

Verjetnostni račun in statistika

Zapiski predavanj 2019/20

Zadnji popravek: 15. januar 2020

Kazalo

1	Uvod v verjetnost	3
1.1	Kratka zgodovina verjetnostnega računa	3
1.1.1	Igre na srečo v antiki in srednjem veku	3
1.1.2	Cardano, Pascal, Fermat in Huygens: zametki matematične teorije	3
1.1.3	Bernoulli, De Moivre in Laplace: klasična verjetnost	3
1.1.4	Razvoj sodobne verjetnosti	3
1.2	Definicija verjetnosti	4
1.2.1	Klasična definicija verjetnosti	4
1.2.2	Statistična definicija verjetnosti	5
1.2.3	Aksiomska definicija verjetnosti	5
1.2.4	Osnovne lastnosti verjetnosti	7
1.2.5	Verjetnost na končnem prostoru izidov	8
1.3	Osnove kombinatoričnega preštevanja	9
1.3.1	Pravilo vsote in produkta	9
1.3.2	Razvrstitve in izbori	10
1.3.3	Načelo vključitve in izključitve	12
2	Pogojna verjetnost in neodvisnost dogodkov	13
2.1	Pogojna verjetnost	13
2.2	Produktno pravilo za pogojno verjetnost	15
2.3	Neodvisni dogodki	16
2.4	Nezdružljivi dogodki in popolna verjetnost	17
2.5	Bayesov izrek o verjetnosti hipotez	18
2.6	Zaporedje neodvisnih poskusov	20
3	Diskretne slučajne spremenljivke	20
3.1	Diskretne in zvezne slučajne spremenljivke	20
3.2	Verjetnostna funkcija in pričakovana vrednost	21
3.3	Funkcije slučajnih spremenljivk in njihova pričakovana vrednost	24
3.4	Nekaj standardnih tipov dss	26
3.4.1	Bernoullijeva dss	27
3.4.2	Enakomerna dss	27
3.4.3	Geometrijska dss	28
3.4.4	Binomska dss	28
3.4.5	Pascalova in negativna binomska dss	29
3.4.6	Hipergeometrijska dss	31
3.4.7	Poissonova dss	32
3.5	Aproksimacija binomske s Poissonovo dss	34

4	Zvezne slučajne spremenljivke	35
4.1	Definicija in zgledi zveznih slučajnih spremenljivk	35
4.1.1	Zvezna enakomerna porazdelitev	36
4.1.2	EkspONENTNA porazdelitev	37
4.1.3	Zgledi nestandardnih zveznih porazdelitev	37
4.1.4	Normalna porazdelitev	38
4.1.5	Paretova porazdelitev	43
4.1.6	Cauchyjeva porazdelitev	43
4.2	Srednje vrednosti: Pričakovana vrednost, mediana in modus	43
4.3	Funkcije zveznih slučajnih spremenljivk	45
4.4	Neenakost Čebiševa	47
5	Slučajni vektorji	48
5.1	Diskretni slučajni vektorji	49
5.1.1	Verjetnostne in porazdelitvene funkcije za dsv	49
5.1.2	Neodvisnost in pogojna porazdelitev dss	50
5.1.3	Vsote, produkti in kovarianca dss	51
5.1.4	Vsota neodvisnih dss	54
5.2	Zvezni slučajni vektorji	55
5.2.1	Osnovni pojmi	55
5.2.2	Neodvisnost zss	56
6	Limitni izreki	58
6.1	Zakon velikih števil	58
6.2	Centralni limitni izrek	59

Predgovor

Zapiski povzemajo vsebino predavanj pri predmetu Verjetnostni račun in statistika v študijskem letu 2019/20. Zapiski niso povsem popolni, v njih manjka uvodno predavanje o zgodovini verjetnosti, natančnejša obravnava normalne porazdelitve, in tu in tam kak dokaz ali zgled, ki sem ga dodal sproti. Verjetno pa bo bralec v zapiskih našel tudi kakšno opombo ali dodaten račun, ki ga na tablo nisem zapisal. Žal se bodo v zapiskih gotovo našle tudi kakšne napake, upam, da ne prehude, da bi jih bralci ne zmogli odkriti in popraviti sami. V okviru uvodnega enosemestrskega predmeta tudi ni veliko prostora za inovativnost, zato vsebina v veliki meri sledi vsebinam iz verjetnosti, kot se predavajo na večini slovenskih fakultet. Kljub temu sem občasno zanemaril nekatere tehnične detajle in namesto njih poskusil z zanimivimi, tudi zgodovinskimi zgledi, poskušal prepričati poslušalce, da je verjetnostni račun zares čudovito področje matematike. Upam, da bo kratek seznam literature na koncu zapiskov bralce spodbudil, da bodo na lastno pest o njem še kaj prebrali.

Avtor

1 Uvod v verjetnost

1.1 Kratka zgodovina verjetnostnega računa

1.1.1 Igre na srečo v antiki in srednjem veku

O tem, kako so starodavna ljudstva razumela in uporabljala naključja, ni povsem jasnih dokazov. Zdi se, da je bil žreb, s katerim so določili izbranega nagrajenca ali žrtev, razumljen predvsem kot izraz volje višje sile.

1.1.2 Cardano, Pascal, Fermat in Huygens: zametki matematične teorije

1.1.3 Bernoulli, De Moivre in Laplace: klasična verjetnost

1.1.4 Razvoj sodobne verjetnosti

1.2 Definicija verjetnosti

Osnovna objekta verjetnosti sta poskus in dogodek. **Poskus** je proces, ki ga lahko (vsaj načeloma) v nespremenjenih okoliščinah velikokrat ponovimo. **Dogodek** je pojav, ki se zgodi kot neposredna posledica poskusa. **Gotov dogodek** se zgodi pri vsaki ponovitvi poskusa, **nemogoč dogodek** se ne zgodi nikoli, **slučajni dogodek** pa se včasih zgodi, včasih pa ne.

Zgled. Pri poskusu "met kocke" je "pade sodo število pik" slučajen dogodek, "pade 7 pik" nemogoč dogodek, "začne deževati" pa ni dogodek.

V vsakdanjem življenju se pogosto opiramo na **subjektivno razumevanje verjetnosti**, ki je odvisno od osebnih izkušenj posameznika. To ponazarjajo naslednje izjave:

- Zgrešil je 100-odstotno priložnost!
- Danes bom gotovo opravil izpit.
- Če danes izgubim, bom naslednjič zmagal.

Osebnne izkušnje so pogosto vezane na specifične okoliščine, zaradi katerih ponovitve poskusov in dogodkov med seboj niso primerljive. Subjektivna ocena verjetnosti je zato lahko le izhodišče za resnejšo obravnavo, pri kateri izhajamo bodisi iz matematične bodisi iz statistične definicije verjetnosti.

1.2.1 Klasična definicija verjetnosti

Klasična definicija verjetnosti¹ izhaja iz predpostavke, da je poskus ponovljiv proces, za katerega lahko vnaprej predvidimo vse možne izide, ki so rezultat poskusa, vsi izidi pa so *enako verjetni*. Tedaj je verjetnost dogodka število med 0 in 1, ki je enako razmerju

$$P = \frac{\text{število ugodnih izidov}}{\text{število vseh izidov}}.$$

Klasična definicija predstavlja na prvi pogled idealno matematično izhodišče, vendar ima precejšnje omejitve: možnih izidov mora biti končno in vsi morajo biti enako verjetni.

Zgled. Pri metu poštene kocke je vsak od šestih izidov enako verjeten. Po klasični definiciji je verjetnost dogodka, da pade sodo število pik, enaka $\frac{3}{6}$.

Zgled. Kakšna je verjetnost, da pri metu dveh poštenih kovancev pade natanko 1 grb? Možni so trije bistveno različni izidi: padeta 2 grba, 1 grb ali 0 grbov. Vendar ti trije izidi niso enako verjetni, zato odgovor ni $1/3$. Za pravilno obravnavo je potrebno kovanca v mislih ločiti na levega in desnega (ali rdečega in modrega). Potem imamo štiri enako verjetne izide: GG, CC, CG ali GC, ugodna pa sta dva od njih, torej je verjetnost enega grba enaka $1/2$.

Zgled. Kovanec mečemo do prvega grba, izid poskusa je število potrebnih metov. Možnih je neskončno izidov G, CG, CCG, CCCG, ..., njihove verjetnosti pa so različne. V tem primeru klasična definicija odpove.

Zgled. Pri metu astragalija padeta izida koza in ovca z verjetnostjo približno 0.1, izida kamela in konj pa približno 0.4. Tudi te preproste igre ne moremo opisati s klasično definicijo.

¹ Definicijo je v tej obliki zapisal Laplace, že pred njim pa je verjetnost na podoben način razumel Bernoulli. Zgodnejši avtorji, na primer Pascal ali Huygens, so namesto o verjetnosti govorili o pričakovanih dobičkih pri stavi.

1.2.2 Statistična definicija verjetnosti

Zaradi zavedanja omejitev klasične definicije je v 19. stoletju v praktični uporabi začela prevladovati statistična definicija verjetnosti. Denimo, da lahko neki poskus velikokrat ponovimo. Posamična ponovitev je ugodna za dogodek A , če se dogodek A pri tej ponovitvi zgodi. **Relativna frekvenca dogodka** f_n je razmerje med številom vseh za dogodek ugodnih ponovitev in številom vseh ponovitev poskusa:

$$f_n = \frac{\text{število ugodnih ponovitev poskusa}}{\text{število vseh ponovitev poskusa, torej } n}.$$

Pri velikem številu ponovitev poskusa ostane to razmerje običajno skoraj konstantno, vselej pa ima vrednost med vključno 0 in 1. To je osnova za naslednjo definicijo.

Statistična (empirična) definicija verjetnosti pravi, da je verjetnost dogodka število, pri katerem se ustali relativna frekvenca dogodka pri dovolj velikem številu poskusov:

$$P = \lim_{n \rightarrow \infty} f_n.$$

Zgled. Pri 1000 metih astragalija je padlo 411 konj, 398 kamel, 99 koz in 92 ovc. Zato ocenjujemo, da je verjetnost konja približno 0.41, kar uporabljamo pri napovedovanju prihodnjih dogodkov.

Zgled. V neki porodnišnici se je lani rodilo 1023 novorojencev, od tega 515 moškega spola. Verjetnost, da bo naslednji tam rojeni otrok fant, je približno $515/1023 \doteq 0.503$.

Statistična definicija verjetnosti je zelo prilagodljiva - izidi lahko imajo različne verjetnosti in sploh ni potrebno poznati vseh, saj se lahko osredotočimo le na dogodke, ki nas zanimajo. Vseeno pa ima svoje omejitve. Kaj pomeni, da se relativna frekvenca ustali? Koliko je dovolj veliko število poskusov? Kaj, če poskusa ne moremo velikokrat ponoviti? Definicija je torej odvisna od vsakega posameznega poskusa in natančnosti, s katero želimo določiti iskano verjetnost. Kljub temu je za praktično rabo v statistiki najprimernejša.

1.2.3 Aksiomska definicija verjetnosti

Aksiomska definicija verjetnosti v jeziku teorije množic je podal Kolmogorov (1933) in predstavlja osnovo sodobne matematične verjetnosti, saj razreši vse omejitve klasične definicije.

Verjetnostni prostor je urejena trojica (S, \mathcal{F}, P) , za katero velja:

- (i) Množica S je poljubna neprazna množica, ki jo imenujemo **prostor izidov** ali **vzorčni prostor**, njene elemente pa **izidi**.
- (ii) Množica \mathcal{F} je taka družina podmnožic množice S , da zanjo velja
 - (a) $S \in \mathcal{F}$.
 - (b) Če je $A \in \mathcal{F}$, je tudi komplement $\bar{A} \in \mathcal{F}$.
 - (c) Če so $A_1, A_2, \dots \in \mathcal{F}$, potem je tudi $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Družino \mathcal{F} imenujemo **σ -algebra** ali **prostor dogodkov**, njene elemente pa **dogodki**.

- (iii) Preslikava $P: \mathcal{F} \rightarrow \mathbb{R}$ ima naslednje lastnosti:

- (a) $P(A) \geq 0$ za vse $A \in \mathcal{F}$.
- (b) $P(S) = 1$.

(c) P je **šteвно aditivna**: če so $A_1, A_2, \dots \in \mathcal{F}$ paroma disjunktne množice, potem je $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Preslikavo P , ki ustreza tem zahtevam, imenujemo **verjetnost**.

Za našo uporabo bomo zapisano definicijo nekoliko poenostavili² in privzeli, da je prostor dogodkov \mathcal{F} vselej potenčna množica $\mathcal{P}(S)$, torej družina vseh podmnožic prostora izidov. Opazovanemu poskusu bomo torej priredili prostor izidov S , prostor dogodkov $\mathcal{P}(S)$ in verjetnost $P: \mathcal{P}(S) \rightarrow \mathbb{R}$, ustrezeni verjetnostni prostor pa označili kot par (S, P) . Tako dobimo naslednji slovar izrazov:

Oznaka	Verjetnost	Množice
S	prostor izidov	neprazna množica
$a, b, c, \dots \in S$	izidi	elementi množice
$\mathcal{P}(S)$	prostor dogodkov	potenčna množica od S
$A, B, C, \dots \subseteq S$	dogodki	podmnožice
$\{a\} \subseteq S$	elementarni dogodek	enoelementna podmnožica (singleton)
$B \subseteq A$	poddogodek dogodka A	podmnožica v A
\emptyset	prazen dogodek	prazna množica
\bar{A} ali A^C ali $S \setminus A$	nasprotni dogodek od A	komplement množice
$A \cup B$ ali $A + B$	unija ali vsota dogodkov	unija podmnožic
$A \cap B$ ali AB	preseki ali produkt dogodkov	preseki podmnožic
$A \cap B = \emptyset$	nezdružljiva ali ločena dogodka	disjunktne podmnožice

Opomba. Pri elementarnih dogodkih namesto $P(\{a\})$ pogosto zapišemo kar $P(a)$.

Zgled. Naj prostor izidov $S = \{G, C\}$ vsebuje dva izida. Tedaj vsebuje prostor dogodkov $\mathcal{P}(S) = \{\emptyset, \{G\}, \{C\}, \{G, C\}\}$ štiri dogodke. Enakosti $P(G) = P(C) = 1/2$, $P(\emptyset) = 0$ in $P(S) = 1$ določajo neko preslikavo P , ki zadošča vsem trem aksiomom za verjetnost. Verjetnostni prostor $(S, \mathcal{P}(S), P)$ oziroma (S, P) opisuje verjetnosti pri enem metu poštenega kovanca.

Zgled. Na istem prostoru izidov in dogodkov kot prej lahko definiramo tudi drugačno verjetnost P' , s katero bi opisali nepošten kovanec z lastnostjo $P'(G) = 1/3$.

Zgled. Kovanec mečemo do prvega grba. Prostor izidov $S = \{G, CG, CCG, CCCG, \dots\}$ je števno neskončna množica. Če je A_i elementarni dogodek, da prvi grb pade v i -tem metu, definiramo njegovo verjetnost kot $P(A_i) = (1/2)^i$, verjetnost poljubnega dogodka A pa kot $P(A) = \sum_{i \in A} (1/2)^i$. Za preslikavo P očitno velja prvi aksiom verjetnosti, drugega pa dobimo iz enakosti

$$P(S) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{\infty} (1/2)^i = 1.$$

Z nekaj dodatnega truda bi lahko preverili še tretji aksiom.

Zgled. Denimo, da želimo na slepo izbrati naključno naravno število tako, da bodo vse izbire enako verjetne. Potem bi obstajalo tako število $\alpha \in \mathbb{R}$, da bi veljalo $P(n) = \alpha$ za vse $n \in \mathbb{N}$. Toda tedaj bi po lastnostih verjetnosti veljalo $P(\mathbb{N}) = \sum_{i=1}^{\infty} \alpha = \infty$, protislovje. Tak poskus ni možen! Če želimo izbrati naključno naravno število, potem ne moremo vseh števil izbrati z enako verjetnostjo.

Včasih lahko verjetnostni prostor opišemo geometrijsko, tako da je prostor izidov neka množica v ravnini ali prostoru, verjetnost pa je povezana s ploščino ali prostornino te množice.

²Ker je potenčna množica $\mathcal{P}(S)$ pri neskončni množici S zelo velika, je pri študiju zahtevnejših problemov pogosto nujno izhajati iz manjšega prostora dogodkov \mathcal{F} , vendar tukaj takih primerov ne bomo srečali.

Zgled. Kovanec s premerom 2 cm vržemo na tla, ki so tlakovana s kvadratnimi ploščicami z robom 10 cm. Kakšna je verjetnost, da kovanec prekriva kakšno mejno črto med dvema ploščicama? Prostor izidov je v tem primeru ravnina \mathbb{R}^2 , pri čemer vsaka točka določa možen položaj središča kovanca. Iz skice razberemo, da je odgovor $19/100$, glede na površino ugodne lege središča kovanca v primerjavi s celotno površino posamezne ploščice.

1.2.4 Osnovne lastnosti verjetnosti

V definiciji verjetnosti P smo zahtevali 3 aksiome. Iz njih lahko neposredno izpeljemo naslednje preproste posledice.

Trditev. Naj bo $P: \mathcal{P}(S) \rightarrow \mathbb{R}$ neka verjetnost na prostoru izidov S . Potem velja:

- (a) $P(\emptyset) = 0$.
- (b) Za nezdružljiva dogodka $A, B \subseteq S$, $A \cap B = \emptyset$, velja $P(A \cup B) = P(A) + P(B)$. Splošneje, če so $A_1, \dots, A_n \subseteq S$ paroma disjunktni dogodki, potem je $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.
- (c) $P(\bar{A}) = 1 - P(A)$.
- (d) Za vse $A \subseteq S$ velja $0 \leq P(A) \leq 1$.
- (e) $P(A \setminus B) = P(A) - P(A \cap B)$.
- (f) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- (g) Če je $A \subseteq B$, potem je $P(A) \leq P(B)$.

