eredménye meglepő, hiszen egy látszólag jól működő teszt esetén, annak a valószínűsége, hogy egy sportoló tényleg doppingolt, feltéve, hogy a teszt eredménye pozitív, $1/2$. Világos, hogy ilyen tesztelés mellett nem vehetjük el senkitől az olimpiai aranyérmét. A hiba onnan jön, hogy ha 100 sportolóból 1 doppingol, akkor a teszt ezt az 1-et nagy valószínűséggel kimutatja, viszont a 99 becsületes sportoló közül is kb. egyet tévesen a doppingolók közé sorol. Így kb. két pozitív teszteredmény lesz, de a két sportoló közül csak az egyik doppingol.

4.8. *Példa.* Egy hallgató p valószínűséggel tudja a választ egy kérdésre. Ha nem tudja, akkor az n lehetséges válasz közül véletlenül választ egyet. Mennyi legyen a lehetséges válaszok n száma, hogy az oktató legalább $0,9$ valószínűséggel következtethessen arra a hallgató jó válaszából, hogy a hallgató tudta a választ?

Jelölje H azt az eseményt, hogy a hallgató jól válaszol, T pedig azt az eseményt, hogy tudja a választ. Ekkor olyan n értéket keresünk, melyre teljesül a $\mathbf{P}(T|H) > 0,9$ egyenlőtlenség. Tehát a

$$\mathbf{P}(T|H) = \frac{\mathbf{P}(H|T)\mathbf{P}(T)}{\mathbf{P}(H|T)\mathbf{P}(T) + \mathbf{P}(H|T^c)\mathbf{P}(T^c)} = \frac{1 \cdot p}{1 \cdot p + \frac{1}{n} \cdot (1 - p)} \geq 0,9,$$

egyenlőtlenséget kell megoldanunk n -re. Rövid számolás után adódik, hogy $n \geq 9(1 - p)/p$. Azaz $p = 1/2$ esetén az oktátónak legalább 9 lehetséges választ, míg $p = 0,7$ esetén legalább 4 lehetséges választ kell megadnia.

Vegyük észre, hogy ahogy p tart 0-hoz, az oktató bizonyosságához szükséges lehetséges válaszok száma tart végtelenbe. Gondoljuk meg mért természetes ez.

4.9. *Példa.* Szindbádnak jogában áll N háremhölgy közül egyet kiválasztania oly módon, hogy az előtte egyenként elvonuló hölgyek valamelyikére rámutat. Tegyük fel, hogy egyértelmű szigorúan monoton szépségi sorrendet tud felállítani, és a háremhölgyek bármely elvonulási sorrendje egyformán valószínű. Szindbád k hölgyet elenged, majd kiválasztja az elsőt, aki szebb az összes előtte elvonultnál. Mennyi a valószínűsége, hogy a legszebb hölgyet választja ki? Milyen k esetén lesz ez a valószínűség a legnagyobb, ha N elég nagy?

Jelölje A_i azt az eseményt, hogy a i -edik lány a legszebb, $i = 1, 2, \dots, N$, és legyen B az az esemény, hogy Szindbád a legszebb lányt választja. Ekkor a teljes valószínűség tétele szerint

$$\mathbf{P}(B) = \sum_{i=1}^N \mathbf{P}(A_i) \mathbf{P}(B|A_i).$$

Világos, hogy $\mathbf{P}(A_i) = N^{-1}$ minden i -re, és $\mathbf{P}(B|A_i) = 0$ ha $i \leq k$. Ha $i > k$, akkor Szindbád pontosan akkor választja ki a legszebb háremhölgyet, ha az első $i-1$ lány közül a legszebb az első k -ban volt. Tehát $\mathbf{P}(B|A_i) = k/(i-1)$. Összegezve

$$\mathbf{P}(B) = \sum_{i=1}^N \mathbf{P}(A_i) \mathbf{P}(B|A_i) = \sum_{i=k+1}^N \frac{1}{N} \frac{k}{i-1}.$$

5. Függetlenség

Két esemény függetlensége intuitívan azt jelenti, hogy bekövetkezéseik nem befolyásolják egymást. Tekintsünk egy adott kísérlethez tartozó A és B eseményt. Ismételjük n -szer a kísérletet. Ekkor $S_n(A)/n$ az A esemény relatív gyakorisága az n kísérlet során. Most figyeljük csak azokat a kísérleteket, ahol B bekövetkezett, ezek száma $S_n(B)$. Ezek közül $S_n(A \cap B)$ azon kísérletek száma, ahol A is bekövetkezett, így a megfelelő relatív gyakoriság $S_n(A \cap B)/S_n(B)$. Az, hogy A és B nem befolyásolják egymást, azt jelenti, hogy ez a két relatív gyakoriság kb. megegyezik, azaz

$$\frac{S_n(A)}{n} = \frac{S_n(A \cap B)}{S_n(B)} = \frac{S_n(A \cap B)/n}{S_n(B)/n}.$$

A bal oldal kb. $\mathbf{P}(A)$, a jobb oldal pedig $\mathbf{P}(A \cap B)/\mathbf{P}(B)$, vagyis azt kaptuk, hogy

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

Ez a függetlenség definíciója.

Másképpen, a B esemény bekövetkezése nem befolyásolja az A bekövetkezését, azaz $\mathbf{P}(A|B) = \mathbf{P}(A)$, ahonnan $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

5.1. Definíció. Az A és B események **függetlenek**, ha $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

A definícióból világos, hogy a függetlenség szimmetrikus. Továbbá, a biztos ill. a lehetetlen eseménytől minden esemény független.

5.2. *Példa.* Francia kártyapakliból véletlenszerűen húzunk egy lapot. Jelölje D azt az eseményt, hogy dámát húzunk, K pedig azt, hogy kőrt. Ekkor $D \cap K$ az az esemény, hogy a kőr dámát húztuk ki, így $\mathbf{P}(D \cap K) = 1/52$. Ugyanakkor $\mathbf{P}(D) = 4/52 = 1/13$ és $\mathbf{P}(K) = 13/52 = 1/4$, azaz a két esemény független.

5.3. *Példa.* Földobunk n -szer egy szabályos érmét. Legyen A az az esemény, hogy legfeljebb egy fejet dobunk, B pedig az, hogy legalább egy fejet és egy írást dobunk.

Ekkor $\mathbf{P}(A) = (n + 1)/2^n$, $\mathbf{P}(B) = 1 - 2/2^n$, és $\mathbf{P}(A \cap B) = n/2^n$. Azaz A és B pontosan akkor függetlenek, ha

$$\mathbf{P}(A) \cdot \mathbf{P}(B) = \frac{n + 1}{2^n} \frac{2^n - 2}{2^n} = \frac{n}{2^n} = \mathbf{P}(A \cap B)$$

teljesül. Innen kis számolgatással kapjuk, hogy A és B függetlenek, ha $n = 3$, különben pedig nem azok.

5.4. Definíció. Az A, B, C események **függetlenek**, ha $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$, $\mathbf{P}(A \cap C) = \mathbf{P}(A)\mathbf{P}(C)$, $\mathbf{P}(B \cap C) = \mathbf{P}(B)\mathbf{P}(C)$, és $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C)$ teljesül. Továbbá, az A, B, C események **páronként függetlenek**, ha bármely kettő független.

5.5. *Példa.* Válasszunk egyenletes eloszlás szerint egy pontot a $[0, 1]^2$ egység-négyzetben. Legyen A az az esemény, hogy a választott pont a $[0, 1] \times [0, 1/2]$ téglalapba esik, B az az esemény, hogy a választott pont az $[1/2, 1] \times [0, 1]$ téglalapba esik, C pedig az az esemény, hogy a választott pont a $[0, 1/2]^2 \cup [1/2, 1]^2$ halmazba esik. Könnyen ellenőrizhető, hogy A, B, C páronként függetlenek, de *nem* függetlenek.

5.6. Definíció. Az A_1, A_2, \dots, A_n események **függetlenek**, ha bármely $k \in \{2, 3, \dots, n\}$ és $1 \leq i_1 < i_2 < \dots < i_k \leq n$ esetén

$$\mathbf{P}(A_{i_1} \cap \dots \cap A_{i_k}) = \mathbf{P}(A_{i_1}) \dots \mathbf{P}(A_{i_k}).$$

Végtelen sok esemény akkor független, ha közülük bármely véges sok független.

5.7. Állítás. Ha az A_1, \dots, A_n események függetlenek, akkor tetszőleges $k \in \{1, 2, \dots, n\}$ esetén az $\{A_1, \dots, A_k\}$ eseményekből ill. az $\{A_{k+1}, \dots, A_n\}$ eseményekből alkotott események függetlenek.

Ezt nem bizonyítjuk. Az állítás szerint például ha A, B, C, D független események, akkor $A \cup B$ és $C \cap D$ is függetlenek.

5.8. Állítás. Független események közül ha néhányat kicserélünk a komplementerére, akkor is független eseményeket kapunk.

Bizonyítás. Legyenek A_1, A_2, \dots, A_n függetlenek. Nyilván elég megmutatni, hogy A_1^c, A_2, \dots, A_n is függetlenek. Hiszen ekkor egyesével kicserélhetünk akárhány eseményt. A definíciót elég az $1 = i_1 < i_2 < \dots < i_k \leq n$ esetben ellenőrizni, hiszen ha A_1^c nincs a kiválasztott események közt, akkor a feltevés szerint teljesül a függetlenség. Ekkor viszont, előbb a mérték tulajdonsága, majd a függetlenség miatt

$$\begin{aligned} \mathbf{P}(A_1^c \cap A_{i_2} \cap \dots \cap A_{i_k}) &= \mathbf{P}(A_{i_2} \cap \dots \cap A_{i_k}) - \mathbf{P}(A_1 \cap A_{i_2} \cap \dots \cap A_{i_k}) \\ &= \mathbf{P}(A_{i_2}) \dots \mathbf{P}(A_{i_k}) - \mathbf{P}(A_1) \mathbf{P}(A_{i_2}) \dots \mathbf{P}(A_{i_k}) \\ &= [1 - \mathbf{P}(A_1)] \mathbf{P}(A_{i_2}) \dots \mathbf{P}(A_{i_k}) \\ &= \mathbf{P}(A_1^c) \mathbf{P}(A_{i_2}) \dots \mathbf{P}(A_{i_k}), \end{aligned}$$

amit igazolni kellett. □

Kísérletek függetlenségéről akkor beszélünk, ha a hozzájuk tartozó események függetlenek.

5.1. Craps játék

A craps játékot és annak változatait jelenleg is játsszák kaszinókban. A játék az Egyesült Államokban népszerű, 1820 körül terjedt el New Orleansban.

A játékos két dobókockával dob. Ha az első dobásnál a dobott számok összege 7 vagy 11, akkor azonnal nyer, ha 2,3 vagy 12 akkor veszít. Különben folytatja a dobásokat, és akkor nyer, ha hamarabb dobja meg azt az összeget, amit elsőre dobott, mint a 7-et.

A következőkben meghatározzuk a nyerés valószínűségét.

Jelölje A azt az eseményt, hogy nyerünk, A_i pedig azt, hogy az első dobás eredménye i , és nyerünk. Világos, hogy $\mathbf{P}(A_2) = \mathbf{P}(A_3) = \mathbf{P}(A_{12}) = 0$, továbbá

$$\mathbf{P}(A_7) = \frac{6}{36} = \frac{1}{6}, \quad \text{és} \quad \mathbf{P}(A_{11}) = \frac{2}{36} = \frac{1}{18},$$

hiszen ezekben az esetekben a játék az első dobás után véget ér. Ha az első dobásnál az összeg 4, 5, 6, 8, 9 vagy 10 akkor a dolog érdekesebb. Jelölje $A_{i,n}$

azt az eseményt, hogy az első dobásnál az összeg i és pontosan az n -edik dobásnál nyerünk. Nyilván $A_i = \cup_{n=2}^{\infty} A_{i,n}$, és az unió diszjunkt.

Tekintsük az $A_{4,n}$ eseményt. Ekkor az első dobásnál az összeg 4, ami 3 féleképpen következhet be $((1, 3), (2, 2), (3, 1))$, és mivel nyertünk, az utolsó dobásnál is 4 az összeg. A közbülső $n - 2$ dobás során nem dobtunk 4-et, és 7-et, hiszen ekkor véget ért volna a játék korábban. Így $3 + 6$ esetet zártunk ki. Ezek szerint

$$\mathbf{P}(A_{4,n}) = \frac{3 \cdot (36 - 9)^{n-2} \cdot 3}{36^n} = \frac{9}{36^2} \left(\frac{27}{36}\right)^{n-2}.$$

Innen pedig geometria sort összegezve, kapjuk

$$\mathbf{P}(A_4) = \sum_{n=2}^{\infty} \mathbf{P}(A_{4,n}) = \frac{9}{36^2} \left(\frac{27}{36}\right)^{n-2} = \frac{1}{36}.$$

A többi eset hasonlóan megy,

$$\begin{aligned} \mathbf{P}(A_{4,n}) &= \mathbf{P}(A_{10,n}) = \frac{9}{36^2} \left(\frac{27}{36}\right)^{n-2} \\ \mathbf{P}(A_{5,n}) &= \mathbf{P}(A_{9,n}) = \frac{16}{36^2} \left(\frac{26}{36}\right)^{n-2} \\ \mathbf{P}(A_{6,n}) &= \mathbf{P}(A_{8,n}) = \frac{25}{36^2} \left(\frac{25}{36}\right)^{n-2}, \end{aligned}$$

majd a megfelelő geometriai sorokat összegezve

$$\begin{aligned} \mathbf{P}(A_4) &= \mathbf{P}(A_{10}) = \frac{1}{36}, \\ \mathbf{P}(A_5) &= \mathbf{P}(A_9) = \frac{2}{45}, \\ \mathbf{P}(A_6) &= \mathbf{P}(A_8) = \frac{25}{396}. \end{aligned}$$

Végül azt kapjuk, hogy

$$\mathbf{P}(A) = \sum_{i=2}^{12} \mathbf{P}(A_i) = \frac{244}{495} \approx 0,493.$$

6. Véletlen változók

6.1. Definíció. Tekintsünk egy $(\Omega, \mathcal{A}, \mathbf{P})$ valószínűségi mezőt. Az

$$X : \Omega \mapsto \mathbb{R}$$

függvényeket *véletlen változónak* nevezzük, ha a

$$X^{-1}((-\infty, a]) = \{\omega : X(\omega) \leq a\}$$

inverzkép \mathcal{A} -beli tetszőleges $a \in \mathbb{R}$ esetén.

Már sok példát láttunk véletlen változóra. Ilyen például a dobókockával dobott szám értéke, vagy ha három kockával dobunk, akkor a legkisebb dobott szám. Ilyen az ötös lottón kihúzott legnagyobb szám, vagy az egy szelvényen elért találatok száma. Véletlen változó az is, hogy a ropi hol törik el, vagy az egységnégyzetben egyenletesen választott pont milyen távol van a négyzet határától, stb.

6.2. Definíció. Az X véletlen változó *eloszlásfüggvénye* az

$$F(x) = \mathbf{P}(X \leq x) = \mathbf{P}(\{\omega : X(\omega) \leq x\}), \quad x \in \mathbb{R},$$

függvény.

6.3. *Példa.* Dobókockával dobunk. Jelölje X a dobott értéket. Ekkor a lehetséges kimenetek halmaza $\Omega = \{1, 2, \dots, 6\}$, az események halmaza $\mathcal{A} = 2^\Omega$. Szabályos a kockánk, ezért minden értéket $1/6$ -od valószínűséggel dobunk, azaz $\mathbf{P}(A) = \frac{|A|}{6}$, (klasszikus valószínűségi mező). A dobott szám $X : \Omega \rightarrow \mathbb{R}$, $\omega \mapsto \omega$, azaz az identikus leképezés. Ezért

$$\{X \leq x\} = \{\omega : X(\omega) \leq x\} = \begin{cases} \emptyset, & \text{ha } x < 1, \\ \{1, 2, \dots, [x]\}, & \text{ha } 1 \leq x \leq 6, \\ \{1, 2, \dots, 6\}, & \text{ha } x > 6. \end{cases}$$

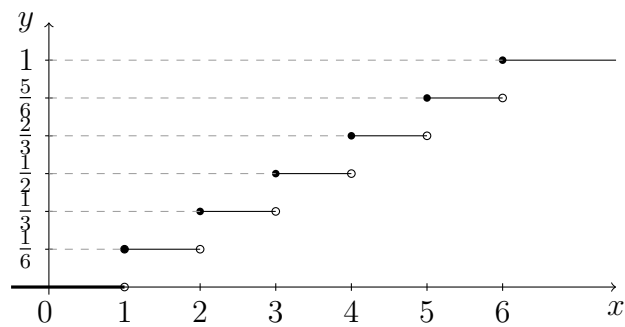
Így az eloszlásfüggvény

$$F(x) = \mathbf{P}(X \leq x) = \mathbf{P}(\{\omega : X(\omega) \leq x\}) = \begin{cases} 0, & x < 1, \\ \frac{[x]}{6}, & 1 \leq x \leq 6, \\ 1, & x \geq 6. \end{cases}$$

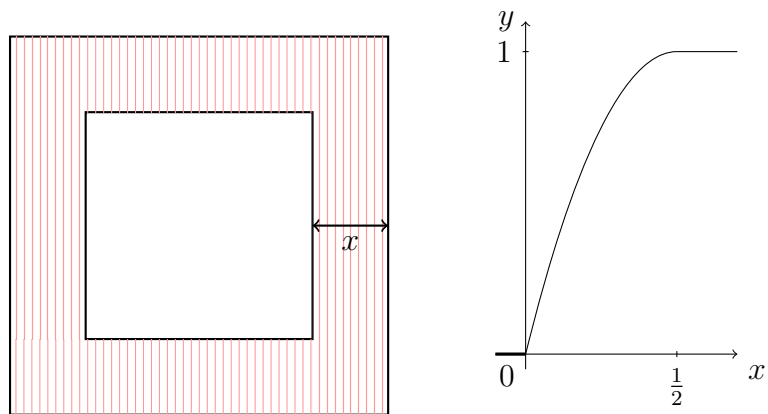
6.4. *Példa.* Egységnégyzetben választunk egyenletes eloszlás szerint egy pontot. Adjuk meg a pont a négyzet határától vett távolságának eloszlását!

