

VERJETNOST - ZAPISKI

1 Osnove verjetnosti

1.1 Izidi, dogodki, verjetnosti

Zgled 1.1. Za zgled vzemimo igro na srečo s tremi kockami, pri kateri igralci stavijo na vsoto pik na kockah. V 17. stoletju, ko je bila ta igra popularna, so kockarji pogosto stavili na vsoto 9 ali 10. Kockarji so namreč imeli teorijo, po kateri sta bili ti stavi enakovredni – napisali so vse trojice v nepadajočem vrstnem redu, ki se seštejejo v 9 oziroma 10, in ugotovili da sta števili trojic v obeh primerih enaki. Vendar pa so iz igralniške prakse vedeli, da ti stavi nista ekvivalentni. To razhajanje je pojasnil Galileo Galilej, ki je ugotovil, da je trojic, ki se seštejejo v 9, 25, in trojic, ki se seštejejo v 10, 27.

Vsako razmišljanje v verjetnosti se začne z množico vseh možnih izidov. Označimo jo tipično z grško črko Ω .

Zgled 1.2.

- Pri Galileju je $\Omega = \{(i, j, k) : 1 \leq i, j, k \leq 6\} = \{1, 2, 3, 4, 5, 6\}^3$.
- Če kovanec vržemo stokrat, je $\Omega = \{G, Š\}^{100}$. Splošneje, če kovanec vržemo n -krat, potem je $\Omega = \{G, Š\}^n$.
- Recimo, da mešamo 52 kart. Na koncu dobimo karte v nekem vrstnem redu – to pomeni, da so karte permutirane. V tem primeru je $\Omega = S_{52}$. Bolj splošno je $\Omega = S_n$.
- Kaj pa, če kovanec vržemo neskončnokrat? V tem primeru je $\Omega = \{G, Š\}^{\mathbb{N}}$.
- Recimo, da imamo posodo rdečih in modrih kroglic. Enih je R , drugih pa B . Izberemo podmnožico velikosti $n < B + R$ kroglic. Tukaj je $\Omega = C_{B+R}^n$.

V vsakdanjem jeziku govorimo o dogodkih in njihovih verjetnostih. V matematičnem smislu bodo dogodki podmnožice množice Ω . Dogodke bomo označili z velikimi tiskanimi črkami z začetka abecede, kot recimo A, B, A_1, B_1, \dots

Definicija 1.3. Družina \mathcal{F} dogodkov (podmnožic Ω) ima naslednje lastnosti:

- $\Omega \in \mathcal{F}$ in $\emptyset \in \mathcal{F}$.
- Če je $A \in \mathcal{F}$, potem je $A^c \in \mathcal{F}$.
- Če so $A_1, A_2, \dots \in \mathcal{F}$, je $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Opomba. Družini množic z zgornjimi lastnostmi pravimo σ -algebra.

Vsakemu dogodku dodelimo verjetnost, kar pomeni, da mu dodelimo število na $[0, 1]$. Za dodeljevanje zahtevamo določene lastnosti, ki jim rečemo aksiomi Kolmogorova.

Aksiom 1.1. Verjetnost je funkcija $P : \mathcal{F} \rightarrow [0, 1]$. Pri tem velja:

- $p(\Omega) = 1$ in $P(\emptyset)$
- Če so dogodki A_1, A_2, \dots disjunktni, je $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Opomba. Drugi lastnosti iz zgornjega aksioma rečemo števna aditivnost. Ta lastnost velja tudi za končne unije disjunktnih dogodkov.

Opomba. Če privzamemo, da je Ω končna in so vsi izidi enako verjetni, je $P(A) = \frac{|A|}{|\Omega|}$. V končnih primerih ponavadi naredimo to predpostavko in dodelimo verjetnost posameznim izidom.

Zgled 1.4. Oglejmo si primer iz genetike: naj bo $\Omega = S_n$. Permutaciji δ dodelimo verjetnost $P(\delta) = \frac{\theta^k}{(\theta)_n}$, kjer je k število ciklov permutacije δ , $(\theta)_n$ pa Pochhammerjev simbol. V primeru $\theta = 1$ dobimo kar običajno verjetnostno funkcijo. Hitro se da dokazati, da je res $\sum_{\delta \in S_n} \frac{\theta^k}{(\theta)_n} = 1$.

Navedimo nekaj preprostih posledic aksiomov Kolmogorova.

Posledica 1.5. Ker sta si A in A^c disjunktni in velja $A \cup A^c = \Omega$, imamo $P(A^c) = 1 - P(A)$.

Posledica 1.6. Naj bosta A in B disjunktna. Potem lahko zapišemo

$$A \cup B = (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B)$$

kot unijo treh disjunktnih dogodkov. Če razpišemo, od tod dobimo

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Formulo posplošimo:

$$\begin{aligned} P(A \cup B \cup C) &= P((A \cup B) \cup C) \\ &= P(A \cup B) + P(C) - P((A \cup B) \cap C) \\ &= P(A) + P(B) + P(C) - P(A \cap B) \\ &\quad - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

Izrek 1.7 (Formula za vključitve in izključitve).

Naj bodo A_1, \dots, A_n dogodki. Potem velja

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) - \sum_{1 \leq k < l \leq n} P(A_k \cap A_l) + \dots + (-1)^{n-1} P(A_1 \cap \dots \cap A_n).$$

Dokaz. Formula velja za $n = 2$. Privzemimo induksijsko predpostavko, da formula velja za poljubne

unije n dogodkov. Računamo:

$$\begin{aligned}
P\left(\bigcup_{k=1}^{n+1} A_k\right) &= P\left(\bigcup_{k=1}^n A_k \cup A_{n+1}\right) \\
&= P\left(\bigcup_{k=1}^n A_k\right) + P(A_{n+1}) - P\left(\bigcup_{k=1}^n A_k \cap A_{n+1}\right) \\
&= \sum_{k=1}^n P(A_k) - \sum_{1 \leq k < l \leq n} P(A_k \cap A_l) + \cdots + (-1)^{n-1} P(A_1 \cap \cdots \cap A_n) + P(A_{n+1}) \\
&\quad - \sum_{k=1}^n P(A_k \cap A_{n+1}) + \sum_{1 \leq k < l \leq n} P(A_k \cap A_l \cap A_{n+1}) - \cdots + (-1)^n P(A_1 \cap \cdots \cap A_n \cap A_{n+1}) \\
&= \sum_{k=1}^{n+1} P(A_k) - \sum_{1 \leq k < l \leq n+1} P(A_k \cap A_l) + \cdots + (-1)^n P(A_1 \cap \cdots \cap A_n). \quad \square
\end{aligned}$$

Zgled 1.8. Denimo, da n parov gre na ples. Ko odhajajo, zmanjka toka in v temi vsaka ženska naključno izbere moškega. Kolikšna je verjetnost, da nobena ne bo izbrala svojega partnerje? Če pare oštevilčimo, si izid lahko predstavljamo kot naključno permutacijo. Če ženska k izbere svojega partnerja, je k negibna točka permutacije. Za $k = 1, 2, \dots, n$ definiramo A_k kot množico vseh permutacij s fiksno točko k , z A pa označimo množico permutacij brez negibne točke. Velja $A^c = \bigcup_{k=1}^n A_k$. Računamo

$$\begin{aligned}
P(A^c) &= P\left(\bigcup_{k=1}^n A_k\right) \\
&= \sum_{k=1}^n P(A_k) - \sum_{1 \leq k < l \leq n} P(A_k \cap A_l) + \cdots \\
&= n \cdot \frac{1}{n} - \binom{n}{2} \frac{(n-2)!}{n!} + \binom{n}{3} \frac{(n-3)!}{n!} - \cdots + (-1)^{n-1} \frac{1}{n!} \\
&= 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1} \frac{1}{n!}
\end{aligned}$$

Od tod sledi

$$\begin{aligned}
P(A) &= 1 - P(A^c) \\
&= \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} \\
&\approx e^{-1}.
\end{aligned}$$

Izrek 1.9.

1. Naj bodo $A_1 \subseteq A_2 \subseteq \dots$ naraščajoči dogodki. Potem velja $P(\bigcup_{k=1}^{\infty} A_k) = \lim_{n \rightarrow \infty} P(A_n)$.
2. Naj bodo $A_1 \supseteq A_2 \supseteq \dots$ padajoči dogodki. Potem velja $P(\bigcap_{k=1}^{\infty} A_k) = \lim_{n \rightarrow \infty} P(A_n)$.

Dokaz. Dovolj je, če dokažemo le prvo točko, saj druga točka sledi po De Morganovih formulah.

Najprej opazimo, da je

$$\bigcup_{k=1}^{\infty} A_k = A_1 \cup \bigcup_{k=1}^{\infty} (A_2 \setminus A_1),$$

kjer je izraz na levi unija disjunktih dogodkov. Po aksiomih Kolmogorova je

$$\begin{aligned} P\left(\bigcup_{k=1}^{\infty} A_k\right) &= P(A_1) + \sum_{k=2}^{\infty} P(A_k \setminus A_{k-1}) \\ &= P(A_1) + \sum_{k=2}^{\infty} (P(A_k) - P(A_{k-1})) \\ &= \lim_{n \rightarrow \infty} \left(P(A_1) + \sum_{k=2}^n (P(A_k) - P(A_{k-1})) \right) \\ &= \lim_{n \rightarrow \infty} P(A_n). \end{aligned}$$

□

Izrek 1.10 (Borel-Cantellijeva lema).

Naj bodo A_1, A_2, \dots dogodki. Naj bo

$$A = \{\omega \in \Omega : \omega \text{ je vsebovan v neskončno mnogo } A_k\} = \bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k \right).$$

Če je $\sum_{k=1}^{\infty} P(A_k) < \infty$, je $P(A) = 0$.

Dokaz. Unije $\bigcup_{k=n}^{\infty} A_k$ so padajoči dogodki. Po izreku 1.9 je $P(A) = \lim_{n \rightarrow \infty} P\left(\bigcup_{k=n}^{\infty} A_k\right)$. Pri dokazu tega izreka smo tudi dobili oceno

$$\begin{aligned} P\left(\bigcup_{k=n}^{\infty} A_k\right) &= P(A_n) + \sum_{k=n+1}^{\infty} P(A_k \setminus A_{k-1}) \\ &\leq P(A_n) + \sum_{k=n+1}^{\infty} P(A_k) \\ &= \sum_{k=n}^{\infty} P(A_k). \end{aligned}$$

Sedaj pa imamo

$$\lim_{n \rightarrow \infty} P\left(\bigcup_{k=n}^{\infty} A_k\right) \leq \lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} P(A_k) = 0.$$

□

1.2 Pogojne verjetnosti

Vrnimo se h Galilejevemu primeru, ko je $\Omega = \{1, 2, 3, 4, 5, 6\}^3$. Denimo, da na prvi kocki padejo štiri pike. V zgornjem zgledu smo gledali dogodek $A = \{\text{vsota je } 9\}$ in ugotovili, da je $P(A) = \frac{25}{216}$. Označimo $B = \{\text{prva kocka kaže } 4\}$ in potem $P(B) = \frac{36}{216}$. Trojice, ki se začnejo s 4 in imajo vsoto 9, so $(4, 1, 4), (4, 2, 3), (4, 3, 2), (4, 4, 1)$. Rečemo, da je $\frac{4}{36}$ pogojna verjetnost dogodka A pri pogoju B . To verjetnost označimo s $P(A \mid B)$. Razpišemo:

$$P(A \mid B) = \frac{4}{36} = \frac{\frac{4}{216}}{\frac{36}{216}} = \frac{P(A \cap B)}{P(B)}.$$

Definicija 1.11. Naj bo B dogodek s $P(B) > 0$. Pogojna verjetnost dogodka A glede na dogodek B je definirana kot $P(A | B) = \frac{P(A \cap B)}{P(B)}$.

Pogojno verjetnost dobimo tako, da množico izidov zoožamo na B in nato verjetnosti normiramo s tem, da delimo s $P(B)$. Vse verjetnosti so premo sorazmerne z začetnimi.

Zgled 1.12. Delimo karte z dobro premešanega kupa. Naj bo $A = \{\text{druga karta je as}\}$ in $B = \{\text{prva karta je as}\}$. Uganemo lahko, da je verjetnost $P(A | B) = \frac{3}{51}$. Formalno pa to izračunamo tako, da zapišemo $P(B) = \frac{4}{52}$ in $P(A \cap B) = \frac{4}{52} \cdot \frac{3}{51}$ in nato po definiciji sledi želen rezultat. Tipično je, da je pogojne verjetnosti lažje „uganiti“ kot pa izračunati s formulami.

Zgled 1.13. V družini z dvema otrokoma so možnosti $\{MM, M\check{Z}, \check{Z}M, \check{Z}\check{Z}\}$. Privzemimo, da so vsi izidi enako verjetni. Naj bo sedaj $A = \{\text{oba otroka sta moškega spola}\}$ in $B = \{\text{prvi otrok je moškega spola}\}$. Po definiciji je

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}.$$

Definicija 1.14. Števen nabor dogodkov $\{H_1, H_2, \dots\}$ je particija Ω , če je $\bigcup_i H_i = \Omega$ in $H_i \cap H_j = \emptyset$ za vse $i \neq j$.

Izrek 1.15 (Formula za popolno verjetnost).

Naj bo $\{H_1, H_2, \dots\}$ particija in A dogodek. Velja

$$P(A) = \sum_j P(A | H_i) \cdot P(H_i).$$

Dokaz.

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P(A \cap \bigcup_i H_i) \\ &= P(\bigcup_i (A \cap H_i)) \\ &= \sum_i P(A \cap H_i) \cdot \frac{P(H_i)}{P(H_i)} \\ &= \sum_i P(A | H_i) P(H_i). \end{aligned}$$

□

Opomba. Pri dokazu smo privzeli, da je $P(H_i) > 0$ za vsak i .

Posledica 1.16.

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

Zgled 1.17. Na internetu lahko najdemo naslednjo igro na srečo: imamo 12 ploščic, od katerih so štiri 1, dve 2, ena 3, ena D (double) in štiri S (stop). Kartice se nato obrnejo, da ne vidimo vrednosti, in naključno permutirajo med sabo. Igralec nato obrača ploščice od leve proti desni, dokler ne naleti na ploščico S . Izplačilo igre je vsota števk, ki jo še pomnožimo z 2, če vidimo D .

Kolikšna je verjetnost, da bomo videli D ? Definiramo particijo z množicami

$$H_k = \{\text{prvi } S \text{ se pojavi na } k\text{-tem mestu}\},$$

kjer k teče od 1 do 9. Takoj lahko vidimo, da je $P(A_1) = \frac{4}{12}$, $P(H_2) = \frac{8}{12} \cdot \frac{4}{11}$ in $P(H_3) = \frac{8}{12} \cdot \frac{7}{11} \cdot \frac{4}{10}$. V splošnem lahko torej prepoznamo vzorec kot

$$P(H_k) = \frac{8 \cdot 7 \cdots (8 - k + 2) \cdot 4}{12 \cdot 11 \cdots (12 - k + 1)} = \frac{\binom{8}{k+1}}{\binom{12}{k+1}} \cdot \frac{4}{12 - k + 1}.$$

Izračunajmo še

$$P(A \mid H_k) = \frac{\binom{7}{k-2}}{\binom{8}{k-1}} = \frac{k-1}{8}.$$

Formula za popolno verjetnost nam da:

$$\begin{aligned} p(A) &= \sum_{k=1}^9 P(A \mid H_k) \cdot P(H_k) \\ &= \sum_{k=1}^9 \frac{\binom{8}{k+1}}{\binom{12}{k+1}} \frac{4}{12 - k + 1} \frac{k-1}{8} \\ &= \cdots = \frac{1}{5}. \end{aligned}$$

Do rezultata lahko pridemo tudi z hitrim razmislekom: vsaka permutacija kartic S in D je enako verjetna, zato je $\frac{1}{5}$ možnosti, da D pade prva.

Zgled 1.18. V posodi imamo m belih in n rdečih kroglic. Izbiramo kroglice naključno brez vračanja, dokler ne izberemo kroglice, ki je različne barve kot prva ali dokler ne izberemo zadnje kroglice. Postopek ponavljamo, dokler res ne izberemo zadnje kroglice. Kakšna je verjetnost, da je bela? Izkáže se, da je odgovor vedno kar ena polovica, kar dokažemo z indukcijo. Definiramo particijo $H_i = \{\text{izberemo } i \text{ belih in rdečo}\}$ za $i = 1, 2, \dots, m$ in $K_j = \{\text{izberemo } j \text{ rdečih in belo}\}$ za $j = 1, 2, \dots, n$. Indukcijska predpostavka je, da če je število kroglic v posodi strogo manjše od $m+n$, trditev velja. Definirajmo dogodek $A = \{\text{zadnja kroglica je rdeča}\}$. Sedaj imamo po indukcijski predpostavki pogojne verjetnosti $P(A \mid H_i) = \frac{1}{2}$ za $i < m$ in $P(A \mid H_m) = 0$ in podobno $P(A \mid K_j) = \frac{1}{2}$ za $j < n$ in $P(A \mid K_n) = 1$. Po formuli za popolno verjetnost je

$$\begin{aligned} P(A) &= \sum_{i=1}^m P(A \mid H_i)P(H_i) + \sum_{j=1}^n P(A \mid K_j)P(K_j) \\ &= \frac{1}{2} \cdot \sum_{i=1}^{m-1} P(H_i) + \frac{1}{2} \cdot \sum_{j=0}^{n-1} P(K_j) + P(K_n) \\ &= \frac{1}{2} \cdot \sum_{i=1}^m P(H_i) + \frac{1}{2} \cdot \sum_{j=0}^n P(K_j) + \frac{1}{2}P(K_n) - \frac{1}{2}P(H_m) \\ &= \frac{1}{2} + \frac{1}{2}P(K_n) - \frac{1}{2}P(H_m). \end{aligned}$$

Potem imamo

$$P(H_m) = \frac{m}{n+1} \cdot \frac{m-1}{m+n-1} \cdots \frac{1}{n+1} \cdot \frac{m}{n} = \frac{m!n!}{(m+n)!}$$

in če zamenjamo vlogi a in b , dobimo $P(K_n) = \frac{m!n!}{(m+n)!}$. S tem je indukcijski korak zaključen.

Zgled 1.19 (Jetniški paradoks). *Jetniki A, B, C so v mračni deželi obsojeni na smrt. Na državni praznik vladar naključno izbere enega od njih in ga pomilosti. Stražar že ve, kateri jetnik je srečnejši. Poglejmo si dve izjavi.*

A: *Stražar, če mi poveš, kdo od ostalih dveh bo usmrčen, mi ne daš nobene informacije, saj vem, da eden od njiju bo.*

Stražar: *Če ti povem, ostaneta samo dva. Potem je tvoja verjetnost preživetja $\frac{1}{2}$.*

Kdo ima prav? Stražarja moramo vključiti v množico izidov:

$$\Omega = \{(A, B \mid B), (A, C \mid C), (B, C \mid B), (B, C \mid C)\}.$$

Iz besedila naloge vemo, da je verjetnost prvega izida $\frac{1}{3}$. Enako seveda velja tudi za drugi dogodek, torej imata tretji in četrti dogodek skupaj verjetnost $\frac{1}{3}$. Vendar pa iz besedila naloge ne izhaja, kako bo ta verjetnost med dogodkoma porazdeljena. Recimo torej, da je verjetnost tretjega $\frac{x}{3}$ in četrtega $\frac{1-x}{3}$ za nek $x \in [0, 1]$. Označimo $A = \{A \text{ pomiloščen}\}$ in $B = \{\text{stražar reče } B\}$. Potem imamo

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{x}{1+x}.$$

Kot pogojne verjetnosti pridejo v poštev vse med 0 in $\frac{1}{2}$. Tako ima A prav, če je $x = \frac{1}{2}$ in stražar prav, če je $x = 1$. Sicer pa noben izmed njiju.

Če je $P(A \mid B) = P(A)$, potem zožitev na B ne spremeni verjetnosti. Rečemo lahko, da sta A in B neodvisna. Kdaj bi potem rekli, da so trije dogodki A, B, C neodvisni? Domnevamo, da morajo biti tudi v parih neodvisni, torej A je neodvisen od B in $A \cap B$ neodvisen od C . Enakost $P(A \mid B) = P(A)$ lahko prepišemo v $P(A \cap B) = P(A) \cdot P(B)$.

Definicija 1.20.

1. Dogodka A in B sta neodvisna, če velja $P(A \cap B) = P(A) \cdot P(B)$.

2. Dogodki A_1, A_2, \dots, A_n so neodvisni, če velja

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m}) = P(A_{i_1}) \dots P(A_{i_m})$$

za vse nabore $1 \leq i_1 < \dots < i_m \leq n$.

3. Dogodki $\{A_i\}_{i \in J}$ so neodvisni, če so neodvisni dogodki v vsaki končni poddružini dogodkov.

Opomba. Neodvisnost včasih izhaja iz same podelitve verjetnosti dogodkov, pogosto pa je privzetek.

Zgled 1.21 (Paradoks Viteza de Méréja). *Imamo dve igri na srečo.*

1. *Kocko vržemo štirikrat. Zmagamo, če vsaj enkrat pride šestica.*

2. *Vržemo dve kocki štiriindvajsetkrat. Zmagamo, če vsaj enkrat pride dvojna šestica.*

Rešimo najprej prvo. Definiramo dogodke $A = \{\text{zmagamo}\}$ in $A_i = \{\text{v } i\text{-tem metu ne dobimo šestice}\}$. Ker so meti kocke ločeni, so A_1, A_2, A_3, A_4 neodvisni.

$$\begin{aligned} P(A^c) &= P(A_1 \cap A_2 \cap A_3 \cap A_4) \\ &= P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot P(A_4) = \left(\frac{5}{6}\right)^4. \end{aligned}$$

Od tod sledi, da je $P(A) = 1 - P(A^c) = 1 - \left(\frac{5}{6}\right)^4 \doteq 0.5177$. Podobno dobimo seveda tudi pri drugi igri: tam bi za verjetnost zmage dobili verjetnost $P(A) = 1 - \left(\frac{35}{36}\right)^{24} \doteq 0.49$

Zgled 1.22 (Problem rojstnih dni). Denimo, da imamo m posameznikov in vsi si „izberejo“ tojstni dan v letu neodvisno en od drugega. Kolikšna je verjetnost, da nobena dva (ali več) posameznika nimata rojstnega dne na isti dan? Posplošimo in recimo, da ima leto n dni in je $m < n$. Definiramo dogodek

$$A_{i,j} = \{\text{oseba } i \text{ ima rojstni dan na dan } j\}$$

za $i = 1, \dots, m$ in $j = 1, \dots, n$. S tem dobimo matriko dogodkov:

$$\begin{array}{cccc} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ A_{2,1} & A_{2,2} & \dots & A_{2,n} \\ \vdots & & & \\ A_{m,1} & A_{m,2} & \dots & A_{m,n} \end{array}$$

Privzetek je ta, da so dogodki $A_{1,i_1}, A_{2,i_2}, \dots, A_{m,i_m}$ neodvisni za katerokoli n -terico i_1, i_2, \dots, i_m . Zanima nas dogodek

$$A = \bigcup_{i_1, \dots, i_m \text{ različni}} \bigcap A_{1,i_1} \cap A_{2,i_2} \cap \dots \cap A_{m,i_m}.$$

Verjetnost preseka je $P(A_{1,i_1}) \dots P(A_{m,i_m}) = \frac{1}{n^m}$. Torej je

$$P(A) = n(n-1) \dots (n-m+1) \cdot \frac{1}{n^m} = \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{m-1}{n}\right).$$

Izrek 1.23 (Druga Borel-Cantellijeva lema).

Naj bodo dogodki A_1, A_2, \dots neodvisni in naj velja $\sum_{i=1}^{\infty} P(A_i) = \infty$. Definiramo

$$A = \{\omega : \omega \text{ je vsebovan v neskončno mnogo } A_i\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Potem je $P(A) = 1$.