Dokaz.

- (a) Denimo, da je $P(\emptyset) = \alpha$ za neko število $\alpha \in \mathbb{R}$. Za zaporedje paroma disjunktnih množic $A_i = \emptyset$ velja $\bigcup_{i=1}^{\infty} A_i = \emptyset$ in zato

$$P(\emptyset) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{\infty} \alpha,$$

kar je konvergentno le za $\alpha = 0$.

- (b) Trditev za n množic sledi iz prejšnje trditve in tretjega aksioma verjetnosti, če zaporedje A_1, \dots, A_n dopolnimo s praznimi množicami $A_{n+1} = A_{n+2} = \dots = \emptyset$.
- (c) Ker sta A in \bar{A} nezdružljiva dogodka, po prejšnji točki sledi

$$1 = P(S) = P(A \cup \bar{A}) = P(A) + P(\bar{A}).$$

- (d) Ker je $P(A)$ in $P(\bar{A}) \geq 0$ po prvem aksiomu verjetnosti, trditev sledi iz prejšnje točke.
- (e) Premisli sam.
- (f) Premisli sam.
- (g) Iz $A \subseteq B$ sledi $A \cap B = A$, zato je po eni prejšnjih točk $P(A) = P(A \cap B) = P(B) - P(B \setminus A) \leq P(B)$.

□

Zgled. Naj za dogodka A in B velja $P(A) = 0.3$, $P(B) = 0.5$ in $P(A \cap B) = 0.2$. Potem je verjetnost, da se zgodi vsaj eden od dogodkov A ali B , enaka $P(A \cup B) = 0.3 + 0.5 - 0.2 = 0.6$. Verjetnost, da se ne zgodi noben od dogodkov A ali B je $P(\overline{A \cup B}) = 1 - P(A \cup B) = 0.4$. Verjetnost, da se zgodi le A , pa je $P(A \setminus B) = P(A) - P(A \cap B) = 0.1$.

1.2.5 Verjetnost na končnem prostoru izidov

Kadar je prostor izidov našega poskusa končen, si lahko verjetnost predstavljamo na posebej enostaven način.

Izrek (Točkasta verjetnost). Naj bo $S = \{a_1, a_2, \dots, a_n\}$ končen ali števno neskončen prostor izidov. Tedaj je vsaka verjetnost na S natanko določena z verjetnostmi izidov $P(a_i)$. Natančneje, za vsak nabor števil p_1, \dots, p_n z lastnostjo $0 \leq p_i \leq 1$ in $\sum_{i=1}^n p_i = 1$ obstaja natanko ena verjetnost P z lastnostjo $P(a_i) = p_i$ za vse i .

Dokaz. Pri danih pogojih lahko verjetnost dogodka $A \subseteq S$ definiramo kot končno vsoto $P(A) = \sum_{i: a_i \in A} p_i$. Potem je očitno $P(A) \geq 0$ in $P(S) = 1$, torej sta izpolnjena prva dva aksioma iz definicije verjetnosti. Naj bo A_1, A_2, \dots števno zaporedje disjunktnih množic z unijo A . Po definiciji P sledi

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = P(A) = \sum_{i: a_i \in A} p_i = \sum_{j=1}^{\infty} \left(\sum_{i: a_i \in A_j} p_i \right) = \sum_{j=1}^{\infty} P(A_j),$$

torej je P res verjetnost. Če je P' neka druga verjetnost, ki prav tako ustreza pogoju $P'(a_i) = p_i$ za vse i , potem mora za vsak dogodek A veljati $P'(A) = \sum_{i: a_i \in A} p_i$, kar pomeni, da je $P = P'$. \square

Opomba. Gornji izrek lahko primerjamo z izrekom iz linearne algebre, ki pravi, da je vsaka linearna preslikava natanko določena s slikami baznih vektorjev. Izrek lahko v podobni obliki dokažemo tudi za števno neskončen prostor izidov S .

Zgled. Če je $S = \{a, b, c\}$, potem zahteve $P(a) = 0.3$, $P(b) = 0.6$ in $P(c) = 0.1$ natanko določajo verjetnostno funkcijo P .

Zgled. Pri poskusu z dvema izidoma $S = \{a, b\}$ je verjetnost natanko določena z verjetnostjo enega od izidov: če je $P(a) = p$, je $P(b) = 1 - p$. Tak poskus je denimo met kovanca.

Zgled. Verjetnost pri metu astragalija natanko določimo s pogoji $P(a_1) = P(a_2) = 0.4$ in $P(a_3) = P(a_4) = 0.1$.

Ob koncu razdelka se vprašajmo še, kako je aksiomska definicija usklajena s klasično definicijo verjetnosti.

Trditev. Klasična definicija verjetnosti je poseben primer aksiomatsko definirane verjetnosti na končnem prostoru izidov.

Dokaz. Če za končen prostor izidov $S = \{a_1, \dots, a_n\}$ definiramo $P(a_i) = \frac{1}{|S|}$ za vsak i , so vsi izidi enako verjetni in verjetnost dogodka A po gornji definiciji bo enaka $P(A) = |A|/|S|$, kar je ravno razmerje med številom ugodnih in številom vseh izidov. \square

Zgled. Vržemo 3 kocke hkrati. Kakšna je verjetnost, da pade vsaj ena šestica? Če kocke ločimo, je vseh različnih izidov $6 \cdot 6 \cdot 6 = 216$ in vsi so enako verjetni. Neugodni so izidi brez šestice, takih je $5 \cdot 5 \cdot 5 = 125$. Verjetnost vsaj ene šestice je torej $1 - \frac{125}{216}$.

Zgled. Kockar in pisatelj Chevalier de Meré se je leta 1654 obrnil na matematika Blaisa Pascala z vprašanjem, zakaj pri stavi na vsaj 1 šestico v štirih metih kocke dobiva, pri stavi na vsaj 1 dvojno šestico pri 24 metih dveh kock pa izgublja. Pascal je izračunal, da dogodka nista enako verjetna, kakor je menil de Meré. Verjetnost je približno 0.517, verjetnost drugega pa 0.491. Zakaj?

1.3 Osnove kombinatoričnega preštevanja

1.3.1 Pravilo vsote in produkta

Pri reševanju številnih elementarnih problemov verjetnostnega računa je bistveni del naloge preštevanje izidov oziroma elementov končnih množic. Osnovni pravili, ki ju pri tem uporabljamo, izhajata iz teorije množic.

Trditev. Naj bosta A in B poljubni končni množici. Potem velja:

(a) *Pravilo vsote:* če je $A \cap B = \emptyset$, potem je $|A \cup B| = |A| + |B|$.

(b) *Pravilo produkta:* $|A \times B| = |A| \cdot |B|$.

Dokaz. Pravili sta na pogled očitni. Za dokaz, zgrajen iz aksiomov teorije množic, pa preglejte zapiske lanskih predavanj iz predmeta Logika in množice. \square

Opomba. Obe pravili lahko z indukcijo posplošimo tudi na končne množice A_1, \dots, A_n , $n \in \mathbb{N}$.

Pravilo vsote uporabljamo takrat, kadar izbiramo posamezni element iz unije dveh ločenih podmnožic, pravilo produkta pa takrat, kadar izbiramo pare, sestavljene iz po enega elementa vsake od dveh množic.

Zgled. Za sladico imamo na razpolago 3 vrste pite in 2 vrsti kompota. Če moramo izbrati pito ali kompot, imamo $3+2=5$ možnih izbir. Če lahko izberemo pito in kompot, pa $3 \cdot 2 = 6$ možnih izbir.

Zgled. Janez ima v omari 5 hlač, 4 majice in 6 srajc, 3 kravate in 2 pokrivali. Vsak dan si obleče hlače, zraven pa včasih srajco in kravato, drugič majico in pokrivalo. Na koliko različnih načinov se lahko obleče?

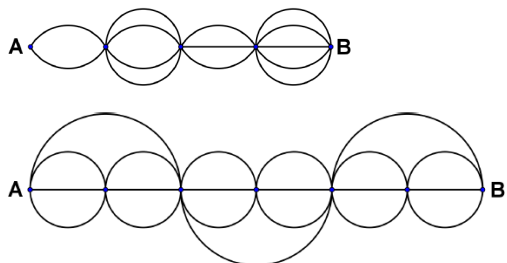
Če se Janez obleče elegantno, izbere torej po en element iz vsake od množic H, S, K . Če pa se obleče za sprehod, izbere po en element iz vsake od množic H, M, P . Njegovo garderobo lahko torej opišemo kot

$$(H \times S \times K) \cup (H \times M \times P) = H \times (S \times K \cup M \times P),$$

moč te množice pa je po pravilu vsote in pravilu produkta enaka

$$|H| \cdot (|S| \cdot |K| + |M| \cdot |P|) = 5 \cdot (6 \cdot 3 + 4 \cdot 2) = 130.$$

Zgled. Na koliko načinov lahko pridemo od točke A do točke B , če se lahko po narisanih poteh pomikamo le v smeri od leve proti desni?



Odgovor v prvem primeru je 120, v drugem pa 1000.

1.3.2 Razvrstitve in izbori

Iz pravila produkta, ki mu nekateri pravijo tudi Osnovni izrek kombinatorike, sledi večina osnovnih pravil kombinatoričnega preštevanja, ki smo jih spoznali že v srednji šoli.

Trditev. Naslednja tabela podaja pravila za preštevanje vseh razvrstitev (permutacij), urejenih izborov (variacij) in neurejenih izborov (kombinacij) elementov dane množice.

Opis	Oznaka in algebrski izraz
<i>Permutacije brez ponavljanja:</i> Število razvrstitev n različnih elementov v vrsto.	$P_n = n! = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$
<i>Permutacije s ponavljanjem:</i> Število razvrstitev r različnih elementov v vrsto dolžine n , pri čemer je k_i število ponovitev elementa a_i .	$P_n^{k_1, k_2, \dots, k_r} = \frac{n!}{k_1! \cdot \dots \cdot k_r!}$
<i>Variacije brez ponavljanja:</i> Število urejenih izborov r različnih elementov izmed n elementov.	$V_n^r = \frac{n!}{(n-r)!} = n(n-1) \cdot \dots \cdot (n-r+1) = n^{[r]}$
<i>Variacije s ponavljanjem:</i> Število urejenih izborov r elementov izmed n različnih.	${}^{(p)}V_n^r = n \cdot n \cdot \dots \cdot n = n^r$
<i>Kombinacije brez ponavljanja:</i> Število neurejenih izborov r različnih elementov izmed n različnih.	$C_n^r = \frac{n!}{r!(n-r)!} = \binom{n}{r}$
<i>Kombinacije s ponavljanjem:</i> število neurejenih izborov r elementov izmed n različnih.	${}^{(p)}C_n^r = \binom{n+r-1}{r}$

Dokaz. O številu razvrstitev naj bralec razmisli sam, skupaj pa razmislimo, od kod izrazi o številu urejenih in neurejenih izborov.

- Urejeni izbori brez ponavljanja: Če izmed n možnih izbiramo r različnih elementov in pri tem pazimo na vrstni red izbranih elementov, potem imamo za prvi element n možnosti, za drugega $n-1$, in za r -tega $n-r+1$ možnosti. Po pravilu produkta je vseh možnosti $n(n-1)\dots(n-r+1)$, kar označimo z znakom za padajočo potenco $n^{[r]}$ ali zapisom $\frac{n!}{(n-r)!}$.
- Urejeni izbori s ponavljanjem: Če izmed n možnih elementov izbiramo r elementov, ki niso nujno različni, pomemben pa je vrstni red izbora, potem imamo za prvi element n možnosti, za drugega prav tako n in tako dalje. Vseh možnosti je n^r .
- Neurejeni izbori brez ponavljanja: Izbiramo r različnih elementov izmed n . Če bi upoštevali vrstni red, bi bilo vseh izborov $n^{[r]} = \frac{n!}{(n-r)!}$, ker pa ta ni pomemben, lahko to število delimo s številom različnih razvrstitev r elementov, kar je enako $r!$. Dobimo izraz $\frac{n!}{(n-r)!r!} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-r+1)}{1 \cdot 2 \cdot \dots \cdot r}$, ki ga označimo z binomskim simbolom $\binom{n}{r}$.
- Neurejeni izbori s ponavljanjem: V tej situaciji nas zanima le, kolikokrat smo izbrali vsakega od možnih elementov x_1, \dots, x_n . Vsak izbor r elementov si zato predstavljamo kot da bi med r kroglic v vrsti postavili $n-1$ pregrad, in tako dobili n predalčkov (nekateri

so lahko prazni). Ker je v dobljeni vrsti skupaj $r + n - 1$ pregrad in kroglic, je možnih postavitev pregrad $\binom{r+n-1}{n-1}$, kar je enako kot $\binom{r+n-1}{r}$

□

Zgled. V posodi je 5 belih in 3 črne kroglice. Izvlečemo 2 kroglici. Če izvlečenih kroglic ne vračamo, je verjetnost, da sta obe črni, enaka $\binom{3}{2}/\binom{8}{2} = 3/28 \doteq 0.10$. Če pa smo prvo kroglico vrnili in nato izvlekli drugo, se verjetnost poveča na $3^2/8^2 \doteq 0.14$.

Zgled. Izmed kart od 1 do 10 izvlečemo 2 karti. Verjetnost, da je vsota števil na kartah liha, če kart ne vračamo, je enaka $\frac{10 \cdot 5}{10 \cdot 9} = 5/9$. Če prvo karto vrnemo in šele potem izvlečemo drugo, pa $\frac{10 \cdot 5}{10^2} = 1/2$.

Zgled. Naj za neki množici velja $|A| = m$ in $|B| = n$. Bralec naj sam premisli, zakaj je število vseh funkcij $f: A \rightarrow B$ enako n^m , zakaj je njimi $m^{\lfloor n \rfloor}$ injektivnih funkcij, če je $m \leq n$, in $n!$ bijektivnih funkcij, če je $m = n$. Surjektivne funkcije pa bomo s pomočjo Stirlingovih števil preštevali pri Diskretni matematiki ali kakem drugem predmetu.

Omenimo še naslednjo situacijo, ko urejeno izberemo skupaj k (ne nujno različnih) elementov izmed n različnih elementov, pri tem pa element x_i izberemo k_i -krat, tako da je $k_1 + \dots + k_n = k$. Tedaj je število vseh izborov enako

$$\binom{k}{k_1} \binom{k-k_1}{k_2} \dots \binom{k-k_1-\dots-k_{n-1}}{k_n} = \frac{k!}{k_1!k_2!\dots k_n!}.$$

Ta izraz imenujemo **multinomski koeficient** in označimo z $\binom{k}{k_1, k_2, \dots, k_n}$.

Zgled. Učitelj je z žrebom razvrstil 10 učencev v petčlansko ekipo za košarko, tričlansko za odbojko in par za badminton. Kakšna je verjetnost, da je Janezek v ekipi za odbojko? Vseh možnih razvrstitev je $\binom{10}{5,3,2} = \frac{10!}{5!3!2!} = 2120$. Če vemo, da Janezek igra odbojko, pa lahko preostalih devet učencev razvrstimo na $\binom{9}{5,2,2} = \frac{9!}{5!2!2!}$ načinov. Torej je verjetnost $\frac{3}{10}$. Do iste ugotovitve bi sicer lahko prišli tudi brez preštevanja vseh razvrstitev: predstavljajmo si, da Janezek šport izbere tako, da iz vreče s 5 črnimi, 3 modrimi in 2 rdečima kroglicama izvleče modro, ki predstavlja odbojko, njegovi sošolci pa kroglice vlečejo za njim.

Opomba (Stirlingova formula). Pri kombinatoriki se pogosto pojavlja izraz $n!$, ki z večanjem n raste izjemno hitro. Poskusimo denimo brez kalkulatorja izračunati ali vsaj oceniti, na koliko načinov lahko razvrstimo od leve proti desni 25 učencev v razredu. Če računamo rekurzivno $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$ in tako dalje, bomo porabili kar nekaj časa, da bomo prišli do vrednosti

$$25! = 15511210043330985984000000.$$

Kaj pa, če nas zanima število možnih razvrstitev 52 kart iz igralnega kompleta? Tega raje ne poskušajte računati rekurzivno, če nimate nekaj dni časa. Abraham de Moivre je okoli 1733 prvi opazil, da je število $n!$ za velike n sorazmerno izrazu $C \cdot \left(\frac{n}{e}\right)^n \sqrt{n}$, James Stirling pa je nato določil vrednost konstante $C = \sqrt{2\pi}$. S tako imenovano **Stirlingovo formulo**

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

so nato ocenjevali izraz $n!$ vse do prevlade računalnikov.

Če želimo oceniti $52!$ na nekaj mest natančno po Stirlingovi formuli, najprej izračunamo $52/e = 19.12973098$ in $\sqrt{2\pi \cdot 52} = 2\sqrt{\pi} \cdot \sqrt{26} = 18.07554591$. Nato s kvadriranjem določimo približne vrednosti $(52/e)^2$, $(52/e)^4$, $(52/e)^8$, $(52/e)^{16}$, $(52/e)^{32}$ in še

$$(52/e)^{52} = (52/e)^{32} \cdot (52/e)^{16} \cdot (52/e)^4 = 4.4551343 \cdot 10^{66},$$

od koder sledi končen rezultat

$$52! \approx 4.4551343 \cdot 10^{66} \cdot 18.07554591 = 8.0529 \cdot 10^{67}.$$

V 18. stoletju je spreten matematik za takšen izračun verjetno porabil manj kot pol ure.