Jelölje X a távolságot. Ekkor geometriai valószínűségi mezőn vagyunk, $\Omega = [0, 1]^2$, $\mathcal{A} = \mathcal{B}([0, 1]^2)$, és $\mathbf{P}(A) = |A|$, ahol $|\cdot|$ a terület. Könnyen látható, hogy $X : \Omega \rightarrow \mathbb{R}$, $(u, v) \mapsto \min\{u, v, 1 - u, 1 - v\}$, így

$$\begin{aligned} \{X \leq x\} &= \{\omega : X(\omega) \leq x\} \\ &= \begin{cases} \emptyset, & x < 0, \\ \{(u, v) : \min\{u, v, 1 - u, 1 - v\} \leq x\}, & 0 \leq x \leq 1/2, \\ [0, 1]^2, & x \geq 1/2. \end{cases} \end{aligned}$$



2. ábra. Dobott szám eloszlásfüggvénye



3. ábra. A jó terület és az eloszlásfüggvény

Azaz

$$F(x) = \mathbf{P}(X \leq x) = \mathbf{P}(\{\omega : X(\omega) \leq x\})$$

$$= \begin{cases} 0, & \text{ha } x < 0, \\ 4x(1-x), & \text{ha } 0 \leq x \leq 1/2, \\ 1, & \text{ha } x \geq 1/2, \end{cases}$$

6.5. Tétel. Legyen $F(x)$ egy X véletlen változó eloszlásfüggvénye. Ekkor

- (i) F monoton nemcsökkenő;
- (ii) $\lim_{x \rightarrow \infty} F(x) = 1$ és $\lim_{x \rightarrow -\infty} F(x) = 0$;
- (iii) F jobbról folytonos.

Bizonyítás. (i) Ha $x_1 < x_2$ akkor $\{X \leq x_1\} \subset \{X \leq x_2\}$ és így a mérték monotonitása miatt $F(x_1) = \mathbf{P}(X \leq x_1) \leq \mathbf{P}(X \leq x_2) = F(x_2)$.

(ii) A monotonitásból következik, hogy $\lim_{x \rightarrow \infty} F(x)$ létezik, így elég belátni, hogy $\lim_{n \rightarrow \infty} F(n) = 1$. Tekintsük az $A_n = \{X \leq n\} = \{\omega : X(\omega) \leq n\}$ halmazokat. Ekkor $F(n) = \mathbf{P}(A_n)$. Világos, hogy (A_n) monoton bővülő halmzsorozat, azaz $A_n \subset A_{n+1}$. Ugyanakkor $\cup A_n = \{X < \infty\} = \Omega$, és ezért

$$\lim_{n \rightarrow \infty} \mathbf{P}(A_n) = \mathbf{P}(\cup_{n=1}^{\infty} A_n) = \mathbf{P}(\Omega) = 1.$$

A másik határérték igazolásához is elég részsorozaton dolgozni a monotonitás miatt. Legyen $B_n = \{X \leq -n\}$. Ekkor $F(-n) = \mathbf{P}(B_n)$, a (B_n) halmzsorozat monoton csökkenő, és $\cap B_n = \{X \leq -\infty\} = \emptyset$. Ezért (ismét a mértékek folytonossági tétele szerint)

$$\lim_{n \rightarrow \infty} \mathbf{P}(B_n) = \mathbf{P}(\cap_{n=1}^{\infty} B_n) = \mathbf{P}(\emptyset) = 0.$$

Végül, a (iii) pont belátásához is hasonlóan okoskodunk. Jelölje $F(x+)$ az x pontban vett jobboldali határértéket. Ez megint létezik a monotonitás miatt. Tekintsük a $C_n = \{X \leq x + n^{-1}\}$ halmazokat. Ekkor (C_n) csökkenő halmzsorozat, és $\cap C_n = \{X \leq x\}$. Így ismét a folytonossági tétel szerint

$$F(x+) = \lim_{n \rightarrow \infty} F(x + n^{-1}) = \lim_{n \rightarrow \infty} \mathbf{P}(C_n) = \mathbf{P}(\cap_{n=1}^{\infty} C_n) = \mathbf{P}(X \leq x) = F(x).$$

□

Vegyük észre, hogy F monotonitásából következik az $F(x-)$ baloldali határérték létezése is, azonban általában az $F(x) = F(x-)$ egyenlőség nem teljesül. Az előzőekhez hasonlóan látható, hogy $F(x-) = \mathbf{P}(X < x)$, továbbá

$$F(x) = \mathbf{P}(X \leq x) = \mathbf{P}(X < x) + \mathbf{P}(X = x) = F(x-) + \mathbf{P}(X = x).$$

Ez pedig éppen azt jelenti, hogy F pontosan akkor folytonos az x pontban, ha $\mathbf{P}(X = x) = 0$.

A definícióból adódik, hogy $a < b$ esetén $\mathbf{P}(a < X \leq b) = F(b) - F(a)$.

6.1. Diszkrét véletlen változók

6.6. Definíció. Egy véletlen változó *diszkrét*, ha értékészlete megszámlálható (azaz véges vagy megszámlálhatóan végtelen). Ha egy diszkrét véletlen változó lehetséges értékei x_1, x_2, \dots , akkor $p_i = \mathbf{P}(X = x_i) > 0$ a változó eloszlása.

Ha (p_i) eloszlás, akkor $\sum_i p_i = 1$. Az eloszlásfüggvény $F(x) = \sum_{i: x_i \leq x} p_i$.

6.7. Példa. Legyen $X = I_A$ az A esemény indikátorváltozója. Azaz

$$I_A(\omega) = \begin{cases} 0, & \text{ha } \omega \notin A, \\ 1, & \text{ha } \omega \in A. \end{cases}$$

Ekkor X lehetséges értékei 0 és 1, és $\mathbf{P}(X = 1) = \mathbf{P}(A) = p = 1 - \mathbf{P}(X = 0)$. Ő a p paraméterű Bernoulli eloszlás.

6.8. Példa. Legyen $X = S_n$, egy kísérlet n -szeri ismétlése során az A esemény bekövetkezéseinek a száma. Ekkor X lehetséges értékei $0, 1, 2, \dots, n$, és $\mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, ahol $p = \mathbf{P}(A) \in (0, 1)$. Ő az (n, p) paraméterű binomiális eloszlás.

6.9. Példa. Egy kísérletet addig ismétlünk, amíg egy adott A esemény be nem következik. Legyen X az elvégzett kísérletek száma. Ekkor X lehetséges értékei $1, 2, \dots$, és $\mathbf{P}(X = k) = p(1-p)^{k-1}$. Ő a p paraméterű geometria eloszlás.

6.2. Folytonos véletlen változók

Egy véletlen változó értékészlete nem feltétlenül megszámlálható. A ropi például bárhol eltörhet. Vagy gondolhatunk tetszőleges mérés eredményére, élettartamra, ... Ilyenkor a változó kontinuum sok értéket vehet fel, mindegyiket 0 valószínűséggel. Ez a mese, a definíció a következő.

6.10. Definíció. Egy X véletlen változó *folytonos eloszlású*, ha létezik egy nemnegatív f függvény, melyre

$$F(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^x f(y) dy, \quad x \in \mathbb{R}.$$

Az $f(x)$ függvény az X véletlen változó sűrűségfüggvénye.

A definícióból világos, hogy $\mathbf{P}(X \in (a, b)) = \mathbf{P}(X \in (a, b]) = \int_a^b f(y) dy$, $-\infty \leq a \leq b \leq \infty$. Speciálisan

$$\int_{-\infty}^{\infty} f(x) dx = \mathbf{P}(X \in \mathbb{R}) = 1, \quad \text{és} \quad \mathbf{P}(X = x) = \int_x^x f(y) dy = 0.$$

6.11. *Példa.* A standard normális eloszlás sűrűségfüggvénye $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$. Az eloszlásfüggvény $\Phi(x) = \int_{-\infty}^x \varphi(y)dy$.

6.3. Véletlen vektorváltozók

Egy kísérletnél sokszor több a kísérlet eredményét leíró adatra vagyunk kíváncsiak. Például testtömeg, testmagasság, vérnyomás, pulzus, ...

6.12. Definíció. Az $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ függvény véletlen vektorváltozó, ha minden komponense véletlen változó. Az X eloszlásfüggvénye

$$F(x_1, \dots, x_n) = \mathbf{P}(X_1 \leq x_1, \dots, X_n \leq x_n).$$

Az (X_1, \dots, X_n) véletlen vektorváltozó diszkrét, ha értékészlete megszámlálható, és folytonos, ha van olyan f nemnegatív n -változós függvény, melyre

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(y_1, \dots, y_n) dy_n \dots dy_1$$

teljesül minden $(x_1, \dots, x_n) \in \mathbb{R}^n$ esetén. Ilyenkor az f függvényt az X vektorváltozó sűrűségfüggvényének nevezzük.

Az X_i , $i = 1, 2, \dots, n$, változók eloszlását, peremeloszlásnak, vagy margiális eloszlásnak nevezzük.

Folytonos esetében a definícióból világos, hogy az egyváltozós eset analógiájára

$$\frac{\partial^n}{\partial x_1 \dots \partial x_n} F(x_1, \dots, x_n) = f(x_1, \dots, x_n)$$

teljesül. Az $x_i \rightarrow \infty$, $i = 1, 2, \dots, n$ határátmenettel azt is látjuk, hogy

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_n \dots dx_1 = 1,$$

mint az egyváltozós esetben.

6.13. Állítás. Legyen $X = (X_1, \dots, X_n)$ véletlen vektorváltozó. Az X_i eloszlásfüggvénye

$$F_i(x) = \lim_{x_1 \rightarrow \infty, \dots, x_{i-1} \rightarrow \infty, x_{i+1} \rightarrow \infty, \dots, x_n \rightarrow \infty} F(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n).$$

Bizonyítás. Tekintsük az

$$A_m = \{\omega : X_j(\omega) \leq m, j \neq i, X_i \leq x\}, \quad m \geq 1,$$

halmazokat. Ekkor az (A_m) halmazzorozat monoton bővülő, és

$$\cup_{m=1}^{\infty} A_m = \{\omega : X_i \leq x\}.$$

Tehát

$$\begin{aligned} \lim_{m \rightarrow \infty} F(m, \dots, m, x, m, \dots, m) &= \lim_{m \rightarrow \infty} \mathbf{P}(A_m) \\ &= \mathbf{P}(X_i \leq x) = F_i(x). \end{aligned}$$

A koordinátánkénti monotonitásból az állítás következik. □

6.14. Állítás. Legyen $X = (X_1, \dots, X_n)$ folytonos véletlen vektorváltozó f sűrűségfüggvénnyel. Ekkor X_i , $i = 1, 2, \dots, n$, folytonos véletlen változó, melynek sűrűségfüggvénye

$$f_i(x_i) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(y_1, \dots, y_{i-1}, x_i, y_{i+1}, \dots, y_n) dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_n,$$

azaz az i -edik változón kívül minden változót kiintegrálunk \mathbb{R} -en.

Bizonyítás. Legyen f_i az állításban szereplő függvény. Ekkor a szukcesszív integrálásra vonatkozó tétel szerint

$$\begin{aligned} \int_{-\infty}^x f_i(y) dy &= \int_{-\infty}^x \left[\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(y_1, \dots, y_{i-1}, y, y_{i+1}, \dots, y_n) \right. \\ &\quad \left. dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_n \right] dy \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^x \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(y_1, \dots, y_{i-1}, y, y_{i+1}, \dots, y_n) \\ &\quad dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_n \\ &= \lim_{x_1 \rightarrow \infty, \dots, x_{i-1} \rightarrow \infty, x_{i+1} \rightarrow \infty, x_n \rightarrow \infty} F(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) \\ &= \mathbf{P}(X_i \leq x), \end{aligned}$$

ahol az utolsó előtti egyenlőségénél fölhasználtuk a 6.13 Állítást. Tehát

$$\int_{-\infty}^x f_i(y) dy = \mathbf{P}(X_i \leq x),$$

ami éppen a bizonyítandó állítás. □

6.15. Állítás. Legyen $f(u, v)$ az (X, Y) véletlen vektor sűrűségfüggvénye. Ekkor X és Y is folytonos véletlen változók $\int_{\mathbb{R}} f(u, v) dv$ ill. $\int_{\mathbb{R}} f(u, v) dv$ sűrűségfüggvénnyel.

Bizonyítás. A sűrűségfüggvény definíciója szerint

$$\mathbf{P}(X \leq x) = \mathbf{P}(X \leq x, Y < \infty) = \int_{-\infty}^x \left(\int_{-\infty}^{\infty} f(u, v) dv \right) du.$$

□

6.4. Véletlen változók függetlensége

Legyenek X_1, \dots, X_n az $(\Omega, \mathcal{A}, \mathbf{P})$ valószínűségi mezőn értelmezett véletlen változók.

6.16. Definíció. Az X_1, \dots, X_n függetlenek, ha minden $x_1, \dots, x_n \in \mathbb{R}$ esetén

$$\mathbf{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \mathbf{P}(X_1 \leq x_1) \dots \mathbf{P}(X_n \leq x_n)$$

teljesül. Vagyis az együttes eloszlásfüggvény az egyes eloszlásfüggvények szorzata.

Megjegyzés. Megmutatható, hogy ha X_1, \dots, X_n függetlenek, akkor tetszőleges B_1, \dots, B_n véges vagy végtelen intervallumok esetén

$$\mathbf{P}(X_1 \in B_1, \dots, X_n \in B_n) = \mathbf{P}(X_1 \in B_1) \dots \mathbf{P}(X_n \in B_n)$$

teljesül.

A diszkrét, illetve a folytonos esetben ez a karakterizáció tovább egyszerűsíthető.

6.17. Állítás. *Legyenek X_1, \dots, X_n diszkrét véletlen változók úgy, hogy X_i lehetséges értékei $x_1^{(i)}, x_2^{(i)}, \dots, i = 1, 2, \dots, n$. Ekkor X_1, \dots, X_n pontosan akkor függetlenek, ha*

$$\mathbf{P}(X_1 = x_{i_1}^{(1)}, \dots, X_n = x_{i_n}^{(n)}) = \mathbf{P}(X_1 = x_{i_1}^{(1)}) \dots \mathbf{P}(X_n = x_{i_n}^{(n)})$$

teljesül tetszőleges i_1, \dots, i_n indexekre.

Legyenek X_1, \dots, X_n együttesen folytonos véletlen változók f együttes sűrűségfüggvénnyel. Ekkor X_1, \dots, X_n pontosan akkor függetlenek, ha

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n),$$

ahol f_{X_i} az X_i sűrűségfüggvénye.

Bizonyítás. A diszkrét esetben csak ki kell írni a függetlenség definícióját.

A folytonos esetben csak $n = 2$ esetén bizonyítunk. Az általános eset ugyanez, csak macerásabb a jelölés. Ha a sűrűségfüggvény faktorizálódik, akkor

$$\begin{aligned} \mathbf{P}(X \leq x, Y \leq y) &= \int_{-\infty}^x \int_{-\infty}^y f_X(u) f_Y(v) du dv \\ &= \int_{-\infty}^x f_X(u) du \int_{-\infty}^y f_Y(v) dv = \mathbf{P}(X \leq x) \mathbf{P}(Y \leq y), \end{aligned}$$

azaz a változók függetlenek. Megfordítva, tegyük fel, hogy a változók függetlenek. Ekkor

$$\begin{aligned} \mathbf{P}(X \leq x, Y \leq y) &= \mathbf{P}(X \leq x) \mathbf{P}(Y \leq y) \\ &= \int_{-\infty}^x f_X(u) du \int_{-\infty}^y f_Y(v) dv \\ &= \int_{-\infty}^x \int_{-\infty}^y f_X(u) f_Y(v) du dv, \end{aligned}$$

azaz $f_X(u) f_Y(v)$ az együttes sűrűségfüggvény, amint állítottuk. □

6.5. Függetlenség és geometriai valószínűség

6.18. Definíció. Legyen $T = [a_1, b_1] \times \dots \times [a_n, b_n]$ egy n -dimenziós téglá, ahol $-\infty < a_i < b_i < \infty$, $i = 1, 2, \dots, n$. Az $X = (X_1, \dots, X_n) : \Omega \rightarrow T$ egyenletes eloszlású véletlen változó T -n, ha tetszőleges $S = [c_1, d_1] \times \dots \times [c_n, d_n]$ résztéglájára T -nek

$$\mathbf{P}(X \in S) = \frac{(d_1 - c_1) \dots (d_n - c_n)}{(b_1 - a_1) \dots (b_n - a_n)}.$$

Ekkor X indukál egy geometriai valószínűségi mezőt.

6.19. Állítás. Az $X = (X_1, \dots, X_n)$ véletlen vektorváltozó pontosan akkor egyenletes eloszlású a $T = [a_1, b_1] \times \dots \times [a_n, b_n]$ téglán, ha minden i -re X_i egyenletes eloszlású $[a_i, b_i]$ -n, és X_1, \dots, X_n függetlenek.