Dokaz. Po De Morganovih pravilih je $A^c = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c$. Dokazati moramo, da imajo vsi preseki verjetnost 0. Če so A_1, A_2, \dots neodvisni, so neodvisni komplementi. Vedno je

$$P\left(\bigcap_{k=n}^{\infty} A_k^c\right) \leq P\left(\bigcap_{k=n}^N A_k^c\right) = \prod_{k=n}^N P(A_k^c) = \prod_{k=n}^N (1 - P(A_k)).$$

Iz analize 1 vemo, da je $e^{-x} \geq 1 - x$, torej je

$$\prod_{k=n}^N (1 - P(A_k)) \leq \prod_{k=n}^N e^{-P(A_k)} = e^{-\sum_{k=n}^N P(A_k)} \rightarrow 0 \text{ kot } N \rightarrow \infty.$$

□

Zgled 1.24 (Kockarjev bankrot). Kockarja A in B začneta z m oziroma n zlatnikov. V vsaki rundi igre vržeta kovanec. Če je grb, da B igralcu A zlatnik, sicer pa da A igralcu B zlatnik. Mati kovanca so neodvisni in verjetnost grba je p . Igra se konča, ko eden od igralcev ostane brez denarja. Kolikšna je verjetnost, da bo bankrotiral B ? Definirajmo $A = \{\text{bankrotira } B\}$ in

označimo verjetnost s $p_{m,n}$. Sedaj vpeljimo še $H = \{v \text{ prvi igri zmaga } A\}$ in $q = 1 - p$.

$$\begin{aligned} P(A) &= P(A \mid H)P(H) + P(A \mid H^c)P(H^c) \\ &= P(A \mid H) \cdot p + P(A \mid H^c) \cdot q \\ &= p_{m+1,n-1} \cdot p + p_{m-1,n+1} \cdot p \end{aligned}$$

Igra se po prvem metu resetira z drugačnimi začetnimi pogoji. Označimo $\pi_k := p_{k,m+n-k} = p_{N-k}$, kjer je $N := m + n$. Eačbo prepíšemo v

$$(p + q)\pi_k = \pi_{k+1} \cdot p + \pi_{k-1} \cdot q.$$

To je linearna diferenčna enačba, ki jo znamo rešiti. To enačbo lahko preuredimo v

$$p(\pi_k - \pi_{k+1}) = q(\pi_{k-1} - \pi_k)$$

in zato $\pi_{k+1} - \pi_k = \frac{q}{p}(\pi_k - \pi_{k-1})$, iz besedila pa sledi $\pi_0 = 0$.

$$\begin{aligned} 1 &= \pi_N = (\pi_N - \pi_{N-1}) + (\pi_{N-1} - \pi_{N-2}) + \cdots + (\pi_1 - \pi_0) + \pi_0 \\ &= \pi_1 \left(\left(\frac{q}{p} \right)^{N-1} + \left(\frac{q}{p} \right)^{N-2} + \cdots + 1 \right) \end{aligned}$$

Sledi $\pi_1 = \frac{1}{\left(\frac{q}{p} \right)^{N-1} + \left(\frac{q}{p} \right)^{N-2} + \cdots + 1}$ in zapišemo

$$\begin{aligned} \pi_k &= (\pi_k - \pi_{k-1}) + \cdots + (\pi_1 - \pi_0) + \pi_0 \\ &= \left(\frac{q}{p} \right)^{k-1} + \cdots + \pi_1 \\ &= \left(\left(\frac{q}{p} \right)^{k-1} + \cdots + 1 \right) \pi_1. \end{aligned}$$

Končni rezultat je

$$\pi_k = p_{k,N-k} = \frac{\left(\frac{q}{p} \right)^{k-1} + \cdots + 1}{\left(\frac{q}{p} \right)^{N-1} + \cdots + 1}.$$

Če je kovanec pošten, torej je $p = q = \frac{1}{2}$, dobimo $p_{m,n} = \frac{m}{n+m}$. Kako pa vemo, da se kockar-jev bankrot sploh konča? V tem primeru je $\Omega = \{G, C\}^{\mathbb{N}}$. Neodvisnost pomeni, da so dogodki $A_i = \{r\text{-ti met je grb}\}$ neodvisni. Na vprašanje, ali lahko definiramo ustrezno družino (σ -algebro) dogodkov in jim dodelimo verjetnost v skladu z aksiomi Kolmogorova, odgovarja Carathéodori-jev izrek iz teorije mere. Predstavljajmo si zaporedje metov kovanca, kjer razdelimo mete v bloke po N metov:

$$B_i = \{v \text{ } i\text{-tem bloku dobimo } N \text{ grbov zapovrstjo}\}.$$

Dogodki B_i so neodvisni in imajo vsi enako verjetnost. Izrek 1.23 pravi, da se bo z verjetnostjo 1 zgodilo neskončno B_i -jev. To pomeni, da se igra zagotovo ustavi, torej ima dogodek

$$\{\text{množica vseh } \omega \in \Omega, \text{ ki ne vodijo do konca igre}\}$$

verjetnost 0.

Definicija 1.25. Druina dogodkov P je π -sistem, če je za $A, B \in P$ tudi $A \cap B \in P$.

Opomba. Če je P π -sistem, mu lahko dodamo \emptyset in Ω in bo še vedno π -sistem.

Izrek 1.26.

Naj bo družina $\{B_1, B_2, \dots, B_n\}$ π -sistem. Naj bo A dogodek, ki je neodvisen od vsakega $B \in \{B_1, \dots, B_n\}$. Potem je A neodvisen od vsakega dogodka, ki ga lahko sestavimo iz družine dogodkov $\{B_1, \dots, B_n\}$ s komplementi, preseki in unijami.

- Opomba.*
1. To je elementarna verzija Dynkinove leme.
 2. V izreku se lahko omejimo na končne unije in preseke.
 3. Vsak π -sistem je zaprt za končne preseke.

Dokaz. Vsak dogodek, ki ga sestavimo iz dogodkov $\{B_1, B_2, \dots, B_n\}$, je dsijunktna unija dogodkov oblike $B_1^* \cap B_2^* \cap \dots \cap B_n^*$, kjer je $*$ $\in \{\cdot, c\}$. Dovolj je dokazati, da je A neodvisen od dogodkov te oblike, ker je potem tudi neodvisna od disjunktne unij teh dogodkov. Ker je π -sistem zaprt za preseke, se dogodki oblike $B_1^* \cap B_2^* \cap \dots \cap B_n^*$ prevedejo v dogodke tipa $B_1^c \cap \dots \cap B_{n+1}^c$ (namesto indeksov $1, 2, \dots, m$ lahko vzamemo poljuben nabor m indeksov), neodvisnost pa oreverimo po definiciji.

$$\begin{aligned}
P\left(A \cap \bigcap_{i=1}^m B_i^c \cap B_{m+1}\right) &= P\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^c \cap B_{m+1}\right) \\
&= P(A \cap B_{m+1}) - P\left(\bigcup_{i=1}^m B_i \cap B_{m+1} \cap A\right) \\
&= P(A)P(B_{m+1}) - P\left(\bigcup_{i=1}^m (B_i \cap B_{m+1} \cap A)\right) \\
&= P(A)P(B_{m+1}) - \sum_{i=1}^m P(A \cap B_i \cap B_{m+1}) \\
&\quad + \sum_{1 \leq i < j \leq m} P(B_i \cap B_j \cap B_{m+1} \cap A) - \dots \\
&= P(A)P(B_{m+1}) - \sum_{i=1}^m P(A)P(B_i \cap B_{m+1}) \\
&\quad + \sum_{1 \leq i < j \leq m} P(B_i \cap B_j \cap B_{m+1})P(A) - \dots \\
&= P(A)\left(P(B_{m+1}) - P\left(\bigcup_{i=1}^{\infty} (B_i \cap B_{m+1})\right)\right) \\
&= P(A)\left(P(B_{m+1}) - P\left(\left(\bigcup_{i=1}^{\infty} B_i\right) \cap B_{m+1}\right)\right) \\
&= P(A)P(B_1^c \cap B_2^c \cap \dots \cap B_m^c \cap B_{m+1}). \quad \square
\end{aligned}$$

Zapišimo še posledico.

Izrek 1.27.

Naj bosta $\{A_1, A_2, \dots, A_m\}$ in $\{B_1, B_2, \dots, B_n\}$ π -sistema dogodkov. Vsak $A \in \{A_1, \dots, A_m\}$ naj bo neodvisen od vsakega $B \in \{B_1, \dots, B_n\}$. Potem sta neodvisna vsaka dogodka, ki jim lahko sestavimo s komplementiranjem, končnimi unijami ali preseki iz dogodkov v prvem oziroma drugem π -sistemu.

Dokaz. Dvakrat uporabimo prejšnji izrek.

□

2 Slučajne spremenljivke

2.1 Diskretne slučajne spremenljivke

Za motivacijo se vrnimo h Galileju in primeru meta dveh kock.

Zgled 2.1. Množica izidov je $\Omega = \{1, 2, 3, 4, 5, 6\}^3$. Kockarjev ni zanimal notranji ustroj izida ω , temveč število pik na vseh treh kockah. Z metom treh kock in seštevanjem pik nastane „slučajno število“. Možne vrednosti so $k = 3, 4, \dots, 18$ in za vsak možen k se lahko vprašamo, kolikšna je možnost, da bo vsota pik enaka k . Vsota vseh teh verjetnosti je 1, torej celoto razdelimo med vse možne k . Rečemo, da te verjetnosti določajo porazdelitev.

V matematičnem jeziku bomo slučajnim številom rekli slučajne spremenljivke in jih označili z velikimi tiskanimi črkami s konca abecede: X, Y, Z, \dots

Definicija 2.2. Slučajna spremenljivka X je funkcija $X : \Omega \rightarrow \mathbb{R}$, tako da je $X^{-1}((a, b])$ dogodek v \mathcal{F} za vsak interval $(a, b]$.

Opomba. 1. Iz definicije sledi, da je $X^{-1}(\{x\})$ dogodek za vsak $x \in \mathbb{R}$.

2. „Slučajno število“ nastane tako, da naključno izberemo $\omega \in \Omega$ in ga vstavimo v funkcijo X .

Definicija 2.3. Slučajna spremenljivka X je diskretna, če je zaloga vrednosti diskretna (končna ali števna) množica.

Definicija 2.4. Porazdelitev slučajne spremenljivke X je dana z naborom verjetnosti $P(X^{-1}((a, b]))$ za vse $a < b$.

Opomba. Za diskretno spremenljivko X z zalogo vrednosti $\{X_1, X_2, \dots\}$ je porazdelitev dana z verjetnostmi $P(X^{-1}(\{x_k\}))$. Za diskretne X je ta definicija ekvivalentna splošni.

Zaradi lastnosti praslík so dogodki tudi praslike $X^{-1}(A)$, kjer A sestavimo iz intervalov $(a, b]$ s števnimi unijami, preseki in komplementiranjem. Takim množicam rečemo Borelove množice.

Opomba. Za diskretne spremenljivke bomo namesto $P(X^{-1}(\{x\}))$ pisali $P(X = k)$.

Hiper-geometrijska porazdelitev

Zgled 2.5. V posodi imamo B belih in R rdečih kroglic. Označimo $N = B + R$. Naključno izberemo podmnožico $n \leq N$ kroglic. Vse podmnožice velikosti n so enako verjetne. Naj bo X enako številu belih kroglic med izbranimi. Očitno je X diskretna slučajna spremenljivka. Možne vrednosti so $\max(0, n - R) \leq k \leq \min(B, n)$. Kakšna je porazdelitev X ? Z računom dobimo

$$P(X = k) = \frac{\binom{B}{k} \binom{R}{n-k}}{\binom{N}{n}}.$$

S tem smo navedli porazdelitev slučajne spremenljivke X . Rečemo, da ima X hiper-geometrijsko porazdelitev s parametri n, B, N . To na kratko zapišemo $X \sim \text{HiperGeom}(n, B, N)$. Porazdelitev X ponazorimo s histogramom.

Za hipergeometrijsko porazdelitev izračunamo $\frac{P(X=k)}{P(X=k-1)} > 1$ natanko tedaj, ko je $k < \frac{(B+1)(n+1)}{N+2}$. Če $\frac{(B+1)(n+1)}{N+2}$ ni celo število, potem je stolpec pri $\left\lfloor \frac{(B+1)(n+1)}{N+2} \right\rfloor$ največji in stolpci pred njim naraščajo, pod njim pa padajo. Če pa $k = \frac{(B+1)(n+1)}{N+2}$ je celo število, potem si stolpca $k-1$ in k delita prvo mesto po višini. Stolpci pred njima naraščajo, za njima pa padajo.

Zgled 2.6 (Loterija Slovenije). Na loterijskem listku si izmed 39 izberemo m števil, pri čemer je $8 \leq m \leq 17$. Na žrebanju izmed devetintridesetih števil naključno izberejo 7 števil. Dobitek (recimo sedmica) je odvisen od pravilno izbranih števil, število pravilno izbranih izbranih števil pa je slučajna spremenljivka

$$X \sim \text{HiperGeom}(7, m, 39).$$

Najbolj nas seveda zanima $P(X=7)$.

Binomska porazdelitev

Zgled 2.7. Kovanec mečemo n -krat, kjer so meti neodvisni. Verjetnost grba je $p \in [0, 1]$. Označimo z X število grbov v n metih. Možne vrednosti za X so $0 \leq k \leq n$. Ker je $X : \Omega \rightarrow \{0, 1, 2, \dots, n\}$, je porazdelitev dana z verjetnostmi $P(X=k)$. Ker ima vsako zaporedje s k grbi verjetnost $p^k(1-p)^{n-k}$ in je takih zaporedij $\binom{n}{k}$, sledi

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n.$$

Tedaj rečemo, da ima X binomsko porazdelitev s parametroma n, k in označimo $X \sim \text{Bin}(n, p)$.

Za $k \geq 1$ velja

$$\frac{P(X=k)}{P(X=k-1)} = \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{k-1} (1-p)^{n-k+1}} = \frac{(n-k+1) \cdot p}{k \cdot (1-p)}.$$

Ta kvocient je večji od 1 natanko tedaj, ko je $(n+1)p > k$. Sedaj ponovno ločimo dva primera; če $(n+1)p$ ni celo število, je najvišji stolpec $\lfloor (n+1)p \rfloor$, sicer pa sta najvišja stolpca $(n+1)p-1$ in $(n+1)p$.

Opomba. V zgornjem zgledu neodvisnost pomeni, da zaporedju elementov $\Omega = \{G, C\}$ dodelimo verjetnost

$$p^{\#\text{grbov}} \cdot (1-p)^{\#\text{cifera}}.$$

S tem smo dogodkom in izidom na Ω dodelili verjetnost. Po definiciji lahko preverimo, da so za tako dodeljeno verjetnost meti kovancev res neodvisni.

Geometrijska porazdelitev

Zgled 2.8. Mečemo kovance in čakamo na prvi grb. Meti so neodvisni, verjetnost grba pa je ves čas enaka $p \in [0, 1]$. Oglejmo si

$$P(X=k) = P(\underbrace{CC \cdots C}_{k-1} G) = (1-p)^{k-1} p.$$

Rečemo, da ima X geometrijsko porazdelitev s parametrom p in označimo $X \sim \text{Geom}(p)$.

Opomba. V zgornjem zgledu je ustrezen $\Omega = \{G, C\}^{\mathbb{N}}$. Iz teorije mere izhaja, da lahko definiramo ustrezno družino dogodkov in jim dodelimo verjetnosti v skladu z aksiomi Kolmogorova.

Zgled 2.9. Recimo, da igralec na ruleto vedno stavi na 17. Privzemimo, da so igre neodvisne in je vseh 37 izidov enako verjetnih. Kolikšna je verjetnost, da med prvimi 37 igrami pade 17? V tem primeru je slučajna spremenljivka $X \sim \text{Geom}\left(\frac{1}{37}\right)$. Sprašujemo se torej po verjetnosti

$$\begin{aligned} P(X \leq 37) &= 1 - P(X > 37) = 1 - \sum_{k=38}^{\infty} \frac{1}{37} \left(1 - \frac{1}{37}\right)^{k-1} \\ &= 1 - \frac{1}{37} \cdot \left(1 - \frac{1}{37}\right)^{37} \frac{1}{1 - \left(1 - \frac{1}{37}\right)} \\ &= 1 - \left(1 - \frac{1}{37}\right)^{37} \doteq 0.63. \end{aligned}$$

Negativna binomska porazdelitev

Zgled 2.10. Namesto da čakamo na en sam grb, lahko čakamo na m grbov. Število potrebnih metov je torej $k \geq m$. Računamo:

$$\begin{aligned} P(X = k) &= P(\underbrace{** \cdots **}_{k-1 \text{ grbov}} G) \\ &= \binom{k-1}{m-1} p^{m-1} (1-p)^{(k-1)-(m-1)} p \\ &= \binom{k-1}{m-1} p^m (1-p)^{k-m}. \end{aligned}$$

Pri izračunu smo upoštevali, da je izid na k -tem mestu neodvisen od izbire od bloka metov do $k-1$ -tega mesta. To je posledica leme o π -sistemih. Rečemo, da ima X negativno binomsko porazdelitev s parametroma m, p in označimo $X \sim \text{NegBin}(m, p)$.

Tokrat bomo preverili, da je verjetnost neskončnega števila metov enaka nič. Privzemimo, da je $\sum_{k=m}^{\infty} P(X = k) = 1$. Analiza 1 nam pove, da je za $|x| < 1$

$$\begin{aligned} (1-x)^{-m} &= \sum_{k=0}^{\infty} \binom{-m}{k} (-x)^k \\ &= \sum_{k=0}^{\infty} \frac{(-m)(-m-1)\cdots(-m-k+1)}{k!} \cdot (-x)^k \\ &= \sum_{k=0}^{\infty} \frac{m(m+1)\cdots(m+k-1)}{k!} x^k \\ &= \sum_{k=0}^{\infty} \binom{m+k-1}{m-1} x^k \\ &= \sum_{l=m}^{\infty} \binom{l-1}{k-1} x^{l-m}, \end{aligned}$$

torej imamo

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

S tem pa smo dobili želen rezultat:

$$\sum_{l=m}^{\infty} P(X = m) = \sum_{l=m}^{\infty} \binom{l-1}{k-1} q^{l-m} p^m = 1.$$

Zgled 2.11 (Banachove cigarete). *Poljski matematik Stefan Banach (1892-1945) je bil strasten kadilec, zato je imel v obeh žepih plašča škatlico vžigalic z n vžigalicami v vsaki. Ko si je želel prižgati cigareto, je naključno segel v enega od žepov in iz škatlice vzel vžigalico. Če je vzel zadnjo vžigalico v škatlici, tega ni opazil. Nekoč bo Banach segel v žep in iz njega potegnil prazno škatlico. Ker je naključno segal v žep, bo v drugem žepu še X vžigalic, kjer lahko X zavzame vrednosti od 0 do n . Kakšna je porazdelitev X ? Dogodek $\{X = k\}$ je unija dveh disjunktnih dogodkov*

$$X_1 = \{X = k\} \cap \{\text{Banach je prazno škatlico potegnil iz levega žepa}\}$$

in

$$X_2 = \{X = k\} \cap \{\text{Banach je prazno škatlico potegnil iz desnega žepa}\}.$$

Če želimo, da se zgodi prvi od teh dveh dogodkov, mora Banach kaditi $(n + (n - k) + 1)$ -to cigareto in pri tem točno $(n + 1)$ -tič seči v levi žep. To spominja na negativno binomsko porazdelitev, zato lahko zapišemo

$$P(X_2) = \binom{(2n - k + 1) - 1}{(n + 1) - 1} \left(\frac{1}{2}\right)^{n+1} \left(\frac{1}{2}\right)^{n-k} = \binom{2n - k}{n} \left(\frac{1}{2}\right)^{2n-k+1}.$$

Zaradi simetrije je

$$P(X = k) = \binom{2n - k}{n} \left(\frac{1}{2}\right)^{2n-k}.$$

Poissonova porazdelitev

Kaj se zgodi z binomsko porazdelitvijo, če povečujemo število metov ($n \rightarrow \infty$) in hkrati zmanjšujemo verjetnost, da pade grb. Privzemimo, da bo vedno $n \cdot p = \lambda > 0$. Vemo, da je za fiksno k in $n \geq k$ verjetnost

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

Izračunajmo limito:

$$\begin{aligned} & \lim_{n \rightarrow \infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \frac{n(n-1) \cdots (n-k+1)}{n \cdot n \cdots n} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k} \\ &= \frac{\lambda^k}{k!} e^{-\lambda}. \end{aligned}$$

Za slučajno spremenljivko X z vrednostmi $k = 0, 1, \dots$, za katero je $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ za $\lambda > 0$. Rečemo, da ima X Poissonovo porazdelitev s parametrom λ , in označimo $X \sim \text{Po}(\lambda)$.

Definicija 2.12. Pochhammerjev simbol definiramo za vsak $z \in \mathbb{C}$ s predpisom $(z)_0 = 1$ in

$$(z)_n = z(z+1)(z+2) \cdots (z+n-1)$$

za $n \geq 1$.

Alternativno lahko zapišemo

$$\begin{aligned} \frac{\Gamma(z+n)}{\Gamma(z)} &= \frac{(z+n-1)\Gamma(z+n-1)}{\Gamma(z)} \\ &= \cdots = \frac{(z+n-1)(z+n-2) \cdots z\Gamma(z)}{\Gamma(z)} \\ &= (z)_n. \end{aligned}$$

Sedaj lahko za $|x| < 1$ izpeljemo tudi

$$(1-x)^{-a} = \sum_{k=0}^{\infty} \binom{-a}{k} (-1)^k x^k \quad (1)$$

$$= \sum_{k=0}^{\infty} \frac{(-a)(-a-1)\dots(-a-k+1)}{k!} (-1)^k x^k \quad (2)$$

$$= \sum_{k=0}^{\infty} \frac{a(a+1)\dots(a+k-1)}{k!} x^k \quad (3)$$

$$= \sum_{k=0}^{\infty} \frac{(x)_k}{k!} x^k. \quad (4)$$

Pólyeva porazdelitev

Če za slučajno spremenljivko z vrednostmi $k = 0, 1, \dots$ velja

$$P(X = k) = \frac{\beta^a (a)_k}{k! (1 + \beta)^{a+k}},$$

rečemo, da ima X Pólyevo porazdelitev s parametroma β in a . Predpostavimo $\beta, a > 0$. Vsota (4) nam pove, da je $\sum_{k=0}^{\infty} P(X = k) = 1$, kot bi si želeli.

Pochhammerjevi simboli imajo še naslednjo lastnost:

$$(a+b)_n = \sum_{k=0}^n \binom{n}{k} (a)_k (b)_{n-k}.$$

Vemo, da je $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$. Sedaj poračunamo:

$$\begin{aligned} & \sum_{k=0}^n \binom{n}{k} (a)_k (b)_{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} \frac{\Gamma(a+n)}{\Gamma(a)} \frac{\Gamma(b+n-k)}{\Gamma(b)} \frac{\Gamma(a+b+n)}{\Gamma(a+b+n)} \\ &= \sum_{k=0}^n \binom{n}{k} \frac{\Gamma(a+b+n)}{\Gamma(a)\Gamma(b)} B(a+k, b+n-k) \\ &= \frac{\Gamma(a+b+n)}{\Gamma(a)\Gamma(b)} \sum_{k=0}^n \binom{n}{k} \int_0^1 x^{a+k-1} (1-x)^{b+n-k-1} dx \\ &= \frac{\Gamma(a+b+n)}{\Gamma(a)\Gamma(b)} \int_0^1 x^{a-1} (1-x)^{b-1} \sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} dx \\ &= \frac{\Gamma(a+b+n)}{\Gamma(a)\Gamma(b)} \int_0^1 x^{a-1} (1-x)^{b-1} dx \\ &= \frac{\Gamma(a+b+n)}{\Gamma(a)\Gamma(b)} B(a, b) \\ &= \frac{\Gamma(a+b+n)}{\Gamma(a+b)} \\ &= (a+b)_n. \end{aligned}$$

Zgled 2.13. Imamo n volilcev in r predsedniških kandidatov. Volilci se odločajo neodvisno eden od drugega, vsakega kandidata pa izberejo z verjetnostjo $\frac{1}{r}$. Kolikšna je verjetnost, da bo eden od kandidatov zbral strogo več kot $\frac{n}{2}$ glasov? Označimo

$$X_k = \{\text{kandidat } k \text{ ima strogo več kot } \frac{n}{2} \text{ glasov}\}$$

in iščemo verjetnost $P(\bigcup_{k=1}^n X_k)$. Takoj lahko vidimo, da so dogodki X_i med seboj paroma disjunktni. Tukaj imamo seveda binomsko porazdelitev, zato lahko zaradi simetrije zapišemo

$$P\left(\bigcup_{k=1}^n X_k\right) = r \cdot \sum_{k > \frac{n}{2}} \binom{n}{k} \left(\frac{1}{r}\right)^k \left(1 - \frac{1}{r}\right)^{n-k}.$$

Če imamo dva kandidata, je ta verjetnost velika. Če pa število kandidatov in volilcev večamo, pa se verjetnost hitro zmanjša. V primeru $n = 100$ in $r = 3$ je ta verjetnost recimo 0.0005.