1.3.3 Načelo vključitve in izključitve

Za končni množici A in B velja $|A \cup B| = |A| + |B| - |A \cap B|$, saj smo pri preštevanju elementov iz A in B dvakrat šteli tiste elemente, ki ležijo v obeh množicah. Podobno lahko premislimo, da za tri končne množice A, B, C velja

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

Trditev. Za poljubne končne množice A_1, \dots, A_n velja načelo vključitve in izključitve v obliki

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{1 \leq i \leq n} |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k| + \dots + (-1)^{n+1} |A_1 \cap A_2 \cap \dots \cap A_n|.$$

Dokaz. Dokaz s pomočjo indukcije na število množic je nekoliko nadležna, a koristna vaja za bralca. \square

Če namesto o moči množic govorimo o verjetnosti dogodkov, dobi načelo vključitve in izključitve naslednjo obliko.

Posledica. Naj bo (P, S) verjetnostni prostor s končnim prostorom izidov in naj bodo $A_1, \dots, A_n \subseteq S$ neki dogodki. Potem je verjetnost, da se je zgodil vsaj eden od njih, enaka

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{1 \leq i \leq n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \dots + (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n).$$

Zgled. Kakšna je verjetnost, da naključno izbrano število med 1 in 1000 ni deljivo z 2, 3 in 5? Označimo z A, B, C množice števil med 1 in 1000, ki so deljiva z 2, 3 oziroma 5. Potem je $|A| = 500$, $|B| = 333$, $|C| = 200$, $|A \cap B| = 166$, $|A \cap C| = 100$, $|B \cap C| = 66$ in $|A \cap B \cap C| = 33$. Zato je vseh števil, ki so deljiva vsaj z enim od naštetih, enako

$$|A \cup B \cup C| = 500 + 333 + 200 - 166 - 100 - 66 + 33 = 734.$$

Iskana verjetnost je zato enaka $\frac{1000-734}{1000} = 0.266$.

Zgled. Napisali smo 4 pisma in ovojnice z naslovi, tajnica pa je pisma naključno vstavila v ovojnice. Kakšna je verjetnost, da nobeno pismo ni v pravi ovojnici?

Zamislimo si, da so pisma in ovojnice označena s številkami od 1 do 4. Vseh načinov, kako vstavimo pisma v štiri dane ovojnice, je $4! = 24$. Če jih izpišemo in izločimo neustrezne, jih ostane 9, to so: 2143, 2341, 2413, 3142, 3412, 3421, 4123, 4312 in 4321, torej je verjetnost enaka $9/24$, saj so vse razporeditve enako verjetne.

Če želimo isti problem rešiti za n pisem, pa moramo najti ustrežnejši postopek. Naj bo A_i dogodek, da je i -to pismo v pravi ovojnici. Potem velja $P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}$, saj je je vseh razporeditev $n!$, od teh pa je $(n-1)!$ takih, pri katerih je i -to pismo v pravi ovojnici. Podobno velja $P(A_i \cap A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$, če $1 \leq i < j \leq n$, pa $P(A_i \cap A_j \cap A_k) = \frac{(n-3)!}{n!}$, če $1 \leq i < j < k \leq n$, in tako dalje vse do $P(A_1 \cap \dots \cap A_n) = \frac{1}{n!}$.

Po načelu vključitve in izključitve je verjetnost dogodka, da je vsaj eno pismo v pravi ovojnici, enaka

$$\begin{aligned} P\left(\bigcup_{1 \leq i \leq n} A_i\right) &= n \cdot \frac{(n-1)!}{n!} - \binom{n}{2} \cdot \frac{(n-2)!}{n!} + \binom{n}{3} \cdot \frac{(n-3)!}{n!} - \dots + (-1)^{n+1} \binom{n}{n} \frac{1}{n!} = \\ &= 1 + \frac{-1}{2!} + \frac{1}{3!} + \dots + \frac{(-1)^{n+1}}{n!}, \end{aligned}$$

verjetnost, da nobeno pismo ni v pravi ovojnici, pa je enaka

$$P(\overline{\bigcup_{i=1}^n A_i}) = 1 - P(\bigcup_{i=1}^n A_i) = 1 - \frac{1}{1!} + \frac{1}{2!} - \dots + (-1)^n \frac{1}{n!} = \sum_{k=0}^n \frac{(-1)^k}{k!}.$$

Zadnji izraz prepoznamo kot vrednost n -tega Taylorjevega polinoma za eksponentno funkcijo $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ v točki $x = -1$. Z večanjem števila pisem iskana verjetnost torej konvergira proti

$$\lim_{n \rightarrow \infty} P(\overline{\bigcup_{i=1}^n A_i}) = e^{-1} \doteq 0.3678.$$

Takšno rešitev je leta 1708 opisal P. R. de Montmort, verjetno pa jo je približno v istem času poznal tudi Jacob Bernoulli.

2 Pogojna verjetnost in neodvisnost dogodkov

2.1 Pogojna verjetnost

Naj bo (S, P) verjetnostni prostor in naj bosta $A, B \subseteq S$ neka dogodka. Kadar je $P(A \cap B) \neq 0$, se lahko dogodka A in B zgodita tudi hkrati. Ob dodatni informaciji, da se je zgodil eden od njiju, na primer B , se lahko vprašamo, kakšna je verjetnost, da se je tedaj zgodil tudi A . **Pogojno verjetnost dogodka A pri pogoju B** z verjetnostjo $P(B) > 0$ definiramo kot razmerje

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

Zgled. Met kocke. Naj bo $A = \{6\}$ dogodek, da je padla šestica, in $B = \{2, 4, 6\}$ dogodek, da je padlo sodo število. Potem je $P(A|B) = \frac{1}{3}$ in $P(B|A) = 1$. Dodajmo še dogodek $C = \{4, 5, 6\}$, da smo vrgli vsaj štiri pike. Potem je $P(B|C) = P(C|B) = 2/3$, $P(A|C) = 1/2$ in $P(C|A) = 1$.

Zgled. Ana in Bine vržeta vsak svojo kocko. Zmaga tisti, ki je vrgel več pik, če je rezultat izenačen, pa met ponovita. Denimo, da je Ana v zadnji igri zmagala. Kakšna je verjetnost, da je vrgla 5 pik?

Če izide zapišemo kot pare (a, b) , je možnih izidov $36 - 6 = 30$, saj lahko izpustimo pare z $a = b$. Vsi izidi so enako verjetni. Za Anino zmago je ugodnih 15 izidov:

$$A = \{(2, 1), (3, 2), (3, 1), (4, 3), (4, 2), (4, 1), (5, 4), (5, 3), (5, 2), (5, 1), (6, 5), (6, 4), (6, 3), (6, 2), (6, 1)\}.$$

Izidov, pri katerih je Ana vrgla 5 pik, pa je 5:

$$B = \{(5, 6), (5, 4), (5, 3), (5, 2), (5, 1)\}.$$

Ker je $|A \cap B| = 4$, sledi $P(B|A) = 4/15$. Splošneje, če z A_i označimo dogodek, da je Ana vrgla i pik, velja $P(A_i|A) = \frac{i-1}{15}$.

Navedena zgleda kažeta, da je definicija pogojne verjetnosti smiselna, kar potrjuje tudi naslednji izrek.

Izrek. Naj bo (S, P) verjetnostni prostor in naj bo $B \subseteq S$ dogodek z neničelno verjetnostjo. Potem je preslikava $Q: \mathcal{P}(S) \rightarrow \mathbb{R}$, ki jo definiramo s predpisom $Q(A) = P(A|B)$, verjetnost na S .

Dokaz. Preveriti moramo tri aksiome za verjetnost. Za poljubno množico $A \subseteq S$ velja $Q(A) = P(A|B) = P(A \cap B)/P(B) \geq 0$. Vrednost $Q(S) = P(S|B) = P(S \cap B)/P(B) = P(B)/P(B)$ je enaka 1. Če pa so $A_1, A_2, \dots \subseteq S$ paroma ločeni dogodki, potem je

$$\begin{aligned} Q\left(\bigcup_{i=1}^{\infty} A_i\right) &= \frac{P\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{i=1}^{\infty} (A_i \cap B)\right)}{P(B)} = \frac{\sum_{i=1}^{\infty} P(A_i \cap B)}{P(B)} = \\ &= \sum_{i=1}^{\infty} \frac{P(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} P(A_i|B) = \sum_{i=1}^{\infty} Q(A_i), \end{aligned}$$

saj so tudi dogodki $A_1 \cap B, A_2 \cap B, \dots$ paroma ločeni in lahko pri tretjem enačanju uporabimo števno aditivnost funkcije P . \square

Zberimo še nekaj osnovnih lastnosti pogojne verjetnosti.

Trditev. Naj bo (S, P) verjetnostni prostor in naj bo $B \subseteq S$ dogodek z neničelno verjetnostjo. Potem za pogojno verjetnost $P(\cdot|B)$ in vse dogodke $A \subseteq S$ velja:

- (a) $P(A \cap B) = P(A|B)P(B)$ (produktno pravilo).
- (b) $P(A^C|B) = 1 - P(A|B)$.
- (c) Če je $A \subseteq B$, potem je $P(A|B) = \frac{P(A)}{P(B)}$ in $P(B|A) = 1$. Posebej, $P(B|B) = 1$.
- (d) Če je $A \cap B = \emptyset$, je $P(A|B) = 0$.

Dokaz trditve prepuščamo bralcu. Mi pa izračunajmo nekoliko zahtevnejši zgled.

Zgled. Iz kompleta 54 igralnih kart najprej izbere svoj par kart vaš soigralec, nato še vi. Definimo, da ima vaš partner dve dami. Kakšna je potem verjetnost, da imate tudi vi dve dami? Kaj pa, da imate vi srčevo damo?

Označimo z A dogodek, da imate vi dve dami, z B dogodek, da ima vaš partner dve dami, in s C dogodek, da imate srčevo damo. Zanimata nas verjetnosti $P(A|B)$ in $P(C|B)$, zato izračunamo verjetnosti $P(B)$, $P(A \cap B)$ in $P(C \cap B)$:

$$P(B) = \frac{\binom{4}{2}}{\binom{54}{2}}, P(A \cap B) = \frac{\binom{4}{2}\binom{2}{2}}{\binom{54}{2}\binom{52}{2}}, P(C \cap B) = \frac{\binom{1}{1}\binom{3}{2}\binom{51}{1}}{\binom{54}{2}\binom{52}{2}}.$$

Zdaj nam preostane le še aritmetika:

$$P(A|B) = \frac{1}{\binom{52}{2}} = \frac{1}{1326}, P(C|B) = \frac{\binom{3}{2}\binom{51}{1}}{\binom{52}{2}\binom{4}{2}} = \frac{1}{52}.$$

Rezultat nas morda malo preseneti. Ali bi morda lahko našli tudi krajšo pot do istega rezultata?

Zgled (Paradoks Montyja Halla). Monty Hall je bil ameriški TV voditelj, ki je leta 1963 v svoji oddaji v živo srečnemu gledalcu ponudil naslednjo igro na srečo. Na odru so tri vrata: za enimi je skrit avto. Igralec izbere naključna vrata. Če je za njimi avto, ga lahko obdrži. Verjetno se vsi strinjamo, da je v primeru poštene igre verjetnost, da gledalec na ta način dobi avto, enaka $1/3$.

Za dodatno dramatičnost pa je potem, ko je gledalec že izbral ena od vrat, odprl ena od neizbranih vrat, za katerimi ni bilo avta, in gledalcu ponudil menjavo vrat. Gledalec je lahko

zdaj izbiro zamenjal, če se je tako odločil. Potem pa je voditelj odprl izbrana vrata in srečnež je dobil avto, ali pa tudi ne.

Na prvi pogled je igra povsem preprosta. Ker gledalec v drugi fazi izbira ena od dveh zaprtih vrat, se zdi, da se je zdaj njegova verjetnost za zadetek povečala na $1/2$, ne glede na to, ali je izbiro zamenjal ali pa je vztrajal pri prvotni izbiri. (Večina gledalcev se sicer iz psiholoških razlogov odloči, da vrat ne zamenja, ker imajo občutek, da jih voditelj nalašč sili k napačni izbiri.) Pa je to res?

Izkaže se, da je tako razmišljanje povsem napačno. Če eksperimentalno izvedemo dovolj veliko število ponovitev igre, bomo opazili, da gledalec osvoji avto v $1/3$ primerov, če vrat ne zamenja, in v $2/3$ primerov, če vrata zamenja. Torej se vrata splača zamenjati!

To lahko dokažemo tudi teoretično z uporabo pogojne verjetnosti. Prostor izidov poskusa bomo opisali takole. Imamo troje vrat. Tista, za katerimi je skrit avto, označimo s črko a , preostala vrata pa poljubno z b in c . Predpostavimo, da igralec vrata v drugem koraku zamenja, in izračunajmo (pogojno) verjetnost, da v tem primeru dobi avto. Posamezni izid poskusa je torej par (x, y) , kjer je x oznaka vrat, ki jih prvotno izbere gledalec, y pa oznaka vrat, ki jih odpre voditelj.

Če gledalec pri prvi izbiri izbere vrata a , lahko voditelj odpre katerakoli od preostalih vrat b ali c . Če pa prvotno izbere prazna vrata b , potem mora voditelj odpreti prazna vrata c , in obratno. Množica vseh izidov igre je torej

$$S = \{(a, b), (a, c), (b, c), (c, b)\}.$$

Ugodni izidi so seveda tisti, pri katerih dobimo avto. To sta (b, c) in (c, b) . Verjetnost $P(b, c)$ je enaka $\frac{1}{3}$, saj je gledalec izbral katerakoli od treh vrat, voditelj pa je odpreti izbrati edina preostala prazna vrata. Enako velja za $P(c, b)$. Torej je verjetnost, da dobimo avto, če izbiro zamenjamo, enaka $\frac{2}{3}$.

Vse skupaj lahko utemeljimo tudi krajše. Verjetnost, da je avto skrit za prvotno izbranimi vrati, je $1/3$. Zato je verjetnost, da je skrit za enimi od neizbranih, enaka $2/3$, kar se ne spremeni, če voditelj odpre neizbrana prazna vrata. Toda s tem se verjetnost, da je avto za neizbranimi neodprtimi vrati, poveča na $2/3$.

2.2 Produktno pravilo za pogojno verjetnost

V prejšnjih zgledih smo izračunali pogojno verjetnost $P(A|B)$ iz verjetnosti $P(A \cap B)$ in $P(B)$. V praksi pa pogosto računamo ravno obratno in verjetnost preseka $P(A \cap B)$ izračunamo iz verjetnosti $P(A|B)$ in $P(B)$ (oziroma $P(B|A)$ in $P(A)$).

Zgled. Hokejsko moštvo se je uvrstilo v polfinale, kjer ima 60% možnosti za zmago. Če zmagajo, bodo dobili v finalu močnega nasprotnika, proti kateremu imajo le 40% možnosti. Kakšna je verjetnost, da osvojijo turnir?

Naj bosta A in B dogodka, da naši hokejisti zmagajo v polfinalu oziroma v finalu. Potem je $A \cap B$ dogodek, da so osvojili turnir, zato velja

$$P(A \cap B) = P(A) \cdot P(B|A) = 0.6 \cdot 0.4 = 0.24.$$

Produktno pravilo zelo pogosto uporabljamo tudi v situacijami z več zaporednimi dogodki.

Trditev (Produktno pravilo za več dogodkov). Naj bodo $A_1, A_2, \dots, A_n \subseteq S$ taki dogodki, da je $P(A_1 \cap \dots \cap A_{n-1}) > 0$. Potem velja

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

Dokaz. Če so verjetnosti ustreznih presekov neničelne, je izraz na desni strani enakosti po definiciji pogojne verjetnosti enak

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

Neničelnost imenovalcev zagotavlja pogoj $P(A_1 \cap \dots \cap A_{n-1}) > 0$, saj velja

$$P(A_1) \geq P(A_1 \cap A_2) \geq \dots \geq P(A_1 \cap \dots \cap A_{n-1}) > 0$$

zaradi monotonosti verjetnostne funkcije P . □

Zgled. V vrečki je b belih in c črnih bonbonov. Vsak dan pojemo enega, ki ga izberemo naključno. Kakšna je verjetnost, da prva dva dni pojemo bel, druga dva dni pa črn bonbon?

Razmišljamo lahko na dva načina. Prvo rešitev dobimo s pomočjo produktne formule. Naj A_i pomeni, da i -ti dan pojemo bel bonbon. Zanima nas

$$\begin{aligned} P(A_1 \cap A_2 \cap \bar{A}_3 \cap \bar{A}_4) &= P(A_1)P(A_2|A_1)P(\bar{A}_3|A_1 \cap A_2)P(\bar{A}_4|A_1 \cap A_2 \cap \bar{A}_3) = \\ &= \frac{b}{b+c} \cdot \frac{b-1}{b+c-1} \cdot \frac{c}{b+c-2} \cdot \frac{c-1}{b+c-3}. \end{aligned}$$

Pri drugem načinu pa vseh $b+c$ bonbonov naključno postavimo v vrsto na $\binom{b+c}{b}$ načinov (predstavljamo si, da v celotni vrsti izberemo b mest za bele bonbone). Pri ugodnih razvrstitvah so prva štiri mesta že zapolnjena. Preostalih $b+c-4$ bonbonov razvrstimo na $\binom{b+c-4}{b-2}$ načinov. Iskana verjetnost je torej enaka

$$\frac{\binom{b+c-4}{b-2}}{\binom{b+c}{b}} = \dots = \frac{b(b-1)c(c-1)}{(b+c)(b+c-1)(b+c-2)(b+c-3)}.$$

2.3 Neodvisni dogodki

V dosedanjih zgledih smo videli, da se včasih verjetnost enega dogodka bistveno spremeni, če vemo, da se je zgodil neki drugi dogodek. Kadar to ni res, govorimo o neodvisnih dogodkih.