Bizonyítás. \Leftarrow : Legyen $S = [c_1, d_1] \times \dots \times [c_n, d_n]$. Ekkor

$$\begin{aligned} \mathbf{P}(X \in S) &= \mathbf{P}(X_1 \in [c_1, d_1], \dots, X_n \in [c_n, d_n]) \\ &= \prod_{i=1}^n \mathbf{P}(X_i \in [c_i, d_i]) \\ &= \prod_{i=1}^n \frac{d_i - c_i}{b_i - a_i} \\ &= \frac{\prod_{i=1}^n (d_i - c_i)}{\prod_{i=1}^n (b_i - a_i)}. \end{aligned}$$

\Rightarrow : Legyen $i \in \{1, \dots, n\}$, $a_i \leq c_i < d_i \leq b_i$ tetszőleges. Ekkor

$$\begin{aligned} \mathbf{P}(X_i \in [c_i, d_i]) &= \mathbf{P}(X_1 \in [a_1, b_1], \dots, X_i \in [c_i, d_i], \dots, X_n \in [a_n, b_n]) \\ &= \mathbf{P}(X \in [a_1, b_1] \times \dots \times [c_i, d_i] \times \dots \times [a_n, b_n]) \\ &= \frac{(d_i - c_i) \prod_{j \neq i} (b_j - a_j)}{\prod_{j=1}^n (b_j - a_j)} \\ &= \frac{d_i - c_i}{b_i - a_i}. \end{aligned}$$

Hasonlóan igazolható a függetlenség az intervallumok esetén, ahonnan pedig következik általánosan. \square

7. Várható érték

Egy kísérletet n -szer függetlenül ismétlünk és minden alkalommal megfigyeljük X értékét: X_1, X_2, \dots, X_n . Ezen értékek átlagai egy számhoz tartanak, ez lesz $\mathbf{E}X$.

Motiváció

7.1. Definíció. Ha X diszkrét véletlen változó x_1, x_2, \dots lehetséges értékekkel, akkor az X várható értéke

$$\mathbf{E}X = \sum_i x_i \mathbf{P}(X = x_i),$$

ha $\sum_i |x_i| \mathbf{P}(X = x_i) < \infty$.

Ha X folytonos véletlen változó $f(x)$ sűrűségfüggvénnyel, akkor az X várható értéke

$$\mathbf{E}X = \int_{-\infty}^{\infty} yf(y)dy,$$

ha $\int_{-\infty}^{\infty} |y|f(y)dy < \infty$.

Folytonos eset magyarázata h hosszúságú intervallumokkal.

7.1. Várható érték tulajdonságai

7.2. Állítás. Legyenek X, Y véletlen változók, $g : \mathbb{R} \rightarrow \mathbb{R}$, $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ olyan függvények, melyekre az állításokban szereplő várható értékek léteznek. Ekkor

$$\mathbf{E}g(X) = \sum_{i=1}^{\infty} g(x_i) \mathbf{P}(X = x_i), \text{ ill. } \mathbf{E}g(x) = \int_{-\infty}^{\infty} g(y)f(y)dy,$$

ahol $f(x)$ az X sűrűségfüggvénye, és

$$\begin{aligned} \mathbf{E}h(X, Y) &= \sum_i \sum_j h(x_i, y_j) \mathbf{P}(X = x_i, Y = y_j), \text{ ill.} \\ \mathbf{E}h(X, Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) dx dy, \end{aligned}$$

ahol $f(x, y)$ az (X, Y) folytonos véletlen vektorváltozó sűrűségfüggvénye.

Bizonyítás. Csak az első állítást igazoljuk és csak diszkrét esetben. Mivel X diszkrét, ezért $g(X)$ is diszkrét y_1, y_2, \dots lehetséges értékekkel. Ezért

$$\begin{aligned} \mathbf{E}g(X) &= \sum_j y_j \mathbf{P}(g(X) = y_j) \\ &= \sum_j y_j \sum_{i:g(x_i)=y_j} \mathbf{P}(X = x_i) \\ &= \sum_j \sum_{i:g(x_i)=y_j} y_j \mathbf{P}(X = x_i) \\ &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i) \mathbf{P}(X = x_i) \\ &= \sum_i g(x_i) \mathbf{P}(X = x_i). \end{aligned}$$

□

7.3. Állítás. A következőkben a, b valós konstansok, X, Y, X_1, \dots, X_n véletlen változók.

(i) A várható érték lineáris, azaz tetszőleges $a, b \in \mathbb{R}$ állandókra

$$\mathbf{E}aX + b = a\mathbf{E}X + b.$$

(ii) Ha $a \leq X \leq b$, akkor $a \leq \mathbf{E}X \leq b$ tetszőleges $a, b \in \mathbb{R}$ számok esetén.

(iii) $\mathbf{E}(X + Y) = \mathbf{E}X + \mathbf{E}Y$.

(iv) Ha X_1, X_2, \dots, X_n véletlen változók, akkor

$$\mathbf{E} \sum_{i=1}^n X_i = \sum_{i=1}^n \mathbf{E}X_i.$$

(v) Ha X és Y függetlenek, akkor $\mathbf{E}g_1(X)g_2(Y) = \mathbf{E}g_1(X)\mathbf{E}g_2(Y)$. Speciálisan, ha X és Y függetlenek, akkor $\mathbf{E}XY = \mathbf{E}X\mathbf{E}Y$.

Bizonyítás. (i) Az előző állítást $g(x) = ax + b$ függvénnyel felírva

$$\mathbf{E}(aX + b) = \sum_i (ax_i + b)\mathbf{P}(X = x_i) = a\mathbf{E}X + b.$$

Folytonosra ugyanígy.

(ii)

$$a = a \sum_i \mathbf{P}(X = x_i) \leq \sum_i x_i \mathbf{P}(X = x_i) = \mathbf{E}X \leq \sum_i b \mathbf{P}(X = x_i) = b.$$

Folytonosra ugyanígy.

(iii) A 7.2 Állítást $h(x, y) = x + y$ függvénnyel felírva

$$\begin{aligned} \mathbf{E}(X + Y) &= \sum_i \sum_j (x_i + y_j) \mathbf{P}(X = x_i, Y = y_j) \\ &= \sum_i x_i \sum_j \mathbf{P}(X = x_i, Y = y_j) + \sum_j y_j \sum_i \mathbf{P}(X = x_i, Y = y_j) \quad \text{tvt} \\ &= \sum_i x_i \mathbf{P}(X = x_i) + \sum_j y_j \mathbf{P}(Y = y_j) \\ &= \mathbf{E}X + \mathbf{E}Y. \end{aligned}$$

Csak diszkrét esetben bizonyítunk.

(iv) Következik (iii)-ból teljes indukcióval.

(v) A 7.2 Állítást $h(x, y) = g_1(x)g_2(y)$ függvénnyel felírva

$$\begin{aligned} \mathbf{E}(g_1(X)g_2(Y)) &= \sum_i \sum_j g_1(x_i)g_2(y_j)\mathbf{P}(X = x_i, Y = y_j) \\ &= \sum_i \sum_j g_1(x_i)g_2(y_j)\mathbf{P}(X = x_i)\mathbf{P}(Y = y_j) \quad \text{függetlenség} \\ &= \sum_i g_1(x_i)\mathbf{P}(X = x_i) \sum_j g_2(y_j)\mathbf{P}(Y = y_j) \\ &= \mathbf{E}(g_1(X))\mathbf{E}(g_2(Y)). \end{aligned}$$

A folytonos esetben

$$\begin{aligned} \mathbf{E}(g_1(X)g_2(Y)) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(x)g_2(y)f(x, y)dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(x)g_2(y)f_1(x)f_2(y)dx dy \quad \text{függetlenség} \\ &= \int_{-\infty}^{\infty} g_1(x)f_1(x)dx \int_{-\infty}^{\infty} g_2(y)f_2(y)dy \quad \text{szukcesszív integrálás} \\ &= \mathbf{E}(g_1(X))\mathbf{E}(g_2(Y)). \end{aligned}$$

□

7.4. *Példa.* Csodaország munka törvénykönyve szerint egy cég minden munkása fizetett szabadságot kap azokon a napokon, amikor legalább az egyiküknek születésnapja van. Ezen napok kivételével azonban az év minden napján mindenkinek dolgoznia kell. Minden munkás 1 TV-készüléket készít egy nap alatt. Hány alkalmazottat vegyen fel a cégtulajdonos, ha azt akarja, hogy a gyártott TV-készülékek számának a várható értéke maximális legyen?

Legyen n az alkalmazottak száma. Jelölje X_n az egy évben gyártott TV-k számát, és legyen Y_i az i -edik napon gyártott TV-k száma, $i = 1, 2, \dots, 365$. Világos, hogy

$$X = Y_1 + Y_2 + \dots + Y_{365}.$$

Másrészt Y_i -k azonos eloszlásúak, lehetséges értékeik n vagy 0, és

$$\mathbf{P}(Y_i = n) = \mathbf{P}(\text{nincs születésnap az } i\text{-edik napon}) = \left(\frac{364}{365}\right)^n.$$

Tehát

$$\mathbf{E}(X_n) = 365 n \left(\frac{364}{365}\right)^n.$$

Egyszerű számolással kapjuk, hogy

$$\frac{\mathbf{E}(X_n)}{\mathbf{E}(X_{n+1})} \leq 1 \Leftrightarrow n \leq 364,$$

azaz a maximum az $n = 364$ és $n = 365$ helyeken vétetik fel.

7.5. Definíció. Az X véletlen változó k -adik momentuma $\mathbf{E}(X^k)$, és k -adik centrális momentuma $\mathbf{E}[(X - \mathbf{E}X)^k]$, $k = 1, 2, \dots$. A 7.2 Állítás szerint

$$\mathbf{E}(X^k) = \begin{cases} \sum_i x_i^k \mathbf{P}(X = X_i), & \text{ha } X \text{ diszkrét,} \\ \int_{-\infty}^{\infty} x^k f(x) dx, & \text{ha } X \text{ folytonos.} \end{cases}$$

7.2. Szórás, kovariancia, korreláció

7.6. Definíció. Az X véletlen változó szórása $\mathbf{D}(X) = \sqrt{\mathbf{E}(X - \mathbf{E}(X))^2}$.

A szórás annak a mérőszáma, hogy a változó mennyire tér el a várható értékétől. Mivel $\mathbf{E}(X - \mathbf{E}(X)) = 0$, $\mathbf{E}|X - \mathbf{E}(X)|$ pedig nehezen kezelhető (nem differenciálható az $|\cdot|$ függvény), ezért ez a legegyszerűbb ilyen.

7.7. Állítás. Tetszőleges X véletlen változó és a, b valós számok esetén

- (i) $\mathbf{D}^2(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2$;
- (ii) $\mathbf{D}^2(aX + b) = a^2 \mathbf{D}^2(X)$;
- (iii) $\mathbf{D}(X) = 0$ akkor és csak akkor, ha $X = \mathbf{E}(X)$, azaz X konstans véletlen változó.

Bizonyítás. A definíció alkalmazása. □

Véletlen változók függőségének mérőszámai a kovariancia és a korreláció.

7.8. Definíció. Az X és Y véletlen változók *kovarianciája*

$$\mathbf{Cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))],$$

korrelációja

$$\rho(X, Y) = \frac{\mathbf{Cov}(X, Y)}{\mathbf{D}(X)\mathbf{D}(Y)}.$$

A kovariancia egyszerű tulajdonságai:

7.9. Állítás. Tetszőleges $X, X_1, \dots, X_n, Y, Y_1, \dots, Y_m$ véletlen változók és a, b valós számok esetén igazak az alábbiak.

- (i) $\mathbf{Cov}(X, X) = \mathbf{D}^2(X)$;

- (ii) $\mathbf{Cov}(X, Y) = \mathbf{Cov}(Y, X)$;
 (iii) $\mathbf{Cov}(aX, bY) = ab\mathbf{Cov}(X, Y)$;
 (iv) $\mathbf{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m \mathbf{Cov}(X_i, Y_j)$;
 (v) ha X és Y függetlenek, akkor $\mathbf{Cov}(X, Y) = 0$.

Bizonyítás. Egyszerű számolás. □

7.10. Állítás. (i) $|\mathbf{Cov}(X, Y)| \leq \mathbf{D}(X)\mathbf{D}(Y)$ (Bunyakovszkij–Cauchy–Schwarz),
 ahonnan adódik, hogy $\rho(X, Y) \in [-1, 1]$;

(ii) ha $\rho(X, Y) = 1$, akkor

$$X = \mathbf{E}(X) + \frac{\mathbf{D}(X)}{\mathbf{D}(Y)}(Y - \mathbf{E}(Y));$$

(iii) ha $\rho(X, Y) = -1$, akkor

$$X = \mathbf{E}(X) - \frac{\mathbf{D}(X)}{\mathbf{D}(Y)}(Y - \mathbf{E}(Y)).$$

Bizonyítás. (i): Tekintsük az $U + tV$ véletlen változót, ahol t egy valós szám. Mivel $\mathbf{E}[(U + tV)^2] \geq 0$, ezért a

$$p(t) = \mathbf{E}[(U + tV)^2] = t^2\mathbf{E}(V^2) + 2t\mathbf{E}(UV) + \mathbf{E}(U^2) \quad (1)$$

t -ben másodfokú polinom diszkriminánsa nempozitív. Azaz

$$4[\mathbf{E}(UV)]^2 \leq 4\mathbf{E}(U^2)\mathbf{E}(V^2), \quad (2)$$

amiből következik, hogy

$$|\mathbf{E}(UV)| \leq \sqrt{\mathbf{E}(U^2)\mathbf{E}(V^2)}.$$

Ezt az egyenlőtlenséget az $U = X - \mathbf{E}(X)$ és $V = Y - \mathbf{E}(Y)$ változókra felírva kapjuk az állítást.

(ii) és (iii): Ha $|\rho(X, Y)| = 1$, akkor a (2) egyenlőtlenség $U = X - \mathbf{E}(X)$ és $V = Y - \mathbf{E}(Y)$ változókra egyenlőség, azaz a másodfokú p polinom diszkriminánsa 0. Ezek szerint

$$t_0 = -\frac{\mathbf{E}[(X - \mathbf{E}X)(Y - \mathbf{E}Y)]}{\mathbf{E}[(Y - \mathbf{E}Y)^2]} = -\rho(X, Y) \frac{\mathbf{D}(X)}{\mathbf{D}(Y)}$$

zérushely, vagyis

$$X - \mathbf{E}(X) + t_0(Y - \mathbf{E}(Y)) = 0,$$

ami éppen a bizonyítandó. □

Megjegyzés. Korreláció jelentése. Ha $\rho(X, Y) = 0$, akkor X és Y *korrelálatlanok*. A 7.9 Állítás (v) pontja szerint a függetlenségből következik a korrelálatlanság. Fordítva ez nem igaz, könnyű ellenpéldát gyártani. Mindenesetre, minél kisebb a korreláció annál gyengébb a két változó közötti függés. A 7.10 Állításból pedig azt látjuk, hogy minél közelebb van $|\rho(X, Y)|$ értéke 1-hez, annál erősebb a változók közötti függés.

Ha a korreláció pozitív, akkor ha X nagy, akkor Y is nagy, ha pedig negatív, akkor ha X nagy, akkor Y kicsi, és fordítva. Ezek a megállapítások persze nem tehetők nagyon precízzé, ez a szemléletes jelentés.

7.11. Állítás. *Legyenek X_1, X_2, \dots, X_n páronként független véletlen változók. Ekkor*

$$\mathbf{D}^2 \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \mathbf{D}^2(X_i).$$

Bizonyítás. Egyszerű számolás. □

7.12. *Példa.* Egy szabályos dobókockával n -szer dobunk. Jelölje X a hatosok, Y egyesek számát! Legyen $I_i = 1$, ha az i -edik dobás hatos, különben 0, $J_i = 1$, ha az i -edik dobás egyes, különben 0, $i = 1, 2, \dots, n$. Nyilván

$$X = \sum_{i=1}^n I_i \quad \text{és} \quad Y = \sum_{i=1}^n J_i. \quad (3)$$

Továbbá I_1, \dots, I_n függetlenek, és J_1, \dots, J_n is függetlenek.

Ekkor I_1, \dots, I_n független, *Bernoulli eloszlású* véletlen változók $1/6$ paraméterrel. A megfelelő eloszlás $\mathbf{P}(I_i = 1) = 1/6$, $\mathbf{P}(I_i = 0) = 5/6$. A J -kre hasonlóan. A hatosok száma X *binomiális eloszlású* véletlen változó ($n, 1/6$) paraméterrel. Eloszlása

$$\mathbf{P}(X = k) = b_k(n, 1/6) = \binom{n}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}, \quad k = 0, 1, \dots, n.$$

A várható értéket és a szórást a definíció alapján számolhatjuk. Valóban,

$$\mathbf{E}(I_1) = 1 \cdot \frac{1}{6} + 0 \cdot \frac{5}{6} = \frac{1}{6},$$

és

$$\mathbf{D}^2(I_1) = \mathbf{E}(I_1^2) - (\mathbf{E}(I_1))^2 = \frac{1}{6} - \frac{1}{36} = \frac{5}{36}.$$

És persze $\mathbf{E}(I_i) = \mathbf{E}(J_i) = \frac{1}{6}$, $\mathbf{D}^2(I_i) = \mathbf{D}^2(J_i) = \frac{5}{36}$, $i = 1, \dots, n$. Az $\mathbf{E}(X)$, $\mathbf{D}^2(X)$ értékek meghatározása számolásabb:

$$\begin{aligned} \mathbf{E}(X) &= \sum_{k=0}^n k \cdot \mathbf{P}(X = k) \\ &= \sum_{k=0}^n k \cdot \binom{n}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k} \\ &= \frac{n}{6} \sum_{k=1}^n \binom{n-1}{k-1} \left(\frac{1}{6}\right)^{k-1} \left(\frac{5}{6}\right)^{n-k} \\ &= \frac{n}{6}, \end{aligned}$$

és hasonlóan

$$\begin{aligned} \mathbf{E}(X^2) &= \sum_{k=0}^n k^2 \cdot \mathbf{P}(X = k) \\ &= \sum_{k=0}^n (k(k-1) + k) \frac{n!}{k!(n-k)!} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k} \\ &= \frac{n(n-1)}{6^2} \sum_{k=2}^n \binom{n-2}{k-2} \left(\frac{1}{6}\right)^{k-2} \left(\frac{5}{6}\right)^{n-k} + \frac{n}{6} \\ &= \frac{n(n-1)}{36} + \frac{n}{6}. \end{aligned}$$

Ezért

$$\mathbf{D}^2(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = n \frac{5}{36}.$$

Sok számolást megspórolunk, ha felhasználjuk a (3) egyenletet és a 7.3 (iv) és 7.11 Állításokat. Valóban,

$$\mathbf{E}(X) = \sum_{i=1}^n \mathbf{E}(I_i) = \frac{n}{6},$$

és

$$\mathbf{D}^2(X) = \sum_{i=1}^n \mathbf{D}^2(I_i) = n \frac{5}{36}.$$

Sőt, így az X és Y kovarianciáját is könnyen meghatározhatjuk. Vegyük észre, hogy ha $i \neq j$, akkor I_i és J_j függetlenek, azaz $\mathbf{Cov}(I_i, J_j) = 0$, különben $I_i J_i = 0$, ezért $\mathbf{Cov}(I_i, J_i) = \mathbf{E}(I_i J_i) - \mathbf{E}(I_i) \mathbf{E}(J_i) = -1/36$. Tehát

$$\mathbf{Cov}(X, Y) = \sum_{i=1}^n \sum_{j=1}^n \mathbf{Cov}(I_i, J_j) = \sum_{i=1}^n -\frac{1}{36} = -n \frac{1}{36},$$

korrelációjuk pedig

$$\rho(X, Y) = \frac{\mathbf{Cov}(X, Y)}{\mathbf{D}(X)\mathbf{D}(Y)} = \frac{-n\frac{1}{36}}{n\frac{5}{36}} = -\frac{1}{5}.$$

Látjuk, hogy a korreláció negatív, azaz ha sok hatost dobunk, akkor kevés egyest, és fordítva, ami teljesen természetes.