2.2 Zvezne slučajne spremenljivke

Do sedaj smo obravnavali slučajne spremenljivke, ki so imele večinoma celoštevilске vrednosti. Lahko pa si zamislimo tudi slučajne spremenljivke ali slučajna števila, ki lahko zavzamejo katerokoli vrednost na (a, b) ali celo \mathbb{R} . Primeri so recimo čas do radioaktivnega razpada, življenska doba in tako naprej.

Zgled 2.14. Recimo, da izbiramo naključno število na $[0, 1]$. Če izbiramo nepristransko, bi morale za poljuben $[a, b] \subseteq [0, 1]$ veljati $P(X \in [a, b]) = b - a$. Posledica je torej, da je $P(X = x) = 0$ za vsak $x \in [0, 1]$. Porazdelitve torej ne moremo opisati z verjetnostmi oblike $P(X = x)$.

Lahko si predstavljamo naključna števila, ki zavzamejo vrednosti na nekem intervalu, ki je lahko tudi neskončen. Formalno bomo govorili o slučajnih spremenljivkah kot o funkcijah $f : \Omega \rightarrow \mathbb{R}$. Porazdelitev slučajne spremenljivke je dana z naborom verjetnosti $P(a < X \leq b)$ za vse $a < b$.

Definicija 2.15. Rečemo, da ima slučajna spremenljivka X zvezno porazdelitev, če obstaja neneativna funkcija $f_X(x)$, za katero velja $\int_{-\infty}^{\infty} f_X(x) dx = 1$ in za vsaka $a < b$:

$$P(a < X \leq b) = \int_a^b f_X(x) dx.$$

Funkciji f_X rečemo gostota porazdelitve.

Opomba. • Izraz „zvezna porazdelitev“ izhaja iz teorije mere. Govorili bomo o zveznih slučajnih spremenljivkah, čeprav to nima zveze z zveznostjo X kot funkcije (na Ω sploh ni nujno definirana metrika ali topologija).

- Privzemamo, da je f_X na vsakem intervalu $[a, b]$, $[a, \infty)$ ali $(-\infty, b]$ (Riemannovo) integrabilna – mora v splošnem smislju.

- Velja tudi

$$P(x \in [a, b]) = \int_a^b f_X(x) dx,$$

saj je $P(X \in [a, a]) = 0$.

- Gostota ni povsem enolično določena. Če sta f_X in g_X gostoti iste spremenljivke, se razlikujeta kvečjemu na množici z mero 0.

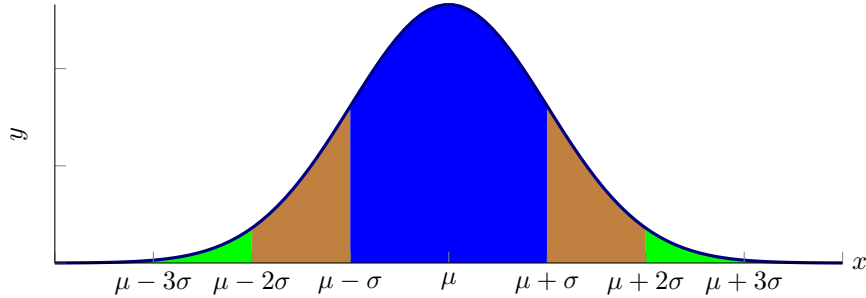
Oglejmo si nekaj standardnih tipov gostot oziroma porazdelitev.

Normalna porazdelitev

Rečemo, da ima X normalno porazdelitev s parametroma μ in σ^2 , če je

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

kjer je $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ ter $\sigma^2 > 0$. Normalno porazdelitev označimo z $X \sim N(\mu, \sigma^2)$. Iz analize 2a



Slika 1: Graf funkcije gostote normalne porazdelitve.

vemo, da je $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$. Z vpeljavo nove spremenljivke $u = \frac{x-\mu}{\sqrt{2}\sigma}$ se prepričamo, da je res $\int_{-\infty}^{\infty} f_X(x) dx = 1$. Iz grafa vidimo, da je μ simetrala gostote, σ pa oddaljenost od prevoja. Večji kot je σ , bolj je gostota nižja in sploščena.

Opomba. Ime normalna porazdelitev je uvedel statistik Adolphe Quetlet. Opazoval je porazdelitev karakteristik telesnih višin in podobnih stvari za velike populacije. Ugotovil je, da porazdelitve sledijo normalni.

Gama in eksponentna porazdelitev

Rečemo, da ima X gama porazdelitev s parametroma $a > 0$ in $\lambda > 0$, če je

$$f_X(x) = \begin{cases} \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}; & x > 0 \\ 0; & \text{sicer} \end{cases}.$$

Parameter a določa obliko gostote, λ pa je stvar izbire enot. Pri tem je

$$\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx$$

običajna Γ funkcija. Hitro preverimo, da je f_X nenegativna in se zintegrira v 1. Uporabljamo oznako $X \sim \Gamma(a, \lambda)$. Posebna vrsta gama porazdelitve je eksponentna porazdelitev (s parametrom λ), ki je definirana z gostoto

$$f_X(x) = \begin{cases} \lambda \cdot e^{-\lambda x}; & x \geq 0 \\ 0; & \text{sicer} \end{cases}$$

in jo označujemo $X \sim \exp(\lambda)$. Eksponentna porazdelitev je pogosto model za življensko dobo.

Enakomerna in beta porazdelitev

Pravimo, da ima X enakomerno porazdelitev na (a, b) , če velja

$$f_X(x) = \begin{cases} \frac{1}{b-a}; & x \in (a, b) \\ 0; & \text{sicer} \end{cases}.$$

Rečemo, da je $X \sim U(a, b)$. Beta porazdelitev pa je definirana z gostoto

$$f_X(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}; & 0 < x < 1 \\ 0; & \text{sicer} \end{cases}.$$

Pri tem sta $a, b > 0$ konstanti. Po definiciji beta funkcije faktor $\frac{1}{B(a,b)}$ povzroči, da velja $\int_0^1 f_X(x) dx = 1$. Označimo $X \sim \text{Beta}(a, b)$.

Definicija 2.16. Porazdelitvena funkcija slučajne spremenljivke X je

$$F_X(x) = P(X \leq x) = P(X \in (-\infty, x])$$

za $x \in \mathbb{R}$.

Po definiciji je

$$\begin{aligned} P(x \in (a, b]) &= P(X \in (-\infty, b] \setminus (-\infty, a]) \\ &= P(x \in (-\infty, b]) - P(x \in (-\infty, a]) \\ &= F_X(b) - F_X(a). \end{aligned}$$

Če torej poznamo funkcijo F_X , imamo enolično določeno porazdelitev X . Če je X zvezno porazdeljena, je

$$F_X(x) = \int_{-\infty}^x f_X(x) dx.$$

Izrek 2.17.

Naj bo F_X porazdelitvena funkcija slučajne spremenljivke X .

1. F_X je nepadajoča.
2. $\lim_{x \rightarrow \infty} F_X(x) = 1$ in $\lim_{x \rightarrow -\infty} F_X(x) = 0$.
3. F_X je desno zvezna.

Dokaz. 1. Za vsaka $x \leq y$ je $F_X(y) - F_X(x) = P(X \in (x, y]) \geq 0$.

2. Definiramo $A_n = \{X \leq n\}$. Potem velja $A_1 \subseteq A_2 \subseteq \dots$ in $\bigcup_{n=1}^{\infty} A_n = \Omega$. Po lemi iz prvega poglavja je

$$1 = P(\Omega) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} F_X(n) = \lim_{x \rightarrow \infty} F_X(x).$$

Simetričen argument in sklic na lemo naredimo tudi za $\lim_{x \rightarrow -\infty} F_X(x)$.

3. Naj gre $x_n \downarrow x$. Dokazati moramo, da velja $F_X(x_n) \downarrow F_X(x)$. Definirajmo $A_n = \{X \leq x_n\}$. Velja $A_1 \supseteq A_2 \supseteq \dots$ in $\bigcap_{n=1}^{\infty} A_n = \{X \leq x\}$. Po isti lemi kot prej sledi, da je

$$\lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = P(X \leq x) = F_X(x). \quad \square$$

Če je X slučajna spremenljivka in f „razumna“ funkcija, recimo zvezna, potem je $Y = f(X)$ spet zvezna spremenljivka. Gledali bomo samo primere, ko je X zvezno porazdeljena.

Opomba. Če je $F_X(x) = \int_{-\infty}^x g(u) du$ in je g v točki x zvezna, je F_X v točki x odvedljiva in velja $F'_X(x) = g(x)$.

Zgled 2.18. Naj bo $X \sim N(\mu, \sigma^2)$ in $Y = aX + b$, kjer je $a > 0$. Računamo

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(aX + b \leq y) \\ &= P\left(X \leq \frac{y-b}{a}\right) \\ &= F_X\left(\frac{y-b}{a}\right). \end{aligned}$$

Analiza 2a zagotavlja odvedljivost F_X , njen odvod je kar f_X . Vemo:

$$f_X(x) = F'_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Potem je

$$f_Y(y) = F'_X\left(\frac{y-b}{a}\right) \frac{1}{a} = \frac{1}{\sqrt{2\pi}a\sigma} e^{-\frac{(y-b-a\mu)^2}{2a^2\sigma^2}}.$$

Opazimo, da velja $Y \sim N(a\mu + b, a^2\sigma^2)$. Poseben primer: če je $X \sim N(\mu, \sigma^2)$, je $Y := \frac{X-\mu}{\sigma} \sim N(0, 1)$. Res, definirajmo porazdelitveno funkcijo porazdelitve $N(0, 1)$:

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du.$$

To lahko preverimo tudi na naslednji način:

$$\begin{aligned} P(Y \leq y) &= P\left(\frac{X-\mu}{\sigma} \leq y\right) \\ &= P(X - \mu \leq \sigma y) \\ &= P(X \leq \sigma y + \mu) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\sigma y + \mu} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &\stackrel{u=\frac{x-\mu}{\sigma}}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{u^2}{2}} du \\ &= \Phi(y) \end{aligned}$$

in zaradi enoličnosti je res $Y \sim N(0, 1)$. Zapišimo še

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right), \end{aligned}$$

torej lahko porazdelitveno funkcijo vsake normalne porazdelitve zapišemo s Φ .

Zgled 2.19. Naj bo $Z \sim N(0, 1)$ in $Y = Z^2$. Računamo za $y > 0$:

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(Z^2 \leq y) \\ &= P(-\sqrt{y} \leq Z \leq \sqrt{y}) \\ &= \Phi(\sqrt{y}) - \Phi(-\sqrt{y}). \end{aligned}$$

Ponovno nam analiza 2a za $y > 0$ zagotavlja odvedljivost.

$$\begin{aligned} F'_Y(y) &= \Phi'(\sqrt{y}) \frac{1}{2\sqrt{y}} + \Phi'(-\sqrt{y}) \frac{1}{2\sqrt{y}} \\ &= \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y}{2}} \cdot \frac{1}{2\sqrt{y}} + \frac{1}{2\sqrt{2\pi y}} \cdot e^{-\frac{y}{2}} \\ &= \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}. \end{aligned}$$

Gauss nam da $Y \sim \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$.

Definicija 2.20. Naj bo X slučajna spremenljivka in $p \in (0, 1)$. Vsakemu številu x_p , za katerega velja $P(X \leq x_p) = p$, rečemo p -ti kvintil porazdelitve X .

Opomba. • Število x_p ne obstaja nujno za vse p , posebej za diskretne porazdelitve. Za zvezne deluje, ni pa nujno enolično določen.

- Za $p = \frac{1}{2}$ govorimo o mediani, za $p = \frac{1}{4}$ ali $p = \frac{3}{4}$ pa govorimo o prvem ali tretjem kvartilu.

3 Slučajni vektorji

3.1 Diskretni slučajni vektorji

Zgled 3.1. Imamo r škatel in vanje mečemo n kroglic. Meti so neodvisni, verjetnost, da zadamo škatlo k , pa je p_k za $k = 1, 2, \dots, r$. Pri tem je $p_1 + p_2 + \dots + p_r = 1$. Števila kroglic v škatlah so slučajne spremenljivke X_1, X_2, \dots, X_r . Te slučajne spremenljivke „kolektivno“ zavzamejo vrednosti k_1, k_2, \dots, k_r , za katere je $k_i \geq 0$ za $i = 1, 2, \dots, r$ in $\sum_{i=1}^r k_i = n$. Izračunajmo verjetnost dogodka

$$A = \{X_1 = k_1\} \cap \{X_2 = k_2\} \cap \dots \cap \{X_r = k_r\}$$

za možno r -terico (k_1, k_2, \dots, k_r) . Pišemo

$$P(A) := P(X_1 = k_1, X_2 = k_2, \dots, X_r = k_r).$$

Tukaj je $\Omega = \{1, 2, \dots, r\}^n$. Posameznemu izidu ω podelimo verjetnost

$$P(\{\omega\}) = p_1^{\# \text{enk v } \omega} p_2^{\# \text{dvojk v } \omega} \dots p_r^{\# \text{r-jev v } \omega}.$$

Dokažemo lahko, da so meti neodvisni in se verjetnosti vseh izidov seštejejo v 1. Dogodek A se zgodi, če izberemo ω , ki ima k_1 enk, k_2 dvojk, in tako naprej do k_r r-jev. Vsak tak ω ima verjetnost $p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}$. Prešteti moramo, koliko je $\omega \in \Omega$ s predpisanimi števili posameznih i -jev. Iz kombinatorike pa vemo, da je iskano število $\frac{n!}{k_1! k_2! \dots k_r!}$, saj iščemo permutacije multimnožice. Potem je

$$P(A) = \frac{n!}{k_1! k_2! \dots k_r!} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}.$$

Iz multinomskega izreka sledi, da je vsota teh verjetnosti po vseh možnih naborih enaka 1.

Komponente X_1, X_2, \dots, X_r strnemo v vektor $\underline{X} = (X_1, X_2, \dots, X_r)$ in govorimo o slučajnem vektorju. Rekli bomo, da ima X iz prejšnjega zgleda multinomsko porazdelitev s parametroma n in $\underline{p} = (p_1, p_2, \dots, p_r)$. To označimo z $X \sim \text{Multinom}(n, \underline{p})$.

Definicija 3.2. Slučajni vektor \underline{X} je funkcija $X : \Omega \rightarrow \mathbb{R}^n$, taka da je $X^{-1}(U)$ dogodek za vsako odprto množico $U \subseteq \mathbb{R}^n$.

Opomba. Odprte množice si izberemo iz praktičnih razlogov. Lahko bi izbrali zaprte ali samo pravokotnike oblike $\prod_{i=1}^r (a_i, b_i]$ in dobili isto definicijo. Od tod izhaja, da je $X^{-1}(A)$ dogodek za vsak $A \subseteq \mathbb{R}^n$, ki ga lahko sestavimo iz odprtih množic s komplementiranjem, števnimi preseki in števnimi unijami, torej Borelovih množic.

Definicija 3.3. Slučajni vektor je diskreten, če je zaloga vrednosti X diskretna množica $\{\underline{x}_1, \underline{x}_2, \dots\}$.

Opomba. Za diskretne slučajne spremenljivke privzamemo, da je $\underline{X}^{-1}(\{\underline{x}_k\})$ dogodek za vsak \underline{x}_k . To je enakovredno splošni definiciji.

Definicija 3.4. Porazdelitev \underline{X} poznamo, če poznamo $P(\underline{X}^{-1}(U))$ za vsako odprto podmnožico $U \subseteq \mathbb{R}^n$.

Opomba. Ta definicija je začasna – prava je v teoriji mere.

Zgled 3.5. Izbiramo slučajno permutacijo π . Vse permutacije so enako verjetne, dolžine permutacija pa je n . Definirajmo x_i kot število ciklov dolžine i za $i = 1, 2, \dots, n$. Oglejmo si slučajni vektor (x_1, x_2, \dots, x_n) . Možne vrednosti so n -terice (k_1, k_2, \dots, k_n) za $k_i \geq 0$ in $\sum_{i=1}^n i k_i = n$. Izračunati moramo

$$P(X_1 = k_1, X_2 = k_2, \dots, X_n = k_n) = \frac{\# \text{ ugodnih permutacij}}{n!}.$$

Ideja je naslednja: denimo, da imamo preslikavo $f : A \rightarrow B$, ki je surjektivna. Recimo, da je za vsak b prasluka $f^{-1}(\{b\})$ enako velika. Potem velja $A = \bigcup_b f^{-1}(\{b\})$ in imamo $|A| = |B| |f^{-1}(\{b\})|$ za poljuben $b \in B$. V našem primeru bo B množica ugodnih permutacij, A pa množica vseh permutacij. Preslikavo f določimo na naslednji način: uredimo seznam oklepajev zelenih dolžin po velikosti, nato pa vtikamo števila $\sigma(i)$ po vrsti v oklepaje. Očitno je f surjektivna. Če vzamemo permutacijo s predpisanim številom ciklov τ , je

$$f^{-1}(\{\tau\}) = k_1! k_2! \dots k_n! \cdot 1^{k_1} \cdot 2^{k_2} \dots n^{k_n}.$$

Torej je število ugodnih permutacij enako

$$\frac{n!}{k_1! \dots k_n! \cdot 1^{k_1} \dots n^{k_n}}$$

in od tod sledi

$$P(X_1 = k_1, X_2 = k_2, \dots, X_n = k_n) = \frac{1}{k_1! \dots k_n! \cdot 1^{k_1} \dots n^{k_n}}$$

za $k_i \geq 0$ in $\sum_{i=1}^n i k_i = n$.

Opomba. Ta porazdelitev se imenuje Ewensova porazdelitev in se pogosto uporablja v matematični biologiji.

Kako iz večrazsežne porazdelitve slučajnega vektorja \underline{X} najdemo porazdelitve komponent? Oglejmo si primer dveh slučajnih spremenljivk. Poznamo $P(X = x, Y = y)$ za vsak možen par (x, y) . Označimo zalogo vrednosti parov (x, y) kot \mathcal{R} . Potem velja

$$\{X = x\} = \bigcup_{y \in (x, y) \in \mathcal{R}} \{X = x, Y = y\}$$

in od tod sledi

$$P(X = x) = \sum_{y: (x,y) \in \mathcal{R}} P(X = x, Y = y)$$

$$P(Y = y) = \sum_{x: (x,y) \in \mathcal{R}} P(X = x, Y = y).$$

Ti dve formuli lahko očitno posplošimo na

$$P(\underline{X} = \underline{x}) = \sum_{\underline{y}: (\underline{x}, \underline{y}) \in \mathcal{R}} P(\underline{X} = \underline{x}, \underline{Y} = \underline{y})$$

$$P(\underline{Y} = \underline{y}) = \sum_{\underline{x}: (\underline{x}, \underline{y}) \in \mathcal{R}} P(\underline{X} = \underline{x}, \underline{Y} = \underline{y}).$$

Zgled 3.6. Naj bo $\underline{X} \sim \text{Multinom}(n, \underline{p})$, torej je

$$P(X_1 = k_1, \dots, X_r = k_r) = \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r}$$

s $k_i \geq 0$ in $\sum_{i=1}^n k_i = n$. Izračunajmo $P(X = k_1)$. Iz prejšnje formule dobimo

$$\begin{aligned} P(X_1 = k_1) &= \sum_{k_2 + \dots + k_r = n - k_1} P(X_1 = k_1) \dots P(X_r = k_r) \\ &= \sum_{k_2 + \dots + k_r} \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r} \\ &= \binom{n}{k_1} p_1^{k_1} (1 - p_1)^{n - k_1} \sum_{k_2 + \dots + k_r} \frac{(n - k_1)!}{k_2! \dots k_r!} \left(\frac{p_2}{1 - p_1} \right)^{k_2} \dots \left(\frac{p_r}{1 - p_1} \right)^{k_r} \\ &= \binom{n}{k_1} p_1^{k_1} (1 - p_1)^{n - k_1}, \end{aligned}$$

kjer smo uporabili dejstvo

$$\text{Multinom} \left(n - k_1, \frac{(p_2, \dots, p_r)}{1 - p_1} \right) = 1.$$

S tem smo dobili $X_1 \sim \text{Bin}(n, p_1)$.

Dogodka A, B sta neodvisna, če velja $P(A \cap B) = P(A)P(B)$. Kaj bi bila definicija neodvisnosti za slučajni spremenljivki X, Y ? Ideja je, da če naj bi bili X, Y neodvisni, bi morala biti neodvisna tudi vsaka dogodka, „povezana“ z X oziroma Y . Vsi taki dogodki pa so oblike $\{X \in A\} = X^{-1}(A)$ oziroma $\{X \in B\} = X^{-1}(B)$.

Definicija 3.7.

1. Slučajni spremenljivki X, Y sta neodvisni, če je

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

za vsaki odprti množici A in B .

2. Slučajne spremenljivke X_1, X_2, \dots, X_n so neodvisne, če velja

$$P\left(\bigcap_{i=1}^{\infty} \{X_i \in A_i\}\right) = \prod_{i=1}^{\infty} P(X_i \in A_i)$$

za vsako izbiro odprtih množic A_1, A_2, \dots, A_n .

3. Slučajne spremenljivke v družini $\{X_i\}_{i \in I}$ so neodvisna, če so neodvisne spremenljivke v vsaki končni poddružini.

4. Slučajna vektorja \underline{X} in \underline{Y} sta neodvisna, če velja

$$P(\underline{X} \in A, \underline{Y} \in B) = P(\underline{X} \in A)P(\underline{Y} \in B)$$

za vsaki odprti A, B .

Opomba. 1. Izbira odprtih množic ni edina. Lahko bi med drugim izbrali zaprte množice ali pa kvadre, in dobili ekvivalentno definicijo. Če definicija velja za odprti A, B , velja tudi za vsaka A', B' , ki ju lahko sestavimo s komplementiranjem, števnimi unijami in preseki (torej Borelovimi množicami).

2. Če za navedene A_i izberemo \mathbb{R} , v naši definiciji dogodek $\{X_i \in A_i\}$ izgine na levi in desni strani. Dogodki $\{X_i \in A_i\}$ so neodvisni po definiciji iz prvega poglavja.

Če sta X, Y diskretni, lahko vzamemo $A = \{x\}$ in $B = \{y\}$ in velja

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

Velja pa tudi obratno. Spremenljivke X_1, X_2, \dots, X_n so neodvisne natanko tedaj, ko velja

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i)$$

za vse možne nabore (x_1, x_2, \dots, x_n) .

Opomba. Definicija je splošna za diskretne in zvezne spremenljivke X, Y . Za zvezne bomo pokazali, da je dovolj preveriti $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$. To velja tudi za diskretne slučajne spremenljivke.

Izrek 3.8.

Naj za diskretni slučajni spremenljivki X, Y velja, da je

$$P(X = x, Y = y) = f(x)g(y)$$

za vsaka $x \in \mathcal{R}(x)$ in $y \in \mathcal{R}(y)$ in funkciji $f : \mathcal{R}(x) \rightarrow [0, \infty)$ in $g : \mathcal{R}(y) \rightarrow [0, \infty)$. Potem sta X, Y neodvisni.

Dokaz. Po formuli za robne porazdelitve je

$$\begin{aligned} P(X = x) &= \sum_y P(X = x, Y = y) \\ &= \sum_y f(x)g(y) \\ &= f(x) \sum_y g(y) = c_1 f(x), \end{aligned}$$

kjer je $C_1 > 0$. Podobno je $P(Y = y) = c_2 g(y)$ za nek $c_2 > 0$. Tako dobimo

$$\begin{aligned} \sum_{(x,y) \in \mathcal{R}(x) \times \mathcal{R}(y)} P(X = x, Y = y) &= \frac{1}{c_1 c_2} \sum_{(x,y) \in \mathcal{R}(x) \times \mathcal{R}(y)} P(X = x) P(Y = y) \\ &= \frac{1}{c_1 c_2} \left(\sum_x P(X = x) \right) \cdot \left(\sum_y P(Y = y) \right) \\ &= \frac{1}{c_1 c_2}. \end{aligned}$$

Od tod sledi $c_1 c_2 = 1$ in imamo neodvisnost. \square

Dokaz velja popolnoma enako, če sta \underline{X} in \underline{Y} neodvisna vektorja.

Opomba. V prejšnjem koraku smo množili neskončni vsoti. Ali smo to smeli narediti? K temu vprašanju se bomo še vrnili.