Dogodka $A, B \subseteq S$ imenujemo **neodvisna dogodka**, če velja $P(A \cap B) = P(A)P(B)$, torej, če je verjetnost, da sta se zgodila oba, enaka produktu njunih verjetnosti.

Zgled. *Met kocke. Dogodek A je "pade sodo pik", dogodek B "padeta največ 2 piki", dogodek C pa "pade 1 pika". Potem sta A in B neodvisna dogodka, da je $P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{3}$, kar je enako kot $P(A \cap B) = P(C) = 1/6$. Dogodka A in C pa nista neodvisna, saj produkt verjetnosti $P(A) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{6}$ ni enak verjetnosti preseka $P(A \cap C) = 0$.*

Trditev. Naj bo $P(A), P(B) > 0$. Potem velja:

- (a) A in B sta neodvisna natanko tedaj, ko je $P(A|B) = P(A)$ in $P(B|A) = P(B)$.
- (b) Če sta A in B nezdružljiva, potem sta odvisna.

Dokaz. Prvi del trditve naj bralec preveri sam. Za drugega se spomnimo, da sta A in $B \subseteq S$ nezdružljiva, če je $A \cap B = \emptyset$. Če velja $P(A), P(B) > 0$, potem je pogojna verjetnost $P(A|B) = \frac{P(A \cap B)}{P(B)} = 0$, torej se gotovo ne more zgoditi A , če vemo, da se je zgodil B . □

Včasih lahko že iz opisa poskusa razberemo, kateri dogodki so neodvisni. Če denimo vremo dve pošteni kocki, bo število pik na prvi neodvisno od števila pik na drugi kocki. Če pa brez vračanja izvlečemo dve karti iz kompleta, bo dogodek, da je prva karta srce, zmanjšal verjetnost dogodka, da je druga karta srce, zato sta dogodka odvisna. Pogosto pa je o neodvisnosti težko presoditi brez izračuna.

Zgled. Denimo, da je verjetnost moškega spola pri vsakem posameznem novorojencu enaka $1/2$. Neka družina ima 3 otroke. Naj bo A dogodek, da imajo največ eno deklico, in B dogodek, da imajo otroke obeh spolov. Ali sta A in B neodvisna dogodka? Kaj pa, če ima družina 4 otroke?

Za družino s tremi otroki je prostor izidov

$$S = \{FFF, DFF, FDF, FFD, DDF, DFD, FDD, DDD\},$$

vs izidi pa so enako verjetni. Zato je

$$P(A) = P(\{FFF, DFF, FDF, FFD\}) = 4/8,$$

$$P(B) = 1 - P(\{FFF, DDD\}) = 6/8,$$

$$P(A \cap B) = P(\{DFF, FDF, FFD\}) = 3/8.$$

Ker je $\frac{4}{8} \cdot \frac{6}{8} = \frac{3}{8}$, sta dogodka A in B neodvisna.

Če pa so otroci štirje, na podoben način izračunamo $P(A) = 5/16$, $P(B) = 14/16$ in $P(A \cap B) = 4/16$, torej sta dogodka v tem primeru odvisna. Zgolj iz opisa dogodkov A in B v tem primeru torej ne moremo intuitivno sklepati, ali sta dogodka neodvisna, saj je to odvisno tudi od števila vseh otrok v družini.

Naj bo $\mathcal{D} = \{A_1, \dots, A_n\} \subseteq \mathcal{P}(S)$ neka končna družina dogodkov. Potem so dogodki iz družine \mathcal{D} **paroma neodvisni**, če je neodvisen vsak par različnih dogodkov, torej, če velja

$$P(A_i \cap A_j) = P(A_i)P(A_j) \text{ za vse } 1 \leq i < j \leq n.$$

Družina dogodkov \mathcal{D} pa je **(povsem) neodvisna**, če velja produktno pravilo v obliki

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i) \text{ za vsako podmnožico } I \subseteq \{1, \dots, n\}.$$

Zgled. Vrzimo dva različno velika kovanca in naj bodo dogodki $A \equiv$ "na manjšem kovancu pade grb", $B \equiv$ "na večjem kovancu pade cifra" in $C \equiv$ "pade 1 cifra in 1 grb". Bralec bo zlahka preveril, da so A, B, C paroma neodvisni, vendar niso povsem neodvisni, saj $P(A \cap B \cap C) \neq P(A)P(B)P(C)$.

Zgled. Vrzimo 10 kovancev in naj bo A_i dogodek, da je na i -tem kovancu padel grb. Potem je družina $\{A_i \mid i = 1, \dots, 10\}$ v celoti neodvisna.

Trditev. Če sta dogodka A in $B \subseteq S$ neodvisna, potem so neodvisni tudi pari dogodkov $\{A, \bar{B}\}$, $\{B, \bar{A}\}$ in $\{\bar{A}, \bar{B}\}$.

Dokaz. Ker je

$$(A \cap \bar{B}) = P(A - B) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\bar{B}),$$

sta A in \bar{B} neodvisna. Zaradi simetrije sta potem neodvisna tudi \bar{A} in B ter \bar{A} in \bar{B} . □

2.4 Nezdružljivi dogodki in popolna verjetnost

Naj bo (S, P) verjetnostni prostor. **Popoln sistem dogodkov** je taka družina dogodkov

$$\mathcal{D} = \{A_i \subseteq S \mid i \in I\},$$

za katero velja:

- Družina \mathcal{D} ima končno ali števno neskončno elementov: $|\mathcal{D}| \leq |\mathbb{N}|$.
- Dogodki v \mathcal{D} so nezdružljivi: $A_i \cap A_j = \emptyset$ za vse $i, j \in I, i \neq j$.
- Unija družine \mathcal{D} je prostor izidov: $\cup_{i \in I} A_i = S$.
- Dogodki iz \mathcal{D} imajo pozitivno verjetnost: $P(A_i) > 0$ za vsak $i \in I$.

Spomnimo se, da je družina podmnožic \mathcal{D} **razbitje** ali **particija množice** S , če so njeni elementi disjunktni, njihova unija pa enaka S . Z drugimi besedami lahko torej povemo: popoln sistem dogodkov je števno razbitje prostora izidov na slučajne dogodke. V praksi to pomeni, da se pri kateremkoli izidu poskusa zgodi natanko eden od teh dogodkov.

Zgled. Iz vreče s 5 rumenimi, 4 zelenimi in 2 modrimi kroglicama izvlečemo eno kroglico. Potem dogodka A_1 , da smo izvlekli rumeno, A_2 , da smo izvlekli zeleno, oziroma A_3 , da smo izvlekli modro kroglico, predstavljajo popoln sistem dogodkov $\{A_1, A_2, A_3\}$.

Zgled. Pri naključno izbrani osebi z A, B, AB in O označimo dogodke, da ima ta oseba ustrezno krvno skupino. Potem je $\{A, B, AB, O\}$ popoln sistem dogodkov, saj se vedno zgodi natanko eden od njih.

Izrek (Izrek o popolni verjetnosti). Naj bo (S, P) verjetnostni prostor in $\mathcal{D} = \{A_i \subseteq S \mid i \in I\}$ popoln sistem dogodkov. Potem za poljuben dogodek $B \subseteq S$ velja

$$P(B) = \sum_{i \in I} P(B|A_i)P(A_i).$$

Dokaz. Ker je $B = B \cap S = B \cap (\cup_{i \in I} A_i) = \cup_{i \in I} (B \cap A_i)$ števna unija disjunktnih dogodkov A_i , sledi

$$P(B) = P(\cup_{i \in I} (B \cap A_i)) = \sum_{i \in I} P(B \cap A_i) = \sum_{i \in I} P(B|A_i)P(A_i)$$

iz števne aditivnosti verjetnosti P in definicije pogojne verjetnosti. □

Zgled. Porazdelitev krvnih skupin A, B, O, AB v celotni populaciji je naslednja:

$$O: 44\%, A: 42\%, B: 10\%, AB: 4\%.$$

Denimo, da je verjetnost, da oseba s krvno skupino O zboli za boleznijo X , enaka 0.05 , pri ostalih krvnih skupinah pa 0.15 . Potem je verjetnost, da naključno izbrana oseba zboli, enaka

$$\begin{aligned} P(X) &= P(X|O)P(O) + P(X|A)P(A) + P(X|B)P(B) + P(X|AB)P(AB) \\ &= 0.05 \cdot 0.44 + 0.15 \cdot (0.42 + 0.10 + 0.04) \doteq 0.11. \end{aligned}$$

Zgled (Dvofazni poskus). Iz žare s 5 modrimi in 3 rdečimi kroglicami naključno izvlečemo kroglico in jo damo v drugo žaro, v kateri so 3 modre in 5 rdečih kroglic. Nato iz druge žare izvlečemo eno kroglico. Kakšna je verjetnost, da je ta kroglica modra?

Dogodka $A_1 \equiv$ "prvič smo izvlekli modro" in $A_2 = \bar{A}_1 \equiv$ "prvič smo izvlekli rdečo kroglico" sestavljata popoln sistem dogodkov. Zato je verjetnost dogodka $B \equiv$ "drugič smo izvlekli modro" enaka

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) = \frac{4}{9} \cdot \frac{5}{8} + \frac{3}{9} \cdot \frac{3}{8} = \frac{29}{72} \doteq 0.40.$$

Opomba. Poskuse v več zaporednih fazah, pri katerih vmesni izidi vplivajo na potek poskusa v naslednji fazi, imenujemo **večfazni** ali **relejni poskusi**.

2.5 Bayesov izrek o verjetnosti hipotez

Pravijo³, da je Bayesov izrek za verjetnost tako pomemben kot Pitagorov izrek za geometrijo. Ta ključni izrek je neposredna posledica enakosti

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A).$$

Izrek (Bayesov izrek). Naj bosta $A, B \subseteq S$ poljubna dogodka z neničelno verjetnostjo. Potem je

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

Splošneje, če je $\{A_i \mid i \in I\}$ popoln sistem dogodkov in $B \subseteq S$ poljuben dogodek, potem je

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i \in I} P(B|A_i)P(A_i)} = \frac{P(B \cap A_j)}{\sum_{i \in I} P(B \cap A_i)}.$$

Dokaz. Prva enakost je že dokazana. Za drugo enakost pa vstavimo v prvo namesto A dogodek A_j in $P(B) = \sum_{i=1}^n P(B|A_i)P(A_i)$ po Izreku o popolni verjetnosti. \square

Opomba. Vsebino izreka lahko poskusimo povzeti tudi opisno. V poskusu, ki ga opazujemo, je dogodek B odvisen od hipotetičnih dogodkov A_i , ki so se zgodili pred njim, njihove verjetnosti $P(A_i)$ pa so načeloma znane vnaprej. Leva stran enakosti v izreku pa nam omogoči izračunati verjetnost hipotetičnega dogodka $P(A_j|B)$ na podlagi tega, ali se je zgodil kasnejši dogodek B .

Zgled. Baterije za mobilne telefone izdelujejo v 3 različnih tovarnah. Pri tem je verjetnost okvare 1% v prvi, 2% v drugi in 3% v tretji tovarni, število vseh izdelanih baterij pa je 1000 v prvi, 3000 v drugi in 6000 v tretji tovarni. Izmed 10000 izdelanih telefonov izberemo naključnega. Kakšna je verjetnost, da je bil izdelan v prvi tovarni, če ima okvarjeno baterijo?

Če je A_i dogodek, da je telefon izdelan v i -ti tovarni, in B , da ima okvarjeno baterijo, potem je

$$\begin{aligned} P(A_1|B) &= \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)} = \\ &= \frac{0.01 \cdot 0.1}{0.01 \cdot 0.1 + 0.02 \cdot 0.3 + 0.03 \cdot 0.6} = 4\% \end{aligned}$$

Zgled. Prisotnost nekaterih bolezni v medicini ugotavljajo z diagnostičnimi testi. Načeloma je dober test takšen, ki je z veliko verjetnostjo pozitiven v primeru prisotne bolezni, in z majhno verjetnostjo pozitiven, če nismo bolni.

Z mamogramom preverjamo prisotnost raka na dojki, za katerim zboli približno 1 procent žensk, starih med 40 in 50 let. Naj bo B dogodek, da smo bolni, in A dogodek, da je test pozitiven. Za mamogramski test, velja, da je $P(A|B) = 90\%$ in $P(A|\bar{B}) = 10\%$.

Ker test ni povsem zanesljiv, je verjetnost, da je test negativen, čeprav smo bolni, enaka $P(\bar{A}|B) = 1 - P(A|B) = 10\%$. Kakšna pa je verjetnost, da smo zares bolni, če je test pozitiven? Upoštevajoč, da je $\{B, \bar{B}\}$ popoln sistem dogodkov, lahko po Bayesu izračunamo

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} = \frac{0.9 \cdot 0.01}{0.9 \cdot 0.01 + 0.1 \cdot 0.99} = \frac{0.009}{0.009 + 0.099} \doteq 0.083 \doteq 8\%.$$

Torej je verjetnost, da smo bolni, če je test tako pokazal, komaj 8%? Presenetljivi rezultat je posledica tega, da je bolezen razmeroma redka, verjetnost pozitivnega izida pri zdravih ženskah pa razmeroma velika. Če bi sistematično testirali 1000 kandidat, bi bilo v povprečju bolnih

³Izvirno izjavo se pripisuje statistiku H. Jeffreysu.

10, od katerih bi jih 9 dobilo pozitiven izvid testa, med 990 zdravimi pa bi dobili še 99 lažno pozitivnih izidov, torej je resnično bolnih med pozitivnimi izidi ravno $9/(9+99) \doteq 8\%$.

Zato je testiranje v primeru pozitivnega izida nujno ponoviti. V tem primeru nas zanima verjetnost $P(B|A_1 \cap A_2)$, kjer sta A_i dogodka, da je i -ti test pozitiven. Denimo, da lahko privzamemo, da je $P(A_1 \cap A_2|B) = P(A_1|B)P(A_2|B)$ in $P(A_1 \cap A_2|\bar{B}) = P(A_1|\bar{B})P(A_2|\bar{B})$ (temu se reče **pogojna neodvisnost** dogodkov A_1, A_2 pri pogoju B oziroma \bar{B}). Potem je

$$P(B|A_1 \cap A_2) = \frac{P(A_1 \cap A_2|B)P(B)}{P(A_1 \cap A_2|B)P(B) + P(A_1 \cap A_2|\bar{B})P(\bar{B})} = \frac{0.9 \cdot 0.9 \cdot 0.01}{0.9 \cdot 0.9 \cdot 0.01 + 0.1 \cdot 0.1 \cdot 0.99} \doteq 45\%,$$

kar je bistveno več, a še vedno daleč od zanesljive diagnoze. Zavedati pa se moramo, da gornji izračuni veljajo le ob predpostavki, da je bolezen neodvisno in enakomerno porazdeljena po celotni populaciji. Če testirana oseba spada v katero izmed rizičnih skupin, bo njena bolezen v primeru pozitivnega testa bistveno verjetnejša.

Zgled (Javnomnenjske raziskave). Kandidata X podpira 55% volilcev, kandidata Y pa 45%. Kaj pokažejo predvolilne ankete, če je nanje pripravljenih odgovarjati 30% volivcev kandidata X in 50% volivcev kandidata Y , ostali pa sodelovanje odklonijo.

Imamo $P(X) = 0.55$, $P(Y) = 0.45$, $P(A|X) = 0.3$, $P(A|Y) = 0.5$. Pri tem je $\{X, Y\}$ popoln sistem dogodkov. Po Bayesovem izreku je verjetnost, da naključni anketiranec podpira X , enaka

$$P(X|A) = \frac{P(A|X)P(X)}{P(A|X)P(X) + P(A|Y)P(Y)} = \frac{0.165}{0.165 + 0.225} = 0.42.$$

Torej, napovedi kažejo, da bo zmagal Y z 58% glasov.

Zgled (Prometna nesreča). V mestu je 85% taksijev rumenih, ostali so beli. Ponoči neki taksi povzroči prometno nesrečo in zbeži. Očividec nesreče trdi, da je bil taksi bel. Vendar laboratorijski preizkus vida pokaže, da očividec barvo taksija v nočnih razmerah pravilno določi le v 80% primerov. Kakšna je verjetnost, da je nesrečo res povzročil beli taksi?

Namig: Poznate $P(OR|R)$, $P(OB|B)$, $P(R)$, $P(B)$. Iščete $P(B|OB)$.

Zgled (Goljufivi kovanec). Imate 1000 kovanecov. 999 je poštenih, eden pa je goljufiv, na njem vedno pade grb. Izberete naključni kovanec in ga vržete 5x zapored. Vsakič pade grb. Kakšna je verjetnost, da je ta kovanec goljufiv? Ker vam to ne zadošča, vržete še 5x in vsakič pade grb. Kakšna je zdaj verjetnost, da je goljufiv? Koliko padlih grbov potrebujete, da lahko z 99% gotovostjo trdite, da ste našli goljufiv kovanec?

Namig: Brez težav izračunate $P(n \text{ grbov} | \text{pošten})$, $P(n \text{ grbov} | \text{goljufiv})$, $P(\text{pošten})$, $P(\text{goljufiv})$. Iščete $P(g | n \text{ grbov})$.

2.6 Zaporedje neodvisnih poskusov

V praksi se pogosto srečamo z zaporedjem več poskusov, ki so nekako povezani med seboj. Obravnava takih primerov je nasploh zahtevna. Nekoliko enostavnejša je, kadar so poskusi med seboj neodvisni, torej je katerikoli dogodek enega poskusa neodvisen od kateregakoli dogodka drugega poskusa.