8. Nevezetes eloszlások

8.1. Bernoulli-eloszlás

Az X véletlen változó p paraméterű Bernoulli-eloszlású, $X \sim \text{Bernoulli}(p)$, $p \in [0, 1]$, ha lehetséges értékei $0, 1$, és $\mathbf{P}(X = 1) = p = 1 - \mathbf{P}(X = 0)$. Várható értéke $\mathbf{E}(X) = p$, szórásnégyzete $\mathbf{D}^2(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = p - p^2 = p(1 - p)$.

Tipikus példa egy A esemény I_A indikátorváltozója.

8.2. Binomiális eloszlás

Az X véletlen változó (n, p) paraméterű binomiális eloszlású, $X \sim \text{Bin}(n, p)$, $n \in \{1, 2, \dots\}$, $p \in [0, 1]$, ha lehetséges értékei $0, 1, \dots, n$, és $\mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$, $k = 0, 1, \dots, n$.

Ez tényleg eloszlás, hiszen a binomiális tétel szerint

$$\sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1.$$

Várható értéke $\mathbf{E}(X) = np$, szórásnégyzete $\mathbf{D}^2(X) = np(1 - p)$. Ez ugyanúgy igazolható, mint a fenti példában.

Tipikus példa: egy p valószínűségű A esemény bekövetkezéseinek a számát vizsgáljuk n független kísérlet során. Ekkor, ha

$$I_j = \begin{cases} 1, & \text{ha a } j\text{-edik kísérletnél } A \text{ bekövetkezett,} \\ 0, & \text{különben,} \end{cases}$$

akkor $I_j \sim \text{Bernoulli}(p)$, és $X = \sum_{i=1}^n I_i \sim \text{Bin}(n, p)$. Ebből az előállításból gyorsan adódik a várható értékre és a szórásnégyzetre adott formula.

8.3. Poisson-eloszlás

Az X véletlen változó λ paraméterű Poisson-eloszlású, $X \sim \text{Poisson}(\lambda)$, $\lambda \geq 0$, ha X lehetséges értékei $0, 1, 2, \dots$, és

$$\mathbf{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

Ez valóban eloszlás, hiszen

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}.$$

Várható értéke

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

Második momentuma hasonlóan számolható

$$\mathbf{E}(X^2) = \lambda^2 + \lambda,$$

így szórásnégyzete

$$\mathbf{D}^2(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = \lambda.$$

Poisson-eloszlás a binomiális eloszlás határeloszlásaként áll elő. Legyen $p = p_n = \lambda/n$, valamely $\lambda > 0$ számra. Ha $X_n \sim \text{Bin}(n, p_n)$, akkor némi számolás után

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{P}(X_n = k) &= \binom{n}{k} p_n^k (1 - p_n)^{n-k} \\ &= \frac{\lambda^k}{k!} e^{-\lambda}. \end{aligned}$$

Ezek alapján azt látjuk, hogy akkor lép fel Poisson-eloszlás, ha egy kis valószínűségű eseményt sokszor „ismételünk”:

- téves telefonhívások száma;
- autóbalesetek száma;
- nyomdahubák száma egy oldalon;

- földrengések száma;
- csillagok száma egy adott térrészben;
- mazsolák száma a pudingban.

Az első példa Ladislaus Bortkiewicz (1868–1931) orosz közgazdásztól (statistikus) származik: halálos lórúgások száma egy év alatt a porosz hadseregben (20 évig figyelt 14 lovas ezredet). 1898: A kis számok törvénye (Bortkiewicz-eloszlás).

8.4. Geometriai eloszlás

Az X véletlen változó p paraméterű geometriai eloszlású, $X \sim \text{Geo}(p)$, ha a lehetséges értékek $1, 2, \dots$ és

$$\mathbf{P}(X = k) = (1 - p)^{k-1}p, \quad k = 1, 2, \dots$$

Ez tényleg eloszlás, hiszen

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

Mivel

$$\begin{aligned} \sum_{k=1}^{\infty} kx^{k-1} &= \sum_{k=1}^{\infty} (x^k)' = \left(\sum_{k=1}^{\infty} x^k \right)' \\ &= \left(\frac{1}{1-x} - 1 \right)' = \frac{1}{(1-x)^2}, \end{aligned}$$

ezért a várható érték

$$\mathbf{E}(X) = \sum_{k=1}^{\infty} kp(1 - p)^{k-1} = \frac{1}{p}.$$

A második momentum hasonlóan számolható

$$\mathbf{E}(X^2) = \frac{2 - p}{p^2},$$

és így

$$\mathbf{D}^2(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = \frac{1 - p}{p^2}.$$

Tipikus példa: addig ismétlünk egy kísérletet, amíg a vizsgált A esemény be nem következik.

A geometriai eloszlás a diszkrét örökifjú eloszlás, hiszen ha $k, \ell \in \mathbb{N}$, akkor

$$\begin{aligned} \mathbf{P}(X > k + \ell | X > k) &= \frac{\mathbf{P}(X > k + \ell)}{\mathbf{P}(X > k)} \\ &= \frac{q^{k+\ell}}{q^k} = q^\ell = \mathbf{P}(X > \ell). \end{aligned}$$

8.1. *Példa. Kuponyűjtő probléma.* Egy N különböző elemből álló sokaságból visszatevések mintát veszünk. Jelölje S_r azt a véletlen számot, ahány elemet kellett húznunk, hogy kapjunk r különböző elemet. Határozzuk meg S_r várható értékét, szórását, majd adjunk ezekre kezelhető aszimptotikus egyenlőséget.

Vezessük be az $X_k = S_{k+1} - S_k$ változót, $S_0 = 0$. Ekkor X_k geometriai eloszlású, ahol a siker valószínűsége $p_k = (N - k)/N$.

8.5. Egyenletes eloszlás

Az X véletlen változó *egyenletes eloszlású az (a, b) intervallumon*, $X \sim \text{Egy}(a, b)$, $-\infty < a < b < \infty$, ha sűrűségfüggvénye

$$f(y) = \begin{cases} \frac{1}{b-a}, & \text{ha } y \in (a, b), \\ 0, & \text{különben.} \end{cases}$$

Ez tényleg sűrűségfüggvény, hiszen $f \geq 0$, és $\int_{-\infty}^{\infty} f(y) dy = 1$.

Vegyük észre, hogy ez ugyanaz a definíció, mint korábban, a geometriai valószínűségi mezőnél. Valóban, ha $(c, d) \subset (a, b)$ egy tetszőleges részintervallum, akkor

$$\mathbf{P}(X \in (c, d)) = \int_c^d f(y) dy = \frac{d - c}{b - a}.$$

Eloszlásfüggvénye

$$F(x) = \int_{-\infty}^x f(y) dy = \begin{cases} 0, & \text{ha } x \leq a, \\ \frac{x-a}{b-a}, & \text{ha } x \in [a, b], \\ 1, & \text{ha } x \geq b. \end{cases}$$

Momentumai, $k \geq 1$

$$\begin{aligned} \mathbf{E}(X^k) &= \int_{-\infty}^{\infty} y^k f(y) dy \\ &= \int_a^b y^k \frac{1}{b-a} dy \\ &= \frac{b^{k+1} - a^{k+1}}{(k+1)(b-a)}. \end{aligned}$$

Speciálisan

$$\mathbf{E}(X) = \frac{a+b}{2}, \quad \mathbf{D}^2(X) = \frac{(b-a)^2}{12}.$$

8.6. Exponenciális eloszlás

Az X véletlen változó λ -paraméterű exponenciális eloszlású, $X \sim \text{Exp}(\lambda)$, $\lambda > 0$, ha sűrűségfüggvénye

$$f(y) = \begin{cases} \lambda e^{-\lambda y}, & y \geq 0, \\ 0, & y < 0. \end{cases}$$

Ez tényleg sűrűségfüggvény. A megfelelő eloszlásfüggvény

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(y) dy \\ &= \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & x \leq 0. \end{cases} \end{aligned}$$

Momentumai

$$\begin{aligned} \mathbf{E}(X^k) &= \int_{-\infty}^{\infty} y^k f(y) dy = \int_0^{\infty} y^k \lambda e^{-\lambda y} dy \\ &= \lambda^{-k} \int_0^{\infty} z^k e^{-z} dz = \lambda^{-k} \Gamma(k+1) = \frac{k!}{\lambda^k}. \end{aligned}$$

Itt fölhasználtuk, hogy a

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy, \quad \alpha > 0,$$

Gamma-függvényre teljesül, hogy $\Gamma(k) = (k-1)!$, azaz a függvény a faktoriális folytonos kiterjesztése. Ez az azonosság következik a

$$\Gamma(\alpha+1) = \alpha \Gamma(\alpha)$$

azonosságból, ami parciális integrálással könnyen adódik.

Ezek szerint

$$\mathbf{E}(X) = \frac{1}{\lambda}, \quad \mathbf{D}^2(X) = \frac{1}{\lambda^2}.$$

Az exponenciális eloszlás karakterizálja az ún. örökifjú tulajdonság, vagy emlékezet nélküliség. Ez azt jelenti, hogy tetszőleges $x, y > 0$ esetén

$$\mathbf{P}(X \geq x+y | X \geq x) = \mathbf{P}(X \geq y). \quad (4)$$

Ez valóban azt jelenti, hogy az eloszlás nem öregszik.

Ha $X \sim \text{Exp}(\lambda)$, akkor ez teljesül, hiszen

$$\begin{aligned} \mathbf{P}(X \geq x + y | X \geq x) &= \frac{\mathbf{P}(X \geq x + y)}{\mathbf{P}(X \geq x)} \\ &= e^{-\lambda y} = \mathbf{P}(X \geq y), \end{aligned}$$

ami éppen (4). A fordított irány, logaritmust véve, a Cauchy-féle függvény-egyenlet megoldásából következik.

Tipikus példák: telefonhívás hossza, várakozási idő, alkatrészek élettartama, üvegpothár élethossza.

8.7. Normális eloszlás

Az X véletlen változó *normális eloszlású* μ és σ^2 paraméterekkel, $X \sim N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$, ha sűrűségfüggvénye

$$f_{\mu, \sigma}(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}.$$

A $\mu = 0$ és $\sigma = 1$ paraméterekhez tartozó eloszlást *standard normális eloszlásnak* nevezzük. A normális eloszlást nevezik Gauss-eloszlásnak is.

Könnyen látható, hogy $f_{\mu, \sigma}$ függvény μ -re szimmetrikus, azaz $f_{\mu, \sigma}(\mu + y) = f_{\mu, \sigma}(\mu - y)$, $y \in \mathbb{R}$, μ -ben van a maximuma, és $\mu \pm \sigma$ inflexiós pontok.

Ahhoz, hogy belássuk, hogy $f_{\mu, \sigma}$ sűrűségfüggvény, először az alábbi lemmát igazoljuk.

8.1.1. Lemma. *Az $\int_{-\infty}^{\infty} e^{-t^2/2} dt$ integrál létezik mint improprius Riemann-integrál és értéke $\sqrt{2\pi}$.*

Bizonyítás. Legyen $I = \int_{-\infty}^{\infty} e^{-t^2/2} dt$ és $I_n = \int_{-n}^n e^{-t^2/2} dt$, $n \in \mathbb{N}$. Ekkor $I_n \rightarrow I$, amint $n \rightarrow \infty$. Jelölje $R_n = \{(x, y) : |x| \leq n, |y| \leq n\}$ a $2n$ élhosszúságú négyzetet és $B_n = \{(x, y) : \sqrt{x^2 + y^2} \leq n\}$ az n sugarú körlelapot. A szukcesszív integrálás szabálya szerint

$$I_n^2 = \iint_{R_n} e^{-\frac{x^2+y^2}{2}} dx dy.$$

Vezessük be a $J_n^2 = \iint_{B_n} e^{-(x^2+y^2)/2} dx dy$ jelölést. Ekkor $J_n^2 \leq I_n^2 \leq J_{2n}^2$, hiszen $B_n \subset R_n \subset B_{2n}$, ezért elegendő belátni, hogy $J_n^2 \rightarrow 2\pi$, amint $n \rightarrow \infty$. Áttérve polárkoordinátákra az $x = r \cos \theta$, $y = r \sin \theta$ helyettesítéssel a

$$J_n^2 = \int_0^n \int_0^{2\pi} e^{-r^2/2} r dr d\theta = 2\pi \int_0^n r e^{-r^2/2} dr = 2\pi (1 - e^{-n^2/2})$$

egyenlőséget kapjuk, amiből $\lim_{n \rightarrow \infty} J_n^2 = 2\pi$ adódik. □

Az $f_{\mu,\sigma}$ függvény nemnegatív. A $t = (y - \mu)/\sigma$ helyettesítéssel

$$\begin{aligned}\int_{-\infty}^{\infty} f_{\mu,\sigma}(y)dy &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = 1,\end{aligned}$$

ahol az utolsó egyenlőségnél a 8.1.1 Lemmát használtuk. Azaz $f_{\mu,\sigma}$ valóban sűrűség.

A várható érték

$$\begin{aligned}\mathbf{E}(X) &= \int_{-\infty}^{\infty} y f_{\mu,\sigma}(y)dy \\ &= \frac{\sigma}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} \frac{y - \mu}{\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \frac{1}{\sigma} dy + \int_{-\infty}^{\infty} \frac{\mu}{\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \frac{1}{\sigma} dy \right) \\ &= \mu,\end{aligned}$$

a szórásnégyzet pedig

$$\begin{aligned}\mathbf{D}^2(X) &= \mathbf{E}((X - \mu)^2) \\ &= \int_{-\infty}^{\infty} (y - \mu)^2 f_{\mu,\sigma}(y)dy \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (y - \mu) \cdot \frac{y - \mu}{\sigma^2} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\ &= \left[-\frac{\sigma}{\sqrt{2\pi}} (y - \mu) e^{-\frac{(y-\mu)^2}{2\sigma^2}} \right]_{-\infty}^{\infty} + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\ &= \sigma^2.\end{aligned}$$

Tehát a definícióban szereplő két paraméter az a várható érték és a szórásnégyzet.

Az X eloszlásfüggvénye a következőképpen számolható:

$$\begin{aligned}F(x) &= \mathbf{P}(X \leq x) \\ &= \int_{-\infty}^x f_{\mu,\sigma}(y)dy \\ &= \int_{-\infty}^{(x-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &= \Phi((x - \mu)/\sigma),\end{aligned}\tag{5}$$

ahol

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy$$

a standard normális eloszlás eloszlásfüggvénye. Ebből a számolásból világos, hogy elég a Φ függvény értékeit ismerni, és ebből tetszőleges paraméterű normális eloszlás eloszlásfüggvénye számolható.

Ugyancsak (5) egyszerű következménye az alábbi állítás.

8.2. Állítás. *Ha $X \sim N(\mu, \sigma^2)$, akkor $(X - \mu)/\sigma \sim N(0, 1)$.*

Sőt, ez kicsit általánosabban is igaz: ha $X \sim N(\mu, \sigma^2)$, és a, b valós állandók, $a \neq 0$, akkor $aX + b \sim N(a\mu + b, a^2\sigma^2)$.