Zgled 3.9. Denimo, da je v družini $N \sim \text{Po}(\lambda)$. Vsak otrok je fant z verjetnostjo $p = \frac{1}{2}$ neodvisno od ostalih. Lahko si mislimo, da gledamo neodvisne mete kovancev, kjer je N neodvisen od matov kovancev. Naj naključna spremenljivka X označuje število fantov med otroci, Y pa število punc. Potem imamo

$$\begin{aligned} P(X = k, Y = l) &= P(X = k, Y = l, N = k + l) \\ &= P(X = k, Y = l \mid N = k + l) P(N = k + l) \\ &= \binom{k+l}{k} \left(\frac{1}{2}\right)^{k+l} \frac{e^{-\lambda} \lambda^{k+l}}{(k+l)!} \\ &= \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!} \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^l}{l!} \\ &= f(k)g(l), \end{aligned}$$

zato sta spremenljivki X in Y neodvisni.

Zgled 3.10 (The Art of Computer Programming). Pri generatorju naravnih števil naletimo na naslednje vprašanje: eden od testov je test naraščajočih zaporedij. Generator slučajnih števil generira neodvisne enako porazdeljene slučajne spremenljivke ξ_1, ξ_2, \dots , enakomerno porazdeljene na $\{1, 2, \dots, m\}$. Definiramo:

Y_1 = dolžina segmenta strogo naraščajočih ξ_1, ξ_2, \dots
 Y_2 = zadnje število v naraščajočem zaporedju
 Y_3 = naslednje število
 Y_4 = drugo naslednje število.

Kako je z neodvisnostjo (Y_1, Y_2) in Y_3 , ali (Y_1, Y_2) in Y_4 ? Izračunajmo

$$P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3, Y_4 = y_4) = \binom{y_2}{y_1} \cdot \frac{1}{y_2!} \cdot \frac{y_2}{m} \cdot \frac{1}{m}.$$

za $y_1 \leq m$, $y_2 \geq y_1$ in $y_3 \leq y_2$. Iz porazdelitve (Y_1, Y_2, Y_3, Y_4) sledi

$$\begin{aligned} P(Y_1 = y_1, Y_2 = y_2, Y_4 = y_4) &= \sum_{y_3=1}^{y_2} \binom{y_2}{y_1} \frac{1}{y_2!} \frac{1}{m} \\ &= \binom{y_2}{y_1} \cdot \frac{1}{(y_2 - 1)!} \cdot \frac{1}{m} \\ &= f(y_1, y_2)g(y_4), \end{aligned}$$

torej sta ta dve vektorja res neodvisna.

Opomba. Za neodvisnost slučajnega vektorja \underline{X} in slučajne spremenljivke Y v diskretnem primeru zadošča preveriti $P(\underline{X} = \underline{x}, Y = y)$ za vse $(\underline{x}, y) \in \mathcal{R}(\underline{x}) \times \mathcal{R}(y)$.

3.2 Pričakovana vrednost

Zgled 3.11. Ogledali smo si že internetno igrico z 12 ploščicami, od katerih so 4 enice, 2 dvojki, 1 trojka, 1 karta D (double) in 4 karte S (stop). V igri se ploščice obrnejo in slučajno permutirajo, nato pa jih igralec obrača, dokler ne naleti na stop. Izplačilo je vsota števk, ki se pomnoži z 2, če je igralec dobil karto D . Izplačilo pa je tudi slučajna spremenljivka X , katere vrednosti so

$$\Omega = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 16, 18, 20, 22\}.$$

Koliko smo pripravljeni največ plačati za igranje te igre? Privzemimo, da igro ponavljamo v nedogled in dobimo izplačila x_1, x_2, \dots . Po velikem številu iger izračunamo povprečje izplačil

$$\frac{x_1 + x_2 + \dots + x_n}{n} = \sum_{k \in \Omega} k \frac{\# \text{ pojavitev v } n}{n} \approx \sum_{k \in \Omega} k \cdot P(X = k).$$

Definicija 3.12. Naj bo X diskretna slučajna spremenljivka z vrednostmi $\{x_1, x_2, \dots\}$. Pričakovana vrednost slučajne spremenljivke X je število

$$E(X) = \sum_{x_k} x_k \cdot P(X = x_k).$$

Rečemo, da $E(X)$ obstaja, če ta vrsta konvergira oziroma obstaja $\sum_{x_k} |x_k| \cdot P(X = x_k)$.

Opomba. 1. Vsote oblike $\sum_{i \in I} a_i$ za števne množice I , ki niso nujno urejene na naraven način, bomo obravnavali posebej. Primer je recimo $I = \mathbb{Z}^d$.

2. Definicija $E(X)$ je odvisna od porazdelitve X , ne od ponavljanj.

Vemo, da je $Y = f(X)$ tudi slučajna spremenljivka. Kako izračunamo $E(Y)$? Recimo, da ponavljamo X , torej dobimo zaporedje x_1, x_2, \dots . Potem dobimo tudi zaporedje $f(x_1), f(x_2), \dots$. Potem računamo $E(Y) = \frac{f(x_1) + \dots + f(x_n)}{n} = \sum_{x_k} f(x_k) P(X = x_k)$. Ta razmislek bomo kasneje formalno utemeljili.

Zgled 3.13. Naj bo $X \sim \text{Bin}(n, p)$, torej

$$\begin{aligned}
 E(X) &= \sum_{k=0}^n k \cdot P(X = k) \\
 &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \sum_{k=0}^n n \binom{n-1}{k-1} p \cdot p^{k-1} \cdot q^{(n-1)-(k-1)} \\
 &= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)} \\
 &= np.
 \end{aligned}$$

Sedaj poračunamo še

$$\begin{aligned}
 E(X^2) &= \sum_{k=0}^n k^2 P(X = k) \\
 &= \sum_{k=0}^n (k(k-1) + k) \binom{n}{k} p^k q^{n-k} \\
 &= \sum_{k=0}^n k(k-1) \binom{n}{k} p^k q^{n-k} + np \\
 &= \sum_{k=2}^n n(n-1) \binom{n-2}{k-2} p^2 p^{k-2} q^{(n-2)-(k-2)} + np \\
 &= n(n-1) p^2 \binom{n-2}{k-2} p^{k-2} q^{(n-2)-(k-2)} + np \\
 &= n(n-1) p^2 + np \\
 &= n^2 p^2 + npq.
 \end{aligned}$$

Zgled 3.14. Naj bo $X \sim \text{HiperGeom}(n, B, N)$. Interpretiramo $\binom{a}{b} = 0$, če je $b < 0$ ali $b > a$. Sedaj računamo

$$\begin{aligned}
 E(X) &= \sum_k k \cdot P(X = k) \\
 &= \sum_k k \cdot \frac{\binom{B}{k} \binom{R}{n-k}}{\binom{N}{n}} \\
 &= \sum_k \frac{B \binom{B-1}{n-1} \binom{R}{(n-1)-(n-k)}}{\frac{N}{n} \binom{N-1}{n-1}} \\
 &= \frac{n \cdot B}{N} \sum_k \frac{\binom{N-1}{k-1} \binom{R}{(n-1)-(k-1)}}{\binom{N-1}{n-1}} \\
 &= \frac{n \cdot B}{N}.
 \end{aligned}$$

Zgled 3.15. Izračunajmo še pričakovano vrednost naključne spremenljivke $X \sim \text{NegBin}(m, p)$.

$$\begin{aligned}
 E(X) &= \sum_{k=m}^{\infty} k \binom{k-1}{n-1} p^m q^{k-m} \\
 &= \sum_{k=m}^{\infty} \binom{(k+1)-1}{(n+1)-1} \frac{m}{p} p^{m-1} q^{(k+1)-(m+1)} \\
 &= \frac{m}{p} \sum_{k=m}^{\infty} \binom{(k+1)-1}{(m+1)-1} \cdot p^{m+1} q^{(k+1)-(m+1)} \\
 &= \frac{m}{p},
 \end{aligned}$$

saj je faktor na desni enak vsoti vseh verjetnosti porazdelitve $\text{NegBin}(m+1, p)$. Še kvadrat:

$$\begin{aligned}
 E(X^2) &= \sum_{k=m}^{\infty} k^2 \binom{k-1}{m-1} p^m q^{k-m} \\
 &= \sum_{k=m}^{\infty} (k(k+1) - k) \binom{k-1}{m-1} p^m q^{k-m} \\
 &= \sum_{k=m}^{\infty} k(k+1) \binom{k-1}{m-1} p^m q^{k-m} - \frac{m}{p} \\
 &= \sum_{k=m}^{\infty} \binom{(k+2)-1}{(m+2)-1} \frac{m(m+1)}{p^2} p^{m+2} q^{(k+2)-(m+2)} - \frac{m}{p} \\
 &= \frac{m(m+1)}{p^2} - \frac{m}{p} \\
 &= \frac{m^2}{p^2} + \frac{mq}{p^2}.
 \end{aligned}$$

Kaj pa, če imamo slučajni vektor \underline{X} , funkcijo $f: \mathbb{R}^n \rightarrow \mathbb{R}$ in je $Y = f(\underline{X})$. Potem po istem argumentu kot prej „uganemo“ formulo $E(f(\underline{X})) = \sum_{\underline{x}} f(\underline{x}) P(\underline{X} = \underline{x})$. Tudi to bomo kasneje formalno utemeljili.

Zgled 3.16. Naj bo $X \sim \text{Multinom}(n, \underline{p})$ in $f(\underline{x}) = x_1 \cdot x_2$. Potem dobimo

$$\begin{aligned}
 E(f(X)) &= E(X_1 \cdot X_2) \\
 &= \sum_{k_1 + \dots + k_r = n} k_1 \cdot k_2 \cdot P(X_1 = k_1, \dots, X_r = k_r) \\
 &= \sum_{k_1, k_2 \geq 1, k_1 + \dots + k_r = n} k_1 \cdot k_2 \cdot \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r} \\
 &= \sum_{k_1, k_2 \geq 1, k_1 + \dots + k_r = n} \frac{n(n-1)(n-2)!}{(k_1-1)!(k_2-2)!k_3! \dots k_r!} p_1^{k_1-1} p_2^{k_2-1} p_3^{k_3} \dots p_r^{k_r} p_1 p_2 \\
 &= n(n-1) \cdot p_1 \cdot p_2
 \end{aligned}$$

V verjetnosti moramo seštevati vsote oblike $\sum_{i \in I} a_i$, kjer je I števna množica. Tipično I ne bo urejena na kakšen naraven način (recimo \mathbb{Z}^d). Kako razumeti take vsote? Predpostavimo najprej, da je $a_i \geq 0$ za vse $i \in I$. V tem primeru je smiselno definirati

$$\sum_{i \in I} a_i = \sup \left\{ \sum_{i \in K} a_i : K \subseteq I, K \text{ končna} \right\}.$$

Rečemo, da vrsta obstaja, če je supremum določen.

Opomba. Za običajne vrste z nenegativnimi členi definiciji sovpadata.

Če členi niso negativni, pa definiramo $a^+ = \frac{|a_i|+a_i}{2}$ in $a^- = \frac{|a_i|-a_i}{2}$. Očitno sta $a_i^+ \geq 0$ in $a_i^- \geq 0$ ter velja $a_i^+ - a_i^- = a_i$.

Definicija 3.17. Rečemo, da vrsta $\sum_{i \in I} a_i$ obstaja, če obstajata $\sum_{i \in I} a_i^+$ in $\sum_{i \in I} a_i^-$. Teda j definiramo $\sum_{i \in I} a_i = \sum_{i \in I} a_i^+ - \sum_{i \in I} a_i^-$.

Opomba. Za vrste je ta definicija ekvivalentna absolutni konvergeni.

Napišimo nekaj lastnosti takih vrst.

Lema 3.18. Vrsta $\sum_{i \in I} a_i$ obstaja, če in samo če obstaja $\sum_{i \in I} |a_i|$.

Lema 3.19. Če obstajata $\sum_{i \in I} a_i$ in $\sum_{i \in I} b_i$, potem obstaja $\sum_{i \in I} (\alpha a_i + \beta b_i)$ in velja $\sum_{i \in I} (\alpha a_i + \beta b_i) = \alpha \sum_{i \in I} a_i + \beta \sum_{i \in I} b_i$.

Za nas bo pomembno naslednje vprašanje: če je $\{I_j : j \in J\}$ particija I , torej $\bigcup_{j \in J} I_j = I$ in $I_j \cap I_{j'} = \emptyset$ za $j \neq j'$, ali velja $\sum_{i \in I} a_i = \sum_{j \in J} \sum_{i \in I_j} a_i$?

Lema 3.20. Če so vsi členi a_i nenegativni, potem sta dve vsoti ali končni in enaki ali pa ne obstajata.

Lema 3.21. Vsota $\sum_{i \in I} a_i$ obstaja, če in samo če obstaja $\sum_{j \in J} \sum_{i \in I_j} |a_i|$. Če vsoti obstajata, velja $\sum_{i \in I} a_i = \sum_{j \in J} \sum_{i \in I_j} a_i$.

Zgled 3.22. Naj bosta I_1, I_2 števni. Seštevamo

$$\begin{aligned} \sum_{(i_1, i_2) \in I_1 \times I_2} a_i b_i &= \sum_{i_1 \in I_1} a_{i_1} \left(\sum_{i_2 \in I_2} b_{i_2} \right) \\ &= \left(\sum_{i_1 \in I_1} a_{i_1} \right) \left(\sum_{i_2 \in I_2} b_{i_2} \right). \end{aligned}$$

Zadnja lema nam pove, da $\sum_{(i_1, i_2) \in I_1 \times I_2} |a_{i_1}| |a_{i_2}|$ obstaja natanko tedaj, ko obstaja $\sum_{i_1 \in I_1} \left(\sum_{i_2 \in I_2} |a_{i_1} b_{i_2}| \right)$, ta vrsta pa obstaja natanko tedaj, ko obstajata $\sum_{i_1 \in I_1} |a_{i_1}|$ in $\sum_{i_2 \in I_2} |b_{i_2}|$. S tem pa smo utemeljili naš postopek.

Zgled 3.23. Naj bo \underline{X} slučajni vektor z vrednostmi $\{\underline{x}_1, \underline{x}_2, \dots\}$. Naj bo $Y = f(\underline{X})$, kjer je

$f : \mathbb{R}^n \rightarrow \mathbb{R}$. Naj bodo $\{y_1, y_2, \dots\}$ možne vrednosti. Po definiciji je

$$\begin{aligned} E(Y) &= \sum_y y \cdot P(Y = y) \\ &= \sum_y y \cdot \sum_{\underline{x} \in \mathcal{R}(\underline{X}), f(\underline{x})=y} P(\underline{X} = \underline{x}) \\ &= \sum_y \sum_{\underline{x} \in \mathcal{R}(\underline{X}), f(\underline{x})=y} f(\underline{x}) P(\underline{X} = \underline{x}) \\ &= \sum_{\underline{x} \in \mathcal{R}(\underline{x})} f(\underline{x}) P(\underline{X} = \underline{x}). \end{aligned}$$

Po zadnji lemi vsoti obstajata hkrati in sta enaki, če obstajata.

V zgledu smo dokazali vsoto

$$E(f(X)) = \sum_{\underline{x}} f(\underline{x}) P(\underline{X} = \underline{x}).$$

Poglejmo si posledico. Vzamemo $f(x, y) = x$ ali $f(x, y) = x + y$. Potem velja

$$\begin{aligned} E(\alpha X + \beta Y) &= \sum_{(x,y) \in \mathcal{R}(x,y)} (\alpha x + \beta y) P(X = x, Y = y) \\ &= \alpha \sum_{(x,y) \in \mathcal{R}(x,y)} x P(X = x, Y = y) + \beta \sum_{(x,y) \in \mathcal{R}(x,y)} y P(X = x, Y = y) \\ &= \alpha E(X) + \beta E(Y). \end{aligned}$$

Izrek 3.24.

Če obstajata $E(X)$ in $E(Y)$, potem obstaja tudi $E(\alpha X + \beta Y)$ in velja $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$.

Temu pravimo linearnost pričakovane vrednosti. Ta velja tudi v več spremenljivkah.

$$E\left(\sum_{k=1}^n \alpha_k X_k\right) = \sum_{k=1}^n \alpha_k E(X_k).$$

To dejstvo lahko že takoj uporabimo v metodi indikatorjev.

Definicija 3.25. Slučajna spremenljivka I je indikator ali Bernoullijeva slučajna spremenljivka, če zavzame le vrednosti v $\{0, 1\}$. Če označimo $p = P(I = 1)$, potem označimo $I \sim \text{Bernoulli}(p)$.

Formalno je $I : \Omega \rightarrow \{0, 1\}$. To pomeni, da je I indikator dogodka $I^{-1}(\{1\})$. Velja pa tudi obratno, vsakemu dogodku A pripada indikator 1_A , ki ga označimo z 1_A . Seveda velja $1_A \sim \text{Bernoulli}(P(A))$. Velja $E(I) = 0 \cdot P(I = 0) + 1 \cdot P(I = 1) = P(I = 1)$, zato sledi $E(1_A) = P(A)$. Ideja metode indikatorjev je, da slučajno spremenljivko X zapišemo kot linearno kombinacijo indikatorjev in uporabimo linearnost za izračun $E(X)$.

Zgled 3.26. Naj bo $X \sim \text{Bin}(n, p)$. Mislamo si lahko, da je $X : \{G, C\}^n \rightarrow \{0, 1, \dots, n\}$. Definiramo

$$I_k = \begin{cases} 1; & \text{če je } k\text{-ti met grb} \\ 0; & \text{sicer.} \end{cases}$$

Zapišimo $I = I_1 + I_2 + \dots + I_n$. Ker je na vsakem mestu verjetnost grba enaka p , je $I_k \sim$

Bernoulli(p) za $k = 1, 2, \dots, n$, torej je $E(I_k) = p$. Od tod sledi

$$E(X) = E(I_1 + I_2 + \dots + I_n) = E(I_1) + \dots + E(I_n) = np.$$

Zgled 3.27. Naj bo $X \sim \text{HiperGeom}(n, B, N)$. Mislamo si, da kroglice izbiramo naključno in brez vračanja eno po eno, dokler jih ne izberemo n . Zaradi simetrije so vsi izbori enako verjetni. Definiramo

$$I_k = \begin{cases} 1; & \text{če je } k\text{-ta kroglica bela} \\ 0; & \text{sicer} \end{cases}.$$

Velja $X = I_1 + I_2 + \dots + I_n$. Zaradi linearnosti je

$$E(X) = E(I_1 + \dots + I_n) = E(I_1) + \dots + E(I_n).$$

Seveda velja $E(I_1) = P(I_1 = 1) = \frac{B}{N}$. Zaradi simetrije bo k -ta izbrana kroglica naključno izbrana izmed vseh. To pomeni, da je $I_k \sim \text{Bernoulli}\left(\frac{B}{N}\right)$, zato je $E(I_k) = \frac{B}{N}$ in $E(X) = \frac{nB}{N}$.

Zgled 3.28. Oglejmo si spet internetno igrico z 12 ploščicami. Igralec obrača ploščica+e od leve proti desni, dokler ne naleti na kartico S , izplačilo pa je vsota števk (pomnoženo z dva, če zadanemo še kartico D). Naj bo X izplačilo. Ideja je, da če za vsako števko vemo prispevek h končnemu izplačilu, je X vsota teh prispevkov. Označimo kartice z indeksi od 1 do 12 (od 1 do 4 so enke, od 5 do 6 dvojke, itd.). Definiramo

$$\begin{aligned} A_{i,1} &= \{\text{vidimo } 1 \text{ in ne vidimo } D\}, & A_{i,2} &= \{\text{vidimo } 1 \text{ in } D\}, & i &= 1, \dots, 4 \\ B_{i,1} &= \{\text{vidimo } 2 \text{ in ne vidimo } D\}, & B_{i,2} &= \{\text{vidimo } 2 \text{ in } D\}, & i &= 5, 6 \\ C_1 &= \{\text{vidimo } 3 \text{ in ne vidimo } D\}, & C_2 &= \{\text{vidimo } 3 \text{ in } D\} \end{aligned}$$

Sedaj velja

$$\begin{aligned} X &= \sum_{i=1}^4 1_{A_{i,1}} + \sum_{i=1}^4 2 \cdot 1_{A_{i,2}} + \sum_{i=1}^2 2 \cdot 1_{B_{i,1}} \\ &\quad + \sum_{i=1}^2 4 \cdot 1_{B_{i,2}} + 3 \cdot 1_{C_1} + 6 \cdot 1_{C_2}. \end{aligned}$$

in od tod

$$\begin{aligned} E(X) &= \sum_{i=1}^4 E(1_{A_{i,1}}) + \sum_{i=1}^4 2 \cdot E(1_{A_{i,2}}) + \sum_{i=1}^2 2 \cdot E(1_{B_{i,1}}) \\ &\quad + \sum_{i=1}^2 4 \cdot E(1_{B_{i,2}}) + 3 \cdot E(1_{C_1}) + 6 \cdot E(1_{C_2}) \\ &= 11 \cdot E(1_{A_{i,1}}) + 22 \cdot E(1_{A_{i,2}}). \end{aligned}$$

Preostane nam še izračun $P(A_{i,1})$. Mislamo si, da ploščice obračamo do konca. Če odstranimo vse ploščice razen tiste, označene z i , D , in vse kartice S . S tem smo inducirali naključno permutacijo 6 ploščic. Zaradi simetrije so vse inducirane permutacije enako verjetne, zato imamo $P(A_{i,1}) = \frac{1}{6} \cdot \frac{4}{5}$ in $P(A_{i,2}) = \frac{1}{6} \cdot \frac{1}{5} + \frac{1}{6} \cdot \frac{1}{5}$. S tem smo dobili $E(X) = \frac{44}{15}$.

3.3 Zvezni slučajni vektorji

V eni dimenziji je ideja, da je $P(X \in (a, b]) = \int_a^b f_X(x) dx$. Kakšna je posplošitev na več dimenzij? Predstavljamo si, da naključno izbiramo točko v ravnini. Kako bi opisali porazdelitev te naključno izbrane točke v \mathbb{R}^2 .

Definicija 3.29. Porazdelitev slučajnega vektorja \underline{X} je dana z vsemi verjetnostmi $P(\underline{X} \in U)$ in odprte množice $U \subseteq \mathbb{R}^n$.

Opomba. 1. Ta definicija je splošna. Za diskretne slučajne vektorje je dovolj navesti $P(\underline{X} = x)$ za vse možne vrednosti.

2. Izbira odprtih množic je arbitrarna. Lahko bi izbrali zaprte, ali kvadre oblike $\prod_{k=1}^n (a_k, b_k]$ in bi dobili ekvivalentno definicijo. Iz odprtih množic lahko sestavimo s števnimi unijami, preseki in komplementi Borelove množice. Na koncu je prava definicija uvedba $P(\underline{X} \in U)$ za vse Borelove množice.

Definicija 3.30. Slučajni vektor \underline{X} ima zvezno porazdelitev, če obstaja funkcija $f_{\underline{X}}(\underline{x})$, tako da za vsako (razumno) množico $A \subseteq \mathbb{R}^n$ velja

$$P(\underline{X} \in A) = \int_A f_{\underline{X}}(\underline{x}) d\underline{x}.$$

Opomba. 1. Za razumne množice v zgornji definiciji bomo vzeli Jordanove množice.

2. Privzamemo, da je $f_{\underline{X}}$ Riemannovo integrabilna, mogoče v splošnem smislu, na vsaki Jordanovi množici.

Zgled 3.31. Ena od bivariantnih normalnih gostot je dana z

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left[-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)} \right].$$

Preverimo, da je ta funkcija res gostota.

$$\begin{aligned} \iint_{\mathbb{R}^2} f_{X,Y}(x, y) dx dy &= \frac{1}{2\pi\sqrt{1-\rho^2}} \iint_{\mathbb{R}^2} \exp \left[-\frac{(y-\rho x)^2}{2(1-\rho^2)} - \frac{x^2}{2} \right] dx dy \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} \exp \left[-\frac{(y-\rho x)^2}{2(1-\rho^2)} - \frac{x^2}{2} \right] dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \cdot \frac{1}{\sqrt{2\pi(1-\rho^2)}} \int_{-\infty}^{\infty} e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1, \end{aligned}$$

ker je to integral $N(0, 1)$ gostote. Izračunajmo še $P(X \geq 0, Y \geq 0)$. Po definiciji je

$$\begin{aligned} P(X \geq 0, Y \geq 0) &\geq \iint_{[0, \infty)^2} f_{X,Y}(x, y) dx dy \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_0^{\infty} e^{-\frac{x^2}{2}} dx \cdot \int_0^{\infty} e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}} dy. \end{aligned}$$

V notranjem integralu uvedemo $u = \frac{y-\rho x}{\sqrt{1-\rho^2}}$ in zato $du = \frac{dy}{\sqrt{1-\rho^2}}$.

$$\begin{aligned} &= \frac{1}{2\pi} \int_0^\infty e^{-\frac{x^2}{2}} dx \int_{-\rho x/\sqrt{1-\rho^2}}^\infty e^{-\frac{u^2}{2}} du \\ &= \frac{1}{2\pi} \iint_{x \geq 0, y \geq -\rho x/\sqrt{1-\rho^2}} e^{-\frac{x^2+y^2}{2}} dx dy. \end{aligned}$$

Integriramo rotacijsko simetrično funkcijo, ki se na \mathbb{R}^2 zintegrira v 1. Integral po področju med dvema premicama je sorazmeren kotu. V našem primeru je kot enak $\frac{\pi}{4} + \arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right)$. Od tod sledi:

$$\begin{aligned} &= \frac{1}{2\pi} \left(\frac{\pi}{2} + \arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right) \right) \\ &= \frac{1}{4} + \frac{1}{2\pi} \arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right). \end{aligned}$$

Opomba. Rezultat velja tudi za negativne ρ .