Zgled. Pri zaporedju n metov kovanca je izid pri posameznem metu neodvisen od ostalih metov.

Poskus z le dvema možnima izidoma A in \bar{A} imenujemo **Bernoullijev poskus**. Če označimo verjetnost prvega izida s $p = P(A)$, potem je verjetnost drugega izida natanko določena, $q = P(\bar{A}) = 1 - p$. Zaporedje enakih, a neodvisnih Bernoullijevih poskusov imenujemo **Bernoullijevo zaporedje**. Osnovno vprašanje v zvezi z Bernoullijevim zaporedjem je določiti verjetnosti za k ponovitev izida A v n ponovitvah poskusa.

Trditev. Naj bo $p = P(A)$ verjetnost dogodka A v nekem Bernoullijevem poskusu. Potem velja:

(i) Verjetnost, da se je dogodek A v n ponovitvah poskusa zgodil natanko k -krat, je enaka

$$P(n, p, k) = \binom{n}{k} p^k q^{n-k}. \quad (\text{Bernoullijev obrazec})$$

(ii) Verjetnost, da se je dogodek A zgodil k -tič natanko v n -ti ponovitvi, je enaka

$$P(n-1, p, k-1) \cdot p = \binom{n-1}{k-1} p^k q^{n-k}. \quad (\text{Pascalov obrazec}).$$

Posebej, verjetnost, da se je dogodek A prvič zgodil v n -ti ponovitvi, je enaka $p q^{n-1}$.

Dokaz. Naj bo A_i dogodek, da se je v i -to ponovitvi poskusa zgodil dogodek A . Potem je

$$A_1 \cap A_2 \cap \dots \cap A_k \cap \bar{A}_{k+1} \cap \dots \cap \bar{A}_n$$

dogodek, da se je v n ponovitvah poskusa najprej k -krat zgodil A , nato pa $(n-k)$ -krat \bar{A} . Zaradi neodvisnosti je verjetnost tega dogodka enaka produktu $p^k q^{n-k}$. Toda k ponovitev dogodka A bi se lahko zgodilo tudi z drugačnim vrstnim redom dogodkov A in \bar{A} . Vseh možnih razporeditev je $\binom{n}{k}$, saj izbiramo k od možnih n mest. Ker so načini med seboj nezdružljivi, lahko njihove verjetnosti seštejemo in rezultat je Bernoullijev obrazec.

Za Pascalov obrazec pa lahko razmišljam o rekurzivno. Dogodek A se k -tič ponovi natanko v n -ti ponovitvi poskusa, če se je ponovil $(k-1)$ -krat v prvih $(n-1)$ ponovitvah in nato še v n -ti. Ker je zadnja ponovitev neodvisna od prejšnjih, ustrezni verjetnosti zmnožimo in dobimo željeni rezultat. Zadnji obrazec pa je seveda le poseben primer prejšnjega za $k=1$. \square

Zgled. Verjetnost, da v 5 metih kovanca padeta 2 grba, je enaka $P(5, p, 2) = \binom{5}{2} p^2 (1-p)^3$, kar je približno 31%, če je $p = 1/2$. Verjetnost, da je drugi grb padel v petem metu, pa je enaka $\binom{4}{1} p^2 (1-p)^3$ oziroma približno 12%, če je $p = 1/2$.

3 Diskretne slučajne spremenljivke

3.1 Diskretne in zvezne slučajne spremenljivke

Naj bo (S, P) neki verjetnostni prostor. **Slučajna spremenljivka** je vsaka⁴ preslikava $X: S \rightarrow \mathbb{R}$, ki vsakemu izidu priredi neko realno število. Zaloga vrednosti oziroma slika slučajne spremenljivke je torej neka podmnožica $X(S) = \{X(s) \mid s \in S\} \subseteq \mathbb{R}$.

Zgled. Vržemo dve kocki. Naj bo X vsota števila padlih pik. Njena zaloga vrednosti je $X(S) = \{2, 3, \dots, 12\}$, torej končna.

Zgled. Kovanec mečemo do prvega grba. Naj bo X potrebno število metov. Zaloga vrednosti $X(S) = \mathbb{N}$ je števno neskončna.

Zgled. Puščico izstrelimo v okroglo tarčo na neskončni steni. Naj bo X razdalja zadetka od središča tarče. Zaloga vrednosti $X(S) = [0, \infty)$ je neštevno neskončna.

⁴Ta definicija je nekoliko poenostavljena, a je smiselna, ker smo privzeli, da je algebra dogodkov $P(S)$ kar potenčna množica množice izidov S . V zahtevnejših tekstih se običajno zahteva, da je slučajna spremenljivka *merljiva* preslikava $X: S \rightarrow E$, kjer je E poljuben *merljiv* prostor v smislu abstraktne teorije mere.

Glede na kardinalnost zaloge vrednosti ločimo dva osnovna tipa slučajnih spremenljivk. **Diskretna** oziroma **diskretno porazdeljena slučajna spremenljivka** ima končno ali števno zalogo vrednosti. **Zvezna** oziroma **zvezno porazdeljena slučajna spremenljivka** pa ima neštevno neskončno zalogo vrednosti. Natančnejšo definicijo zveznih slučajnih spremenljivk bomo obravnavali kasneje.

Označimo z $(X = x)$ dogodek, da slučajna spremenljivka X zavzame vrednost $x \in \mathbb{R}$. Temu dogodku ustreza množica izidov $\{s \in S \mid X(s) = x\} = X^{-1}(x)$, torej praslina elementa x glede na X . Podobno lahko z $(x < X < y)$ označimo dogodek, da X zavzame vrednosti iz intervala (x, y) . Temu dogodku ustreza množica $\{s \in S \mid x < X(s) < y\} = X^{-1}(x, y)$.

Osnovno vprašanje bo seveda, kakšna je verjetnost teh dogodkov, pa tudi, kakšna je "povprečna" vrednost neke slučajne spremenljivke in kakšna odstopanja od tega povprečja lahko pričakujemo. Pri matematični verjetnosti lastnosti slučajnih spremenljivk napovedujemo teoretično, statistiki pa prepustimo, da z ustreznimi slučajnimi spremenljivkami opiše empirično opazovane pojave. Zaradi enostavnejše obravnave bomo v tem besedilu diskretne in zvezne slučajne spremenljivke obravnavali ločeno.

3.2 Verjetnostna funkcija in pričakovana vrednost

Najosnovnejša pojma v zvezi z diskretnimi slučajnimi spremenljivkami sta njena verjetnostna funkcija in pričakovana vrednost.

Zgled. Denimo, da 5-krat vržemo pošten kovanec. Kakšna je verjetnost, da pade 0, 1, 2, 3, 4, 5 grbov? Koliko pa je povprečno število grbov, ki jih lahko pričakujemo v velikem primeru ponovitev tega poskusa?

Ustrezne verjetnosti izračunamo z Bernoullijevim obrazcem $P(n, 1/2, x)$ in dobimo tabelo:

x	0	1	2	3	4	5
$P(X = x)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

Izračunane verjetnosti pomenijo, da se bo v 32 ponovitvah poskusa v povprečju 1-krat zgodilo, da bo padlo 0 grbov, 5-krat, da bo padel 1 grb in tako dalje. Zato je smiselno "povprečno" število grbov definirati kot

$$0 \cdot \frac{1}{32} + 1 \cdot \frac{5}{32} + 2 \cdot \frac{10}{32} + 3 \cdot \frac{10}{32} + 4 \cdot \frac{5}{32} + 5 \cdot \frac{1}{32} = \frac{5}{2} = 2.5,$$

kar se tudi sklada z našo predstavo o "povprečju".

Za poljubno diskretno slučajno spremenljivko $X: S \rightarrow \mathbb{R}$ lahko definiramo naslednje pojme. **Verjetnostna funkcija** (angl. probability mass function - pmf) diskretne slučajne spremenljivke X je funkcija $p_X: \mathbb{R} \rightarrow [0, 1]$, podana s predpisom

$$p_X(x) = P(X = x).$$

(Kumulativna) porazdelitvena funkcija (angl. cumulative distribution function - cmf) diskretne slučajne spremenljivke X je funkcija $F_X: \mathbb{R} \rightarrow [0, 1]$, podana s predpisom

$$F_X(x) = P(X \leq x).$$

Pričakovana vrednost (angl. expectation) ali **matematično upanje** diskretne slučajne spremenljivke X je vrednost $E(X) \in \mathbb{R}$, podana z vsoto

$$E(X) = \sum_{x \in X(S)} x P(X = x),$$

če je ta vsota absolutno konvergentna. V nasprotnem primeru rečemo, da pričakovana vrednost za X ne obstaja.

Zgled. Vrnimo se k metu kovanca iz prejšnjega zgleda. V prejšnji tabeli smo že izračunali vrednosti verjetnostne funkcije p_X , prav tako smo že izračunali pričakovano vrednost $E(X) = 2.5$.

x	0	1	2	3	4	5
$P_X(x)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

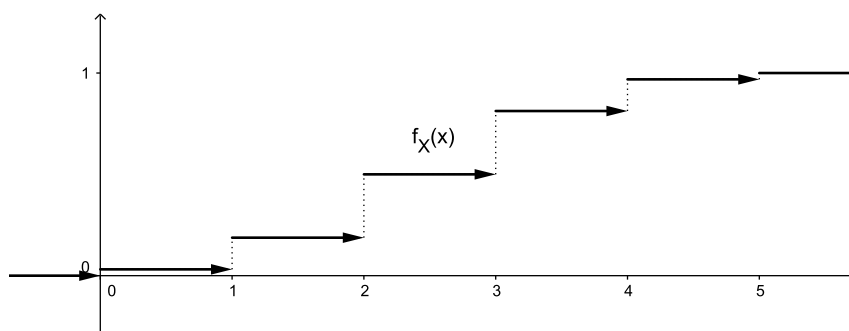
Dogovorimo se še, da lahko verjetnostno funkcijo za diskretno porazdeljeno X podamo tudi v obliki tabele, v kateri so zgoraj izidi, spodaj pa verjetnosti posameznih izidov. V tem primeru pišemo

$$X \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ \frac{1}{32} & \frac{5}{32} & \frac{10}{32} & \frac{10}{32} & \frac{5}{32} & \frac{1}{32} \end{pmatrix}.$$

Ostane nam še kumulativna porazdelitvena funkcija za X , ki je enaka

$$f_X(x) = \begin{cases} 0, & x < 0, \\ 1/32, & 0 \leq x < 1, \\ 6/32, & 1 \leq x < 2, \\ 16/32, & 2 \leq x < 3, \\ 26/32, & 3 \leq x < 4, \\ 31/32, & 4 \leq x < 5, \\ 1 & 5 \leq x. \end{cases}$$

Graf porazdelitvene funkcije ima torej stopničasto obliko:



Zgled (Sichermanove kocke). Ana in Bine stavita, kdo od njiju bo vrgel večjo vsoto pik na dveh kockah. Ana meče običajen par kock s ploskvami od 1 do 6 pik. Bine pa meče nenavadni kocki z oznakami 1,3,4,5,6,8 na prvi in 1,2,2,3,3,4 na drugi kocki. Ali je igra poštena?

Označimo z X vsoto pik pri Ani in z Y vsoto pik pri Binetu. Potem velja $P(X = 2) = P(Y = 2) = 1/36$, saj lahko Bine in Ana vsoto 2 dobita le na en način. Podobno velja $P(X = 3) = P(Y = 3) = 2/36$, saj lahko Ana vsoto 3 dobi na dva načina (1 na prvi kocki in 2 na drugi, ali obratno), prav tako Bine (1 na prvi kocki, ter dve različni možnosti za 2 na drugi kocki). Podobno velja, da je $P(X = k) = P(Y = k)$ za vse $k = 2, \dots, 12$. Igra je torej poštena, saj sta slučajni spremenljivki X in Y enako porazdeljeni:

$$X \sim Y \sim \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \end{pmatrix}.$$

Bralec je gotovo opazil, da je vsota števil v spodnji vrstici tabele, ki opisuje porazdelitev za X , enaka 1. To seveda ni naključje.

Trditev. Naj bo (S, P) verjetnostni prostor in $X: S \rightarrow \mathbb{R}$ diskretna slučajna spremenljivka z verjetnostno funkcijo $p_X: S \rightarrow [0, 1]$. Potem velja:

- (i) $\sum_{x \in X(S)} p_X(x) = 1$.
- (ii) $p_X(x) = 0$ za vse $x \in \mathbb{R} \setminus X(S)$.
- (iii) Predpis $P'(x) = p_X(x)$ določa verjetnost P' na prostoru izidov $X(S)$.

Dokaz.

- (i) Ker je X diskretna slučajna spremenljivka, je $X(S) = \{x_1, x_2, \dots\}$ končna ali števna množica. Zato zapišemo

$$\begin{aligned} \sum_{x \in X(S)} p_X(x) &= \sum_{i \in \mathbb{N}} P(X = x_i) = \sum_{i \in \mathbb{N}} P(\{s \in S \mid X(s) = x_i\}) = \\ &= P\left(\bigcup_{i \in \mathbb{N}} \{s \in S \mid X(s) = x_i\}\right) = P(S) = 1. \end{aligned}$$

Pri tem smo upoštevali, da so množice $\{s \in S \mid X(s) = x_i\}$ paroma disjunktne, in uporabili števno aditivnost verjetnosti P .

- (ii) Za $x \notin X(S)$ velja $\{s \in S \mid X(s) = x\} = \emptyset$ in zato $p_X(x) = 0$.
- (iii) Ker je $X(S) = \{x_1, x_2, \dots\}$ števna množica, števila $p_i = p_X(x_i)$ ustrezajo pogojem izreka o točkasti verjetnosti iz razdelka 1.2.5. Manjkajoče podrobnosti prepuščamo v razmislek bralcu.

□

Zadnji del prejšnje trditve nam pove, da slučajna spremenljivka X dani verjetnostni prostor (S, P) preoblikuje v nov verjetnostni prostor $(X(S), P')$. Lahko pa razmišljamo tudi drugače: če vnaprej predpišemo možne vrednosti slučajne spremenljivke X in njihove verjetnosti $p_X(x)$, smo s tem določili tako spremenljivko X kot tudi primeren verjetnostni prostor (S, P) .

Posledica. Naj bo $x_1, x_2, \dots \in \mathbb{R}$ zaporedje različnih realnih števil in naj bo $p_1, p_2, \dots \in [0, 1]$ zaporedje ne nujno različnih števil z lastnostjo $\sum_{i \in \mathbb{N}} p_i = 1$. Tedaj obstajata verjetnostni prostor (S, P) in slučajna spremenljivka $X: S \rightarrow \mathbb{R}$ z lastnostjo $X(S) = \{x_1, x_2, \dots\}$ in $p_X(x_i) = p_i$ za vse $i \in \mathbb{N}$.

Dokaz. Vzamemo $S = \mathbb{N}$ in $P(i) = p_i$. Potem je (S, P) verjetnostni prostor po izreku 1.2.5 (različica za števno množico izidov), slučajna spremenljivka X na S pa je definirana s predpisom $X(i) = x_i$. □

Opomba. Gornja trditev in posledica seveda veljata tudi, če namesto o števni govorimo o končni množici oziroma $X(S)$. V dokazu pa smo se sklicali na števno verzijo izreka o točkasti verjetnosti, ki je sicer nismo dokazali. Bralca vabim, da poskusi sam dokazati števno verzijo izreka.

Zgled. Porazdelitvena shema

$$X \sim \begin{pmatrix} -2 & 0 & \sqrt{7} \\ 0.1 & 0.4 & 0.5 \end{pmatrix}$$

določa slučajno spremenljivko X , četudi verjetnostnega prostora (S, P) eksplicitno ne poznamo.

Trditev. Naj bo $X: S \rightarrow \mathbb{R}$ diskretna slučajna spremenljivka z zalogo vrednosti $X(S) = \{x_1, x_2, \dots\}$, kjer je $x_i < x_{i+1}$, in naj bo $F_X(x) = P(X \leq x)$ njena porazdelitvena funkcija. Potem velja:

$$(i) F_X(x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} p_X(x_i) \text{ za vsak } x \in \mathbb{R}.$$

(ii) F_X je odsekoma konstantna, naraščajoča in desno zvezna.

$$(iii) \lim_{x \rightarrow -\infty} F_X(x) = 0 \text{ in } \lim_{x \rightarrow \infty} F_X(x) = 1.$$

Dokaz.

(i) Ker je $X(S)$ števna množica, lahko dogodek $(X \leq x)$ zapišemo kot unijo poddogodkov, med katerimi je nepraznih le števno mnogo:

$$(X \leq x) = \bigcup_{t \leq x} (X = t) = \left(\bigcup_{x_i \leq x} (X = x_i) \right) \cup \left(\bigcup_{t \leq x, t \notin X(S)} (X = t) \right) = \bigcup_{x_i \leq x} (X = x_i).$$

Rezultat sledi zaradi števne aditivnosti verjetnosti P .

(ii) Iz prejšnje točke neposredno sledi, da je F_X odsekoma konstantna in zato desno zvezna na intervalih $[x_i, x_{i+1})$. Ker so vrednosti $P(X = x_i)$ nenegativne, je F_X tudi naraščajoča.