A normális eloszlás nagyon erősen koncentrálódik a várható értéke körül. Valóban, ha $X \sim N(\mu, \sigma^2)$ és $Z \sim N(0, 1)$, akkor

$$\begin{aligned} \mathbf{P}(|X - \mu| \leq \lambda\sigma) &= \mathbf{P}(|Z| \leq \lambda) \\ &= \Phi(\lambda) - \Phi(-\lambda) \\ &= 2\Phi(\lambda) - 1 \\ &= \begin{cases} 0,6827, & \lambda = 1, \\ 0,9545, & \lambda = 2, \\ 0,9973, & \lambda = 3, \\ 0,9999, & \lambda = 4. \end{cases} \end{aligned}$$

9. Véletlen változók konvergenciája

9.1. Markov és Csebisev egyenlőtlenségei

9.1. Tétel. *Markov-egyenlőtlenség* Legyen X egy véletlen változó $(\Omega, \mathcal{A}, \mathbf{P})$ valószínűségi mezőn, melynek véges a várható értéke. Ekkor tetszőleges pozitív c konstansra

$$\mathbf{P}(|X| \geq c) \leq \frac{\mathbf{E}(|X|)}{c}.$$

Bizonyítás. Ha X diszkrét x_1, x_2, \dots lehetséges értékekkel, akkor

$$\begin{aligned} \mathbf{P}(|X| \geq c) &= \sum_{i:|x_i| \geq c} \mathbf{P}(X = x_i) \leq \sum_{i:|x_i| \geq c} \frac{|x_i|}{c} \mathbf{P}(X = x_i) \\ &\leq \sum_i \frac{|x_i|}{c} \mathbf{P}(X = x_i) = \frac{\mathbf{E}(|X|)}{c}. \end{aligned}$$

Ha X folytonos f sűrűségfüggvénnyel, akkor

$$\begin{aligned} \mathbf{P}(|X| \geq c) &= \int_{|y| \geq c} f(y) dy \leq \int_{|y| \geq c} \frac{|y|}{c} f(y) dy \\ &\leq \int_{-\infty}^{\infty} \frac{|y|}{c} f(y) dy = \frac{\mathbf{E}(|X|)}{c}. \end{aligned}$$

□

A Markov-egyenlőtlenség egyszerű alkalmazásával adódik a

9.2. Tétel. *Csebisev-egyenlőtlenség* Legyen X egy véletlen változó $(\Omega, \mathcal{A}, \mathbf{P})$ valószínűségi mezőn, melynek véges a szórása. Ekkor tetszőleges pozitív c konstansra

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq c) \leq \frac{\mathbf{D}^2(X)}{c^2}.$$

Bizonyítás. A Markov-egyenlőtlenség szerint

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq c) = \mathbf{P}((X - \mathbf{E}(X))^2 \geq c^2) \leq \frac{\mathbf{D}^2(X)}{c^2}.$$

□

9.3. *Példa.* Most a Csebisev-egyenlőtlenség analízisbeli alkalmazására adunk egy szép példát. Weierstrass approximációtétele szerint a polinomok szuprémum normában sűrűn vannak a zárt intervallumon folytonos függvények terében. Az alábbiakban erre adunk egy konstruktív bizonyítást. Legyen f folytonos függvény a $[0, 1]$ intervallumon. A hozzátartozó n -edik *Bernstein-polinom* $B_n(f)(x) = \sum_{k=0}^n f(k/n) \binom{n}{k} x^k (1-x)^{n-k}$. Ekkor $B_n(f)$ egyenletesen konvergál az f függvényhez, azaz

$$\lim_{n \rightarrow \infty} \sup_{x \in [0,1]} |B_n(f)(x) - f(x)| = 0.$$

Ennek igazolásához vegyük észre, hogy $B_n(f)(x) = \mathbf{E}f(S_n/n)$, ahol $S_n = X_1 + \dots + X_n$, és X_1, \dots, X_n független azonos eloszlású Bernoulli(x) véletlen változók (azaz $\mathbf{P}(X_1 = 1) = x = 1 - \mathbf{P}(X_1 = 0)$). A Csebisev-egyenlőtlenség szerint $\mathbf{P}(|S_n/n - x| > c) \leq \mathbf{D}^2(S_n)/(n^2 c^2) = x(1-x)/(nc^2)$. Legyen $\varepsilon > 0$ rögzített. Mivel folytonos függvény zárt intervallumon egyenletesen folytonos, ezért létezik olyan $\delta = \delta(\varepsilon) > 0$, hogy $|u-v| \leq \delta$ esetén $|f(u) - f(v)| \leq \varepsilon$. Így

$$\begin{aligned} |f(x) - B_n(f)(x)| &= |\mathbf{E}[f(x) - f(S_n/n)]| \leq 2M\mathbf{P}(|S_n/n - x| > \delta) + \varepsilon \\ &\leq 2M \frac{\mathbf{D}^2(S_n)}{n^2 \delta^2} + \varepsilon \leq \frac{2Mx(1-x)}{n\delta^2} + \varepsilon, \end{aligned}$$

ahol M az $|f|$ maximuma a $[0, 1]$ intervallumon. A kapott becslés x -ben egyenletes, ezért az állítást beláttuk.

9.2. Nagy számok gyenge törvénye

9.4. Tétel. *Csebisev-féle nagy számok gyenge törvénye* Legyenek X_1, X_2, \dots páronként független, véges szórású véletlen változók, melyek közös várható értéke μ és szórásnégyzete σ^2 . Ekkor tetszőleges $\varepsilon > 0$ esetén

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left| \frac{X_1 + \dots + X_n}{n} - \mu \right| > \varepsilon \right) = 0.$$

Bizonyítás. A páronkénti függetlenség miatt

$$\mathbf{D}^2(X_1 + \dots + X_n) = n\sigma^2.$$

A Csebisev-egyenlőtlenséget az $X = X_1 + \dots + X_n$ változóra fölírva kapjuk, hogy

$$\begin{aligned} \mathbf{P} \left(\left| \frac{X_1 + \dots + X_n}{n} - \mu \right| > \varepsilon \right) &\leq \frac{\mathbf{D}^2(X_1 + \dots + X_n)}{n^2\varepsilon^2} \\ &\leq \frac{\sigma^2}{n\varepsilon^2}, \end{aligned}$$

ami tart 0-hoz. □

A bizonyításból látjuk, hogy a páronkénti függetlenség helyett elég korrelálatlanságot feltenni.

Speciális esetként adódik a

9.5. Tétel. *Bernoulli-féle nagy számok gyenge törvénye (1713)* Jelölje S_n egy p valószínűségű A esemény bekövetkezéseinek a számát egy kísérlet n független ismétlése során. Ekkor tetszőleges $\varepsilon > 0$ esetén

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left| \frac{S_n}{n} - p \right| > \varepsilon \right) = 0.$$

A tétel szerint a relatív gyakoriságok a fenti értelemben konvergálnak az igazi valószínűséghez. Mivel a valószínűség definícióját a relatív gyakoriságok tulajdonságai motiválták (additivitás), ezért a fenti tétel szerint a valószínűség tényleg az, amit akarunk.

A fenti tételekben szereplő konvergencia a sztochasztikus konvergencia, melynek általános definíciója a következő.

9.6. Definíció. Az $(X_n)_{n \in \mathbb{N}}$ véletlen változók sorozata *sztochasztikusan konvergál* X -hez, ha minden $\varepsilon > 0$ esetén

$$\lim_{n \rightarrow \infty} \mathbf{P} (|X_n - X| > \varepsilon) = 0.$$

Megjegyzés. A fenti tételekben szereplő gyenge jelző arra utal, hogy a konvergencia sztochasztikusan teljesül. Erős konvergenciáról akkor beszélünk, ha a véletlen változók majdnem biztosan konvergálnak. Pontosabban, az $(X_n)_{n \in \mathbb{N}}$ véletlen változók sorozata *majdnem biztosan, vagy egy valószínűséggel konvergál X -hez*, ha

$$\mathbf{P} \left(\left\{ \omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) \right\} \right) = \mathbf{P}(\lim_{n \rightarrow \infty} X_n = X) = 1.$$

Itt persze már az is magyarázatra szorul, hogy az $\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}$ halmaz valóban esemény, azaz eleme a megfelelő σ -algebrának. Ez a σ -algebra tulajdonságaiból következik. Erre részletesebben nem térünk ki.

A nagy számok erős törvénye a következő.

9.7. Tétel. *Nagy számok erős törvénye* Legyenek X, X_1, X_2, \dots független, azonos eloszlású véletlen változók, véges $\mathbf{E}(X)$ várható értékkel. Ekkor

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n X_i}{n} = \mathbf{E}(X) \quad \text{majdnem biztosan.}$$

9.3. Centrális határeloszlás-tétel

A nagy számok törvénye azt állítja, hogy független, azonos eloszlású véletlen változók átlagai közel vannak a várható értékhez. Az alábbiakban ezt a közelséget tesszük precízzé.

9.8. Tétel. *Centrális határeloszlás-tétel* Legyenek X, X_1, X_2, \dots független, azonos eloszlású véletlen változók közös $\mathbf{E}(X) = \mu$ várható értékkel, és véges $\mathbf{D}(X) = \sigma$ szórással. Ekkor tetszőleges $x \in \mathbb{R}$ esetén

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\frac{\sum_{i=1}^n (X_i - \mu)}{\sqrt{n}\sigma} \leq x \right) = \Phi(x),$$

ahol

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy$$

a standard normális eloszlás eloszlásfüggvénye.

A tétel bizonyítása már komolyabb eszközökkel, a karakterisztikus függvények módszerével történik.

A tétel indikátorváltókra vonatkozó speciális esete a

9.9. Tétel. *de Moivre–Laplace tétel* Jelölje S_n egy p valószínűségű A esemény bekövetkezéseinek a számát egy kísérlet n független ismétlése során. Ekkor tetszőleges $x \in \mathbb{R}$ esetén

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq x \right) = \Phi(x).$$

Valóban, korábban láttuk, hogy a p -paraméterű Bernoulli-eloszlás várható értéke p és szórása $\sqrt{p(1-p)}$. A speciális eset bizonyítása a binomiális együtthatók pontos aszimptotikájának meghatározásával történhet.

9.10. *Példa.* A kakucsretyegek polgármesterválasztáson két jelölt van: A és B. Kakucsretyege 40000 szavazója egymástól függetlenül, $1/2-1/2$ valószínűséggel szavaz a két jelölt egyikére. A feszült politikai helyzet miatt a szavazatok újraszámolását rendelik el, ha a két jelöltre leadott szavazatok száma között legfeljebb 100 a különbség. Mi a valószínűsége, hogy újraszámolásra kerül sor?

Ekkor tehát $n = 40000$, $p = \mathbf{P}(\text{A-ra szavaz valaki}) = 1/2$. Legyen S_n az A-ra szavazók száma, ekkor $n - S_n$ a B-re szavazók száma. A kérdés $\mathbf{P}(|S_n - (n - S_n)| \leq 100)$. A CHT-ban előforduló mennyiségek $np = 20000$ és $\sqrt{np(1-p)} = 100$. Így a CHT szerint

$$\begin{aligned} \mathbf{P}(|S_n - (n - S_n)| \leq 100) &= \mathbf{P}(-100 \leq 2S_n - n \leq 100) \\ &= \mathbf{P}\left(-0,5 \leq \frac{S_n - np}{\sqrt{npq}} \leq 0,5\right) \\ &\approx \Phi(0,5) - \Phi(-0,5) \\ &= 2\Phi(0,5) - 1 \approx 0,38. \end{aligned}$$

Galton deszkája. *Sir Francis Galton (1822–1911): polihisztor, Darwin unokatestvére.* A centrális határeloszlás szemléltetése. Az első sorban 1 ék van, alatta 2, ..., az n -edik sorban n . Az n -edik éksor alatt van $n+1$ tartály, 0-tól n -ig sorszámozva. Egy golyót elindítunk az első éknél, és a golyót minden ék $1/2-1/2$ valószínűséggel téríti el jobbra vagy balra. Annak a valószínűsége, hogy a golyó a k -edik tartályban landol $= \frac{1}{2^n} \cdot$ azon útvonalak száma, ahol a golyó k -szor megy jobbra és $(n-k)$ -szor balra $= \frac{1}{2^n} \binom{n}{k}$. Másként, ha S_n a golyó jobbra eltérítéseinek száma, akkor S_n binomiális eloszlású $(n, 1/2)$ paraméterekkel. Ha sok golyót engedünk le, akkor a haranggörbe rajzolódik ki a tartályokban.

10. Statisztikai alapfogalmak

Az $(\Omega, \mathcal{A}, \mathcal{P})$ hármast *statisztikai mezőnek* nevezzük, ha Ω tetszőleges halmaz, az eseménytér, \mathcal{A} az események halmaza, \mathcal{P} pedig valószínűségi mértékek halmaza.

A statisztikai problémák során a megfelelő valószínűségi mérték kiválasztása, illetve ennek valamilyen tulajdonságának meghatározása a feladat.

Független, azonos eloszlású véletlen változók egy X, X_1, X_2, \dots sorozatát *statisztikai mintának* nevezzük. Mivel statisztikai problémák során a közös

eloszlást általában nem ismerjük, ezért a közös eloszlást *háttéreloszlásnak* is nevezik. Ha a minta véges, akkor n -elemű mintáról beszélünk. A minta egy adott realizációját x_1, \dots, x_n jelöli. A minta egy T függvényét *statisztikának* nevezzük.

10.1. Alapstatisztikák

Legyen X_1, X_2, \dots, X_n egy n -elemű minta. A várható érték, szórásnégyzet és kovariancia empirikus megfelelői az alábbiak. Ezek a megfelelő elméleti mennyiségek természetes becslései.

10.1. Definíció. Az

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

a *mintaátlag*. Az *empirikus szórásnégyzet*

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Az $(X_1, Y_1), \dots, (X_n, Y_n)$ minta *empirikus kovarianciája*

$$C_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n) = \frac{1}{n} \sum_{i=1}^n X_i Y_i - \bar{X}_n \bar{Y}_n.$$

Az alábbi egyszerű állítás megkönnyíti az empirikus szórásnégyzet számítását.

10.2. Tétel. *Steiner-tétel* Tetszőleges x_1, \dots, x_n értékekre és c valós számra

$$\frac{1}{n} \sum_{i=1}^n (x_i - c)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 + (\bar{x}_n - c)^2.$$

Bizonyítás. Egyszerű számolás. □

Ezek alapján a $c = 0$ választással

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2.$$

10.3. Definíció. Az X_1, \dots, X_n minta *empirikus eloszlásfüggvénye*

$$F_n(x) = \frac{1}{n} |\{i : X_i \leq x\}|.$$

Az empirikus eloszlásfüggvény indikátorváltozókkal is definiálható:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x).$$

A függetlenség miatt $I(X_1 \leq x), \dots, I(X_n \leq x)$ független Bernoulli-eloszlású véletlen változók $p = F(x)$ paraméterrel, ahol F a közös elméleti eloszlásfüggvény. Innen következik, hogy $nF_n(x) \sim \text{Binom}(n, F(x))$. Ezzel beláttuk az alábbi.

10.4. Állítás. Legyen X_1, X_2, \dots az F háttéreloszlásból származó minta, és tekintsük ennek az F_n empirikus eloszlásfüggvényét. Ekkor

$$\mathbf{E}(F_n(x)) = F(x), \quad \mathbf{D}^2(F_n(x)) = \frac{F(x)(1 - F(x))}{n}.$$

Továbbá, tetszőleges $x \in \mathbb{R}$ esetén, tetszőleges $\varepsilon > 0$ számra

$$\lim_{n \rightarrow \infty} \mathbf{P}(|F_n(x) - F(x)| > \varepsilon) = 0.$$

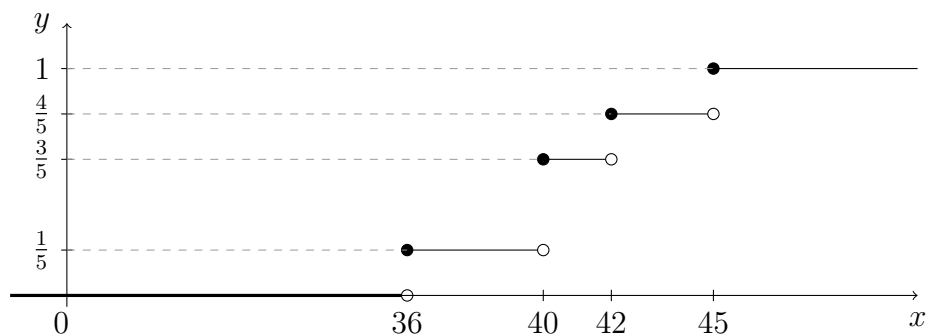
Az utolsó állítás a nagy számok Csebisev-féle gyenge törvénye.

Megjegyzés. A gyenge törvények kapcsán megjegyeztük, hogy a megfelelő egy valószínűségű, vagy majdnem biztos állítások is teljesülnek. A következő, statisztika alaptételének is nevezett eredmény egy ennél lényegesen erősebb állítást fogalmaz meg.

10.5. Tétel. Glivenko–Cantelli-tétel Legyen X_1, X_2, \dots az F háttéreloszlásból származó minta, és jelölje F_n az n -elemű minta empirikus eloszlásfüggvényét. Ekkor

$$\mathbf{P} \left(\limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| = 0 \right) = 1.$$

10.6. Példa. Az $x_1 = 40, x_2 = 45, x_3 = 40, x_4 = 42, x_5 = 36$ minta empirikus eloszlásfüggvénye:



11. Pontbecslések

11.1. Német tank probléma

A II. világháborúban a szövetségesek a havi német tanktermelést akarták megbecsülni. Erre a kilőtt tankok sorozatszámából tudtak következtetni. A matematikai modell a következő. Az $1, 2, \dots, N$ számok közül visszatevés nélkül kivesszünk k számot (kilőtt tankok sorozatszáma). Egy adott megfigyelés (statisztikai minta) alapján adjunk becslést N értékére.

Legyen a mintánk nagyság szerint rendezve $1 \leq X_1 < \dots < X_k \leq N$, és ennek egy adott realizációja $1 \leq x_1 < \dots < x_k \leq N$. A k érték nagyjából $(k+1)$ egyenlő blokkra osztja föl az $\{1, 2, \dots, N\}$ halmazt. Ezek szerint az utolsó kimaradó blokk hossza, $N - X_k$ nagyjából megegyezik az

$$\frac{1}{k} [(X_1 - 1) + (X_2 - X_1 - 1) + \dots + (X_k - X_{k-1} - 1)] = \frac{X_k}{k} - 1$$

átlaggal. Ezért N becslésére természetes jelölt a

$$\hat{N}_1 = \frac{k+1}{k} X_k - 1.$$

Hasonlóan, az legkisebb megfigyelt érték előtt várhatóan annyi elem van, mint a legnagyobb megfigyelt érték után, ezért ugyancsak természetes jelölt N becslésére

$$\hat{N}_2 = X_k + X_1 - 1.$$

Úgy is gondolkodhatunk, hogy a megfigyelt értékek átlaga nagyjából $(N+1)/2$ kell, hogy legyen, és ebből N becslésére

$$\hat{N}_3 = 2 \frac{X_1 + \dots + X_k}{k} - 1$$

adódik.

Melyik becslés jó ezek közül? Ehhez szükségünk van a becslések tulajdonságainak vizsgálatára, különböző becslések összehasonlítására.