Zgled 3.32. Naj bo $\underline{\mu} \in \mathbb{R}^n$ in $\underline{\Sigma}(n \times n)$ pozitivno definitna matrika. Funkcija

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \underline{\Sigma}}} \exp\left(-\frac{1}{2}(\underline{x} - \underline{\mu})^\top \underline{\Sigma}^{-1}(\underline{x} - \underline{\mu})\right)$$

je večrazsežna normalna gostota. Pokažimo, da se zintegrira v 1. Če je $\underline{\Sigma}$ pozitivno definitna, je taka tudi $\underline{\Sigma}^{-1}$. Po Choleskem je $\underline{\Sigma}^{-1} = \underline{A}^\top \underline{A}$ za neko $n \times n$ matriko. Uvedemo novo spremenljivko $\underline{y} = \underline{A}(\underline{x} - \underline{\mu})$. Imamo $\underline{x} = \underline{A}^{-1}\underline{y} + \underline{\mu}$ in zato $d\underline{x} = \det(\underline{A}^{-1}) \cdot d\underline{y}$. Sedaj poračunamo:

$$\begin{aligned} \int_{\mathbb{R}^n} f_{\underline{X}}(\underline{x}) d\underline{x} &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \underline{\Sigma}}} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}(\underline{x} - \underline{\mu})^\top \underline{\Sigma}^{-1}(\underline{x} - \underline{\mu})\right) d\underline{x} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \underline{\Sigma}}} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\underline{y}^\top \underline{y}\right) \det(\underline{A}^{-1}) \cdot d\underline{y} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\|\underline{y}\|^2\right) d\underline{y} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-\frac{y_i^2}{2}} dy_i = 1. \end{aligned}$$

Naj bo (X, Y) slučajni vektor. Računamo:

$$\begin{aligned} P(X \in [a, b]) &= P((X, Y) \in [a, b] \times \mathbb{R}) \\ &= \int_{[a, b] \times \mathbb{R}} f_{X,Y}(x, y) dx dy \\ &= \int_a^b dx \cdot \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x, y) dy}_{g(x)} \\ &= \int_a^b g(x) dx. \end{aligned}$$

Funkcija $g(x)$ je gostota.

Opomba. 1. Privzeli smo, da je g Riemannovo integrabilna.

2. Privzeli smo tudi veljavnost Fubinijevega izreka.

Na podoben način izpeljemo tudi bolj splošno formulo za robno gostoto; če je $\underline{X} \in \mathbb{R}^n$ in $\underline{Y} \in \mathbb{R}^m$, potem je

$$f_{\underline{X}}(\underline{x}) = \int_{\mathbb{R}^m} f_{\underline{X}, \underline{Y}}(\underline{x}, \underline{y}) d\underline{y}.$$

Splošna definicija neodvisnosti \underline{X} in \underline{Y} pravi, da sta ti spremenljivki neodvisni, če velja $P(\underline{X} \in A, \underline{Y} \in B) = P(\underline{X} \in A) \cdot P(\underline{Y} \in B)$ za odprti množici A, B . Za $\underline{X}, \underline{Y}$, ki imata gostoto, to pomeni (v dveh dimenzijah)

$$\begin{aligned} P(X \in [a, b])P(Y \in [c, d]) &= P((X, Y) \in [a, b] \times [c, d]) \\ &= \int_{[a, b] \times [c, d]} f_{X, Y}(x, y) dx dy. \end{aligned}$$

Hkrati pa velja tudi

$$\begin{aligned} P(X \in [a, b])P(Y \in [c, d]) &= \int_a^b f_X(x) dx \cdot \int_c^d f_Y(y) dy \\ &= \int_{[a, b] \times [c, d]} f_X(x) f_Y(y) dx dy. \end{aligned}$$

Analiza 2 pove, da če je

$$\int_{[a, b] \times [c, d]} f(x, y) dx dy = \int_{[a, b] \times [c, d]} g(x, y) dx dy$$

za vse pravokotnike $[a, b] \times [c, d]$, potem je $f(x, y) = g(x, y)$ razen morda na množici z mero 0. Od tod sledi, da za neodvisni X, Y velja $f_{X, Y}(x, y) = f_X(x) \cdot f_Y(y)$ skoraj povsod. Velja pa tudi obratno: če je $f_{X, Y}(x, y) = f_X(x) f_Y(y)$, je

$$\begin{aligned} P(X \in [a, b], Y \in [c, d]) &= \int_{[a, b] \times [c, d]} f_X(x) f_Y(y) dx dy \\ &= \int_a^b f_X(x) dx \cdot \int_c^d f_Y(y) dy \\ &= P(X \in [a, b])P(Y \in [c, d]). \end{aligned}$$

S tem smo dokazali naslednji izrek.

Izrek 3.33.

Slučajna vektorja $\underline{X}, \underline{Y}$ z gostoto $f_{\underline{X}, \underline{Y}}(\underline{x}, \underline{y})$ sta neodvisna, če in samo če velja $f_{\underline{X}, \underline{Y}}(\underline{x}, \underline{y}) = f_{\underline{X}}(\underline{x}) f_{\underline{Y}}(\underline{y})$ skoraj povsod.

Zgled 3.34. Naj bo

$$f_{X, Y}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right)$$

za $-1 < \rho < 1$. Potem je.

$$\begin{aligned}
 f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy \\
 &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(y-\rho x)^2}{2(1-\rho^2)}\right) e^{-\frac{x^2}{2}} dy \\
 &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(y-\rho x)^2}{2(1-\rho^2)}\right) dy \\
 &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},
 \end{aligned}$$

torej je $X \sim N(0,1)$. Po simetriji je tudi $Y \sim N(0,1)$. Če želimo, da sta X, Y neodvisni, bi moralo veljati $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$ oziroma

$$\frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) = \frac{1}{2\pi} e^{-\frac{x^2}{2} - \frac{y^2}{2}}.$$

Če je $\rho \neq 0$, je $f_{X,Y}(x,y) > f_X(x)f_Y(y)$ za $(x,y) \in [-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon]$. V tem primeru X, Y nista neodvisni.

Izrek 3.35.

Naj bo $f_{X,Y}(x,y) = f(x) \cdot g(y)$ za funkciji f, g in vse $(x,y) \in \mathbb{R}^2$. Potem sta X, Y neodvisni.

Dokaz. Če velja $f_{X,Y} = f \cdot g$, potem računamo

$$\begin{aligned}
 f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy \\
 &= \int_{-\infty}^{\infty} f(x) \cdot g(y) dy \\
 &= f(x) \int_{-\infty}^{\infty} g(y) dy \\
 &= C_1 f(x).
 \end{aligned}$$

Podobno je tudi $f_Y(y) = g(y) \int_{-\infty}^{\infty} f_X(x) dx = C_2 g(y)$. Od tod sledi

$$f_{X,Y}(x,y) = \frac{f_X(x)}{C_1} \frac{f_Y(y)}{C_2} = \frac{1}{C_1 \cdot C_2} f_X(x) f_Y(y).$$

Sedaj obe strani enačbe integriramo po \mathbb{R}^2 :

$$\begin{aligned}
 1 &= \int_{\mathbb{R}^2} f_{X,Y}(x,y) dx dy = \frac{1}{C_1 \cdot C_2} \int_{\mathbb{R}^2} f_X(x) f_Y(y) dx dy \\
 &= \frac{1}{C_1 \cdot C_2} \int_{-\infty}^{\infty} f_X(x) dx \cdot \int_{-\infty}^{\infty} f_Y(y) dy \\
 &= \frac{1}{C_1 \cdot C_2} 1 \cdot 1.
 \end{aligned}$$

S tem in prejšnjim izrekom sledi zelena trditev. □

Opomba. 1. Iz $P(X \in [a,b], Y \in [c,d]) = P(X \in [a,b]) \cdot P(Y \in [c,d])$ smo sklepali, da sta X in Y neodvisni. Splošna definicija zahteva $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ za odprti A, B .

Splošna trditev sledi, ker je vsaka odprta podmnožica \mathbb{R} disjunktna unija odprtih intervalov. V \mathbb{R}^n pa moramo bolj razmisliti.

2. Če hočemo sklepati iz $f_{X,Y} = f \cdot g$ na neodvisnost, mora enakost veljati za vse $(x, y) \in \mathbb{R}^2$.

3.4 Funkcije slučajnih vektorjev

V diskretnem primeru se omejimo na celoštevilске slučajne spremenljivke in $f(x, y) = x + y$. Iz porazdelitve tega vektorja (X, Y) bi želeli ugotoviti porazdelitev $Z = X + Y$. Računamo:

$$\begin{aligned} P(Z = n) &= P\left(\bigcup_k \{X = k, Y = n - k\}\right) \\ &= \sum_k P(X = k, Y = n - k). \end{aligned}$$

Če sta X, Y še nenegativni, je

$$P(Z = n) = \sum_{k=0}^n P(X = k, Y = n - k).$$

Če sta X, Y neodvisni, pa je še

$$P(Z = n) = \sum_k P(X = k)P(Y = n - k).$$

Zgled 3.36. Naj bosta X, Y neodvisni, $X \sim \text{Po}(\lambda)$, $Y \sim \text{Po}(\mu)$ in $Z = X + Y$. Sedaj poračunamo:

$$\begin{aligned} P(Z = n) &= \sum_{k=0}^n P(X = k, Y = n - k) \\ &= \sum_{k=0}^n \frac{e^{-\lambda} \lambda^k}{k!} \frac{e^{-\mu} \mu^{n-k}}{(n-k)!} \\ &= \frac{e^{-(\lambda+\mu)}}{n!} \sum_{k=0}^n \binom{n}{k} \lambda^k \mu^{n-k} \\ &= \frac{e^{-(\lambda+\mu)}}{n!} (\lambda + \mu)^n, \end{aligned}$$

torej je $Z \sim \text{Po}(\lambda + \mu)$.

Zgled 3.37. Naj bosta X, Y neodvisni in

$$P(X = k) = \frac{\beta^n (a)_k}{k! (1 + \beta)^{n+k}}, \quad P(Y = l) = \frac{\beta^b (b)_l}{l! (1 + \beta)^{b+l}}$$

za $k = 0, 1, \dots$ in $l = 0, 1, \dots$ Ponovno računamo:

$$\begin{aligned}
 P(Z = n) &= \sum_{k=0}^n P(X = k)P(Y = n - k) \\
 &= \sum_{k=0}^n \frac{\beta^{a+b}(a)_k(b)_{n-k}}{k!(b-k)!(1+\beta)^{a+b+n}} \cdot \frac{n!}{n!} \\
 &= \frac{\beta^{a+b}}{(1+\beta)^{a+b+n}n!} \sum_{k=0}^n \binom{n}{k} (a)_k(b)_{n-k} \\
 &= \frac{\beta^{a+b}(a+b)_n}{(1+\beta)^{a+b+n}n!}
 \end{aligned}$$

za $k = 0, 1, \dots$

Zgled 3.38. Naj bosta X, Y neodvisni, $X \sim \text{Bin}(m, p)$, $Y \sim \text{Bin}(n, p)$ in $Z = X + Y$.

$$\begin{aligned}
 P(Z = l) &= \sum_k P(X = k) \cdot P(Y = l - k) \\
 &= \sum_k \binom{m}{k} p^k \cdot q^{m-k} \cdot \binom{n}{l-k} \cdot p^{l-k} \cdot q^{n-(l-k)} \\
 &= p^n \cdot q^m \sum_k \binom{m}{k} \binom{n}{l-k} \\
 &= p^l \cdot q^{m+n-l} \cdot \binom{m+n}{l},
 \end{aligned}$$

torej je $Z \sim \text{Bin}(m+n, p)$, kot bi pričakovali. pri računu smo uporabili dejstvo, da če je $U \sim \text{HiperGeom}(n, B, N)$, je

$$\sum_k \frac{\binom{B}{k} \binom{R}{n-k}}{\binom{N}{n}} = 1$$

ali alternativno

$$\sum_k \binom{B}{k} \binom{R}{n-k} = \binom{N}{n}.$$

Sedaj se posvetimo zveznemu primeru. Naj bo (X, Y) slučajni vektor in $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Označimo $(U, V) = \Phi(X, Y)$. Privzemimo, da je Φ bijektivna. Velja naslednje:

$$\begin{aligned}
 P((X, Y) \in A) &= \int_A f_{X,Y}(x, y) dx dy \\
 &\approx V(A) \cdot f_{X,Y}(x, y),
 \end{aligned}$$

po drugi strani pa

$$\begin{aligned}
 P((U, V) \in \Phi(A)) &= \int_{\Phi(A)} f_{U,V}(u, v) du dv \\
 &\approx V(\Phi(A)) \cdot f_{U,V}(u, v)
 \end{aligned}$$

Sedaj „uganemo“ formulo

$$f_{U,V}(u, v) = \frac{V(A)}{V(\Phi(A))} \cdot f_{X,Y}(x, y).$$

Iz analize 3 pa vemo: $\frac{V(\Phi(A))}{V(A)} \approx |J\Phi(x, y)|$. Od tod sledi

$$\frac{V(A)}{V(\Phi(A))} \approx |J_{\Phi}^{-1}(x, y)| = |J_{\Phi^{-1}}(u, v)|,$$

zato dobimo

$$f_{U,V}(u, v) = f_{X,Y}(\Phi^{-1}(u, v)) |J\Phi^{-1}(u, v)|$$

Izrek 3.39 (Transformacijska formula).

Naj bo \underline{X} slučajni vektor z gostoto $f_{\underline{X}}(\underline{x})$, tak da je $P(\underline{X} \in U) = 1$ za odprto podmnožico $U \subseteq \mathbb{R}^n$. Naj bo $\Phi : U \rightarrow V$ bijektivna, V odprta in Φ ter Φ^{-1} parcialno zvezno odvedljivi. Naj bo $\underline{Y} = \Phi(\underline{X})$. Potem velja

$$f_{\underline{Y}}(\underline{y}) = f_{\underline{X}}(\Phi^{-1}(\underline{y})) |J_{\Phi^{-1}}(\underline{y})|.$$

Opomba. Očitno je $P(\underline{Y} \in V) = 1$, torej je $f_Y(y) = 0$ za $\underline{y} \notin V$.

Dokaz. Naj bo $B \subseteq V$ Jordanovo izmerljiva in omejena. Računamo:

$$\begin{aligned} P(\underline{Y} \in B) &= P(\underline{X} \in \Phi^{-1}(B)) \\ &= \int_{\Phi^{-1}(B)} f_{\underline{X}}(\underline{x}) d\underline{x} \\ &= \int_B f_{\underline{X}}(\Phi^{-1}(\underline{y})) |J_{\Phi^{-1}}(\underline{y})| d\underline{y}. \end{aligned}$$

Ker formula velja za vsako množico B , je izraz dokazan. \square

Zgled 3.40. Naj bosta X_1 in X_2 neodvisni, $X_1 \sim \Gamma(a, \lambda)$ in $X_2 \sim \Gamma(b, \lambda)$. Definirajmo $\Phi(x_1, x_2) = \left(\frac{x_1}{x_1+x_2}, x_1+x_2\right)$ in zato

$$(Y_1, Y_2) = \left(\frac{X_1}{X_1+X_2}, X_1+X_2\right).$$

Definirajmo $U := (0, \infty)^2$ in vidimo, da velja $P((X_1, X_2) \in U) = 1$. Očitno je, da Φ bijektivna preslika U na $V := (0, 1) \times (0, \infty)$. Izračunajmo Φ^{-1} ; za to moramo rešiti enačbi

$$\frac{x_1}{x_1+x_2} = y_1, \quad x_1+x_2 = y_2.$$

Če ti dve enačbi med seboj zmnožimo, dobimo

$$x_1 = y_1 y_2, \quad x_2 = y_2(1 - y_1).$$

Od tod pa že sledi

$$\Phi^{-1}(y_1, y_2) = (y_1 y_2, y_2(1 - y_1))$$

in njen odvod je

$$J_{\Phi^{-1}}(y_1, y_2) = \det \begin{pmatrix} y_2 & y_1 \\ -y_2 & 1 - y_1 \end{pmatrix} = y_2.$$

Zaradi neodvisnosti je

$$f_{\underline{X}}(\underline{x}) = f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$$

za vse pare $(x_1, x_2) \in U$. Od tod tudi dobimo

$$\begin{aligned} f_{X_1, X_2}(x_1, x_2) &= \frac{\lambda^a \cdot \lambda^b}{\Gamma(a)\Gamma(b)} \cdot x_1^{a-1} e^{-\lambda x_1} x_2^{b-1} e^{-\lambda x_2} \\ f_{Y_1, Y_2}(y_1, y_2) &= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} \cdot y_1^{a-1} (1 - y_1)^{b-1} y_2^{a+b-1} e^{-\lambda y_2} \end{aligned}$$

za vse pare $(y_1, y_2) \in V$. Opazimo, da je gostota $f_{Y_1, Y_2}(y_1, y_2)$ je oblike $f(y_1)g(y_2)$, torej sta Y_1, Y_2 neodvisni. Formula velja za $(y_1, y_2) \in V$, zunaj tega območja je $f_{Y_1, Y_2}(y_1, y_2) = 0$. Vemo pa tudi, da je gostota Y_1 proporcionalna $y_1^{a-1}(1-y_1)^{b-1}$ oziroma

$$f_{Y_1}(y_1) = \frac{1}{B(a, b)} y_1^{a-1} (1-y_1)^{b-1}$$

za $y_1 \in (0, \infty)$. Podobno tudi za Y_2 :

$$f_{Y_2} = \frac{\lambda^{a+b}}{\Gamma(a+b)} y_2^{a+b-1} e^{-\lambda y_2}.$$

Če sedaj vse to vstavimo v enačbo $f_{Y_1, Y_2}(y_1, y_2) = f_{Y_1}(y_1) \cdot f_{Y_2}(y_2)$, smo dokazali

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Stranski produkt tega računa pa je tudi dejstvo $Y \sim X_1 + X_2 \sim \Gamma(a+b, \lambda)$.

Zgled 3.41. Naj bo (X, Y) slučajni vektor in $\Phi(x, y) = (x, x+y)$. Potem je $\Phi^{-1}(x, z) = (x, z-x)$ in

$$J_{\Phi^{-1}}(x, z) = \det \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = 1.$$

Transformacijska formula nam da

$$f_{X, Z}(x, z) = f_{X, Y}(x, z-x) \cdot |1|.$$

Po formuli za robno gostoto je

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X, Y}(x, z-x) dx.$$

Če pa sta X in Y še neodvisni, dobimo

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx.$$

Uporabimo formulo za neodvisni spremenljivki $X \sim N(0, \sigma^2)$ in $Y \sim N(0, \tau^2)$, kjer je $\sigma^2 + \tau^2 = 1$. Zanima nas gostota $Z = X + Y$.

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx \\ &= \frac{1}{2\pi\sigma\tau} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} \cdot e^{-\frac{(z-x)^2}{2\tau^2}} dx \\ &= \frac{1}{2\pi\sigma\tau} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-\sigma^2 z)^2 - \sigma^4 z^2 + \sigma^2 z^2}{2\sigma^2\tau^2}\right) dx \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \cdot \frac{1}{\sqrt{2\pi\sigma\tau}} \int_{-\infty}^{\infty} e^{-\frac{(x-\sigma^2 z)^2}{2\sigma^2\tau^2}} dx \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}. \end{aligned}$$

Torej imamo $Z = X + Y \sim N(0, 1)$.

Zgled 3.42. Kot posledico prejšnjega si oglejmo primer, ko sta X, Y neodvisni in $X \sim N(\mu, \sigma^2), Y \sim N(\nu, \tau^2)$ in $Z = X + Y$. Porazdelitev Z sedaj izračunamo:

$$X + Y = \sqrt{\sigma^2 + \tau^2} \left(\frac{X - \mu}{\sqrt{\sigma^2 + \tau^2}} + \frac{Y - \nu}{\sqrt{\sigma^2 + \tau^2}} \right) + \mu + \nu.$$

Vpeljemo $X' := \frac{X - \mu}{\sqrt{\sigma^2 + \tau^2}}$ in $Y' := \frac{Y - \nu}{\sqrt{\sigma^2 + \tau^2}}$ in dobimo

$$X' \sim N\left(0, \frac{\sigma^2}{\sigma^2 + \tau^2}\right), \quad Y' \sim N\left(0, \frac{\tau^2}{\sigma^2 + \tau^2}\right).$$

Ker sta X' in Y' neodvisni, velja po prejšnjem zgledu $X' + Y' \sim N(0, 1) =: Z'$. Torej je

$$X + Y = \sqrt{\sigma^2 + \tau^2} Z' + \mu + \nu \sim N(\mu + \nu, \sigma^2 + \tau^2).$$

Naj ima \underline{X} gostoto $f_{\underline{X}}(\underline{x})$ in naj bo $\underline{Y} = \underline{A}\underline{X} + \underline{\mu}$, kjer je \underline{A} matrika dimenzij $n \times n$. Velja torej $\Phi(\underline{x}) = \underline{A} \cdot \underline{x} + \underline{\mu}$. Potem je

$$\Phi^{-1}(\underline{y}) = \underline{A}^{-1}(\underline{y} - \underline{\mu}) \Rightarrow J\Phi^{-1}(\underline{y}) = \det(\underline{A}^{-1}).$$

Če je $\underline{Y} = \underline{A}\underline{X} + \underline{\mu}$, je

$$f_{\underline{Y}}(\underline{y}) = f_{\underline{X}}(\underline{A}^{-1}(\underline{y} - \underline{\mu})) |\det(\underline{A}^{-1})|.$$

Kot poseben primer si oglejmo $\underline{Z} = (Z_1, \dots, Z_n)$, kjer so komponente med seboj neodvisne in velja $Z_i \sim N(0, 1)$. Potem velja

$$\begin{aligned} f_{\underline{Z}}(\underline{z}) &= \prod_{k=1}^n f_{Z_k}(z_k) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{k=1}^n e^{-\frac{z_k^2}{2}} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{k=1}^n e^{-\frac{(z_1^2 + \dots + z_n^2)}{2}} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{k=1}^n e^{-\frac{\underline{z}^\top \cdot \underline{z}}{2}} \end{aligned}$$

Če je $\underline{X} = \underline{A}\underline{Z} + \underline{\mu}$, je

$$\begin{aligned} f_{\underline{X}}(\underline{x}) &= f_{\underline{Z}}(\underline{A}^{-1}(\underline{X} - \underline{\mu})) \cdot |\det(\underline{A}^{-1})| \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{(\underline{x} - \underline{\mu})^\top (\underline{A}^{-1})^\top \underline{A}^{-1}(\underline{x} - \underline{\mu})}{2}\right) |\det \underline{A}^{-1}| \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}(\underline{x} - \underline{\mu})(\underline{A}\underline{A}^\top)^{-1}(\underline{x} - \underline{\mu})\right) |\det \underline{A}^{-1}|. \end{aligned}$$

Označimo $\underline{\Sigma} = \underline{A} \cdot \underline{A}^\top$. Ta matrika je simetrična in pozitivno definitna, torej imamo

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\underline{\Sigma})}} \exp\left(-\frac{1}{2}(\underline{x} - \underline{\mu})\underline{\Sigma}^{-1}(\underline{x} - \underline{\mu})\right).$$

Gostoti \underline{X} rečemo večrazsežna normalna gostota s parametroma $\underline{\mu}$ in $\underline{\Sigma}$ in jo označujemo z $\underline{X} \sim N(\underline{\mu}, \underline{\Sigma})$.