(iii) Vrednost prve limite sledi iz dejstva, da za vsak $x < x_1$ velja $F_X(x) = 0$. Za drugo pa je potrebno upoštevati, da je $P(S) = 1$, torej (neskončna) vsota $\sum_{i \in \mathbb{N}} P(X = x_i)$ konvergira proti ena. Z drugimi besedami,

$$1 = \lim_{n \rightarrow \infty} \sum_{i=1}^n P(X = x_i) = \lim_{n \rightarrow \infty} \sum_{x_i \leq x_n} P(X = x_i) = \lim_{n \rightarrow \infty} F_X(x_n) \leq \lim_{x \rightarrow \infty} F_X(x) \leq 1.$$

□

Zgled. Za slučajno spremenljivko X iz prejšnjega zgleda velja $F_X(x) = \begin{cases} 0, & x < -2, \\ 0.1, & -2 \leq x < 0, \\ 0.5, & 0 \leq x < \sqrt{7}, \\ 1, & x \geq \sqrt{7}. \end{cases}$

3.3 Funkcije slučajnih spremenljivk in njihova pričakovana vrednost

Spomnimo se, da smo pričakovano vrednost ali matematično upanje diskretne slučajne spremenljivke definirali kot

$$E(X) = \sum_{x \in X(S)} x P(X = x),$$

če ta vrsta konvergira absolutno, kar pomeni, da je $\sum_{x \in X(S)} |x P(X = x)| < \infty$.

Zgled (Francoska ruleta). Na ruleti v igralnici so števila od 0 do 36. Številka 0 je zelena, sicer pa so sode številke rdeče, lihe pa črne. Če stavimo 1 EUR na rdečo, dobimo nazaj 2 EUR v primeru, da se kroglica ustavi ustrezni številki, sicer pa vložek izgubimo. Zasluge lahko torej opišemo s slučajno spremenljivko X , ki ima vrednost 1, če se kroglica ustavi na sodem neničelnem številu, kar se zgodi z verjetnostjo $18/37$, in vrednost -1 , sicer, kar se zgodi z verjetnostjo $19/37$. Povedano drugače,

$$P(X = 1) = \frac{18}{37}, P(X = -1) = \frac{19}{37}, \text{ oziroma } X \sim \begin{pmatrix} 1 & -1 \\ \frac{18}{37} & \frac{19}{37} \end{pmatrix}.$$

Pričakovana vrednost za X je torej

$$E(X) = 1 \cdot \frac{18}{37} + (-1) \cdot \frac{19}{37} = -\frac{1}{37} \doteq -0.027,$$

torej v povprečju pri vsaki igri izgubimo 3% vložka. To se sicer ne zdi veliko, a je dovolj, da igralnici zagotavlja zanesljiv dobiček.

Zgled. V rešitvi znamenitega Baselskega problema je Euler dokazal, da je

$$\frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \dots = \sum_{n \in \mathbb{N}} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Po posledici iz prejšnjega razdelka lahko definiramo slučajno spremenljivko X z zalogo vrednosti \mathbb{N} , za katero velja $P(X = n) = \frac{1}{cn^2}$, kjer je $c = \pi^2/6$, saj je $\sum_{n \in \mathbb{N}} P(X = n) = 1$. Toda njena pričakovana vrednost $E(X)$ ne obstaja, saj je vsota

$$\sum_{n \in \mathbb{N}} nP(X = n) = \sum_{n \in \mathbb{N}} \frac{1}{n} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \dots$$

enaka harmonični vrsti, ki ni konvergentna.

Naj bo $X: S \rightarrow \mathbb{R}$ diskretna slučajna spremenljivka na verjetnostnem prostoru (S, P) in naj bo $g: \mathbb{R} \rightarrow \mathbb{R}$ poljubna funkcija. Kompozitum

$$g \circ X: S \rightarrow \mathbb{R}$$

označimo z $g(X)$ in imenujemo **funkcija slučajne spremenljivke**.

Trditev. Funkcija $g(X)$ je diskretna slučajna spremenljivka s pričakovano vrednostjo

$$E(g(X)) = \sum_{x \in X(S)} g(x)P(X = x),$$

če je vsota absolutno konvergentna.

Dokaz. Očitno je $g \circ X: S \rightarrow \mathbb{R}$ funkcija s števno zalogo vrednosti. Označimo $Y = g(X)$ in izberimo poljuben $y \in Y(S)$. Tedaj je

$$(Y = y) = (g(X) = y) = \bigcup_{x \in X(S): g(x)=y} (X = x)$$

in zato

$$\begin{aligned} E(Y) &= \sum_{y \in Y(S)} yP(Y = y) = \sum_{y \in Y(S)} yP\left(\bigcup_{x \in X(S): g(x)=y} (X = x)\right) = \\ &= \sum_{y \in Y(S)} y \left(\sum_{x \in X(S): g(x)=y} P(X = x) \right) = \sum_{y \in Y(S)} \left(\sum_{x \in X(S): g(x)=y} g(x)P(X = x) \right) = \\ &= \sum_{x \in X(S)} g(x)P(X = x), \end{aligned}$$

saj je dvojna vsota po vseh možnih x, y z lastnostjo $g(x) = y$ enaka kar vsoti po vseh x . □

Zgled. Naj bo X število pik pri metu poštene kocke in naj bo $g(x) = x^2$. Potem je $g(X) = X^2$ slučajna spremenljivka, ki predstavlja kvadrat števila pik, torej velja

$$X^2 \sim \begin{pmatrix} 1 & 4 & 9 & 16 & 25 & 36 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

Njeno matematično upanje je enako $E(X^2) = (1 + 4 + 9 + 16 + 25 + 36)/6 = 91/6$.

Posledica. Naj bo $X: S \rightarrow \mathbb{R}$ diskretna slučajna spremenljivka in $E(X)$ njeno matematično upanje. Potem velja:

- (i) $E(\alpha X) = \alpha E(X)$ za poljubno število $\alpha \in \mathbb{R}$.
- (ii) $E(f(X) + g(X)) = E(f(x)) + E(g(x))$ za poljubni funkciji $f, g: \mathbb{R} \rightarrow \mathbb{R}$.
- (iii) $E(E(X)) = E(X)$.

Dokaz. Za prvi dve enakosti v prejšnji trditvi zamenjamo $g(x)$ z αx oziroma $(f+g)(x)$ in ustrezno preoblikujemo rezultat. Zadnja enakost sledi iz dejstva, da je $E(X)$ konstanta, za konstante pa je $E(a) = a$. □

Če je ustrezna vsota absolutno konvergentna, izraz $E(X^n)$ imenujemo **n -ti moment** slučajne spremenljivke X , izraz $E((X - E(X))^n)$ pa **n -ti centralni moment**. Posebej pomemben je drugi centralni moment, ki ga imenujemo **varianca** ali **dispersija slučajne spremenljivke**, njegov koren pa imenujemo **standardni odklon** ali **standardna deviacija** slučajne spremenljivke X :

$$\text{var}(X) = D(X) = E(X - E(X))^2, \quad \text{stdev}(X) = \sigma(X) = \sqrt{\text{var}(X)}.$$

Pogosto uporabljamo tudi oznaki $\mu_X = E(X)$ za pričakovano vrednost in $\sigma_X = \text{stdev}(X)$ za standardni odklon slučajne spremenljivke X . Če je slučajna spremenljivka ena sama, pišemo kar μ oziroma σ .

Trditev. Naj bo X slučajna spremenljivka, tako da obstajata $E(X)$ in $E(X^2)$. Potem za njeno varianco in standardni odklon veljata zveze:

- (i) $\text{var}(X) = E(X^2) - E(X)^2$.
- (ii) $\text{var}(aX + b) = a^2 \text{var}(X)$ za poljubna $a, b \in \mathbb{R}$.
- (iii) $\sigma(aX + b) = |a| \sigma(X)$ za poljubna $a, b \in \mathbb{R}$.

Dokaz. Iz definicije variance z upoštevanjem lastnosti $E(X)$ dobimo prvo trditev

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

Podobno dokažemo tudi drugo, tretja pa je seveda njena neposredna posledica. □

Zgled. V primeru, ko je X število pik pri metu kocke, smo že izračunali $E(X) = 7/2$ in $E(X^2) = 91/6$. Podobno bi izračunali še tretji moment in dobili $E(X^3) = \frac{1+8+27+64+125+216}{6} = \frac{441}{6}$. Za prvi centralni moment dobimo $E(X - E(X)) = E(X) - E(E(X)) = 0$. Za izračun variance pa uporabimo prvo točko zadnje trditve in sledi

$$\text{var}(X) = E(X^2) - E(X)^2 = \frac{91}{6} - \frac{49}{4} = \frac{105}{36} \doteq 2.92.$$

Zato je $\sigma(X) = \sqrt{\text{var}(X)} \doteq 1.71$.

3.4 Nekaj standardnih tipov dss

V tem razdelku bomo najprej našli nekaj najpogostejših tipov diskretnih slučajnih spremenljivk, nato pa opisali njihove osnovne karakteristike: verjetnostno funkcijo $p_X(x)$, pričakovano vrednost $E(X)$ in varianco $\text{var}(X)$.

3.4.1 Bernoullijeva dss

Bernoullijeva dss $X \sim \text{Ber}(p)$ ima le dve možni vrednosti, 1 in 0, pri čemer parameter p označuje verjetnost, da je $X = 1$.

Zgled. Če je X število grbov pri metu poštenega kovanca, je $X \sim \text{Ber}(1/2)$.

Trditev. Za $X \sim \text{Ber}(p)$ velja

$$p_X(x) = \begin{cases} p, & x = 1, \\ q, & x = 0, \end{cases}$$

$E(X) = p$ in $\text{var}(X) = pq$, kjer je $q = 1 - p$.

Dokaz. Zapišemo lahko $X \sim \begin{pmatrix} 0 & 1 \\ p & q \end{pmatrix}$. Očitno je $\sum_{x \in X(S)} p_X(x) = p + q = 1$, zato je p_X res verjetnostna funkcija. Sledi $E(X) = 1 \cdot p + 0 \cdot q = p$. Podobno je $E(X^2) = 1^2 \cdot p + 0^2 \cdot q = p$ in zato $\text{var}(X) = E(X^2) - E(X)^2 = p - p^2 = pq$. \square

3.4.2 Enakomerna dss

Enakomerna dss $X \sim U(n; a, b)$ ima n možnih vrednosti, ki so vse enako verjetne. Zaloga vrednosti je torej $X(S) = \{a, a + 1, \dots, b\}$, kjer je $a < b$ in $b - a = n - 1$. Če je zaloga vrednosti enaka $X(S) = \{1, \dots, n\}$, govorimo o standardizirani enakomerni dss X , oznaka $X \sim U(n)$.

Zgled. Če je X število pik pri metu poštenih kocke, velja $X \sim U(6; 1, 6)$ oziroma kar $X \sim U(6)$.

Trditev. Za $X \sim U(n; a, b)$ velja $p_X(x) = \frac{1}{n}$ za vsak $x \in \{a, \dots, b\}$, $E(X) = \frac{a+b}{2}$ in $\text{var}(X) = \frac{n^2-1}{12}$.

Dokaz. Ker ima vsak od n možnih izidov enako verjetnost in mora biti vsota enaka 1, je $P_X(x) = \frac{1}{n}$ za vse ustrezne x . Upoštevajoč, da je $b = a + n - 1$, za pričakovano vrednost dobimo

$$\begin{aligned} E(X) &= \sum_{x=a}^{a+n-1} x \frac{1}{n} = \frac{a + (a+1) + \dots + (a+n-1)}{n} = \\ &= \frac{na + \frac{(n-1)n}{2}}{n} = \frac{2a + n - 1}{2} = \frac{a + b}{2}. \end{aligned}$$

Za izračun drugega momenta uporabimo obrazec za vsoto kvadratov $\sum_{x=1}^m x^2 = \frac{m(m+1)(2m+1)}{6}$. Od tod dobimo

$$\begin{aligned} E(X^2) &= \sum_{x=a}^{a+n-1} x^2 \frac{1}{n} = \frac{1}{n} \left(\sum_{x=1}^{a+n-1} x^2 - \sum_{x=1}^{a-1} x^2 \right) = \\ &= \frac{(a+n-1)(a+n)(2a+2n-1) + (a-1)a(2a-1)}{6n} = \\ &= \frac{6a^2 + 6an - 6a + 2n^2 - 3n + 1}{6} \end{aligned}$$

in sledi

$$\begin{aligned}\text{var}(X) &= E(X^2) - E(X)^2 = \\ &= \frac{6a^2 + 6an - 6a + 2n^2 - 3n + 1}{6} - \left(\frac{2a + n - 1}{2}\right)^2 = \\ &= \frac{n^2 - 1}{12},\end{aligned}$$

spet po daljšem preurejanju izrazov. □

3.4.3 Geometrijska dss

Geometrijska dss $X \sim \text{Geo}(p)$ predstavlja število ponovitev Bernoullijevega poskusa do prvega uspeha, zato je njena zaloga vrednosti $X(S) = \{1, 2, \dots\}$.

Zgled. Če je X število metov poštenega kovanca do prvega grba, je $X \sim \text{Geo}(1/2)$.

Trditev. Za $X \sim \text{Geo}(p)$ velja $p_X(x) = pq^{x-1}$ za $x \in \{1, 2, \dots\}$, $E(X) = 1/p$ in $\text{var}(X) = q/p^2$.

Dokaz. Najprej preverimo, da je p_X res verjetnostna funkcija:

$$\begin{aligned}\sum_{x \in X(S)} p_X(x) &= \sum_{x=1}^{\infty} pq^{x-1} = p \sum_{k=0}^{\infty} q^k = \\ &= p \frac{1}{1-q} = 1,\end{aligned}$$

kjer smo upoštevali, da je vsota geometrijske vrste $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$, če je $|q| < 1$. Zdaj izračunamo

$$\begin{aligned}E(X) &= \sum_{x=1}^{\infty} x pq^{x-1} = p \sum_{x=0}^{\infty} x q^{x-1} = \\ &= p \frac{d}{dq} \left(\sum_{x=0}^{\infty} q^x \right) = p \frac{d}{dq} (1-q)^{-1} = \\ &= p(1-q)^{-2} = p \cdot p^{-2} = p^{-1},\end{aligned}$$

kjer smo upoštevali, da lahko vrsto odvajamo po členih, če je odvedena vrsta absolutno konvergentna. S podobnim trikom izračunamo še

$$\begin{aligned}E(X^2) &= \sum_{x=1}^{\infty} x^2 pq^{x-1} = p \sum_{x=0}^{\infty} x \cdot x q^{x-1} = \\ &= p \left(\sum_{x=0}^{\infty} (x+1)x q^{x-1} - \sum_{x=0}^{\infty} x q^{x-1} \right) = \\ &= p \left(\frac{d^2}{dq^2} \left(\sum_{x=0}^{\infty} q^{x+1} \right) - \frac{d}{dq} \left(\sum_{x=0}^{\infty} q^x \right) \right) = \\ &= p \left(\frac{d^2}{dq^2} (q(1-q)^{-1}) - \frac{d}{dq} (1-q)^{-1} \right) = \\ &= p(2(1-q)^{-2} + 2q(1-q)^{-3} - (1-q)^{-2}) = \\ &= p(p^{-2} + 2qp^{-3}) = p^{-1} + 2qp^{-2} = \frac{p+2q}{p^2} = \\ &= (2-p)p^{-2}.\end{aligned}$$

Zdaj sledi

$$\text{var}(X) = E(X^2) - E(X)^2 = (2-p)p^{-2} - (p^{-1})^2 = (1-p)p^{-2} = q/p^2.$$

□

3.4.4 Binomska dss

Binomska dss $X \sim \text{Bin}(n, p)$ predstavlja število uspehov v n ponovitvah Bernoullijevega poskusa. Njena zaloga vrednosti je $\{0, 1, \dots, n\}$.

Zgled. Če je X število grbov v 4 metih poštenega kovanca, je $X \sim (4, 1/2)$.

Trditev. Za $X \sim \text{Bin}(n, p)$ velja $p_X(x) = \binom{n}{x} p^x q^{n-x}$, $E(X) = np$ in $\text{var}(X) = npq$.

Dokaz. Izraz za $p_X(x)$ dobimo iz Bernoullijevega obrazca. Po binomskem izreku dobimo

$$\sum_{x=0}^n p_X(x) = \sum_{x=0}^n \binom{n}{x} p^x q^{n-x} = (p+q)^n = 1,$$

zato je p_X res verjetnostna funkcija. Za izračun $E(X)$ najprej preverimo enakost

$$x \binom{n}{x} = x \frac{n!}{x!(n-x)!} = \frac{n(n-1)!}{(x-1)!((n-1)-(x-1))!} = n \binom{n-1}{x-1}.$$

Sledi

$$\begin{aligned} E(X) &= \sum_{x=0}^n x p_X(x) = \sum_{x=0}^n x \binom{n}{x} p^x q^{n-x} = \\ &= \sum_{x=1}^n n \binom{n-1}{x-1} p \cdot p^{x-1} q^{(n-1)-(x-1)} = \\ &= np \sum_{k=0}^m \binom{m}{k} p^k q^{m-k} = np(p+q)^m = np, \end{aligned}$$

kjer smo v zadnji vrstici naredili substitucijo $m = n-1$ in $k = x-1$. Podobno izračunamo

$$\begin{aligned} E(X^2) &= \sum_{x=0}^n x^2 \binom{n}{x} p^x q^{n-x} = \sum_{x=1}^n x \cdot n \binom{n-1}{x-1} p \cdot p^{x-1} q^{n-x} = \\ &= np \left(\sum_{k=0}^m (k+1) \binom{m}{k} p^k q^{m-k} \right) = \\ &= np \left(\sum_{k=0}^m k \binom{m}{k} p^k q^{m-k} + \sum_{k=0}^m \binom{m}{k} p^k q^m \right) = \\ &= np \left(\sum_{k=1}^m m \binom{m-1}{k-1} p \cdot p^{k-1} q^{m-k} + (p+q)^m \right) = \\ &= np \left(mp \sum_{j=0}^r \binom{r}{j} p^j q^{r-j} + 1 \right) = \\ &= np(mp(p+q)^r + 1) = n^2 p^2 - np^2 + np, \end{aligned}$$

od koder sledi

$$\text{var}(X) = E(X^2) - E(X)^2 = np - np^2 = npq.$$

□

3.4.5 Pascalova in negativna binomska dss

Pascalova dss $X \sim \text{Pas}(k, p)$ šteje število ponovitev Bernoullijevega poskusa do vključno k -tega uspeha. Njena zaloga vrednosti je $X(S) = \{k, k+1, k+2, \dots\}$.