11.2. Torzítatlanság

A továbbiakban $(\Omega, \mathcal{A}, \mathcal{P})$ egy statisztikai mező, ahol $\mathcal{P} = \{\mathbf{P}_\theta : \theta \in \Theta\}$. A feladat az ismeretlen θ paraméter, vagy annak valamely függvényének becslése egy minta alapján.

Bevezetjük az $\mathbf{X} = \mathbf{X}_n = (X_1, \dots, X_n)$ jelölést.

11.1. Definíció. A $T(\mathbf{X}_n)$ statisztika $\psi(\theta)$ torzítatlan becslése, ha

$$\mathbf{E}_\theta(T(\mathbf{X}_n)) = \psi(\theta) \quad \text{minden } \theta \in \Theta \text{ esetén.}$$

A $T(\mathbf{X})$ statisztika $\psi(\theta)$ aszimptotikusan torzítatlan becslése, ha

$$\lim_{n \rightarrow \infty} \mathbf{E}_\theta(T(\mathbf{X}_n)) = \psi(\theta) \quad \text{minden } \theta \in \Theta \text{ esetén.}$$

A mintaátlag definíciójából és a várható érték linearitásából azonnal adódik a következő.

11.2. Állítás. *A mintaátlag torzítatlan becslése a várható értéknek, feltéve hogy az létezik.*

Vegyük észre, hogy a minta tetszőleges konvex kombinációja torzítatlan becslése lesz a várható értéknek. Így speciálisan egyetlen mintaelem X_1 is torzítatlan becslés. Tehát a torzítatlanság önmagában nem sokat mond.

11.3. Állítás. *Legyen X, X_1, X_2, \dots független, azonos eloszlású minta, egy olyan eloszlásból, melyre minden $\theta \in \Theta$ esetén $\mathbf{D}_\theta^2(X) < \infty$. Ekkor az empirikus szórásnégyzet nem torzítatlan, csak aszimptotikusan torzítatlan becslése a szórásnégyzetnek. A korrigált empirikus szórásnégyzet*

$$Z_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{n}{n-1} S_n^2$$

torzítatlan becslése a szórásnégyzetnek.

Bizonyítás. A Steiner-tétel és a szórási definíciója alapján

$$\begin{aligned} \mathbf{E}_\theta(S_n^2) &= \frac{1}{n} \sum_{i=1}^n \mathbf{E}_\theta(X_i^2) - \mathbf{E}_\theta((\bar{X}_n)^2) \\ &= \frac{n-1}{n} (\mathbf{E}_\theta(X^2) - (\mathbf{E}_\theta(X))^2) \\ &= \frac{n-1}{n} \mathbf{D}_\theta^2(X), \end{aligned}$$

amiből mindkét állítás adódik. □

11.3. Hatásosság

11.4. Definíció. Legyen T_1, T_2 a $\psi(\theta)$ két torzítatlan becslése. A T_1 hatásosabb mint a T_2 , ha

$$\mathbf{D}_\theta^2(T_1) \leq \mathbf{D}_\theta^2(T_2) \quad \text{minden } \theta \in \Theta \text{ esetén,}$$

és legalább egy $\theta_0 \in \Theta$ esetén szigorú egyenlőtlenség teljesül. Egy torzítatlan becslés hatásos, ha bármely más torzítatlan becslésnél hatásosabb.

Hatásos becslés nem mindig létezik, azonban ha létezik, akkor egyértelmű.

11.5. Tétel. *Ha a háttéreloszlás második momentuma véges, akkor a lineáris becslések között a mintaátlag a leghatásosabb becslése a várható értéknek.*

Bizonyítás. Egy lineáris becslés szükségképpen

$$T(\mathbf{X}) = \sum_{i=1}^n c_i X_i$$

alakú, és mivel a becslés torzítatlan, ezért $\sum_{i=1}^n c_i = 1$. Ekkor a függetlenség miatt

$$\begin{aligned} \mathbf{D}_\theta^2 \left(\sum_{i=1}^n c_i X_i \right) &= \mathbf{D}_\theta^2(X) \sum_{i=1}^n c_i^2 \\ &\geq \mathbf{D}_\theta^2(X) \frac{(\sum_{i=1}^n c_i)^2}{n} \\ &= \frac{\mathbf{D}_\theta^2(X)}{n}, \end{aligned}$$

ahol az egyenlőtlenség a számtani és négyzetes közép közötti egyenlőtlenség miatt teljesül (persze a Cauchy–Schwarz is használható). Pontosan akkor van egyenlőség, ha $c_i = 1/n$, $i = 1, \dots, n$, azaz ha $T(\mathbf{X}) = \bar{X}$, amint állítottuk. \square

Nem mindig a mintaátlag a leghatásosabb becslés a várható értékre. Az egyenletes eloszlás esetén van hatásosabb becslés.

11.6. *Példa.* Legyen X_1, \dots, X_n egy minta a $\text{Egyenletes}(0, \theta)$ háttéreloszlásból, $\theta > 0$. Ekkor \bar{X}_n torzítatlan becslése a $\theta/2$ várható értéknek. Jelölje X_n^* a minta maximális elemét. Ekkor

$$T(\mathbf{X}) = \frac{n+1}{2n} X_n^*$$

is torzítatlan becslése $(\theta/2)$ -nek, és hatásosabb, mint \bar{X} .

11.4. Konzisztencia

11.7. Definíció. A $T(\mathbf{X}_n)$ statisztika *gyengén konzisztens becslése* $\psi(\theta)$ -nak, ha tetszőleges $\theta \in \Theta$ és $\varepsilon > 0$ esetén

$$\lim_{n \rightarrow \infty} \mathbf{P}_\theta (|T(\mathbf{X}_n) - \psi(\theta)| > \varepsilon) = 0.$$

A nagy számok Csebisev-féle gyenge törvénye éppen a mintaátlag gyenge konzisztenciáját jelenti.

11.8. Állítás. *Ha a háttéreloszlás szórásnégyzete létezik, akkor a mintaátlag gyengén konzisztens becslése a várható értéknek.*

Megjegyzés. Hasonlóan definiálható az erős konzisztencia, ami a gyenge konzisztencia definíciójában szereplő sztochasztikus konvergencia helyett majdnem biztos konvergenciát jelent. A nagy számok erős törvénye éppen a mintaátlag erős konzisztenciáját jelenti.

A Csebisev-egyenlőtlenség alkalmazásával kapjuk az alábbi.

11.9. Állítás. *Ha $T(\mathbf{X}_n)$ aszimptotikusan torzítatlan becslés $\psi(\theta)$ -ra, és*

$$\lim_{n \rightarrow \infty} \mathbf{D}_\theta^2(T(\mathbf{X}_n)) = 0 \quad \text{minden } \theta \in \Theta \text{ esetén,}$$

akkor $T(\mathbf{X}_n)$ gyengén konzisztens becslése $\psi(\theta)$ -nak.

Bizonyítás. A Markov-egyenlőtlenség szerint

$$\mathbf{P}_\theta (|T(\mathbf{X}_n) - \psi(\theta)| > \varepsilon) \leq \frac{\mathbf{E}_\theta ((T(\mathbf{X}_n) - \psi(\theta))^2)}{\varepsilon^2}.$$

Tehát elég látni, hogy a számláló 0-hoz tart. Ez pedig következik az aszimptotikus torzítatlanságból és a feltételből, hiszen

$$\begin{aligned} \mathbf{E}_\theta ((T(\mathbf{X}_n) - \psi(\theta))^2) &= \mathbf{E}_\theta (T(\mathbf{X}_n) - \mathbf{E}_\theta(T(\mathbf{X}_n)))^2 + (\mathbf{E}_\theta(T(\mathbf{X}_n)) - \psi(\theta))^2 \\ &= \mathbf{D}_\theta^2(T(\mathbf{X}_n)) + (\mathbf{E}_\theta(T(\mathbf{X}_n)) - \psi(\theta))^2, \end{aligned}$$

ahol mindkét tag 0-hoz tart. □

11.10. Tétel. *Ha $\mathbf{E}_\theta(X^4) < \infty$, minden $\theta \in \Theta$ esetén, akkor mind az empirikus szórásnégyzet, mind a korrigált empirikus szórásnégyzet gyengén konzisztens becslése a szórásnégyzetnek.*

Bizonyítás. A 11.9 Állítás szerint azt kell igazolni, hogy

$$\lim_{n \rightarrow \infty} \mathbf{D}_\theta^2(S_n^2) = 0 \quad \text{és} \quad \lim_{n \rightarrow \infty} \mathbf{D}_\theta^2(Z_n^2) = 0.$$

Ez hosszadalmas, de egyszerű számolás.

Mivel mindkét statisztika változatlan marad, ha a mintaelemeket eltoljuk, azaz X_1, X_2, \dots helyett a $X_1 - c, X_2 - c, \dots$ értékeket tekintjük, ezért föltehető,

hogy $\mathbf{E}_\theta(X) = 0$. Ekkor $\mathbf{E}_\theta(X_i^2) = \sigma_\theta^2$. Az aszimptotikus torzítatlanság bizonyításánál láttuk, hogy

$$\mathbf{E}_\theta(S_n^2) = \frac{n-1}{n}\sigma^2,$$

ezért

$$\mathbf{D}_\theta^2(S_n^2) = \mathbf{E}_\theta(S_n^4) - \left(\frac{n-1}{n}\sigma^2\right)^2.$$

A Steiner-tétel alapján

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2.$$

Így

$$\begin{aligned} \mathbf{E}_\theta(S_n^4) &= \frac{1}{n^2} \mathbf{E}_\theta \left[\left(\sum_{i=1}^n X_i^2 \right)^2 \right] - \frac{2}{n^3} \mathbf{E}_\theta \left[\left(\sum_{i=1}^n X_i^2 \right) \left(\sum_{i=1}^n X_i \right)^2 \right] \\ &\quad + \frac{1}{n^4} \mathbf{E}_\theta \left[\left(\sum_{i=1}^n X_i \right)^4 \right]. \end{aligned}$$

A függetlenség miatt, és mert $\mathbf{E}_\theta(X) = 0$, az $\mathbf{E}_\theta(X_i X_j X_k X_\ell)$ várható érték pontosan akkor nem 0, ha az i, j, k, ℓ indexek között 2-2 megegyezik. Ezért, némi számolás után

$$\begin{aligned} \mathbf{E}_\theta \left[\left(\sum_{i=1}^n X_i^2 \right)^2 \right] &= n \mathbf{E}_\theta(X^4) + n(n-1) \sigma_\theta^4, \\ \mathbf{E}_\theta \left[\left(\sum_{i=1}^n X_i^2 \right) \left(\sum_{i=1}^n X_i \right)^2 \right] &= n \mathbf{E}_\theta(X^4) + n(n-1) \sigma_\theta^4, \\ \mathbf{E}_\theta \left[\left(\sum_{i=1}^n X_i \right)^4 \right] &= n \mathbf{E}_\theta(X^4) + 3n(n-1) \sigma_\theta^4. \end{aligned}$$

Amit visszahelyettesítve

$$\mathbf{E}_\theta(S_n^4) = \sigma^4 \frac{n-1}{n} \left(1 - \frac{2}{n} + \frac{3}{n^2} \right) + \mathbf{E}_\theta(X^4) \left(\frac{1}{n} - \frac{2}{n^2} + \frac{1}{n^3} \right).$$

Végül kapjuk, hogy

$$\mathbf{D}_\theta^2(S_n^2) = \sigma^4 \frac{n-1}{n} \left(-\frac{1}{n} + \frac{3}{n^2} \right) + \mathbf{E}_\theta(X^4) \left(\frac{1}{n} - \frac{2}{n^2} + \frac{1}{n^3} \right),$$

ami $1/n$ rendben tart 0-hoz. Mivel $Z_n^2 = \frac{n}{n-1} S_n^2$, ezért

$$\mathbf{D}_\theta^2(Z_n^2) = \frac{n^2}{(n-1)^2} \mathbf{D}_\theta^2(S_n^2)$$

is tart 0-hoz. Ezzel a tételt beláttuk. \square

11.11. Tétel. *Legyen $(X, Y), (X_1, Y_1), \dots$ olyan statisztikai minta, melyre $\mathbf{E}_\theta(X^4) < \infty$, $\mathbf{E}_\theta(Y^4) < \infty$, minden $\theta \in \Theta$ esetén. Ekkor a*

$$C_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)$$

empirikus kovariancia gyengén konzisztens becslése a kovarianciának.

12. Becslési módszerek

12.1. Maximum likelihood módszer

A maximum likelihood módszer, vagy a legnagyobb valószínűség elve, statisztikában nagyon gyakran használt paraméterbecslési eljárás. Tekintsünk egy $(\Omega, \mathcal{A}, \mathcal{P})$ statisztikai mezőt, ahol $\mathcal{P} = \{\mathbf{P}_\theta : \theta \in \Theta\}$, ahol $\Theta \subset \mathbb{R}^k$ (általában egy-, néha kétdimenziós). Egy adott $\mathbf{x} = (x_1, \dots, x_n)$ realizáció esetén azt a θ paramétert fogadjuk el, mely mellett a legnagyobb a valószínűség az adott realizációnak.

A precíz általános definíció előtt nézzünk egy példát. Tegyük föl, hogy egy Bernoulli-eloszlásból veszünk mintát, ahol a paraméter $\theta \in [0, 1]$ ismeretlen, ezt akarjuk becsülni. Tehát X_1, \dots, X_n független Bernoulli(θ)-eloszlású véletlen változók. A függetlenség miatt

$$\mathbf{P}_\theta((X_1, \dots, X_n) = (x_1, \dots, x_n)) = \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{n - \sum_{i=1}^n x_i}.$$

Innen látjuk, hogy csak az számít a mintából, hogy hányszor következett be a vizsgált esemény. Tehát az (x_1, \dots, x_n) realizációhoz tartozó valószínűség, ha θ a valódi paraméter

$$L_\theta(x_1, \dots, x_n) = \theta^{s_n} (1 - \theta)^{n - s_n},$$

ahol $s_n = \sum_{i=1}^n x_i$. Azt a θ értéket gondoljuk az igazi paraméternek, mely esetén az adott kimenetel a legvalószínűbb. Azaz a θ paraméter becslésére azt a $\hat{\theta}$ értéket választjuk, melyre

$$L_{\hat{\theta}}(\mathbf{x}) = \sup_{\theta \in [0,1]} L_{\theta}(\mathbf{x}).$$

Rövidebben

$$\hat{\theta} = \operatorname{argmax}_{\theta \in [0,1]} \{ \sup_{\theta \in [0,1]} L_{\theta}(\mathbf{x}) \}.$$

Adott s, n értékek mellett keressük a

$$L_{\theta} = \theta^s (1 - \theta)^{n-s}$$

függvény maximumát a $\theta \in [0, 1]$ intervallumon. Lederiválva

$$\frac{d}{d\theta} L_{\theta} = \theta^{s-1} (1 - \theta)^{n-s-1} (s - n\theta).$$

Látjuk, hogy a derivált pozitív a $[0, s/n)$ intervallumon, s/n helyen 0, és negatív az $(s/n, 1]$ intervallumon. Ezek szerint a

$$\hat{\theta}(\mathbf{x}) = \frac{s_n}{n}$$

becslést kapjuk, ami éppen a relatív gyakoriság, vagy empirikus várható érték.

12.1. Definíció. Amennyiben \mathbf{P}_{θ} melletti háttéreloszlás diszkrét minden $\theta \in \Theta$ paraméterre, akkor a *likelihood-függvény*

$$L_{\theta}(\mathbf{x}) = L_{\theta}(x_1, \dots, x_n) = \mathbf{P}_{\theta}(\mathbf{X} = \mathbf{x}) = \prod_{j=1}^n \mathbf{P}_{\theta}(X = x_j).$$

Ha \mathbf{P}_{θ} melletti háttéreloszlás folytonos minden $\theta \in \Theta$ esetén és sűrűségfüggvénye f_{θ} , akkor a *likelihood-függvény*

$$L_{\theta}(\mathbf{x}) = L_{\theta}(x_1, \dots, x_n) = \prod_{j=1}^n f_{\theta}(x_j).$$

A θ paraméter *maximum likelihood becslése* az \mathbf{x} minta alapján

$$\hat{\theta}(\mathbf{x}) = \operatorname{argmax}_{\theta \in \Theta} \{ \sup_{\theta \in \Theta} L_{\theta}(\mathbf{x}) \}.$$

A szorzatalak miatt általában kényelmesebb a likelihood függvény logaritmusával számolni, amit log-likelihood függvénynek nevezünk, azaz

$$\ell_{\theta}(\mathbf{x}) = \log L_{\theta}(\mathbf{x}).$$

Mivel a logaritmus függvény szigorúan monoton nő, ezért L és ℓ maximum-helye megegyezik.

A Cramér–Dugué tétel szerint bizonyos általános feltételek mellett a $\hat{\theta}$ maximum likelihood becslés erősen konzisztens (és akkor persze gyengén is), aszimptotikusan torzítatlan és aszimptotikusan hatásos becslés θ -ra.

12.2. *Példa. Hipergeometrikus eloszlásból vett minta ML becslése.* Egy területen meg akarjuk becsülni az ott élő madarak ismeretlen N számát. Meggyűrűzünk M madarat, majd befogunk $n \leq M$ madarat, melyek közül m van meggyűrűzve. Tegyük föl, hogy $m \geq 1$, különben fogjunk még madarat. Megadjuk N ML becslését.

A likelihood függvény

$$L_N(m) = \mathbf{P}_N(X = m) = \frac{\binom{M}{m} \binom{N-M}{n-m}}{\binom{N}{n}},$$

ahol X jelöli a másodsorra befogottak közül meggyűrűzöttek számát. A feladat meghatározni, hogy ez milyen N esetén lesz maximális. Ez az N lesz a ML becslés. Egyszerű számolással

$$\frac{L_{N+1}(m)}{L_N(m)} = \frac{(N+1-m)(N+1-n)}{(N+1)(N+1-M-n+m)} > 1.$$

pontosan akkor teljesül, ha

$$N < \frac{nM}{m} - 1.$$

Ezek szerint az $(L_N(m))_{N \geq M}$ sorozat monoton nő $[nM/m]$ -ig, ahol $[\cdot]$ az egészrészfüggvény. Ezek szerint N ML becslése

$$\hat{N} = \left\lceil \frac{nM}{m} \right\rceil.$$

Ez persze logikus is.