Opomba. Označujemo $Z \sim N(0, I)$. Hkrati za vsako pozitivno definitno $\underline{\Sigma}$ obstaja razcep $\underline{\Sigma} = \underline{A} \cdot \underline{A}^\top$, zato za vsako pozitivno definitno matriko $\underline{\Sigma}$ je zgornja funkcija funkcija gostote.

Zgled 3.43. Naj ima (X, Y) gostoto $f_{X,Y}(x, y)$ in naj bo $Z = \frac{Y}{X}$. Kakšna je njena gostota? Ideja je, da dopolnimo $\frac{Y}{X}$ še z eno komponento, da dobimo preslikavo iz \mathbb{R}^2 v \mathbb{R}^2 . Ena možnost je, da definiramo $\Phi(x, y) = (x, \frac{y}{x})$, kjer bi bila množica $U = \mathbb{R}^2 \setminus \{0\} \times \mathbb{R}$. Potem je $\Phi^{-1}(x, z) = (x, xz)$ in

$$J\Phi^{-1}(x, z) = \det \begin{pmatrix} 1 & 0 \\ z & x \end{pmatrix} = x.$$

Tako dobimo

$$f_{X,Z}(x, z) = f_{X,Y}(x, xz) \cdot |x|.$$

Po formuli za robno gostoto je

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, xz) |x| dx.$$

Če sta $Z_1, Z_2 \sim N(0, 1)$ neodvisni, lahko definiramo $X = \frac{Z_1}{Z_2}$ in dobimo

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{Z_1}(z_1) \cdot f_{Z_2}(z_1 x) |z_1| dz_1 \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{z_1^2 x^2}{2}} |z_1| dz_1 \\ &= \frac{1}{2\pi} \int_0^{\infty} e^{-\frac{z_1^2}{2}(1+x^2)} \cdot z_1 dz_1 \\ &= \frac{1}{\pi} \left(-\frac{e^{-\frac{z_1^2}{2}(1+x^2)}}{(1+x^2)} \Big|_0^{\infty} \right) \\ &= \frac{1}{\pi(1+x^2)}. \end{aligned}$$

To je Cauchyjeva porazdelitev.

Zgled 3.44. Naj bo $\underline{X} \sim N(\underline{\mu}, \underline{\Sigma})$. Razcepimo $\underline{X} = (\underline{X}^{(1)}, \underline{X}^{(2)})$, kjer je prva dimenzije p , druga dimenzije q in $p + q = n$. Enako razcepimo

$$\underline{\Sigma} = \begin{pmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} \end{pmatrix}, \quad \underline{\mu} = \begin{pmatrix} \underline{\mu}^{(1)} \\ \underline{\mu}^{(2)} \end{pmatrix}.$$

Zanima nas robna gostota $X^{(1)}$. Ta je načeloma enaka

$$f_{\underline{X}^{(1)}}(x^{(1)}) = \int_{\mathbb{R}^2} f_{\underline{X}}(\underline{x}) d\underline{x}^{(2)}.$$

Alternativno pa vemo, kako izračunati gostoto $\underline{A}\underline{X} + \underline{\mu}$. Vzamemo $\underline{\mu} = 0$ in

$$\underline{A} = \begin{pmatrix} \underline{I} & 0 \\ -\underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1} & \underline{I} \end{pmatrix}.$$

Ta matrika je spodnje trikotna in obrnljiva:

$$\underline{Y} = \underline{A}\underline{X} = \begin{pmatrix} \underline{X}^{(1)} \\ \underline{X}^{(2)} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{X}^{(1)} \end{pmatrix}.$$

Vemo, da je $Y \sim N(\underline{A}\underline{\mu}, \underline{A}\underline{\Sigma}\underline{A}^T)$ in

$$\underline{A}\underline{\mu} = \begin{pmatrix} \underline{\mu}^{(1)} \\ \underline{\mu}^{(2)} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{\mu}^{(1)} = \underline{\mu}^{(2)} \end{pmatrix}.$$

Sedaj pomnožimo matrike:

$$\begin{aligned} \underline{A}\underline{\Sigma}\underline{A}^\top &= \begin{pmatrix} \underline{I} & 0 \\ -\underline{\Sigma}_{21}\underline{\Sigma}^{-1} & \underline{I} \end{pmatrix} \begin{pmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} \end{pmatrix} \begin{pmatrix} \underline{I} & -\underline{\Sigma}_{21}\underline{\Sigma}^{-1} \\ 0 & \underline{I} \end{pmatrix} \\ &= \begin{pmatrix} \underline{\Sigma}_{11} & 0 \\ 0 & \underline{\Sigma}_{22} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{\Sigma}_{12} \end{pmatrix} \end{aligned}$$

Tako dobimo gostoto

$$\begin{aligned} f_{\underline{Y}}(\underline{y}) &= \frac{1}{(2\pi)^{\frac{n}{2}} \det \sqrt{\underline{A}\underline{\Sigma}\underline{A}^\top}} \exp \left(-\frac{1}{2} (\underline{y} - \underline{A}\underline{\mu})^\top (\underline{A}\underline{\Sigma}\underline{A}^\top) (\underline{y} - \underline{A}\underline{\mu}) \right) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \det \sqrt{\underline{A}\underline{\Sigma}\underline{A}^\top}} \exp \left(-\frac{1}{2} (\underline{y}^{(1)} - \underline{\mu}^{(1)})^\top (\underline{\Sigma}_{11})^{-1} (\underline{y}^{(1)} - \underline{\mu}^{(1)}) \right) \\ &\quad \exp \left(-\frac{1}{2} (\underline{y}^{(2)} - \underline{\nu}^{(2)})^\top (\underline{\Sigma}_{22} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{\Sigma}_{12}) (\underline{y}^{(2)} - \underline{\nu}^{(2)}) \right) \end{aligned}$$

Od tod sklepamo, da sta $\underline{X}^{(1)}$ in $\underline{X}^{(2)} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{X}^{(1)}$ neodvisna. Gostota $\underline{X}^{(1)}$ je proporcionalna

$$f_{\underline{X}^{(1)}}(\underline{X}^{(1)}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \underline{\Sigma}_{11}}} \exp \left(-\frac{1}{2} (\underline{x}^{(1)} - \underline{\mu}^{(1)})^\top \underline{\Sigma}_{11}^{-1} (\underline{x}^{(1)} - \underline{\mu}^{(1)}) \right),$$

torej je $\underline{X}^{(1)} \sim N(\underline{\mu}^{(1)}, \underline{\Sigma}_{11})$.

4 Pričakovana vrednost

4.1 Definicija in primeri

Definicija 4.1.

1. Za zvezno porazdeljeno slučajno spremenljivko X z gostoto $f_X(x)$ definiramo

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) dx, \quad E[f(X)] = \int_{-\infty}^{\infty} f(x) f_X(x) dx.$$

Rečemo, da $E(X)$ in $E[f(X)]$ obstajata, če obstajata integrala z absolutnimi vrednostmi.

2. Če je \underline{X} slučajni vektor z gostoto $f_{\underline{X}}(\underline{x})$, definiramo

$$E[f(\underline{X})] = \int_{\mathbb{R}^n} f(\underline{x}) f_{\underline{X}}(\underline{x}) d\underline{x}.$$

Pričakovana vrednost obstaja, če obstaja integral z absolutnimi vrednostmi.

Zgled 4.2. Vzemimo $X \sim N(\mu, \sigma^2)$. Vemo, da je $\frac{X-\mu}{\sigma} \sim N(0, 1)$. Izračunamo $Z \sim N(0, 1)$:

$$E(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \cdot e^{-\frac{z^2}{2}} dz = 0.$$

Podobno dobimo tudi

$$\begin{aligned} E(Z^2) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} dz \\ &= \frac{1}{\sqrt{2\pi}} z \cdot e^{-\frac{z^2}{2}} \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz \\ &= 1. \end{aligned}$$

Ker smo označili $Z = \frac{X-\mu}{\sigma}$, potem zaradi linearnosti iz $E\left(\frac{X-\mu}{\sigma}\right) = 0$ sledi $E(X) = \mu$. Podobno je $E(X^2) = \sigma^2 + \mu^2$.

Zgled 4.3. Naj bo $X \sim \Gamma(a, \lambda)$, torej imamo gostoto $f_X(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}$ za $x > 0$. Računamo za $f(x) = x^m$:

$$\begin{aligned} E(x^m) &= \int_0^{\infty} x^m \cdot f_X(x) dx \\ &= \frac{\lambda^a}{\Gamma(a)} \int_0^{\infty} x^m x^{a-1} e^{-\lambda x} dx \\ &= \frac{\lambda^a}{\Gamma(a)} \frac{\Gamma(a+m)}{\lambda^{a+m}} \cdot \frac{\lambda^{a+m}}{\Gamma(a+m)} \int_0^{\infty} x^{m+a-1} e^{-\lambda x} dx \\ &= \frac{\Gamma(a+m)}{\Gamma(a)\lambda^m} \\ &= \frac{(a)_m}{\lambda^m}. \end{aligned}$$

Definicija 4.4. Količina $E(X^m)$ se imenuje m -ti moment slučajne spremenljivke X , količina $E[(X - E(X))^m]$ pa m -ti centralni moment.

Zgled 4.5. Naj bo $\underline{X} \sim N(\underline{\mu}, \underline{\Sigma})$. Zanima nas $E(X_i \cdot X_j)$, torej vzamemo $f(\underline{x}) = x_i \cdot x_j$. Po definiciji je

$$\begin{aligned} E(X_i X_j) &= \int_{\mathbb{R}^n} x_i x_j \cdot f_{\underline{X}}(\underline{x}) d\underline{x} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \underline{\Sigma}}} \int_{\mathbb{R}^n} x_i x_j e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^\top \underline{\Sigma}^{-1}(\underline{x}-\underline{\mu})} d\underline{x} \end{aligned}$$

Ker je $\underline{\Sigma}$ pozitivno definitna matrika, si lahko izberemo simetrično pozitivno definitno matriko $\underline{\Sigma}^{\frac{1}{2}}$. Sedaj v integral uvedemo novo spremenljivko $\underline{y} = \underline{\Sigma}^{-\frac{1}{2}}(\underline{x} - \underline{\mu})$:

$$\begin{aligned} &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} (\underline{\mu} + \underline{\Sigma}^{\frac{1}{2}} \underline{y})_i (\underline{\mu} + \underline{\Sigma}^{\frac{1}{2}} \underline{y})_j e^{-\frac{1}{2} \underline{y}^\top \underline{y}} d\underline{y} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \left(\mu_i \mu_j + \mu_i \cdot \left(\underline{\Sigma}^{\frac{1}{2}} \underline{y} \right)_j + \mu_j \cdot \left(\underline{\Sigma}^{\frac{1}{2}} \underline{y} \right)_i + \left(\underline{\Sigma}^{\frac{1}{2}} \underline{y} \right)_i \cdot \left(\underline{\Sigma}^{\frac{1}{2}} \underline{y} \right)_j \right) e^{-\frac{1}{2} \underline{y}^\top \underline{y}} d\underline{y}. \end{aligned}$$

Ker je $\underline{Y} \sim N(\underline{\mu}, \underline{I})$, je $f_{\underline{Y}}(\underline{y}) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2} \underline{y}^\top \underline{y}}$:

$$\begin{aligned} &= \mu_i \mu_j + \left(\underline{\Sigma}^{\frac{1}{2}} \right)_i^\top \left(\underline{\Sigma}^{\frac{1}{2}} \right)_j \\ &= \mu_i \mu_j + \underline{\Sigma}_{ij}. \end{aligned}$$

Pri tem smo uporabili

$$\int_{\mathbb{R}^n} y_k e^{-\frac{1}{2} \underline{y}^\top \underline{y}} d\underline{y} = 0, \quad \int_{\mathbb{R}^n} y_k y_l e^{-\frac{1}{2} \underline{y}^\top \underline{y}} d\underline{y} = 0$$

in pa

$$\begin{aligned} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \left(\sum_{i=1}^n \alpha_i y_i \right) \left(\sum_{j=1}^n \beta y_j \right) \cdot e^{-\frac{1}{2} \underline{y}^\top \underline{y}} d\underline{y} &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \left(\sum_{i=1}^n \alpha_i \beta_i y_i^2 \right) \cdot e^{-\frac{1}{2} \underline{y}^\top \underline{y}} d\underline{y} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \sum_{i=1}^n \alpha_i \beta_i \int_{\mathbb{R}^n} y_i^2 \cdot e^{-\frac{1}{2} \underline{y}^\top \underline{y}} d\underline{y} \\ &= \sum_{i=1}^n \alpha_i \beta_i. \end{aligned}$$

4.2 Varianca in kovarianca

Denimo, da imamo dve skupini števil. Narišemo si jih na premico. Radi bi izmerili „razmetanost“ teh števil na premici. Ena ideja bi bila, da gledamo povprečno razdaljo od povprečja vrednosti točk. Za bolj „razmetane“ skupine števil je to povprečje večje. Mera razmetanosti bi bila torej $\frac{1}{n} \sum_{k=1}^n |x_k - \bar{x}|$. Gauss pa je namesto absolutnih vrednosti predlagal kvadrat. Predlagal je mero $\sigma^2 = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2$. Vrnimo se k slučajnim spremenljivkam: rečemo, da X „ponavljamo“ in s tem dobimo vrednosti x_1, x_2, \dots . Če izračunamo veliko število ponavljanj, dobimo

$$\begin{aligned} \frac{(x_1 - \bar{x})^2 + \dots + (x_k - \bar{x})^2}{n} &\approx \frac{(x_1 - E(X))^2 + \dots + (x_k - E(X))^2}{n} \\ &\approx E(X - E(X))^2. \end{aligned}$$

Definicija 4.6. Varianca slučajne spremenljivke X je količina

$$\text{var}(X) = E[(X - E(X))^2].$$

Rečemo, da varianca obstaja, če je pričakovana vrednost končna.

Za računanje je bolj priročna oblika

$$\begin{aligned} \text{var}(X) &= E[(X - E(X))^2] \\ &= E[X^2 - 2XE(X) + [E(X)]^2] \\ &= E(X^2) - 2E(X) \cdot E(X) + [E(X)]^2 \\ &= E(X^2) - E(X)^2. \end{aligned}$$

Opomba. Iz te definicije takoj sledi, da je $\text{var}(aX) = a^2 \text{var}(X)$.

Definicija 4.7. Standardni odklon slučajne spremenljivke X je $\text{SD}(X) = \sqrt{\text{var}(X)}$.

Kako je z varianco vsote slučajnih spremenljivk? Po definiciji je

$$\begin{aligned} \text{var}(X + Y) &= E[(X + Y)^2] - [E(X + Y)]^2 \\ &= E[X^2 + 2XY + Y^2] - [E(X) + E(Y)]^2 \\ &= E(X^2) + 2E(XY) + E(Y^2) - [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2 \\ &= \text{var}(X) + \text{var}(Y) + 2(E(XY) - E(X)E(Y)). \end{aligned}$$

Izraza v zadnjem oklepaju v splošnem ne moremo zanemariti, zato ga poimenujemo.

Definicija 4.8. Kovarianca slučajnih spremenljivk X, Y je količina

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y).$$

Opomba. Z uporabo linearnosti lahko izraz za kovarianco prevedemo v

$$\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Kovarianca obstaja, če obstajajo vse pričakovane vrednosti v definiciji.

Zgled 4.9. Če je $\underline{X} \sim \text{Multinom}(m, \underline{p})$, smo že izračunali

$$E(X_i X_j) = n^2 p_i p_j - n p_i p_j.$$

To sledi, ker je $X_i \sim \text{Bin}(n, p_i)$. Potem imamo

$$\text{cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = -n p_i p_j.$$

Zgled 4.10. Denimo sedaj, da je $\underline{X} \sim N(\underline{\mu}, \underline{\Sigma})$. Vemo že, da je $E(X_i X_j) = \underline{\Sigma}_{ij} + \mu_i \mu_j$. Iz naših predhodnih izračunov tudi sledi, da je $X_i \sim N(\mu_i, \underline{\Sigma}_{ii})$, torej je $E(X_i) = \mu_i$. Od tod sledi

$$\begin{aligned} \text{cov}(X_i, X_j) &= E(X_i X_j) - E(X_i)E(X_j) \\ &= \underline{\Sigma}_{ij} + \mu_i \mu_j - \mu_i \mu_j \\ &= \underline{\Sigma}_{ij}. \end{aligned}$$

Izrek 4.11.

Naj bodo X_1, X_2, \dots, X_n slučajne spremenljivke, za katere je $E(X_i^2) < \infty$ za vse $i = 1, 2, \dots, n$. Potem velja

$$\text{var} \left(\sum_{i=1}^n \alpha_i X_i \right) = \sum_{i=1}^n \alpha_i^2 \text{var}(X_i) + \sum_{i \neq j} \alpha_i \alpha_j \text{cov}(X_i, X_j).$$

Dokaz. Po definiciji je

$$\begin{aligned} \text{var} \left(\sum_{i=1}^n \alpha_i X_i \right) &= E \left[\left(\sum_{i=1}^n \alpha_i X_i \right)^2 \right] - \left[E \left(\sum_{i=1}^n \alpha_i X_i \right) \right]^2 \\ &= E \left[\sum_{i=1}^n \alpha_i^2 X_i^2 + \sum_{i \neq j} \alpha_i \alpha_j X_i X_j \right] - \left[E \left(\sum_{i=1}^n \alpha_i X_i \right) \right]^2 \\ &= \sum_{i=1}^n \alpha_i^2 E(X_i^2) + \sum_{i \neq j} \alpha_i \alpha_j E(X_i X_j) - \sum_{i=1}^n E(\alpha_i X_i)^2 - \sum_{i \neq j} E(\alpha_i X_i) E(\alpha_j X_j) \\ &= \sum_{i=1}^n \alpha_i^2 \text{var}(X_i) + \sum_{i \neq j} \alpha_i \alpha_j \text{cov}(X_i, X_j). \end{aligned} \quad \square$$

Iz definicije sledijo naslednje lastnosti kovarianc:

$$1. \text{ cov}(X, Y) = \text{cov}(Y, X),$$

2. $\text{cov}(X, X) = \text{var}(X)$,
3. $\text{cov}(\alpha X, \beta Y) = \alpha\beta \text{cov}(X, Y)$,
4. $\text{cov}\left(\sum_{i=1}^n \alpha_i X_i, \sum_{j=1}^n \beta_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j \text{cov}(X_i, Y_j)$,
5. $\text{cov}(c, Y) = 0$, če je c konstanta.

Zgled 4.12. Slučajno spremenljivko $X \sim \text{HiperGeom}(n, B, N)$ smo zapisali kot vsoto indikatorjev $X = I_1 + I_2 + \dots + I_n$, kjer je

$$I_k = \begin{cases} 1; & \text{če je } k\text{-ta kroglica bela} \\ 0; & \text{sicer} \end{cases}.$$

Sedaj poračunamo:

$$\begin{aligned} \text{var}(X) &= \text{var}(I_1 + I_2 + \dots + I_n) \\ &= \sum_{k=1}^n \text{var}(I_k) + \sum_{k \neq l} \text{cov}(I_k, I_l) \\ &= n \cdot \frac{B}{N} \left(1 - \frac{B}{N}\right) \cdot \frac{N-m}{N-1}. \end{aligned}$$

Pri tem smo uporabili naslednja izračuna: če je $I \sim \text{Bernoulli}(p)$, potem je

$$\begin{aligned} \text{var}(I) &= E(I^2) - (E(I))^2 \\ &= E(I) - (E(I))^2 \\ &= p - p^2 = p(1-p), \end{aligned}$$

in pa sledeče:

$$\begin{aligned} \text{cov}(I_1, I_2) &= E(I_1 I_2) - E(I_1)E(I_2) \\ &= P(I_1 = 1, I_2 = 1) - \frac{B}{N} \cdot \frac{B}{N}. \end{aligned}$$

Upoštevajoč dejstvo, da je $I_1 I_2 \sim \text{Bernoulli}(P(I_1 = 1, I_2 = 2))$:

$$\begin{aligned} &= \frac{B}{N} \frac{B-1}{N-1} - \frac{B^2}{N^2} \\ &= -\frac{B}{N} \left(1 - \frac{B}{N}\right) \frac{1}{N-1}. \end{aligned}$$

Zaradi simetrije so namreč vsi pari (I_k, I_l) enako porazdeljeni. Od tod naravno sledi, da so tudi vse kovariance enake.

K prejšnjemu zgledu dodajmo še končno opazko: izbiranje lahko nadaljujemo do zadnje kroglice in dobimo indikatorje I_1, I_2, \dots, I_N . Razmislek s simetrijo še vedno velja, zato velja tudi $I_1 + I_2 + \dots + I_N = B$, kar je konstanta. Od tod takoj sledi

$$\text{cov}(I_1 + \dots + I_N, I_k) = \text{cov}(I_k, I_k) + \text{cov}(I_k, I_l)(N-1) = 0.$$

Ker je $\text{cov}(I_k, I_k) = \text{var}(I_k) = \frac{B}{N} \left(1 - \frac{B}{N}\right)$, je

$$\text{cov}(I_k, I_l) = \frac{B}{N} \left(1 - \frac{B}{N}\right) \frac{1}{N-1}.$$

Sedaj si oglejmo naslednji primer.

Zgled 4.13. Naj bo $\underline{X} \sim \text{Multinom}(n, \underline{p})$. Vemo že, da je $X_i \sim \text{Bin}(n, p_i)$, $X_j \sim \text{Bin}(n, p_j)$ in pa $X_i + X_j \sim \text{Bin}(n, p_i + p_j)$. Potem je

$$\begin{aligned} n(p_i + p_j)(1 - p_i - p_j) &= \delta(X_i + X_j) \\ &= \text{var}(X_i) + \text{var}(X_j) + 2\text{cov}(X_i, X_j) \\ &= np_i(1 - p_i) + np_j(1 - p_j) + 2\text{cov}(X_i, X_j). \end{aligned}$$

Sedaj smo še na ta način dobili $\text{cov}(X_i, X_j) = -np_i p_j$.

Zgled 4.14. Če sta X in Y neodvisni, je

$$\begin{aligned} E(XY) &= \sum_{x,y} xyP(X=x, Y=y) \\ &= \sum_{x,y} xyP(X=x)P(Y=y) \\ &= \left(\sum_x xP(X=x) \right) \left(\sum_y yP(Y=y) \right) \\ &= E(X)E(Y), \end{aligned}$$

kjer smo množenje neskončnih vrst že dodobra utemeljili. Sedaj pa še za zvezne slučajne spremenljivke:

$$\begin{aligned} E(XY) &= \int_{\mathbb{R}^2} xyf_{X,Y}(x,y) dx dy \\ &= \int_{\mathbb{R}^2} xyf_X(x)f_Y(y) dx dy \\ &= \left(\int_{\mathbb{R}} xf_X(x) dx \right) \left(\int_{\mathbb{R}} yf_Y(y) dy \right) \\ &= E(X)E(Y). \end{aligned}$$

Za neodvisni zvezni slučajni spremenljivki X, Y je $\text{cov}(X, Y) = 0$. Obratno pa ni nujno res; za $Z \sim N(0, 1)$ je $\text{cov}(Z, Z^2) = 0$, kljub temu, da Z in Z^2 nista neodvisni.

Opomba. Za neodvisni slučajni spremenljivki X, Y velja

$$E[f(X)g(Y)] = E[f(X)] \cdot E[g(Y)]$$

z enakim dokazom za poljubni funkciji f, g .

Zgled 4.15. Naj bo $\underline{X} \sim N(\underline{\mu}, \underline{\Sigma})$. Najdemo lahko matriko \underline{A} , da je $\underline{\Sigma} = \underline{A}\underline{A}^\top$ (Cholesky). Če je $\underline{Z} = (Z_1, \dots, Z_n)$ vektor z neodvisnimi komponentami in je $Z_i \sim N(0, 1)$ za vse $i = 1, \dots, n$, je

$$\underline{X} = \underline{A} \cdot \underline{Z} + \underline{\mu} \sim N(\underline{\mu}, \underline{\Sigma}).$$

Sedaj pa preprosto poračunamo

$$\begin{aligned}\operatorname{cov}(X_i, X_j) &= \operatorname{cov}\left(\sum_{k=1}^n a_{ik} Z_k + \mu_i, \sum_{l=1}^n a_{jl} Z_l + \mu_j\right) \\ &= \sum_{k=1}^n \sum_{l=1}^n a_{ik} a_{jl} \operatorname{cov}(Z_k, Z_l) \\ &= \sum_{k=1}^n a_{ik} a_{jk} \\ &= (\underline{\Sigma})_{ij}.\end{aligned}$$

Pri tem smo uporabili dejstvo, da za diskretni neodvisni slučajni spremenljivki X, Y velja $E(X \cdot Y) = E(X) \cdot E(Y)$ in zato je $\operatorname{cov}(X, Y) = 0$. Podobno velja tudi za zvezne slučajne pare (X, Y) .