Zgled. $X \sim (5, 0.5)$ predstavlja število metov poštenega kovanca do pete padle cifre.

Trditev. Za $X \sim \text{Pas}(k, p)$ velja $p_X(x) = \binom{x-1}{k-1} p^k q^{x-k}$, $E(X) = \frac{k}{p}$ in $\text{var}(X) = \frac{kp}{q^2}$.

Dokaz. Obrazec za $p_X(x)$ dobimo z naslednjim razmislekom. Če je x število ponovitev do vključno k -tega uspeha, potem je zadnja ponovitev poskusa gotovo uspela, ostalih $k-1$ uspehov pa je razporejenih med prvih $x-1$ ponovitev poskusa. Vseh razporeditev je torej $\binom{x-1}{k-1}$, verjetnost k uspehov in $x-k$ neuspehov v posamezni razporeditvi pa je zaradi neodvisnosti enaka produktu $p^k q^{x-k}$, sledi

$$p_X(x) = \binom{x-1}{k-1} p^k q^{x-k}.$$

Veljavnost pogoja $\sum_{x=k}^{\infty} p_X(x) = 1$ preverimo z uporabo substitucije $r = x - k$:

$$\sum_{x=k}^{\infty} \binom{x-1}{k-1} p^k q^{x-k} = \sum_{r=0}^{\infty} \binom{r+k-1}{k-1} p^k q^r = p^k \sum_{r=0}^{\infty} \binom{r+k-1}{r} q^r = p^k (1-q)^{-k} = 1,$$

kjer smo v predzadnjem koraku uporabili formulo za binomsko vrsto, torej, Taylorjevo vrsto funkcije $f(q) = (1-q)^{-k}$ za $|q| < 1$.

Izraz za $E(X)$ bi prav tako lahko izračunali kot vsoto vrste. Delo pa si skrajšamo, če opazimo, da je $X = X_1 + \dots + X_k$, kjer je X_1 število ponovitev poskusa do prvega uspeha, X_2 število ponovitev poskusov od prvega do drugega uspeha in tako dalje. Vsaka od neodvisnih spremenljivk X_i je torej geometrijska s parametrom p in pričakovano vrednostjo $1/p$. Zato je

$$E(X) = E(X_1) + \dots + E(X_k) = k \cdot E(X_1) = k/p.$$

Podobno dobimo, da je $\text{var}(X) = k \cdot \text{var}(X_1) = kp/q^2$. □

Opomba. Natančni bralec je verjetno opazil, da je geometrijska slučajna spremenljivka poseben primer Pascalove za $k = 1$. Velja torej $X \sim \text{Pas}(1, p)$ natanko tedaj, ko $X \sim \text{Geo}(p)$.

Opomba. V nekaterih virih tako definirano Pascalovo ss imenujejo negativna binomska ss. Drugi viri pa negativno binomsko ss definirajo na nekoliko drugačen način, ki ga opisujemo v nadaljevanju.

Negativna binomska ss $X \sim \text{NBin}(n, p)$ šteje neuspehe do n -tega uspeha v zaporedju Bernoullievih poskusov. Njena zaloga vrednosti je $\{0, 1, 2, \dots\}$.

Zgled. Pošten kovanec mečemo do petega grba. Če je X število vseh padlih cifre, je $X \sim \text{NBin}(5, 1/2)$.

Trditev. Za $X \sim \text{NBin}(n, p)$ velja $p_X(x) = \binom{n+x-1}{n-1} p^n q^x$, $E(X) = \frac{nq}{p}$ in $\text{var}(X) = \frac{np}{q^2}$.

Dokaz. Najprej utemeljimo izraz za $p_X(x)$. Ker je x število neuspehov do n -tega uspeha, je bilo vseh ponovitev $n+x$. Prvih $n-1$ uspehov se je zgodilo v $n+x-1$ poskusih, kar se lahko zgodi na $\binom{n+x-1}{n-1}$ načinov, vsak od teh pa se zgodi z verjetnostjo $p^{n-1} q^x$. V zadnji, neodvisni ponovitvi, se nato zgodi še en uspeh z verjetnostjo p , kar da gornji izraz.

Premislimo, da je $\sum_{x=0}^{\infty} p_X(x) = 1$. Velja

$$\sum_{x=0}^{\infty} p_X(x) = p^n \left(\sum_{x=0}^{\infty} \binom{n+x-1}{x} q^x \right) = p^n (1-q)^{-n} = 1,$$

kjer smo upoštevali, da je neskončna vsota $\sum_{x=0}^{\infty} \binom{n+x-1}{x} q^x$ ravno enaka Taylorjevemu razvoju funkcije $f(q) = (1-q)^{-n}$ v okolici točke $q = 0$.

Pričakovano vrednost $E(X)$ izračunamo s pomočjo istega Taylorjevega razvoja in še odvajanja:

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \binom{n+x-1}{x} p^n q^x = p^n q \sum_{x=0}^{\infty} x \binom{n+x-1}{x} q^{x-1} = \\ &= p^n q \frac{d}{dq} \left(\sum_{x=0}^{\infty} \binom{n+x-1}{x} q^x \right) = p^n q \frac{d}{dq} (1-q)^{-n} = n p^n q (1-q)^{-n-1} = n q / p, \end{aligned}$$

kjer smo upoštevali, da ustrezna vrsta konvergira absolutno za $|q| < 1$, zato jo lahko členoma odvajamo. Podoben trik nato uporabimo še za izračun $E(X^2)$ in od tod $\text{var}(X)$. \square

Izrek. Če je $X \sim \text{NBin}(k, p)$, potem je njen modus enak $\text{Mod}(X) = \lfloor (k-1)q/p \rfloor$. Če je ta vrednost celo število, sta modusa dva, $(k-1)q/p$ in $(k-1)q/p - 1$.

Zgled (Biatlonski trening). Biatlonec pri vsakem strelu zadane tarčo z verjetnostjo $p = 0.6$. Na treningu ima na voljo neomejeno število strel, zato vsakič strelja tako dolgo, dokler ne zadane vseh 5 tarč. Kolikšno je povprečno število zgrešenih strel do zadetih petih tarč in katero število zgrešenih strel je najverjetnejše?

Označimo z X število zgrešenih strel. Potem je pričakovano število zgrešenih strel

$$E(X) = nq/p = 10/3 = 3.33,$$

najverjetnejše število zgrešenih strel pa je

$$\text{Mod}(X) = \lfloor (k-1)q/p \rfloor = \lfloor 8/3 \rfloor = 2.$$

Verjetnost, da je biatlonec zgrešil natanko 2 strela, je enaka

$$P(X = 2) = \binom{6}{2} \cdot (0.6)^5 \cdot (0.4)^2 \doteq 0.1866.$$

Začetne vrednosti funkcije p_X so razvidne iz tabele in grafa.

x	$p_X(x)$
0	0.0778
1	0.1555
2	0.1866
3	0.1742
4	0.1393
5	0.1003
...	...

3.4.6 Hipergeometrijska dss

Hipergeometrijska dss $X \sim \text{HGeo}(n, N, M)$ šteje ugodne elemente pri naključnem izboru vzorca n elementov (brez vračanja) izmed skupaj N elementov, med katerimi je M ugodnih, $M < N$. Zaloga vrednosti je torej podmnožica v $\{0, 1, 2, \dots, n\}$. Natančneje, če je neugodnih elementov manj kot je velikost vzorca, $N - M < n$, potem ima vzorec vsaj $n - (N - M)$ ugodnih elementov, zato je spodnja meja enaka $\max\{0, n - (N - M)\}$. Če pa je vseh ugodnih elementov manj kot je velikost vzorca, $M < n$, potem vsebuje vzorec največ M elementov, zato je zgornja meja enaka $\min\{n, M\}$. Torej je $X(S) = \{\max\{0, n - (N - M)\}, \dots, \min\{n, M\}\}$.

Zgled. Iz žare s 60 belimi in 40 črnimi kroglicami brez vračanja izberemo 10 kroglic. Če je X število belih med izbranimi, je $X \sim \text{HGeo}(10, 100, 60)$ in zaloga vrednosti je $X(S) = \{0, 1, \dots, 10\}$. Če pa bi izbrali 70 kroglic, bi imela slučajna spremenljivka $Y \sim \text{HGeo}(70, 100, 60)$ zalogo vrednosti $Y(S) = \{30, 31, \dots, 60\}$.

Trditev. Za $X \sim \text{HGeo}(n, N, M)$ velja $p_X(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$, $E(X) = \frac{nM}{N}$ in $\text{var}(X) = \frac{nM}{N} \frac{(N-n)(N-M)}{N(N-1)}$.

Dokaz. Izraz za $p_X(x)$ dobimo s preštevanjem možnih izborov vzorca n elementov. Vseh izborov n elementov izmed N je $\binom{N}{n}$, med temi pa je takih, pri katerih smo izbrali x ugodnih in $n-x$ neugodnih elementov natanko $\binom{M}{x} \binom{N-M}{n-x}$. Ustrezni kvocient je torej iskani izraz za $p_X(x)$.

Da bi preverili, da je $\sum_{x=0}^n p_X(x) = 1$, je potreben daljši račun in upoštevanje Vandermondejeve identitete $\binom{m+n}{r} = \sum_{k=0}^r \binom{m}{k} \binom{n}{r-k}$, ki ga prepustimo bralcu.

Tudi direktna algebraična izpeljava za $E(X)$ zahteva daljšo in zelo spretno manipulacijo z vrstami in binomskimi simboli. Delo si lahko spet olajšamo tako, da spremenljivko X predstavimo kot vsoto $X = X_1 + \dots + X_n$, kjer ima spremenljivka X_i vrednost 1, če je i -ti element vzorca ugoden, in 0 sicer. Torej je $X_i = \text{Ber}(M/N)$, saj je verjetnost $p = M/N$ zaradi simetrije enaka za vse i in je odvisna le od razmerja med ugodnimi in vsemi elementi celotnega nabora. Sledi

$$E(X) = n \cdot E(X_i) = \frac{nM}{N}.$$

Izračun variance preko vsote $X_1 + \dots + X_n$ pa oteži dejstvo, da slučajne spremenljivke X_i niso neodvisne, o čemer bomo sicer govorili šele v poglavju o slučajnih vektorjih. Kljub temu lahko izračunamo vrednost $E(X^2)$. Najprej razčlenimo

$$\begin{aligned} E(X^2) &= E((X_1 + \dots + X_n)^2) = \\ &= E(X_1^2 + \dots + X_n^2 + X_1 X_2 + \dots + X_{n-1} X_n) = \\ &= \sum_{i=1}^n E(X_i^2) + \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(X_i X_j). \end{aligned}$$

Ker je $X_i^2 = 1$ natanko tedaj, ko je $X_i = 1$, je $E(X_i^2) = 1 \cdot p + 0 \cdot 0 = p = M/N$. Po drugi strani za $i \neq j$ velja $X_i X_j = 1$ natanko tedaj, ko je $X_i = X_j = 1$, kar se zgodi z verjetnostjo

$$P(X_i X_j = 1) = P(X_i = 1, X_j = 1) = P(X_i = 1 | X_j = 1) P(X_j = 1) = \frac{M-1}{N-1} \cdot \frac{M}{N},$$

saj se ob pogoju, da je j -ti element vzorca ugoden, spremeni razmerje med preostalimi elementi vzorca. Zato je $E(X_i X_j) = \frac{M-1}{N-1} \cdot \frac{M}{N}$ za $i \neq j$. Sledi

$$E(X^2) = n \cdot \frac{M}{N} + n(n-1) \frac{M-1}{N-1} \cdot \frac{M}{N}$$

in željeni rezultat dobimo po preoblikovanju izraza $\text{var}(X) = E(X^2) - E(X)^2$. □

3.4.7 Poissonova dss

Poissonova dss $X \sim \text{Poiss}(\lambda)$ ima zalogo vrednosti $\{0, 1, \dots\}$ in je definirana s formulo za verjetnostno funkcijo $p_X(x) = \frac{\lambda^x e^{-\lambda}}{x!}$.

Zgled. Če je $X \sim \text{Poiss}(\lambda)$ za $\lambda = 2$, je porazdelitvena shema enaka

$$X \sim \left(\begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & \dots \\ 0.1353 & 0.2707 & 0.2707 & 0.1804 & 0.0902 & 0.0361 & 0.012 & 0.0034 & \dots \end{array} \right).$$

Trditev. Verjetnostna funkcija Poissonove porazdelitve res določa porazdelitev. Če je $X \sim \text{Poiss}(\lambda)$, potem velja $E(X) = \text{var}(X) = \lambda$.

Dokaz. Z uporabo Taylorjevega razvoja $e^\lambda = \frac{1}{0!} + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$ hitro preverimo, da velja

$$\sum_{x=0}^{\infty} p_X(x) = \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^\lambda = 1.$$

Podobno dobimo

$$E(X) = \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda.$$

in tudi

$$\begin{aligned} E(X^2) &= \sum_{x=0}^{\infty} x^2 \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \left(\sum_{x=0}^{\infty} x(x-1) \frac{\lambda^x}{x!} + \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} \right) = \\ &= e^{-\lambda} \left(\lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \right) = \\ &= e^{-\lambda} (\lambda^2 e^\lambda + \lambda e^\lambda) = \lambda^2 + \lambda, \end{aligned}$$

od koder zdaj hitro sledi še $\text{var}(X) = E(X^2) - E(X)^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$. □

Poissonova slučajna spremenljivka opisuje verjetnost, da se razmeroma redek dogodek zgodi x -krat v nekem časovnem intervalu, v katerem se povprečno zgodi λ takih dogodkov. Več o tem bomo izvedeli v naslednjem razdelku, tukaj naredimo le kratek zgled.

Zgled. Spletna stran ima v povprečju 5 obiskovalcev na minuto. Če jo v neki minuti obiše vsaj 12 obiskovalcev, pa se bo strežnik sesul. Kakšna je verjetnost, da se bo strežnik sesul v naslednji minuti?

Predpostavimo, da je število obiskovalcev Poissonova slučajna spremenljivka $X_i \sim \text{Poiss}(5)$. Verjetnost, da se strežnik sesuje v naslednji minuti, je potem enaka

$$p = P(X_i \geq 12) = 1 - \sum_{x=0}^{11} P(X_i = x) \doteq 0.005453.$$

3.5 Aproksimacija binomske s Poissonovo dss

V prejšnjem razdelku smo definirali Poissonovo dss in utemeljili njene lastnosti, nismo pa pojasnili, od kod takšna definicija in kakšne poskuse lahko z njo opisujemo. Nenavadna formula za verjetnostno funkcijo izvira iz računanja limite zaporedja.

Spomnimo se, da binomska porazdelitev opisuje število uspehov v zaporedju neodvisnih ponovitev Bernoullijevega poskusa. Denimo, da je število ponovitev zelo veliko, verjetnost uspeha v posameznem poskusu pa zelo majhna. Potem je Poissonova porazdelitev zelo dober približek za ustrezno binomsko porazdelitev.

Izrek (Zakon redkih dogodkov). Denimo, da je X_1, X_2, \dots zaporedje binomskih slučajnih spremenljivk, $X_n \sim \text{Bin}(n, p_n)$, tako da za zaporedje števil p_n obstaja limita $\lim_{n \rightarrow \infty} n p_n = \lambda$. Potem za slučajno spremenljivko $X \sim \text{Poiss}(\lambda)$ velja

$$P(X = x) = \lim_{n \rightarrow \infty} P(X_n = x).$$

X	$X(S)$	$p_X(x)$	$E(X)$	$\text{var}(X)$
Ber(p)	$\{0, 1\}$	$\begin{cases} p, & x = 1; \\ 1-p, & x = 0. \end{cases}$	p	$p(1-p)$
U($n; a, b$)	$\{a, a+1, \dots, b\}$	$\frac{1}{n}$	$\frac{a+b}{2}$	$\frac{(n^2-1)}{12}$
Geo(p)	$\{1, 2, 3, \dots\}$	pq^{x-1}	$\frac{1}{p}$	$\frac{q}{p^2}$
Bin(n, p)	$\{0, \dots, n\}$	$\binom{n}{x} p^x q^{n-x}$	np	npq
NBin(n, p)	$\{0, 1, \dots\}$	$\binom{n+x-1}{n-1} p^n q^{x-n}$	$\frac{nq}{p}$	$\frac{np}{q^2}$
HGeo(n, N, M)	$\{0, 1, \dots, n\}^*$	$\frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$	$\frac{nM}{N}$	$\frac{nM(N-n)(N-M)}{N^2(N-1)}$
Poiss(λ)	$\{0, 1, 2, \dots\}$	$\frac{\lambda^x e^{-\lambda}}{x!}$	λ	λ

Tabela 1: Karakteristike osnovnih tipov diskretnih slučajnih spremenljivk

Dokaz. Označimo $\lambda_n = np_n$ in izračunajmo

$$\begin{aligned}
 P(X_n = x) &= \binom{n}{x} p_n^x (1-p_n)^{n-x} \\
 &= \frac{n(n-1)\cdots(n-x+1)}{x!} \cdot \frac{\lambda_n^x}{n^x} \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-x} \\
 &= \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-x+1}{n} \cdot \frac{\lambda_n^x}{x!} \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-x}.
 \end{aligned}$$

Po upoštevanju zvez $\lim_{n \rightarrow \infty} \lambda_n = \lambda$, $\lim_{n \rightarrow \infty} (1 - \lambda_n/n) = 1$ in $\lim_{n \rightarrow \infty} (1 - \lambda_n/n)^n = e^{-\lambda}$ sledi

$$\lim_{n \rightarrow \infty} P(X_n = x) = \frac{\lambda^x}{x!} e^{-\lambda} = P(X = x).$$

□

Poissonova porazdelitev se uporablja v številnih statističnih modelih, ki napovedujejo verjetnost razmeroma redkih neodvisnih dogodkov v danem časovnem intervalu. Primerna je, kadar je število dogodkov celo število in je sorazmerno dolžini intervala, dogodki so med seboj neodvisni in se ne moreta zgoditi dva dogodka povsem hkrati.