12.3. *Példa. Poisson-eloszlás paraméterének ML becslése.* Legyenek X_1, \dots, X_n független, Poisson(λ) eloszlású véletlen változók. Adott $\mathbf{x} = (x_1, \dots, x_n)$ realizáció esetén a log-likelihood függvény

$$\ell_{\lambda}(\mathbf{x}) = \log \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = \sum_{i=1}^n x_i \log \lambda - \sum_{i=1}^n \log x_i! - \lambda n.$$

A maximumhelyet deriválással határozhatjuk meg,

$$\frac{d\ell_\lambda(\mathbf{x})}{d\lambda} = \frac{1}{\lambda} \sum_{i=1}^n x_i - n,$$

amiből a

$$\frac{1}{\lambda} \sum_{i=1}^n x_i - n = 0$$

likelihood egyenlet adódik. Az \bar{x} zérushely valóban maximum, tehát

$$\hat{\lambda} = \bar{x}.$$

Azaz az ML becslés éppen a mintaátlag. Láttuk, hogy ez torzítatlan, gyengén konzisztens becslés.

12.4. *Példa. Exponenciális eloszlás.* Legyenek X_1, \dots, X_n független, $\text{Exp}(\lambda)$ eloszlású véletlen változók. Adott $\mathbf{x} = (x_1, \dots, x_n)$ realizáció esetén a log-likelihood függvény

$$\ell_\lambda(\mathbf{x}) = \log \prod_{i=1}^n \lambda e^{-\lambda x_i} = n \log \lambda - \lambda \sum_{i=1}^n x_i.$$

Innen a

$$\frac{n}{\lambda} - \sum_{i=1}^n x_i = 0$$

likelihood egyenlet adódik, aminek megoldása

$$\hat{\lambda} = \frac{1}{\bar{x}_n}.$$

Ez valóban maximumhely.

12.5. *Példa. Normális eloszlás.* Legyenek X_1, \dots, X_n független, $N(\mu, \sigma^2)$ eloszlású véletlen változók. Adott $\mathbf{x} = (x_1, \dots, x_n)$ realizáció esetén a log-likelihood függvény

$$\begin{aligned} \ell_{(\mu, \sigma^2)}(\mathbf{x}) &= \log \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \\ &= -\frac{n}{2} (\log(2\pi) + \log \sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2. \end{aligned}$$

Vegyük észre, hogy tetszőleges σ^2 esetén log-likelihood függvény

$$\hat{\mu} = \bar{x}_n$$

helyen lesz maximális. Azaz a maximumhely μ -ben nem függ σ -tól. Ezt visszahelyettesítve, a

$$-\frac{n}{2}(\log(2\pi) + \log \sigma^2) - \frac{n}{2\sigma^2} S_n^2$$

függvény maximumát keressük σ^2 függvényeként. Lederiválva

$$\frac{d\ell(\bar{x}_n, \sigma^2)}{d\sigma^2} = -\frac{n}{2} \left(\frac{1}{\sigma^2} - \frac{S_n^2}{(\sigma^2)^2} \right).$$

Innen a

$$\hat{\sigma}^2 = S_n^2$$

becslést kapjuk. Könnyen látható, hogy ez valóban maximum hely. Tehát, a ML becslés

$$(\hat{\mu}, \hat{\sigma}^2) = (\bar{x}_n, S_n^2).$$

Láttuk, hogy \bar{x}_n torzítatlan, konzisztens, míg S_n^2 aszimptotikusan torzítatlan konzisztens becslés.

12.6. *Példa. Egyenletes eloszlás.* Legyenek X_1, \dots, X_n független Egyenletes(a, b) véletlen változók. Adott \mathbf{x} realizáció esetén jelölje x_{\min} és x_{\max} a legkisebb és legnagyobb mintaelemet. Ekkor

$$L_{(a,b)}(\mathbf{x}) = \left(\frac{1}{b-a} \right)^n I(a \leq x_{\min}, x_{\max} \leq b).$$

Ez akkor maximális, amikor az indikátor 1, és $b-a$ a lehető legkisebb. Tehát a

$$(\hat{a}, \hat{b}) = (x_{\min}, x_{\max})$$

lesz az (a, b) ML becslése. Könnyen látható, hogy ez aszimptotikusan torzítatlan, konzisztens becslés.

12.2. Momentumok módszere

A módszert általában több paraméter együttes becslésére használják. Tegyük fel, hogy $\theta = (\theta_1, \dots, \theta_k)$, azaz $k \geq 1$ paraméterünk van. Válasszunk k darab momentumot, általában az első k -t, amelyek egyértelműen meghatározzák a $\theta = (\theta_1, \dots, \theta_k)$ paramétert. Vezessük be a

$$m_j = \mathbf{E}_\theta(X^j) = g_j(\theta_1, \dots, \theta_k), \quad j = 1, 2, \dots, k,$$

jelölést. Az, hogy az első k momentum egyértelműen meghatározza a paramétereket, azt jelenti, hogy vannak olyan h_1, \dots, h_k függvények, hogy

$$h_i(m_1, \dots, m_k) = \theta_i, \quad i = 1, 2, \dots, k.$$

12.7. Definíció. A $\theta = (\theta_1, \dots, \theta_k)$ momentum becslése a

$$\hat{\theta}_i = h_i(\hat{m}_1, \dots, \hat{m}_k), \quad i = 1, 2, \dots, k,$$

statisztika, ahol \hat{m}_i az empirikus i -edik momentum, azaz

$$\hat{m}_i = \frac{1}{n} \sum_{j=1}^n x_j^i, \quad i = 1, 2, \dots, k.$$

A nagy számok erős / gyenge törvénye szerint $\lim_{n \rightarrow \infty} \hat{m}_i = m_i$ egy valószínűséggel. Innen pedig egyszerűen adódik, hogy $\hat{\theta}_i$ erősen konzisztens becslése θ_i -nek minden i -re.

A Poisson-, exponenciális és normális eloszlás esetén a paraméter(ek) momentum becslése megegyezik a ML becsléssel.

12.8. *Példa. Poisson-eloszlás.* A Poisson-eloszlás paramétere megegyezik a várható értékkel, azaz $m_1 = \mathbf{E}_\lambda(X) = \lambda$, azaz $h_1(x) = x$, vagyis a λ paraméter momentum becslése

$$\hat{\lambda} = \hat{m}_1 = \bar{x}_n,$$

éppen a mintaátlag.

12.9. *Példa. Exponenciális-eloszlás.* Az exponenciális eloszlás esetén azaz $m_1 = \mathbf{E}_\lambda(X) = \frac{1}{\lambda}$, azaz $h_1(x) = x^{-1}$, vagyis a λ paraméter momentum becslése

$$\hat{\lambda} = \frac{1}{\hat{m}_1} = \frac{1}{\bar{x}_n},$$

éppen a ML becslés.

12.10. *Példa. Normális eloszlás.* Ha $X \sim N(\mu, \sigma^2)$, akkor $m_1 = \mathbf{E}_{(\mu, \sigma^2)}(X) = \mu$ és $m_2 = \mathbf{E}_{(\mu, \sigma^2)}(X^2) = \sigma^2 + \mu^2$. Tehát

$$h_1(m_1, m_2) = m_1, \quad h_2(m_1, m_2) = m_2 - m_1^2.$$

Az empirikus momentumokat behelyettesítve

$$\begin{aligned} \hat{\mu} &= \hat{m}_1 = \bar{x}_n \\ \hat{\sigma}^2 &= \hat{m}_2 - (\bar{x}_n)^2 = s_n^2. \end{aligned}$$

Azaz a mintaátlag és az empirikus szórásnégyzet a megfelelő becslések.

Az egyenletes eloszlás momentum becslése nem egyezik meg a ML becsléssel.

12.11. *Példa. Egyenletes eloszlás.* Ha $X \sim \text{Egyenletes}(a, b)$, akkor $m_1 = \mathbf{E}_{(a,b)}(X) = \frac{a+b}{2}$ és $m_2 = \mathbf{E}_{(a,b)}(X^2) = \frac{a^2+ab+b^2}{3}$. Rövid számolás után kapjuk, hogy

$$a = m_1 - \sqrt{3(m_2 - m_1^2)}, \quad b = m_1 + \sqrt{3(m_2 - m_1^2)}.$$

Mivel $\hat{m}_1 = \bar{x}_n$ és $\hat{m}_2 - (\bar{x}_n)^2 = s_n^2$, így a

$$\hat{a} = \bar{x}_n - \sqrt{3}s_n, \quad \hat{b} = \bar{x}_n + \sqrt{3}s_n$$

momentum becslést kapjuk.

13. Konfidenciaintervallumok

Eddig az ismeretlen paramétert egyetlen számmal becsültük. Most egy egész intervallumot szeretnénk megadni, amibe a paraméter nagy valószínűséggel beleesik.

13.1. Definíció. A $(T_1(\mathbf{X}), T_2(\mathbf{X}))$ statisztikapárral definiált intervallum *legalább* $1 - \varepsilon$ szintű konfidenciaintervallum a $\psi(\theta)$ paraméterre, ha

$$\mathbf{P}_\theta(T_1(\mathbf{X}) < \psi(\theta) < T_2(\mathbf{X})) \geq 1 - \varepsilon, \quad \forall \theta \in \Theta,$$

ahol $\varepsilon > 0$ előre adott kicsi szám.

Amennyiben a fönti \geq helyett $=$ szerepel, akkor $(T_1(\mathbf{X}), T_2(\mathbf{X}))$ *pontosan* $1 - \varepsilon$ szintű konfidenciaintervallum.

Az $1 - \varepsilon$ érték a szignifikanciaszint, általában 95% vagy 99%.

13.1. Konfidenciaintervallum normális eloszlás várható értékére ismert szórás esetén

Legyenek $X_1, \dots, X_n \sim N(\mu, \sigma_0^2)$ független véletlen változók, ahol μ az ismeretlen paraméter, σ_0^2 ismert. Megadunk μ -re pontosan $1 - \varepsilon$ szintű konfidenciaintervallumot.

Mivel \bar{X} torzítatlan, erősen konzisztens becslés μ -re, ezért az intervallumot $(\bar{X} - r_\varepsilon, \bar{X} + r_\varepsilon)$ alakban keressük.

Szükségünk lesz az alábbi állításra, melyet nem bizonyítunk.

13.2. Tétel. *Ha X, Y független normális eloszlású véletlen változók akkor $X + Y$ is normális eloszlású.*

Ebből indukcióval az is következik, hogy független normálisok összege normális, és persze a várható értékek összeadódnak (azok mindig), és a szórásnégyzetek is összeadódnak (hiszen függetlenek az összeadandók). Ezért $\bar{X} \sim N(\mu, \sigma_0^2/n)$. Tehát

$$\begin{aligned} \mathbf{P}_\mu(\bar{X} - r_\varepsilon < \mu < \bar{X} + r_\varepsilon) &= \mathbf{P}_\mu\left(-\frac{r_\varepsilon}{\sigma_0/\sqrt{n}} < \frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} < \frac{r_\varepsilon}{\sigma_0/\sqrt{n}}\right) \\ &= \Phi\left(\frac{\sqrt{nr_\varepsilon}}{\sigma_0}\right) - \Phi\left(-\frac{\sqrt{nr_\varepsilon}}{\sigma_0}\right) \\ &= 2\Phi\left(\frac{\sqrt{nr_\varepsilon}}{\sigma_0}\right) - 1 \\ &= 1 - \varepsilon, \end{aligned}$$

azaz

$$\Phi\left(\frac{\sqrt{nr_\varepsilon}}{\sigma_0}\right) = 1 - \frac{\varepsilon}{2}.$$

Tehát az

$$u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$$

jelöléssel, ahol Φ^{-1} a Φ függvény inverze, a keresett $1 - \varepsilon$ szintű konfidencia-intervallum

$$\left(\bar{X} - \frac{u_{\varepsilon/2}\sigma_0}{\sqrt{n}}, \bar{X} + \frac{u_{\varepsilon/2}\sigma_0}{\sqrt{n}}\right).$$

Vegyük észre, hogy a szignifikanciaszint növelésével, azaz ε csökkenésével, az intervallum hossza nő, a mintaelemszám növelésével pedig csökken.

13.2. Konfidenciaintervallum normális eloszlás várható értékére ismeretlen szórás esetén

A feladat ugyanaz, mint az előbb, csak most nem mondja meg senki a szórást. Tehát a szórást is becsülni kell. A fenti számolásnál az volt a kulcs, hogy

$$\frac{\sum_{i=1}^n (X_i - \mu)}{\sqrt{n}\sigma} \quad \text{standard normális.}$$

Ismeretlen σ esetén ezt becsüljük a korrigált empirikus szórással és kapjuk a

$$\frac{\sum_{i=1}^n (X_i - \mu)}{\sqrt{n}Z_n} \tag{6}$$

statisztikát, ahol

$$Z_n = \left(\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2\right)^{1/2}$$

a korrigált empirikus szórás. A kapott statisztika úgynevezett Student-eloszlású, ami a következő.

13.3. Definíció. Legyenek Y és Y_1, \dots, Y_n független standard normálisok. Ekkor a

$$T_n = \frac{Y}{\sqrt{(Y_1^2 + \dots + Y_n^2)/n}} \quad (7)$$

változó n szabadsági fokú Student-eloszlású (vagy t -eloszlású) véletlen változó. Eloszlásfüggvénye $\Phi_n(x) = \mathbf{P}(T_n \leq x)$.

A Φ_n értékei is táblázatba vannak szedve.

Vegyük észre, hogy a (7) nevezőjében 1 várható értékű független azonos eloszlású véletlen változók átlaga szerepel. Ha n tart a végtelenbe, akkor a nagy számok törvénye szerint ez a hányados a várható értékhez, azaz 1-hez konvergál. Vagyis, ha n nagy, akkor a (7) hányados nevezője 1, így maga a hányados egy standard normális eloszláshoz van közel. A Student-eloszlás táblázatából láthatjuk, hogy az eloszlásfüggvénye nagy n esetén majdnem megegyezik a standard normális eloszlásfüggvénnyel.

Most is kell egy eredmény, amit nem bizonyítunk.

13.4. Tétel. Legyenek X_1, \dots, X_n független normális eloszlású véletlen változók μ várható értékkel és σ szórással. Ekkor az \bar{X}_n mintaátlag és az S_n^2 empirikus szórásnégyzet függetlenek. Továbbá a

$$\sqrt{n} \frac{\bar{X}_n - \mu}{\sqrt{Z_n}}$$

hányados $n - 1$ szabadsági fokú Student-eloszlású.

Ezen az állításon azért meg kell lepődni egy kicsit, nevezetesen a függetlenségen. Az az állítás, hogy az ugyanabból a mintából számolt mintaátlag és empirikus szórásnégyzet függetlenek.

Ezek után a konfidenciaintervallum meghatározása pontosan úgy megy mint korábban, csak nem Φ -ből hanem Φ_{n-1} -ből számolunk kritikus értéket. Tehát keressük a konfidenciaintervallumot $[\bar{X}_n - r_\varepsilon, \bar{X}_n + r_\varepsilon]$ alakban. Ekkor a definíció szerint

$$\begin{aligned} \mathbf{P}_\mu(\bar{X} - r_\varepsilon < \mu < \bar{X} + r_\varepsilon) &= \mathbf{P}_\mu \left(-\frac{r_\varepsilon}{\sqrt{Z_n}/\sqrt{n}} < \frac{\bar{X} - \mu}{\sqrt{Z_n}/\sqrt{n}} < \frac{r_\varepsilon}{\sqrt{Z_n}/\sqrt{n}} \right) \\ &= \Phi_{n-1} \left(\frac{\sqrt{nr_\varepsilon}}{\sqrt{Z_n}} \right) - \Phi_{n-1} \left(-\frac{\sqrt{nr_\varepsilon}}{\sqrt{Z_n}} \right) \\ &= 2\Phi_{n-1} \left(\frac{\sqrt{nr_\varepsilon}}{\sqrt{Z_n}} \right) - 1 \\ &= 1 - \varepsilon, \end{aligned}$$

azaz

$$\Phi_{n-1} \left(\frac{\sqrt{nr_\varepsilon}}{\sqrt{Z_n}} \right) = 1 - \frac{\varepsilon}{2}.$$

Itt még annyit felhasználtunk, hogy a Student-eloszlás is szimmetrikus, azaz $\Phi_{n-1}(-x) = 1 - \Phi_{n-1}(x)$. Tehát az

$$t_{\varepsilon/2} = \Phi_{n-1}^{-1}(1 - \varepsilon/2)$$

jelöléssel, ahol Φ_{n-1}^{-1} a Φ_{n-1} függvény inverze, a keresett $1 - \varepsilon$ szintű konfidenciaintervallum

$$\left(\bar{X} - \frac{t_{\varepsilon/2}\sqrt{Z_n}}{\sqrt{n}}, \bar{X} + \frac{t_{\varepsilon/2}\sqrt{Z_n}}{\sqrt{n}} \right).$$

13.3. Konfidenciaintervallum normális eloszlások várható értékének különbségére

Sokszor találkozunk olyan problémával, ahol különböző adatokat akarunk összehasonlítani. Tegyük fel, hogy azt akarjuk eldönteni, hogy valamilyen vérnyomáscsökkentő hatásos-e. Ekkor van egy X_1, \dots, X_n és egy Y_1, \dots, Y_m mintánk, ahol az X -ek az adott gyógyszert szedők (szisztolés) vérnyomása, míg a Y -ok a kontrollcsoportban a vérnyomásértékek. Azt szeretnénk eldönteni, hogy az X -ek várható értéke kisebb-e, azaz hatásos-e a gyógyszer.