Zgled 4.16 (Sir Francis Galton). Pri reučevanju genetike je angleški statistik Galton zbral 1066 parov očetov in sinov ter izrazil njihove telesne višine. Pare višin je označil z $(x_1, y_1), \dots, (x_n, y_n)$, nato pa je narisal razsejni grafikon. Galtonove zahteve za definicijo „mere povezanosti“ so bila, da mera ne sme biti odvisna od enot in da če povezanosti ni, mora biti mera enaka 0. Definirajmo $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, $\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ ter $\sigma_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$. Galton je nato pretvoril višine v standardne enote, torej $x_i \rightarrow \hat{x}_i = \frac{x_i - \bar{x}}{\sigma_x}$ in $y_i \rightarrow \hat{y}_i = \frac{y_i - \bar{y}}{\sigma_y}$. Standardne enote so neodvisne od izhodiščnih enot (metrov, col, itd.). Galton je za mero povezanosti vzel

$$\begin{aligned}\rho &= \frac{1}{n} \sum_{i=1}^n \hat{x}_i \hat{y}_i \\ &= \sum_{i=1}^n \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} \\ &= \frac{1}{\sigma_x \sigma_y} \left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right).\end{aligned}$$

Pri velikem številu ponovitev lahko izračunamo Galtonov ρ . V oklepaju bo kar

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \approx E[(X - E(X))(Y - E(Y))] = \operatorname{cov}(X, Y).$$

Definicija 4.17. Količini

$$\sigma = \frac{\operatorname{cov}(X, Y)}{\sqrt{\operatorname{var}(X)} \sqrt{\operatorname{var}(Y)}}$$

rečemo koleracijski koeficient med X in Y .

Opomba. Cauchy-Schwartzova neenakost pove, da je $-1 \leq \rho \leq 1$. Pove nam tudi, da sta v primeru $\rho \in \{-1, 1\}$ spremenljivki X in Y v linearni zvezi, torej $aX + bY = 0$ za neka a, b .

Definicija 4.18. Za slučajni vektor \underline{X} definiramo

$$E(\underline{X}) = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{pmatrix}.$$

Rečemo, da pričakovana vrednost obstaja, če obstaja pričakovana vrednost vseh komponent.

Definicija 4.19. Naj bosta $\underline{X}, \underline{Y}$ slučajna vektorja. Kovarianco definiramo kot

$$\text{cov}(\underline{X}, \underline{Y}) = \begin{pmatrix} \text{cov}(X_1, Y_1) & \cdots & \text{cov}(X_1, Y_n) \\ \vdots & & \vdots \\ \text{cov}(X_m, Y_1) & \cdots & \text{cov}(X_m, Y_n) \end{pmatrix}.$$

Po analogiji je $\text{var}(\underline{X}) = \text{cov}(\underline{X}, \underline{X})$.

Izrek 4.20.

Naj bo \underline{X} naključna matrika velikosti $m \times n$. Za fiksni matriki \underline{A} in \underline{B} potem velja

$$E(\underline{A} \cdot \underline{X} \cdot \underline{B}) = \underline{A} \cdot E(\underline{X}) \cdot \underline{B},$$

kjer $E(\underline{X})$ razumemo po komponentah.

Dokaz. Dokaz sledi naravnost iz linearnosti pričakovane vrednosti. □

Ko računamo z matrikami, sledimo dogovoru, da je \underline{X} stolpec. Opazimo, da velja

$$\text{cov}(\underline{X}, \underline{Y}) = E[\underline{X} \cdot \underline{Y}^\top] - E(\underline{X})E(\underline{Y}^\top).$$

Izrek 4.21.

Naj bosta $\underline{X}, \underline{Y}$ slučajna vektorja. Potem velja

$$\text{cov}(\underline{AX}, \underline{BY}) = \underline{A} \cdot \text{cov}(\underline{X}, \underline{Y}) \cdot \underline{B}^\top.$$

Dokaz. Sledi iz prejšnjega izreka. □

Zgled 4.22. Naj bo \underline{X} slučajna vektorja

$$\underline{X} = \begin{pmatrix} \underline{X}^{(1)} \\ \underline{X}^{(2)} \end{pmatrix} \sim N \left(\underbrace{\begin{pmatrix} \underline{\mu}^{(1)} \\ \underline{\mu}^{(2)} \end{pmatrix}}_{\underline{\mu}}, \underbrace{\begin{pmatrix} \underline{\Sigma}_{11}^{(1)} & \underline{\Sigma}_{12}^{(1)} \\ \underline{\Sigma}_{21}^{(2)} & \underline{\Sigma}_{22}^{(1)} \end{pmatrix}}_{\underline{\Sigma}} \right).$$

Naj bo sedaj

$$\underline{Y} = \underbrace{\begin{pmatrix} \underline{I} & \underline{0} \\ -\underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1} & \underline{I} \end{pmatrix}}_{\underline{A}} \begin{pmatrix} \underline{X}^{(1)} \\ \underline{X}^{(2)} \end{pmatrix}.$$

Iz prejšnjega izreka imamo sedaj

$$\text{var}(\underline{Y}) = \underline{A} \cdot \underline{\Sigma} \cdot \underline{A}^\top = \begin{pmatrix} \underline{\Sigma}_{11} & 0 \\ 0 & \underline{\Sigma}_{22} - \underline{\Sigma}\underline{\Sigma}_{11}^{-1}\underline{\Sigma}_{12} \end{pmatrix}.$$

4.3 Pogojna pričakovana vrednost, pogojna varianca

Oglejmo si najprej diskretni primer. Definirali smo pogojno porazdelitev slučajne spremenljivke X glede na dogodek B kot $P(X = x \mid B) = \frac{P(\{X=x\} \cap B)}{P(B)}$. Naravno je pogojno pričakovano vrednost definirati s pogojno porazdelitvijo.

Definicija 4.23.

1. Naj bo X diskretna slučajna spremenljivka in B dogodek z $P(B) > 0$. Pogojno pričakovano vrednost X glede na dogodek B definiramo kot

$$E(X \mid B) = \sum_x xP(X = x \mid B).$$

2. Naj bo f funkcija. Potem velja

$$E[f(X) \mid B] = \sum_x f(x)P(X = x \mid B).$$

Za diskretne slučajne spremenljivke lahko razumemo $E(X \mid B)$ tudi nekoliko drugače:

$$\begin{aligned} E(x \cdot 1_B) &= \sum_x xP(X \cdot 1_B = x) \\ &= \sum_{x \neq 0} x \cdot P(\{X = x\} \cap B) \\ &= \sum_{x \neq 0} xP(X = x \mid B) \cdot P(B) \\ &= E(X \mid B) \cdot P(B), \end{aligned}$$

torej imamo formulo

$$E(X \mid B) = \frac{1}{P(B)} \cdot E(X \cdot 1_B).$$

Iz tega zapisa takoj sledi, da je

$$\begin{aligned} E(\alpha X + \beta Y \mid B) &= \frac{1}{P(B)} E[(\alpha X + \beta Y) \cdot 1_B] \\ &= \frac{1}{P(B)} (\alpha E(X \cdot 1_B) + \beta E(Y \cdot 1_B)) \\ &= \alpha E(X \mid B) + \beta E(Y \mid B), \end{aligned}$$

torej je tudi pogojna pričakovana vrednost linearna. Največkrat bomo uporabili za dogodek B množico $\{Y = k\}$ za neko diskretno slučajno spremenljivko. V tem primeru bomo pisali

$$E(X \mid Y = y) = E(X \mid \{Y = y\}).$$

Pogojno varianco definiramo na povsem naraven način kot

$$\text{var}(X \mid B) = E(X^2 \mid B) - [E(X \mid B)]^2.$$

Definicija 4.24. Naj bo \underline{X} diskreten slučajni vektor in $f : \mathbb{R}^n \rightarrow \mathbb{R}$ funkcija. Naj bo B dogodek z $P(B) > 0$. Pogojno pričakovano vrednost definiramo kot

$$E[f(\underline{X}) \mid B] = \sum_{\underline{x}} f(\underline{x})P(\underline{X} = \underline{x}).$$

Tipično bo dogodek B oblike $B = \{\underline{Y} = \underline{y}\}$ za nek slučajni vektor \underline{Y} . V tem primeru bomo pisali $E(f(\underline{X}) \mid \underline{Y} = \underline{y})$.

Izrek 4.25.

Naj bo $\{H_1, H_2, \dots\}$ particija Ω . Naj bo \underline{X} diskreten slučajni vektor. Potem velja

$$E[f(\underline{X})] = \sum_k E[f(\underline{X}) \mid H_k] P(H_k).$$

Opomba. Formuli rečemo tudi formula za popolno pričakovano vrednost.

Dokaz.

$$\begin{aligned} \sum_k E[f(\underline{X}) \mid H_k] \cdot P(H_k) &= \sum_k \sum_{\underline{x}} f(\underline{x}) P(\underline{X} = \underline{x} \mid H_k) P(H_k) \\ &= \sum_{\underline{x}} \sum_k f(\underline{x}) P(\underline{X} = \underline{x} \mid H_k) P(H_k) \\ &= \sum_{\underline{x}} f(\underline{x}) \sum_k P(\underline{X} = \underline{x} \mid H_k) P(H_k) \\ &= \sum_{\underline{x}} f(\underline{x}) P(\underline{X} = \underline{x}). \end{aligned}$$

□

Zgled 4.26. Igralcema A in B razdelimo po 5 kart z dobro premešanega kupa kart. Naj bo X število asov prvega igralca in B število asov drugega. Vemo, da je

$$Y \mid X = k \sim \text{HiperGeom}(5, 4 - k, 47),$$

zato sledi $E(Y \mid X = k) = 5 \cdot \frac{4-k}{47}$ in

$$\text{var}(Y \mid X = k) = 5 \cdot \frac{4-k}{47} \cdot \left(1 - \frac{4-k}{47}\right) \frac{47-5}{47-1}.$$

Zgled 4.27. Kovanec mečemo, dokler ne dobimo r grbov zapovrstjo, kjer je r dan vnaprej. Naj bo X število potrebnih metov. Za particijo vzamemo

$$H_k = \{\text{prva številka se pojavi na } k\text{-tem metu}\}.$$

Če je p verjetnost grba, je $P(H_k) = p^{k-1} \cdot q$ (geometrijska porazdelitev). Sedaj za $k \geq r+1$ velja $E(X \mid H_k) = r$, za $k \leq r$ pa je $E(X \mid H_k) = k + E(X)$, saj se igra po prvi številki resetira. Po formuli za pričakovano verjetnost je

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} E(X \mid H_k) \cdot P(H_k) \\ &= \sum_{k=1}^r (k + E(X)) P(H_k) + \sum_{k=r+1}^{\infty} E(X \mid H_k) P(H_k) \\ &= \sum_{k=1}^r k p^{k-1} q + E(X) \sum_{k=1}^r p^{k-1} q + r \sum_{k=r+1}^{\infty} p^{k-1} q \\ &= q \cdot \frac{1 - p^r(r+1) + r \cdot p^{r+1}}{(1-p)^2} + E(X) \cdot (1 - p^r) + r p^r, \end{aligned}$$

s tem pa dobimo

$$E(X) \cdot (1 - (1 - p^r)) = q \cdot \frac{1 - p^r(r+1) + rp^{r+1}}{(1-p)^2} + p^r$$

in od tod $E(X) = \frac{1-p^r}{q \cdot p^r}$. Ena izmed rešitev linearne enačbe je $E(X) = \infty$. Razdelimo mete v bloke po r metov, kjer je

$$B_k = \{v \text{ } k\text{-tem bloku dobimo } r \text{ grbov}\}.$$

Dogodki B_k so neodvisni in imajo verjetnost p^r . Sedaj opazimo, da velja

$$\{\text{nikoli ni } r \text{ grbov zapored}\} \subseteq \bigcup_{k=1}^{\infty} B_k^c.$$

Dogodki B_1, B_2, \dots so neodvisni, zato so neodvisni tudi dogodki B_1^c, B_2^c, \dots . Sedaj pa naredimo hiter račun in dobimo

$$\begin{aligned} P\left(\bigcap_{k=1}^{\infty} B_k^c\right) &\leq P\left(\bigcap_{k=1}^N B_k^c\right) \\ &= \prod_{k=1}^N P(B_k^c) \\ &= (1 - p^r)^N. \end{aligned}$$

Ker ta ocena velja za vsak $N \in \mathbb{N}$ in je $p > 0$, je res verjetnost, da nikoli ne vržemo r grbov zapored, enaka 0. Če je Y število prvega bloka s samimi grbi, je $Y \sim \text{Geom}(p^r)$ in imamo $E(X) \leq r \cdot E(Y) < \infty$. Po istem argumentu je za $k \geq r+1$ pričakovana vrednost $E(X^2 \mid H_k) = r^2$ in za $k \leq r$ je $E(X^2 \mid H_k) = E[(k+X)^2]$ in zato

$$E(X^2) = \sum_{k=1}^r E[(X+k)^2]P(H_k) + r^2 \sum_{k=r+1}^{\infty} P(H_k).$$

Iz tega načeloma lahko izračunamo $E(X^2)$. Upoštevamo $E(X^2) \leq r^2 E(Y^2) < \infty$.

Sedaj se lotimo zveznega primera. Naj bo $f_{X,Y}$ gostota slučajnega vektorja (X, Y) . Točko (X, Y) si mislimo naključno izbrano točko v ravnini. Sedaj izvemo njeno x -koordinato. Predvidevamo, da je „pogojna gostota“ proporcionalna funkciji $y \mapsto f_{X,Y}(x, y)$, kjer je x fiksni. Gostote, tudi pogojne, bi se morale zintegrirati v 1. Iz formule za robno gostoto vemo

$$\int_{-\infty}^{\infty} f_{X,Y}(x, y) dy = f_X(x).$$

Definicija 4.28. Za $\underline{X} \in \mathbb{R}^n$, za katerega je $f_{\underline{X}}(\underline{x}) > 0$, definiramo pogojno gostoto \underline{Y} glede na $\{\underline{X} = \underline{x}\}$ kot

$$f_{\underline{Y} \mid \underline{X}=\underline{x}}(\underline{y}) = \frac{1}{f_{\underline{X}}(\underline{x})} f_{\underline{X}, \underline{Y}}(\underline{x}, \underline{y}).$$

Definicija 4.29. Pogojno pričakovano vrednost $f(\underline{Y})$ pogojno na $\{\underline{X} = \underline{x}\}$ za \underline{x} , za katere je $f_{\underline{X}}(\underline{x}) > 0$, definiramo kot

$$E[f(\underline{Y}) \mid \underline{X} = \underline{x}] = \int_{\mathbb{R}^n} f(\underline{y}) f_{\underline{Y} \mid \underline{X}=\underline{x}}(\underline{y}) d\underline{y}.$$

Opomba. Če je $f_{\underline{X}}(\underline{x}) = 0$, pogojne gostote ne poznamo, saj bi pogojevali glede na dogodek, ki se ne zgodi.

Zgled 4.30. Naj bo $f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-q^2}} e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}$. Vemo, da je $f_X(x) = \frac{1}{\sqrt{2\pi}e^{-\frac{x^2}{2}}}$ in zato

$$f_{Y \mid X=x}(y) = \frac{1}{\sqrt{2\pi}\sqrt{1-q^2}} e^{-\frac{(e-\rho x)^2}{(1-q^2)}}.$$

S tem smo dokazali $Y \mid X = x \sim N(\rho x, 1-\rho^2)$. Posledično je $E(Y \mid X = x) = \rho x$ in $\text{var}(Y \mid X = x) = 1 - \rho^2$.

Pri tem smo uporabili

$$\text{var}(Y \mid \underline{X} = \underline{x}) = E(Y^2 \mid \underline{X} = \underline{x}) - [E(Y \mid \underline{X} = \underline{x})]^2$$

in

$$\text{cov}(Y_1, Y_2 \mid \underline{X} = \underline{x}) = E(Y_1 Y_2 \mid \underline{X} = \underline{x}) - E(Y_1 \mid \underline{X} = \underline{x})E(Y_2 \mid \underline{X} = \underline{x}).$$

Zgled 4.31. Naj bo

$$\underline{X} = \begin{pmatrix} \underline{X}^{(1)} \\ \underline{X}^{(2)} \end{pmatrix} \sim N \left(\begin{pmatrix} \underline{\mu}^{(1)} \\ \underline{\mu}^{(2)} \end{pmatrix}, \begin{pmatrix} \underline{\Sigma}_{11}^{(1)} & \underline{\Sigma}_{12}^{(1)} \\ \underline{\Sigma}_{21}^{(1)} & \underline{\Sigma}_{22}^{(1)} \end{pmatrix} \right).$$

Zanima nas $f_{\underline{X}^{(2)} \mid \underline{X}^{(1)}=\underline{x}^{(1)}}(\underline{x}^{(2)})$. Načeloma je to enako

$$f_{\underline{X}^{(2)} \mid \underline{X}^{(1)}=\underline{x}^{(1)}}(\underline{x}^{(2)}) = \frac{f_{\underline{X}}(\underline{x})}{f_{\underline{x}^{(1)}}(\underline{x}^{(1)})},$$

česar pa ne izračunamo zlahka. Vemo pa, da sta $\underline{X}^{(1)}$ in $\underline{X}^{(2)} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{X}^{(1)}$ sta neodvisna in $\underline{X}^{(1)} \sim N(\underline{\mu}^{(1)}, \underline{\Sigma}_1)$ in

$$Y = \underline{X}^{(2)} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{X}^{(1)} \sim N(\underline{\mu}^{(2)} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{\mu}^{(1)}, \underline{\Sigma}_{22} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{\Sigma}_{12}).$$

Izrazimo

$$\begin{pmatrix} \underline{X}^{(1)} \\ \underline{X}^{(2)} \end{pmatrix} = \begin{pmatrix} \underline{X}^{(1)} \\ \underline{Y} + \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{X}^{(1)} \end{pmatrix}.$$

Po transformacijski formuli sledi

$$f_{\underline{X}}(\underline{x}) = f_{\underline{X}^{(1)}, \underline{X}^{(2)}}(\underline{X}^{(1)}, \underline{X}^{(2)}) = f_{\underline{X}^{(1)}}(\underline{X}^{(1)}) \cdot f_{\underline{Y}}(\underline{X}^{(2)} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{X}^{(1)}) \cdot |J_{\Phi}|.$$

Od tod sledi

$$f_{\underline{X}^{(2)} \mid \underline{X}^{(1)}=\underline{x}^{(1)}}(\underline{x}^{(2)}) = f_Y(\underline{y}^{(2)} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{x}^{(1)}).$$

Sedaj pa sklepamo

$$\underline{X}^{(2)} \mid \underline{X}^{(1)} = \underline{x}^{(1)} \sim N(\underline{\mu}^{(2)} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}(\underline{x}^{(1)} - \underline{\mu}^{(1)}), \underline{\Sigma}_{22} - \underline{\Sigma}_{21}\underline{\Sigma}_{11}^{-1}\underline{\Sigma}_{12}).$$

Izrek 4.32.

Velja

$$E[f(\underline{X})] = \int_{\mathbb{R}^n} E[f(\underline{X}) \mid \underline{X} = \underline{x}] f_{\underline{X}}(\underline{x}) d\underline{x}.$$

5 Rodovne funkcije

5.1 Definicije in osnovne lastnosti

Ideja rodovnih funkcij je, da porazdelitev nenegativne, celoštevilске slučajne spremenljivke zapakiramo v funkcijo.

Definicija 5.1. Naj bo X nenegativna celoštevilska slučajna spremenljivka. Njena rodovna funkcija $G_X(s)$ je definirana kot

$$G_X(s) = \sum_{k=0}^{\infty} s^k P(X = k).$$

Opomba. 1. Potenčna vrsta za $s \in [-1, 1]$ konvergira enakomerno, saj je dominirana s konvergentno vrsto $\sum_{k=0}^{\infty} P(X = k) = 1$.

2. Za $s \in (-1, 1)$ je $G_X(s)$ neskončnokrat odvedljiva (analiza 1).

Vemo, da je $E(f(X)) = \sum_{k=0}^{\infty} f(k)P(X = k)$. Vzemimo $f(k) = s^k$ za $s \in [-1, 1]$. Potem dobimo

$$E(f(X)) = E(s^X) = \sum_{k=0}^{\infty} s^k P(X = k) = G_X(s).$$

Zgled 5.2. Naj bo $X \sim \text{Bin}(n, p)$. Potem je

$$\begin{aligned} G_X(s) &= \sum_{k=0}^n s^k P(X = k) \\ &= \sum_{k=0}^n s^k \binom{n}{k} p^k q^{n-k} \\ &= (ps + q)^n. \end{aligned}$$

Zgled 5.3. Sedaj pa vzemimo $X \sim \text{NegBin}(m, p)$. Spomnimo se, da je $(1 - x)^{-a} = \sum_{k=0}^{\infty} \frac{(a)_k}{k!} x^k$ za $|x| < 1$. Računamo

$$\begin{aligned} G_X(s) &= \sum_{k=m}^{\infty} s^k \binom{k-1}{m-1} p^m \cdot q^{k-m} \\ &= (ps)^m \sum_{k=0}^{\infty} \binom{m+k-1}{m-1} q^k s^k \\ &= (ps)^m \cdot \sum_{k=0}^{\infty} \frac{(m)_k}{k!} q^k s^k \\ &= (ps)^m \cdot (1 - qs)^{-m} \\ &= \left(\frac{ps}{1 - qs} \right)^m. \end{aligned}$$

Zgled 5.4. Naj bo $X \sim \text{Po}(\lambda)$. Potem je

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

Zgled 5.5. Naredimo enako še za $X \sim \text{Pólya}(a, \beta)$, kjer je

$$P(X = k) = \frac{\beta^a (a)_k}{k! (1 + \beta)^{a+k}}$$

za vsak $k = 1, 2, \dots$. Potem je

$$\begin{aligned} G_X(s) &= \sum_{k=0}^{\infty} \frac{\beta^a (a)_k}{k! (1 + \beta)^{a+k}} s^k \\ &= \frac{\beta^a}{(1 + \beta)^a} \sum_{k=0}^{\infty} \frac{(a)_k}{k!} s^k (1 + \beta)^{-k} \\ &= \frac{\beta^a}{(1 + \beta)^a} \left(1 - \frac{s}{1 + \beta} \right)^{-a} \\ &= \left(\frac{\beta}{1 + \beta - s} \right)^a. \end{aligned}$$

Izrek 5.6.

Naj bo X nenegativna, celoštevilka slučajna spremenljivka z rodovno funkcijo G_X . Rodovna funkcija enolično določa porazdelitev X .

Dokaz. Iz teorije potenčnih vrst vemo $P(X = k) = \frac{G_X^{(k)}(0)}{k!}$. □

Izrek 5.7.

Če sta X, Y nenegativni celoštevilski neodvisni slučajni spremenljivki, je $G_{X+Y} = G_X \cdot G_Y$.

Dokaz. Vemo, da za neodvisni X, Y velja $E[f(X), g(Y)] = E[f(X)] \cdot E[g(Y)]$. V našem primeru je

$$\begin{aligned} G_{X+Y}(s) &= E[s^{X+Y}] \\ &= E[s^X \cdot s^Y] \\ &= E(s^X) \cdot E(s^Y) \\ &= G_X(s) \cdot G_Y(s). \end{aligned} \quad \square$$

Opomba. Zgornji izrek seveda velja tudi za vsote n neodvisnih spremenljivk X_1, X_2, \dots, X_n .

Zgled 5.8. Naj bosta X, Y neodvisni, $X \sim \text{Bin}(m, p)$ in $Y \sim \text{Bin}(n, p)$. Potem je

$$\begin{aligned} G_{X+Y}(s) &= G_X(s) \cdot G_Y(s) \\ &= (ps + q)^m \cdot (ps + q)^n \\ &= (ps + q)^{m+n}, \end{aligned}$$

torej je $X + Y \sim \text{NegBin}(m + n, p)$.

Zgled 5.9. Naj bosta spet X, Y neodvisni in pa $X \sim \text{Po}(\lambda)$, $Y \sim \text{Po}(\mu)$. Potem je

$$\begin{aligned} G_{X+Y}(s) &= G_X(s) \cdot G_Y(s) \\ &= e^{-\lambda(1-s)} \cdot e^{-\mu(1-s)} \\ &= e^{-(\lambda+\mu)(1-s)} \end{aligned}$$

in posledično $X + Y \sim \text{Po}(\lambda + \mu)$.

Zgled 5.10. Naj bosta X, Y neodvisni, $X \sim \text{Pólya}(a, \beta)$ in $Y \sim \text{Pólya}(b, \beta)$. Potem je

$$\begin{aligned} G_{X+Y}(s) &= G_X(s) \cdot G_Y(s) \\ &= \left(\frac{\beta}{1 + \beta - s} \right)^a \left(\frac{\beta}{1 + \beta - s} \right)^b \\ &= \left(\frac{\beta}{1 + \beta - s} \right)^{a+b}, \end{aligned}$$

torej sledi $X + Y \sim \text{Pólya}(a + b, \beta)$.

Zgled 5.11. Naj bodo X_1, X_2, \dots, X_n neodvisne in $X_k \sim \text{Geom}(p)$. Po enem od prejšnjih zgledov imamo $G_{X_k}(s) = \frac{ps}{1-qs}$, torej

$$G_{X_1+X_2+\dots+X_m}(s) = \frac{ps}{1-qs} \cdots \frac{ps}{1-qs} = \left(\frac{ps}{1-qs} \right)^m.$$

Zato lahko zapišemo $X_1 + \dots + X_n \sim \text{NegBin}(m, p)$.

Če napišemo $G_X(s) = \sum_{k=0}^{\infty} s^k \cdot P(X = k)$ in odvajamo po s za $s \in (-1, 1)$, dobimo

$$G'_X(s) = \sum_{k=0}^{\infty} k \cdot s^{k-1} P(X = k).$$

Kaj se zgodi, ko gre $s \uparrow 1$? Pričakovali bi, da je $\lim_{s \uparrow 1} G'_X(s) = E(X)$.

Izrek 5.12.

Naj bo X nenegativna celoštevilska slučajna spremenljivka, potem velja

- $E(X) = \lim_{s \uparrow 1} G'_X(s)$,
- $E[X(X-1)\dots(X-k+1)] = \lim_{s \uparrow 1} G^{(k)}(s)$.

Dokaz. Dovolj je, da dokažemo prvo točko, saj je dokaz druge popolnoma enak. Velja neenačba za $s \in (0, 1)$ in

$$\sum_{k=0}^N k s^{k-1} P(X = k) \leq G'_X(s) \leq E(X).$$

Sedaj pošljimo $s \uparrow 1$. Ker je G'_X naraščajoča na $(0, s)$ (potenčna vrsta z nenegativnimi koeficienti), potem limita $\lim_{s \uparrow 1} G'_X(s)$ obstaja. Iz tega sledi

$$\sum_{k=0}^N k P(X = k) \leq \lim_{s \uparrow 1} G'_X(s) \leq E(X).$$

če pa sedaj pošljemo $N \rightarrow \infty$, dobimo $E(X) \leq \lim_{s \uparrow 1} G'_X(s) \leq E(X)$ in trditev sledi. Če je $E(X) = \infty$, potem so delne vsote $\sum_{k=0}^N k P(X = k)$ poljubno velike. To pomeni $\lim_{k \rightarrow \infty} k = (X = k) \leq \lim_{s \uparrow 1} G'_X(s)$, torej $\lim_{s \uparrow 1} G'_X(s) = \infty$. \square

V uporabah večkrat naletimo na seštevanje slučajnega števila slučajnih spremenljivk. Kot bomo videli, ta koncept naravno nastane v zavarovalništvu. Formlano naj bodo X_1, X_2, \dots slučajne spremenljivke in $N \geq 0$ slučajna spremenljivka. Definiramo $S : \Omega \rightarrow \mathbb{R}$ z vsoto

$$S(\omega) = \sum_{k=1}^{N(\omega)} X_k(\omega),$$

pri čemer vsoto za $N(\omega) = 0$ interpretiramo kot 0. Za zgornjo vsoto bomo uporabili intuitiven simbol $X_1 + X_2 + \dots + X_N$.

Izrek 5.13.

Naj bodo slučajne spremenljivke N, X_1, X_2, \dots med seboj neodvisne in naj bodo X_1, X_2, \dots enako porazdeljene. Potem velja

$$G_{X_1 + \dots + X_N}(s) = G_N(G_{X_1}(s)).$$

Dokaz. Uporabili bomo dejstvo, da za neodvisni X in Y velja $E(f(Y) \mid X = x) = E[f(Y)]$. Sedaj računamo po fomruli za popolno pričakovano vrednost:

$$\begin{aligned} E(s^{X_1 + \dots + X_N}) &= \sum_{n=0}^{\infty} E(s^{X_1 + \dots + X_n} \mid N = n) P(N = n) \\ &= \sum_{n=0}^{\infty} E(s^{X_1 + \dots + X_n} \mid N = n) P(N = n) \\ &= \sum_{n=0}^{\infty} E(s^{X_1 + \dots + X_n}) P(N = n) \\ &= \sum_{n=0}^{\infty} [G_{X_1}(s)]^n P(N = n) \\ &= G_N(G_{X_1}(s)). \end{aligned} \quad \square$$

Zgled 5.14. *Kokoš znese N jajc, kjer je $N \sim \text{Po}(\lambda)$. Iz vsakega jajca se zleže piščanec z verjetnostjo p , neodvisno od N . Naj bo X končno število piščancev. Kakšna je porazdelitev X ? Imamo torej neodvisne enako porazdeljene indikatorje I_1, I_2, \dots z verjetnostjo $P(I_k = 1) = p$. Gledamo spremenljivko $X = I_1 + I_2 + \dots + I_N$, kjer so vsi indikatorji neodvisni tudi od N . Potem velja*

$$G_X(s) = G_N(G_{I_1}(s)),$$

pri čemer je $G_{I_1}(s) = q + ps$ in $G_N(s) = e^{-\lambda(1-s)}$. Od tod sledi

$$\begin{aligned} G_X(s) &= e^{-\lambda(1-q-ps)} \\ &= e^{-\lambda p(1-s)} \end{aligned}$$

in tako sklepamo $X \sim \text{Po}(\lambda p)$.

Zgled 5.15. Mečemo kovanec in čakamo na r grbov zapovrstjo. Vemo, da ima za $k \leq r$ slučajna spremenljivka X (število potrebnih metov) pogojno na

$$H_k = \{\text{prva številka pade na } k\text{-tem metu}\}$$

pogojno porazdelitev enako porazdelitvi $X + k$. Od tod sledi

$$\begin{aligned} G_X(s) &= \sum_{k=1}^{\infty} E(s^X \mid H_k) P(H_k) \\ &= \sum_{k=1}^r E(s^{X+k}) P(H_k) + \sum_{k=r+1}^{\infty} s^r P(H_k) \\ &= \sum_{k=1}^r s^k G_X(s) \cdot p^{k-1} \cdot q + s^r \cdot n \\ &= \frac{s^r p^n}{1 - sq(1 + sp + \dots + s^{r-1} p^{r-1})}. \end{aligned}$$

Načeloma lahko iz tega izračunamo $\frac{G_X^{(k)}(0)}{k!} = P(X = k)$, vendar je to komplicirano. Poskusimo vsaj za $r = 2$ in $p = \frac{1}{2}$. V tem primeru je

$$G_X(s) = \frac{s^2}{4 - 2s - s^2}.$$

Dovolj je, če razvijemo v vrsto funkcijo

$$f(s) = \frac{1}{4 - 2s - s^2} = \sum_{k=0}^{\infty} c_k s^k.$$

Z množenjem ugotovimo, da je $c_0 = \frac{1}{4}$, $c_1 = \frac{1}{8}$ in

$$4c_k - 2c_{k-1} - c_{k-2} = 0.$$

Nastavek za rešitev te diferenčne enačbe je

$$c_k = \alpha \left(\frac{1 + \sqrt{5}}{4} \right)^k + \beta \left(\frac{1 - \sqrt{5}}{4} \right)^k.$$

Iz začetnih pogojev sledi

$$c_k = \frac{(5 + \sqrt{5})}{40} \cdot \left(\frac{1 + \sqrt{5}}{4} \right)^k + \frac{(5 - \sqrt{5})}{40} \cdot \left(\frac{1 - \sqrt{5}}{4} \right)^k$$

in končno $P(X = k) = c_{k-2}$ za $k = 2, 3, \dots$

5.2 Procesi razvejanja

Zgled 5.16. *Sir Francis Galton (1822 - 1911) je leta 1873 postavil naslednje vprašanje: vzemimo viktorijanskega aristokrata, ki bo imel slučajno število sinov. Nato bo vsak od sinov spet imel slučajno število sinov, in tako dalje. Kolikšna je verjetnost, da bo rodbina izumrla? Odgovor sta našla F. Galton in H.W. Watson leta 1875 v članku *On the probability of extinction of families*, objavljenem v znanstveni reviji *Journal of the Royal Anthropological Institute*.*

Za namene matematične obravnave potrebujemo bolj natančne predpostavke:

- generacije so simultane,
- število sinov vsakega posameznika v dani generaciji je neodvisno od števila sinov ostalih posameznikov v isti generaciji,
- porazdelitev števila sinov je enaka za vse posameznike.

Označimo število posameznikov v n -ti generaciji z Z_n . Po predpostavki je $Z_0 = 1$. Naj bodo $\{\xi_{n,k}\}_{n \geq 1, k \geq 1}$ med sabo neodvisne enako porazdeljene slučajne spremenljivke z rodovno funkcijo G , ki bodo predstavljale število sinov. Sedaj lahko definiramo rekurzivno zvezo

$$Z_{n+1} = \xi_{n+1,1} + \xi_{n+1,2} + \cdots + \xi_{n+1,Z_n}.$$

Iz konstrukcije sledi, da je Z_n neodvisna od $\xi_{n+1,1}, \xi_{n+1,2}, \dots$. Po prejšnjem izreku velja

$$G_{n+1}(s) := G_{Z_{n+1}}(s) = G_{Z_n}(G(s)).$$

Od tod sledi $G_1(s) = G(s)$, $G_2(s) = G_1(G(s))$ in tako naprej:

$$G_n(s) = \underbrace{(G \circ G \circ \cdots \circ G)}_n(s).$$

Iz asociativnosti kompozituma sledi rekurzivna zveza

$$G_{n+1}(s) = G(G_n(s)).$$

Definicija 5.17. Takemu zaporedju slučajnih spremenljivk Z_0, Z_1, \dots pravimo proces razvejanja (angl. *branching process*).

Opomba. V verjetnosti pogosto pravimo zaporedju med seboj odvisnih slučajnih spremenljivk X_0, X_1, \dots proces.

Označimo dogodek $A = \{\text{rodbina izumre}\} = \bigcup_{n=1}^{\infty} \{Z_n = 0\}$. Vemo tudi, da velja $\{Z_n = 0\} \subseteq \{Z_{n+1} = 0\}$. Označimo sedaj $\eta = P(A)$. Računamo:

$$\begin{aligned} \eta = P(A) &= \lim_{n \rightarrow \infty} P(Z_n = 0) \\ &= \lim_{n \rightarrow \infty} G_n(0) \\ &= \lim_{n \rightarrow \infty} G(G_{n-1}(0)) \\ &= G\left(\lim_{n \rightarrow \infty} G_{n-1}(0)\right) \\ &= G(\eta), \end{aligned}$$

pri čemer smo uporabili dejstvo, da je G zvezna na $[-1, 1]$ in je $|G_X(s)| \leq 1$. Verjetnost izumrtja je torej fiksna točka rodovne funkcije G na intervalu $[0, 1]$. Ker je $G(1) = 1$, vsaj ena taka točka obstaja, ni pa nujno edina.

Izrek 5.18.

Verjetnost $\eta = P(A)$ je najmanjša fiksna točka funkcije G na intervalu $[0, 1]$.

Dokaz. Množica točk $\{\bar{\eta}; \bar{\eta} = G(\bar{\eta}), \bar{\eta} \in [0, 1]\}$ je neprazna in zaprta, neprazna in zaprta, zato ima najmanjši element na $[0, 1]$. Naj bo $\bar{\eta}$ neka negibna točka. Vemo, da velja $0 \leq \bar{\eta}$. Ker je G nepadajoča na $[0, 1]$, je

$$G(0) \leq G(\bar{\eta}) = \bar{\eta}.$$

Če na tej enačbi še enkrat uporabimo G , dobimo

$$G(G(0)) \leq G(\bar{\eta}) \leq \bar{\eta}.$$

Z iteracijo dobimo $G_n(0) \leq \bar{\eta}$, od koder sledi

$$\eta = \lim_{n \rightarrow \infty} G_n(0) \leq \bar{\eta}.$$

Pokazali smo, da je verjetnost η fiksna točka G in hkrati manjša ali enaka vsem ostalim fiksnim točkam. S tem je dokaz zaključen. \square

Zgled 5.19. Denimo, da je $G(s) = \frac{1}{4}(1 + s + s^2 + s^3)$, s čimer bi vsak posameznik imel 0, 1, 2 ali 3 sinove z enako verjetnostjo. Rešitve enačbe

$$G(s) = \frac{1}{4}(1 + s + s^2 + s^3) = s$$

so $s = 1$, $s = -1 - \sqrt{2}$ in $s = -1 + \sqrt{2}$. Na intervalu $[0, 1]$ sta fiksni točki $s = 1$ in $s = -1 + \sqrt{2}$, torej je $\eta = -1 + \sqrt{2} = 0.41$.

V praksi je kompozitum $G \circ G \circ \dots \circ G$ težko izračunati. Vzemimo $G(s) = \frac{p}{1-qs}$. Hitro vidimo, da bo

$$G_n(s) = \frac{a_n - b_n s}{c_n - d_n s}.$$

Ker je

$$G_{n+1} = \frac{a_n - b_n \cdot \frac{p}{1-qs}}{c_n - d_n \cdot \frac{p}{1-qs}}.$$

Zmnožimo in sledi

$$a_{n+1} = a_n - pb_n, \quad b_{n+1} = qa_n.$$

Ker je $G_0(s) = 1$, je $a_0 = 0$ in $b_0 = -1$. Sistem sedaj zapišemo v matrični obliki:

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & -p \\ q & 0 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix}.$$

Privzemimo najprej, da velja $p \neq q$. Potem lahko matriko diagonaliziramo:

$$\begin{pmatrix} 1 & -p \\ q & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} p & 1 \\ q & 1 \end{pmatrix}}_A \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix} \underbrace{\frac{1}{p-q} \begin{pmatrix} 1 & -1 \\ -q & p \end{pmatrix}}_{A^{-1}}$$

in jo potenciramo:

$$\begin{pmatrix} 1 & -p \\ q & 0 \end{pmatrix}^n = \begin{pmatrix} p & 1 \\ q & 1 \end{pmatrix} \begin{pmatrix} p^n & 0 \\ 0 & q^n \end{pmatrix} \frac{1}{p-q} \begin{pmatrix} 1 & -1 \\ -q & p \end{pmatrix}.$$

Če vse to skupaj zmnožimo, dobimo

$$\begin{pmatrix} 1 & -p \\ q & 0 \end{pmatrix}^n = \frac{1}{p-q} \begin{pmatrix} p^{n+1} - q^{n+1} & -p^{n+1} + pq^n \\ qp^n - q^{n+1} & -qp^n + pq^n \end{pmatrix}.$$

Od tod pa sledi

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 1 & -p \\ q & 0 \end{pmatrix}^n \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \frac{1}{p-q} \begin{pmatrix} p(p^n - q^n) \\ pq(p^{n-1} - q^{n-1}) \end{pmatrix}$$

in podobno

$$\begin{pmatrix} c_n \\ d_n \end{pmatrix} = \begin{pmatrix} 1 & -p \\ q & 0 \end{pmatrix}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{p-q} \begin{pmatrix} p^{n+1} - q^{n+1} \\ q(p^n - q^n) \end{pmatrix}.$$

Tako dobimo rodovno funkcijo

$$G_n(s) = \frac{p(p^n - q^n - qs(p^{n-1} - q^{n-1}))}{p^{n+1} - q^{n+1} - qs(p^n - q^n)}.$$

Opazimo, da če velja $p > q$, potem je $\lim_{n \rightarrow \infty} G_n(0) = 1$, če pa je $p < q$, pa dobimo

$$\lim_{n \rightarrow \infty} G_n(0) = \frac{p}{q} < 1.$$

Fiksni točki $G(s) = \frac{p}{1-qs} = s$ sta

$$s_{1,2} = \frac{1 \pm |1 - 2p|}{2q}.$$

Če je $p < q$, je $p < \frac{1}{2}$, torej je $\frac{p}{q}$ fiksna točka.

Opomba. Za $p = q = \frac{1}{2}$ dobimo

$$G_n(s) = \frac{n - (n-1)s}{n+1 - ns}$$

in posledično

$$\eta = \lim_{n \rightarrow \infty} \frac{n}{n+1} = 1.$$

Izrek 5.20.

Naj bo Z_0, Z_1, \dots proces razvejanja z rodovno funkcijo G . Naj bo $\mu = \lim_{s \uparrow 1} G'(s)$.

1. Če je $\mu < 1$, je $\eta = 1$.
2. Če je $\mu > 1$, je $\eta \in [0, 1)$.
3. Če je $\mu = 1$ in je $G(s) \neq s$, je $\eta = 1$.

Dokaz. 1. Če je $\mu < 1$, je $G'(s) \leq \mu < 1$ za $s \in (0, 1)$. Po Lagrangevem izreku je za $s \in (0, 1)$:

$$\begin{aligned} 1 - G(s) &= G'(\xi)(1 - s) \\ &\leq \mu(1 - s) \\ &< 1 - s, \end{aligned}$$

torej je $G(0) > s$ in s ni fiksna točka.

2. Obstaja $s_0 \in (0, 1)$, da za $s \in (s_0, 1)$ velja $G'(s) > 1$. Ponovno uporabimo prejšnji argument, da dobimo neenakost

$$\begin{aligned} 1 - G(s_0) &= G'(\xi)(1 - s_0) \\ &= 1 - s_0. \end{aligned}$$

Sledi $G(s_0) < s_0$ in $G(0) \geq 0$. Zaradi zveznosti obstaja fiksna točka $\eta \in [0, s_0)$.

3. Če je $G(s) \neq s$, je ali $G(s) = 1$ ali pa je G strogo konveksna na $(0, 1)$. V prvem primeru je očitno $\eta = 1$. V drugem primeru pa je G' strogo naraščajoča na $(0, 1)$, kar pomeni $G'(s) < 1$

za $s \in (0, 1)$. Od tod sledi

$$\begin{aligned} 1 - G(s) &= G'(\xi)(1 - s) \\ &< 1 - s, \end{aligned}$$

kar pa je ekvivalentno $G(s) > s$ za $s \in (0, 1)$. □

5.3 Panjerjeva rekurzija

V zavarovalništvu nas pogosto zanima porazdelitev vsote $X_1 + X_2 + \dots + X_N$. Spremenljivko N tukaj razumemo kot število škod, slučajne spremenljivke X_1, X_2, \dots pa kot višine škod. Vsota je potem celotna škoda. Za namene upravljanja s tveganjem je treba vedeti porazdelitev te vsote. Če je X_1, X_2, \dots enako porazdeljene in med seboj neodvisne ter neodvisne od N , potem je

$$G_{X_1 + \dots + X_N}(s) = G_N(G_{X_1}(s)).$$

Vendar pa je pogosto težko dobiti koeficiente v analitični obliki. Lahko pa izpeljemo rekurzije.

Definicija 5.21. Slučajna spremenljivka N ima porazdelitev Panjerjevega razreda, če za konstanti $a, b \in \mathbb{R}$ velja

$$P(N = n) = \left(a + \frac{b}{n}\right) P(N = n - 1)$$

za $n = 1, 2, \dots$, pri čemer zahtevamo še $a + b > 0$.

Opomba. V Panjerjev razred spada vrsta znanih porazdelitev:

- Poissonova porazdelitev: $a = 0, b > 0$,
- binomska porazdelitev s parametri p in M : $a = -\frac{p}{q}, b = \frac{(M+1)p}{q}$,
- negativna binomska s parametri p in m : $a = q, b = (m - 1)q$.

Če obe strani zgornje enačbe pomnožimo s s^n in seštejemo od $n = 1$ do ∞ , dobimo

$$\begin{aligned} \sum_{k=1}^{\infty} P(N = n) s^n &= G_N(s) - P(N = 0) \\ &= \sum_{n=1}^{\infty} \left(a + \frac{b}{n}\right) P(N = n - 1) s^n \\ &= as \cdot \sum_{n=1}^{\infty} P(N = n - 1) s^{n-1} + b \cdot \sum_{n=1}^{\infty} P(N = n - 1) \frac{s^n}{n} \\ &= as \cdot G_N(s) + b \cdot \sum_{n=1}^{\infty} P(N = n - 1) \int_0^s u^{n-1} du \\ &= as \cdot G_N(s) + b \cdot \int_0^s \sum_{n=1}^{\infty} P(N = n - 1) u^{n-1} du \\ &= as \cdot G_N(s) + b \int_0^s G_N(u) du. \end{aligned}$$

Vrstni red seštevanja in integriranja za $|s| < 1$ lahko zamenjamo zaradi absolutne in enakomerne konvergence vrste. Od tod sledi

$$G_N(s) - P(N = 0) = as \cdot G_N(s) + b \cdot \int_0^s G_N(u) du.$$

Sedaj za $|s| < 1$ odvajamo in dobimo

$$G'_N(s) = asG'_N(s) + (a+b)G_N(s)$$

ali ekvivalentno

$$G'_N(s)(1-as) = (a+b)G_N(s).$$

Označimo $X = X_1 + X_2 + \dots + X_N$ ter

$$f_r = P(X_1 = r), \quad g_r = P(X = r), \quad p_r = P(N = r)$$

za $r = 0, 1, \dots$ Vemo že, da je

$$G_X(s) = G_N(G_{X_1}(s)).$$

Odvajamo in sledi

$$G'_X(s) = G'_N(G_{X_1}(s))G'_{X_1}(s).$$

Sedaj množimo na levi in desni z $(1 - aG_{X_1}(s))$ in dobimo

$$G'_X(s)(1 - aG_{X_1}(s)) = (a+b)G_X(s)G'_{X_1}(s). \quad (5)$$

Sedaj poračunamo:

$$\begin{aligned} P(X=0) &= P(N=0) + \sum_{n=1}^{\infty} P(N=n) \cdot f_0^n \\ &= G_N(f_0) \\ &= g_0. \end{aligned}$$

Iz analize 1 vemo, da je

$$\left(\sum_{k=0}^{\infty} a_k x^k \right) \left(\sum_{k=0}^{\infty} b_k x^k \right) = \sum_{k=0}^{\infty} c_k x^k,$$

kjer je $c_k = \sum_{l=0}^k a_l b_{k-l}$. V formuli (5) izenačimo koeficiente na levi in desni pri potenci s^n in dobimo

$$(n+1)g_{n+1} - a \sum_{k=0}^n f_k g_{n-k+1} (n-k+1) = (a+b) \sum_{k=0}^n (k+1) f_{k+1} g_{n-k}.$$

Od tod naprej izpeljemo

$$\begin{aligned} (n+1)g_{n+1} - a \cdot f_0 g_{n+1} (n+1) &= a \cdot \sum_{k=1}^n f_k g_{n-k+1} (n-k+1) + (a+b) \cdot \sum_{k=0}^n (k+1) f_{k+1} g_{n-k} \\ &= a \cdot \sum_{k=1}^{n+1} f_k g_{n-k+1} (n-k+1) + (a+b) \cdot \sum_{k=1}^{n+1} k f_k g_{n-k+1} \\ &= a(n+1) \sum_{k=1}^{n+1} f_k g_{n-k+1} + b \sum_{k=1}^{n+1} k f_k g_{n-k+1}. \end{aligned}$$

Sedaj delimo levo in desno stran z $(n+1)(1 - af_0)$ in dobimo

$$g_{n+1} = \frac{1}{1 - af_0} \sum_{k=1}^{n+1} \left(a + \frac{bk}{k+1} \right) f_k g_{n-k+1}. \quad (6)$$

Izpeljeli smo rekurzivno enačbo za g_n – tej rekurziji se reče Panjerjeva rekurzija.

Opomba. Panjerjeva rekurzija se dejansko pogosto uporablja v zavarovalništvu. Verjetnosti f_0, f_1, \dots so približno znane iz prakse, za p_0, p_1, \dots pa se predpostavi katero od znanih porazdelitev, tipično Poissonovo.