Legyenek tehát X_1, \dots, X_n független, normális eloszlású véletlen változók μ_1 várható értékkel és σ szórással, míg Y_1, \dots, Y_m független, normális eloszlású véletlen változók, melyek az X -ektől is függetlenek, és a várható értékük μ_2 , szóráruk σ . Tehát feltesszük, hogy a szórássok egyenlők. A $\mu_1 - \mu_2$ különbségre konstruálunk konfidenciaintervallumot. Nyilván μ_1 becslése \bar{X}_n , míg μ_2 becslése \bar{Y}_m , tehát a $\delta = \mu_1 - \mu_2$ különbséget a

$$\hat{\delta} = \bar{X}_n - \bar{Y}_m$$

mennyiséggel becsüljük. Ennek a szórása a függetlenség miatt

$$\mathbf{D}^2(\hat{\delta}) = \mathbf{D}^2(\bar{X}_n) + \mathbf{D}^2(\bar{Y}_m) = \frac{1}{n}\sigma^2 + \frac{1}{m}\sigma^2 = \sigma^2 \left(\frac{1}{n} + \frac{1}{m} \right).$$

A 13.2 Tétel szerint \bar{X}_n és \bar{Y}_m is normális eloszlásúak, és mivel függetlenek ugyanezen tétel szerint a különbségük is normális eloszlású (na jó itt használtuk, hogy normális eloszlású változó -1 -szerese is normális, de ezt láttuk korábban). Tehát $\hat{\delta}$ normális eloszlású, várható értéke $\mu_1 - \mu_2$, szórásnégyzete pedig $\sigma^2(n^{-1} + m^{-1})$. Tehát

$$\frac{\bar{X}_n - \bar{Y}_m - (\mu_1 - \mu_2)}{\sigma\sqrt{n^{-1} + m^{-1}}}$$

standard normális. Ebből pedig legyárthatjuk a konfidenciaintervallumot a korábban látottak szerint. A konfidenciaintervallumot a szimmetria miatt $(\widehat{\delta} - r_\varepsilon, \widehat{\delta} + r_\varepsilon)$ alakban keresve

$$\begin{aligned} \mathbf{P}(\delta \in [\widehat{\delta} - r_\varepsilon, \widehat{\delta} + r_\varepsilon]) &= \mathbf{P}\left(\frac{|\widehat{\delta} - \delta|}{\sigma\sqrt{n^{-1} + m^{-1}}} \leq \frac{r_\varepsilon}{\sigma\sqrt{n^{-1} + m^{-1}}}\right) \\ &= 2\Phi\left(\frac{r_\varepsilon}{\sigma\sqrt{n^{-1} + m^{-1}}}\right) - 1 = 1 - \varepsilon, \end{aligned}$$

ahonnan kapjuk, hogy

$$r_\varepsilon = \sigma\sqrt{n^{-1} + m^{-1}}\Phi^{-1}(1 - \varepsilon/2).$$

Tehát a konfidenciaintervallum

$$\left(\overline{X}_n - \overline{Y}_m - \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}\Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right), \overline{X}_n - \overline{Y}_m + \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}\Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right)\right).$$

Ez persze csak akkor működik, ha ismerjük a szórást. Ha nem ismerjük azt is becsüljük. Mivel az elméleti szórások megegyeznek, ezért a szórásnégyzet becslése

$$D_{n,m} = \frac{1}{n+m-2} \left[\sum_{i=1}^n (X_i - \overline{X}_n)^2 + \sum_{j=1}^m (Y_j - \overline{Y}_m)^2 \right].$$

Ez torzítatlan becslése a szórásnégyzetnek, hiszen láttuk korábban a 11.3 Állításnál, hogy

$$\mathbf{E} \sum_{i=1}^n (X_i - \overline{X}_n)^2 = (n-1)\sigma^2, \quad \text{és} \quad \mathbf{E} \sum_{j=1}^m (Y_j - \overline{Y}_m)^2 = (m-1)\sigma^2.$$

A 13.4 Tételből következik, hogy

$$\frac{\overline{X}_n - \overline{Y}_m - (\mu_1 - \mu_2)}{\sqrt{D_{n,m}}\sqrt{n^{-1} + m^{-1}}}$$

Student-eloszlású $n+m-2$ szabadsági fokkal. Ezért ugyanúgy számolhatunk konfidenciaintervallumot, mint ismert szórásnál, csak a Student-eloszlásból kell kritikus értéket számolni. Azaz a konfidenciaintervallum

$$\left(\overline{X}_n - \overline{Y}_m - \Delta \Phi_{n+m-2}^{-1}\left(1 - \frac{\varepsilon}{2}\right), \overline{X}_n - \overline{Y}_m + \Delta \Phi_{n+m-2}^{-1}\left(1 - \frac{\varepsilon}{2}\right)\right),$$

ahol

$$\Delta = \sqrt{D_{n,m}}\sqrt{\frac{1}{n} + \frac{1}{m}}.$$

14. Hipotézisvizsgálat

14.1. u-próba

Legyen X normális eloszlású véletlen változó ismert σ_0 szórással, és ismeretlen μ várható értékkel. Azt szeretnénk eldönti egy (X_1, \dots, X_n) független minta alapján, hogy igaz-e hogy a várható érték μ_0 .

A nullhipotézis, amit igaznak teszünk föl az az, hogy a várható érték valóban μ_0 , míg az ellenhipotézis, vagy alternatív hipotézis az az, hogy ez nem teljesül. Azaz

$$H_0 : \mu = \mu_0 \quad \text{vs.} \quad H_1 : \mu \neq \mu_0. \quad (8)$$

Az

$$u = \frac{\bar{X}_n - \mu_0}{\sigma_0} \sqrt{n} \quad (9)$$

statisztika standard normális eloszlású a H_0 fennállása esetén. Nagyon fontos megjegyezni, hogy ha H_1 áll fenn, akkor u *nem* standard normális. Legyen $\varepsilon > 0$ rögzített szignifikanciaszint. Az $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$ jelöléssel,

$$\mathbf{P}_{\mu_0} \left(\mu_0 \in \left(\bar{X}_n - \frac{u_{\varepsilon/2} \sigma_0}{\sqrt{n}}, \bar{X}_n + \frac{u_{\varepsilon/2} \sigma_0}{\sqrt{n}} \right) \right) = \mathbf{P}_{\mu_0}(|u| \leq u_{\varepsilon/2}) = 1 - \varepsilon.$$

u-próba. Ezek alapján a következőképpen járunk el:

1. A minta alapján kiszámoljuk az u próbastatisztikát (lásd (9)).
2. Rögzített $\varepsilon > 0$ szignifikanciaszinthez meghatározzuk $u_{\varepsilon/2}$ értéket.
3. Ha $|u| \leq u_{\varepsilon/2}$ akkor elfogadjuk H_0 -t, különben elvetjük.

Valójában $1 - \varepsilon$ szintű konfidenciaintervallumot szerkesztünk a várható értékre, és akkor *fogadjuk el / vetjük el* a nullhipotézist, ha a μ_0 érték *beleesik / nem esik bele* a konfidenciaintervallumba. Ez minden próba esetén így van.

Ekkor kétféleképpen követhetünk el hibát: *Elsőfajú hiba* esetén H_0 fennáll, mégis elutasítjuk, *másodfajú hiba* esetén pedig H_0 nem áll fenn, mégis elfogadjuk.

Elsőfajú hibát akkor követünk el, ha a valódi paraméter μ_0 , és $|u| \geq u_{\varepsilon/2}$, ennek valószínűsége $\mathbf{P}(|u| \geq u_{\varepsilon/2}) = \varepsilon$, azaz ez éppen az előre megadott ε szignifikanciaszint.

Másodfajú hibát akkor követünk el, ha a valódi paraméter $\mu \neq \mu_0$, azaz H_1 áll fenn, és $|u| \leq u_{\varepsilon/2}$. Ennek valószínűsége $\mathbf{P}_{\mu}(|u| \leq u_{\varepsilon/2})$ függ a valódi paraméter értékétől. A próba *erőfüggvénye*

$$\beta_n(\mu, \varepsilon) = 1 - \mathbf{P}_{\mu}(|u| \leq u_{\varepsilon/2}) = \mathbf{P}_{\mu}(|u| \geq u_{\varepsilon/2}). \quad (10)$$

Legyen $\Delta_n = \frac{\mu - \mu_0}{\sigma_0} \sqrt{n}$. Ha H_1 áll fenn, és a várható érték valódi értéke μ , akkor $\sqrt{n}(\bar{X}_n - \mu)/\sigma_0$ változó lesz standard normális, ezért

$$\begin{aligned} \beta_n(\mu, \varepsilon) &= \mathbf{P}_\mu(|u| \geq u_{\varepsilon/2}) \\ &= \mathbf{P}_\mu \left(\left| \frac{\bar{X}_n - \mu}{\sigma_0} \sqrt{n} + \Delta_n \right| \geq u_{\varepsilon/2} \right) \\ &= 1 - \Phi(u_{\varepsilon/2} - \Delta_n) + \Phi(-u_{\varepsilon/2} - \Delta_n) \\ &= 2 - [\Phi(u_{\varepsilon/2} - \Delta_n) + \Phi(u_{\varepsilon/2} + \Delta_n)]. \end{aligned}$$

Innen látjuk, hogy

$$\lim_{n \rightarrow \infty} \beta_n(\mu, \varepsilon) = 1, \quad \text{tetszőleges } \mu \neq \mu_0 \text{ esetén.}$$

Ez utóbbi állítás próba konzisztenciáját jelenti.

14.1. *Példa.* Azt szeretnénk tesztelni, hogy az 1kg-os cukor tényleg 1kg-e? Persze nem várjuk, hogy minden egyes csomagban halálpontosan 1kg cukor legyen. Azt tesszük fel, hogy egy csomag tömege normális eloszlást követ μ várható értékkel (ezt nem ismerjük), és $\sigma_0 = 0,05$ ismert szórással.

A nullhipotézisünk az, hogy valóban 1kg a várható érték, $\mu = 1$, az ellenhipotézis pedig $\mu \neq 1$, azaz

$$H_0 : \mu = 1 \quad \text{vs.} \quad H_1 : \mu \neq 1. \quad (11)$$

Egy 25 elemű minta alapján a mintaátlagra $\bar{x}_{25} = 0,98$ adódik. Ez kisebb, mint 1, de az eltérés adódhat a véletlenből is, nem feltétlenül csaltak. A próbastatisztika értéke

$$u = \frac{\bar{x}_{25} - 1}{0,05} \sqrt{25} = -2.$$

Válasszuk a szignifikanciaszintet 0,05-nek, azaz $\varepsilon = 0,05$. Ekkor a standard normális eloszlástáblázatából kiolvassuk, hogy $u_{\varepsilon/2} = \Phi^{-1}(0,975) = 1,96$. Mivel $|u| = 2 > 1,96$, ezért elvetjük a nullhipotézist, azaz úgy döntünk, hogy nem 1 a várható érték.

Ugyanakkor a gyártó a következőképpen védekezhet. Annak az esélye, hogy elvetjük a nullhipotézist annak ellenére, hogy az igaz, 0,05. Ez nem is olyan kicsi, 100 esetből 5-ször bekövetkezik. Lehet, hogy most is egy kis valószínűségű esemény következett be. Nézzük meg, hogy $\varepsilon = 0,01$ szignifikanciaszinten is elvetjük-e a nullhipotézist. Ekkor $u_{\varepsilon/2} = \Phi^{-1}(0,995) = 2,57$. Mivel $|u| = 2 < 2,57$, ezért a nullhipotézist elfogadjuk.

A feladat jellegétől függően, sokszor ún. egyoldali próbát használunk. Ebben a példában igazából akkor éri sérelem a vásárlót, ha a valódi várható érték kisebb, mint 1. Azaz, a

$$H_0 : \mu = 1 \quad \text{vs.} \quad H_1 : \mu < 1, \quad (12)$$

egyoldali hipotézisvizsgálatot kell elvégezni. H_0 -t akkor fogadjuk el, ha $u \geq -u_\varepsilon = \Phi^{-1}(1 - \varepsilon)$. Ekkor az elsőfajú hiba (azaz, hogy H_0 fennáll, mégis elvetjük) valószínűsége

$$\mathbf{P}_{\mu_0}(u \leq -u_\varepsilon) = 1 - \Phi(u_\varepsilon) = \varepsilon.$$

Ebben az esetben $\varepsilon = 0,01$ szignifikanciaszinthez tartozó kritikus érték $u_\varepsilon = \Phi^{-1}(0,99) = 2,33$. Azaz ilyen szignifikanciaszinten ebben az esetben is felmentjük a gyártót. (Bár nagyon gyanús.)

14.2. t-próba

A feladat ugyanaz, mint az u-próbánál, csak most nem ismerjük a σ szórást. Persze láttunk már ilyet a konfidenciaintervallum megkonstruálásánál.

Legyen X normális eloszlású véletlen változó *ismeretlen* σ szórással, és ismeretlen μ várható értékkel. Azt szeretnénk eldönti egy (X_1, \dots, X_n) független minta alapján, hogy igaz-e hogy a várható érték μ_0 .

A nullhipotézis, amit igaznak teszünk föl az az, hogy a várható érték valóban μ_0 , míg az ellenhipotézis, vagy alternatív hipotézis az az, hogy ez nem teljesül. Azaz

$$H_0 : \mu = \mu_0 \quad \text{vs.} \quad H_1 : \mu \neq \mu_0. \quad (13)$$

Mivel σ ismeretlen, ezért becsüljük a Z_n korrigált empirikus szórással. Különben minden ugyanaz. Így a tesztstatisztika

$$t = \frac{\bar{X}_n - \mu_0}{Z_n} \sqrt{n} \quad (14)$$

alakú, ami H_0 fennállása esetén Student-eloszlású $n-1$ szabadsági fokkal. Ezt láttuk a konfidenciaintervallumnál. A kritikus értéket is a Student-eloszlásból számoljuk. Persze most is, ha H_1 áll fenn, akkor t *nem* Student. Legyen $\varepsilon > 0$ rögzített szignifikanciaszint. A $t_{\varepsilon/2} = \Phi_{n-1}^{-1}(1 - \varepsilon/2)$ jelöléssel,

$$\mathbf{P}_{\mu_0} \left(\mu_0 \in \left(\bar{X}_n - \frac{t_{\varepsilon/2} Z_n}{\sqrt{n}}, \bar{X}_n + \frac{t_{\varepsilon/2} Z_n}{\sqrt{n}} \right) \right) = \mathbf{P}_{\mu_0}(|t| \leq t_{\varepsilon/2}) = 1 - \varepsilon.$$

t-próba. Ezek alapján a következőképpen járunk el:

1. A minta alapján kiszámoljuk az t próbastatisztikát (lásd (14)).
2. Rögzített $\varepsilon > 0$ szignifikanciaszinthez meghatározzuk $t_{\varepsilon/2}$ értéket.
3. Ha $|t| \leq t_{\varepsilon/2}$ akkor elfogadjuk H_0 -t, különben elvetjük.

A t -próba azért különösen fontos, mert nagy mintaelemszám esetén akkor is elvégezhető, ha a normalitás nem igaz. Ez a centrális határeloszlás-tétel és a nagy számok törvényéből következik. Ekkor persze lényegében u -próbát végzünk, és az ismeretlen σ helyére beírjuk a becslését.

14.3. Kétmintás u - és t -próba

A kétmintás u -, és t -próba az egymintás esethez hasonlóan tárgyalható, úgy, ahogy azt a konfidenciaintervallum megkonstruálásánál láttuk.

Ekkor van egy $X_1, \dots, X_n \sim N(\mu_1, \sigma^2)$, és egy ettől független $Y_1, \dots, Y_m \sim N(\mu_2, \sigma^2)$ normális eloszlású mintánk. A várható értékeket nem ismerjük, a szórás az u -próbánál ismert, a t -próbánál nem. Fontos, hogy a szórásnégyzetek egyenlőek a két mintában.

A nullhipotézis, hogy a két minta várható értéke ugyanaz, azaz

$$H_0 : \mu_1 = \mu_2 \quad \text{vs.} \quad H_1 : \mu_1 \neq \mu_2. \quad (15)$$

A próbastatisztika ismert σ esetén

$$u_2 = \frac{\bar{X}_n - \bar{Y}_m}{\sigma \sqrt{n^{-1} + m^{-1}}}.$$

Láttuk, hogy H_0 fennállása esetén u_2 standard normális eloszlású. Ezért rögzített $\varepsilon > 0$ szignifikanciaszint esetén a kritikus érték $u_\varepsilon = \Phi(1 - \varepsilon/2)$. Tehát a nullhipotézist akkor fogadjuk el, ha $|u_2| \leq u_\varepsilon$, különben elvetjük.

Most is valójában konfidenciaintervallumot szerkesztettünk. Hasonlóan tesztelhetjük a $\mu_1 - \mu_2 = \delta$ nullhipotézis is.

Ismeretlen σ esetén a szórást is becsljük. Láttuk, hogy

$$D_{n,m} = \frac{1}{n+m-2} \left[\sum_{i=1}^n (X_i - \bar{X}_n)^2 + \sum_{j=1}^m (Y_j - \bar{Y}_m)^2 \right]$$

torzítatlan becslése a szórásnégyzetnek. A σ helyére az u -statisztikába ezt írjuk be és kapjuk a

$$t_2 = \frac{\bar{X}_n - \bar{Y}_m}{\sqrt{D_{n,m}} \sqrt{n^{-1} + m^{-1}}}$$

statisztikát. Láttuk, hogy a nullhipotézis fennállása esetén t_2 Student-eloszlást követ $n+m-2$ szabadsági fokkal. Ekkor innen számolunk kritikus értéket, ami $t_\varepsilon = \Phi_{n+m-2}^{-1}(1 - \varepsilon/2)$. Akkor fogadjuk el a nullhipotézist, ha $|t| \leq t_\varepsilon$, különben elvetjük.