Poissonovo porazdelitev denimo uporabljajo zavarovalnice za napovedovanje verjetnega števila prometnih nesreč, požarov ali poplav v zavarovalnem obdobju. Resnejša uporaba seveda zahteva dobro razumevanje in oceno natančnosti aproksimacije.

Za naše potrebe zadošča naslednja ohlapna trditev, ki je ne bomo dokazali.

Posledica (Poissonova aproksimacija binomske slučajne spremenljivke). Za velike n in majhne p velja, da je $\text{Bin}(n, p) \approx \text{Poiss}(\lambda)$, kjer je $\lambda = np$.

Zgled. Denimo, da 12-krat vržemo par kock. Kakšna je verjetnost, da vržemo natanko 1 dvojno šestico?

Zanima nas verjetnost $P(X = 1)$, kjer je $X \sim \text{Bin}(n, p)$ za $n = 12$ in $p = 1/36$. Pri teh podatkih ni težko izračunati

$$P(X = 1) = \binom{12}{1} \cdot \frac{1}{36} \cdot \left(\frac{35}{36}\right)^{11} \doteq 0.245.$$

Za primerjavo, če je $X' = \text{Poiss}(\lambda)$, kjer je $\lambda = np = 1/3$, dobimo

$$P(X' = 1) = e^{-1/3} \cdot \frac{(1/3)}{1!} = 0.239.$$

Zgled. Skupaj z Janezom je v letniku 220 študentov. Kakšna je verjetnost, da nihče od njih nima rojstnega dneva na isti datum kot Janez, če zanemarimo možnost prestopnega leta?

Predstavljajmo si, da so ostali študentje oštevilčeni s številkami $1, 2, \dots, 219$. Ugotavljanje, ali ima i -ti študent v letniku rojstni dan na isti dan kot Janez, je Bernoullijev poskus, ki uspe z verjetnostjo $p = 1/365$. Ker so rojstni dnevi študentov neodvisni, imamo torej 219 neodvisnih ponovitev poskusa. Za število študentov, ki imajo rojstni dan na isti dan kot Janez, torej velja $X \sim \text{Bin}(n, p)$, kjer je $n = 219$ število študentov in $p = 1/365$.

Iskana verjetnost je

$$P(X = 0) = (1 - p)^{219} = (364/365)^{219} = 0.548.$$

Ker je $np = 219/365 = 0.6$, bi s pomočjo Poissonove aproksimacije $X' \sim \text{Poiss}(0.6)$ dobili približek

$$P(X' = 0) = e^{-0.6} = 0.549.$$

Zgodovinsko je bila praktična vrednost Poissonove aproksimacije v tem, da se lahko z njo izognemo nerodnemu računanju binomskega simbola $\binom{n}{x}$ za velike vrednosti n in x (denimo $\binom{3600}{1200}$). Bralca vabimo, naj primerjavo binomske in Poissonove porazdelitve ponazori tudi z uporabo vgrajenih funkcij in drsnikov za n in p v GeoGebri.

4 Zvezne slučajne spremenljivke

Zgled. Pri metu kopja je dolžina meta pozitivno realno število, ki ga ne moremo izmeriti povsem natančno. Zato merimo na določeno natančnost - izmerjena dolžina meta 71.01 m denimo pomeni, da je metalec kopje vrgel med 71.005 in 71.015 m daleč. Namesto verjetnosti, da je metalec dosegel dolžino x je zato smiselno računati verjetnost, da je dosežena vrednost x leži na nekem intervalu (a, b) . Z manjšanjem intervalov dobimo namesto diskretnih vrednosti verjetnostne funkcije zvezno funkcijo, ki jo imenujemo gostota verjetnosti. Verjetnost, da vrednost leži na izbranem območju je potem sorazmerna ploščini območja pod krivuljo, torej jo računamo s pomočjo integrala.

Zgled. Trije osnovni zgledi porazdelitev dss so diskretna enakomerna, geometrijska in binomska porazdelitev. Vse tri imajo svoje zvezne analogije, ki jih imenujemo zvezna enakomerna, eksponentna in normalna porazdelitev. (Skice grafov p_X in F_X za dss ter f_X in F_X za zss).

4.1 Definicija in zgledi zveznih slučajnih spremenljivk

Naj bo $X: S \rightarrow \mathbb{R}$ poljubna slučajna spremenljivka. Spomnimo se, da smo za diskretno X definirali porazdelitveno funkcijo kot

$$F_X(x) = P(X \leq x) = P(\{s \in S \mid X(s) \leq x\}).$$

Ta definicija je smiselna tudi, če X ni diskretna. Če obstaja nenegativna realna funkcija $f_X: \mathbb{R} \rightarrow \mathbb{R}$ z lastnostjo

$$F_X(x) = \int_{-\infty}^x f_X(u) du,$$

potem rečemo, da je X **zvezna slučajna spremenljivka** in da je funkcija f_X **gostota slučajne spremenljivke** X .

Imeni diskretna in zvezna porazdelitev oziroma slučajna spremenljivka izvirata iz dejstva, da je porazdelitvena funkcija F_X pri dss stopničasta (torej diskretna), pri zss pa zvezna. Nekatere osnovne lastnosti zveznih slučajnih spremenljivk precej spominjajo na lastnosti diskretnih, druge pa so bistveno drugačne.

Trditev. Če je X zvezna slučajna spremenljivka z gostoto f_X , potem velja

(i) $P(X = a) = 0$ za vsak $a \in \mathbb{R}$.

(ii) $P(a < X < b) = P(a \leq X \leq b) = \int_a^b f_X(u) du$.

(iii) $\int_{-\infty}^{\infty} f_X(x) dx = 1$.

(iv) Gostota f_X natanko določa porazdelitveno funkcijo F_X s predpisom $F_X(x) = \int_{-\infty}^x f_X(u) du$.
Obratno, če je F_X odvedljiva, je gostota f_X natanko določena z odvodom $f_X(x) = \frac{d}{dx} F_X(x)$.

(v) Porazdelitvena funkcija $F_X(x)$ je zvezna in zanjo velja

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \quad \text{ter} \quad \lim_{x \rightarrow \infty} F_X(x) = 1.$$

Dokaz trditve zahteva nekaj dodatnih tehničnih konceptov v zvezi z definicijo integrala in ga tukaj ne bomo navedli.

Opomba. Ker je $P(X = a) = 0$, za zss ni smiselno definirati verjetnostne funkcije p_X kot pri dss.

Za diskretne slučajne spremenljivke smo dokazali, da vsako zaporedje števil p_i z lastnostjo $\sum p_i = 1$ natanko določa neko verjetnostno funkcijo p_X in s tem neko diskretno slučajno spremenljivko X na prostoru izidov $\{1, 2, \dots\}$. Podobno velja, da gostota natanko določa zss. Tudi to trditev navajamo brez dokaza.

Trditev. Vsaka nenegativna funkcija $f: \mathbb{R} \rightarrow \mathbb{R}$ z lastnostjo $\int_{-\infty}^{\infty} f(u) du = 1$ predstavlja gostoto neke natanko določene zvezne slučajne spremenljivke X .

4.1.1 Zvezna enakomerna porazdelitev

Spomnimo se, da smo pri diskretni enakomerni porazdelitvi imeli n možnih izidov, verjetnost vsakega izida pa je bila konstantna, torej $1/n$. Isto idejo posnemamo pri zvezni enakomerni porazdelitvi. **Zvezna enakomerna slučajna spremenljivka** $X \sim U(a, b)$ ima konstantno gostoto na omejenem intervalu $(a, b) \subset \mathbb{R}$, sicer pa je njena gostota enaka 0. Iz pogoja

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_a^b c dx = 1$$

zato sledi $c = \frac{1}{b-a}$, torej je

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

Z integracijo gostote dobimo porazdelitveno funkcijo

$$F_X(x) = \int_{-\infty}^x f_X(u) du = \begin{cases} 0, & x \leq a; \\ \int_a^x \frac{1}{b-a} du, & a < x < b; \\ \int_a^b \frac{1}{b-a} du, & x \geq b. \end{cases}$$

Sledi

$$F_X(x) = \begin{cases} 0, & x \leq a; \\ \frac{x-a}{b-a}, & a < x < b; \\ 1, & x \geq b. \end{cases}$$

Pri diskretnih slučajnih spremenljivkah smo pričakovano vrednost izračunali kot

$$E(X) = \sum_{x \in X(S)} x \cdot p_X(x).$$

Pri zveznih zamenjamo vsoto z integralom, verjetnostno funkcijo p_X pa z gostoto verjetnosti f_X . Tako dobimo pričakovano vrednost zvezno enakomerno porazdeljene slučajne spremenljivke

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{b+a}{2}.$$

4.1.2 Eksponentna porazdelitev

Eksponentna porazdelitev oziroma zss je zvezna analogija geometrijske porazdelitve, ki jo poznamo pri dss. Spomnimo se, da je za $X \sim \text{Geo}(p)$ verjetnostna funkcija enaka $p_X(x) = p^x$, torej jo na grafu upodobimo kot padajoče zaporedje, ki konvergira proti 0. To spominja na graf funkcije e^{-x} za $x > 0$. Zato definiramo **eksponentno slučajno spremenljivko** $X \sim \text{Exp}(\lambda)$, kjer je $\lambda > 0$, kot zvezno spremenljivko z gostoto

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

Izračun pokaže, da je funkcija f_X res gostota, saj je njen integral po \mathbb{R} enak 1.

4.1.3 Zgledi nestandardnih zveznih porazdelitev

Zgled. Naj bo $a > 0$. Oglejmo si funkcijo

$$f(x) = \begin{cases} 0, & x < a; \\ 1/x^2, & x \geq a. \end{cases}$$

Gre za nenegativno realno funkcijo. Poskusimo določiti tak a , da bo f gostota neke slučajne spremenljivke, torej mora biti ustrezni integral enak 1. Od tod sledi $a = 1$. Gostota

$$f_X(x) = \begin{cases} 0, & x < 1; \\ 1/x^2, & x \geq 1 \end{cases}$$

torej natanko določa neko zvezno slučajno spremenljivko X . Zdaj lahko izračunamo tudi njeno porazdelitveno funkcijo

$$F_X(x) = \int_{-\infty}^x f_X(u) \, du = \begin{cases} 0, & x < 1; \\ 1 - x^{-1}, & x \geq 1. \end{cases}$$

Če nas zanima verjetnost dogodka ($X < 2$), lahko izračunamo

$$P(X < 2) = F(2) = \int_{-\infty}^2 f_X(u) \, du = 1 - \frac{1}{2} = 1/2.$$

Če pa nas zanima verjetnost dogodka ($2 \leq X \leq 3$), izračunamo

$$P(2 \leq X \leq 3) = \int_2^3 f_X(u) \, du = F(3) - F(2) = (1 - 1/3) - (1 - 1/2) = 1/6.$$

Verjetnost dogodka ($X = 2$) pa je seveda enaka 0.

Zgled. Naj bo porazdelitvena funkcija zvezne slučajne spremenljivke X podana s predpisom

$$F_X(x) = \begin{cases} 0, & x < 0; \\ x^3, & 0 \leq x \leq 1; \\ 1, & x > 1. \end{cases}$$

Ker je F_X odvedljiva, lahko izračunamo gostoto verjetnosti za X z odvodom

$$f_X(x) = \frac{d}{dx} F_X(x) = \begin{cases} 3x^2, & 0 \leq x \leq 1; \\ 0, & \text{sicer.} \end{cases}$$

4.1.4 Normalna porazdelitev

Najprej se spoznajmo z Gaussovo funkcijo in normalno krivuljo, ki predstavlja njen graf.

Definicija. *Osnovna Gaussova funkcija* je realna funkcija

$$g(x) = e^{-\frac{x^2}{2}}.$$

Graf funkcije $g(x)$ lahko prostoročno skiciramo tako, da najprej narišemo kvadratno krivuljo $h(x) = -x^2/2$, nato pa z upoštevanjem lastnosti eksponentne funkcije približno skiciramo še krivuljo $g(x) = e^{h(x)}$, ki ima vrh v točki $(0, 1)$ in prevoja v točkah $(\pm 1, e^{-1/2})$, njena limita v $\pm\infty$ pa je enaka 0.

Splošno Gaussovo funkcijo dobimo s pomiki in raztegi osnovne. Če je μ neko realno število, potem graf funkcije $g(x - \mu)$ predstavlja pomik osnovne krivulje za μ v vodoravni smeri. Podobno, za $\sigma, c > 0$ graf funkcije $g(x/\sigma)$ predstavlja razteg oziroma skrčitev osnovne krivulje v vodoravni smeri, graf funkcije $c \cdot g(x)$ pa razteg osnovne krivulje v navpični smeri. Kombinacijo teh treh transformacij predstavlja funkcija

$$f(x) = c \cdot g\left(\frac{x - \mu}{\sigma}\right) = c \cdot e^{-\frac{(x - \mu)^2}{2\sigma^2}}.$$

Pri danih parametrih μ, σ lahko tretji parameter $c > 0$ določimo tako, da velja

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} c \cdot g\left(\frac{x - \mu}{\sigma}\right) dx = 1.$$

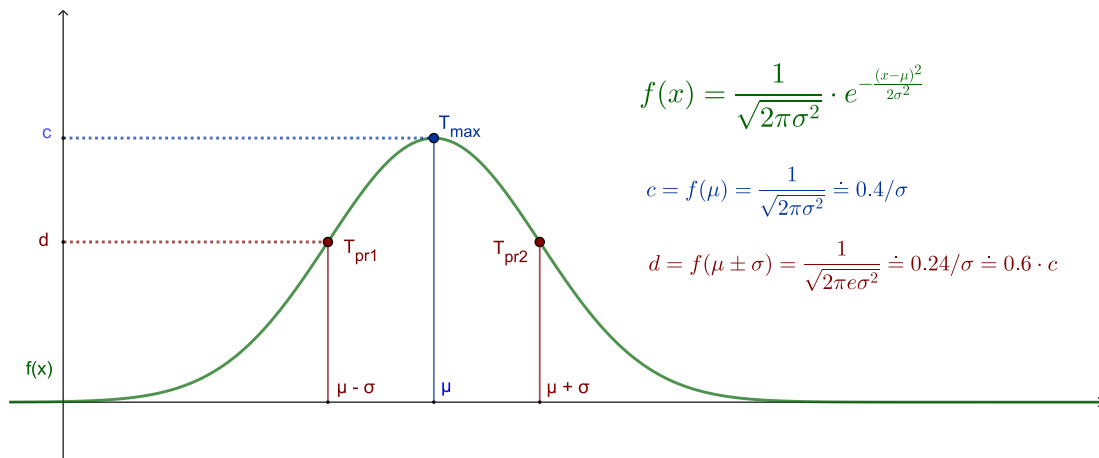
Iz te enakosti z izračunom integrala dobimo

$$c = \frac{1}{\sqrt{2\pi\sigma^2}}.$$

Na podlagi tega definiramo splošno Gaussovo funkcijo.

Definicija. *Gaussova funkcija* s parametroma μ, σ^2 , kjer je $\sigma > 0$, je definirana s predpisom

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x - \mu)^2}{2\sigma^2}}.$$



Trditev. Za Gaussovo funkcijo s parametroma $\mu, \sigma^2 \in \mathbb{R}$, kjer je $\sigma > 0$, velja:

- (i) Funkcija f je zvezna, velja $f(x) > 0$ za vsak $x \in \mathbb{R}$ in $\lim_{x \rightarrow \pm\infty} f(x) = 0$.
- (ii) Velja $f(x - \mu) = f(-x - \mu)$, zato je graf je simetričen glede na premico $x = \mu$.

(iii) Funkcija f ima globalni maksimum v točki $x_m = \mu$ z vrednostjo

$$c = f(x_m) = \frac{1}{\sqrt{2\pi\sigma^2}} \doteq 0.40/\sigma.$$

(iv) Funkcija f ima prevoja v točkah $x_{pr} = \mu \pm \sigma$ z vrednostjo

$$d = f(x_{pr}) = \frac{1}{\sqrt{2\pi e\sigma^2}} \doteq 0.24/\sigma \doteq 0.6c.$$

Na intervalu $(\mu - \sigma, \mu + \sigma)$ je funkcija konkavna, drugje konveksna.

(v) Ploščina območja pod krivuljo je enaka $\int_{-\infty}^{\infty} f(x) dx = 1$.

Dokaz. Razen zadnje lahko ostale trditve zlahka dokažemo z uporabo srednješolskega znanja limit in odvoda, kar prepuščamo bralcu. Pri zadnji trditvi pa je za izračun integrala potrebna zvižaja s prevedbo na dvojni integral in polarne koordinate. Označimo

$$I = \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

in izračunajmo kvadrat tega izraza s pomočjo ustreznih substitucij $x = r \cos \varphi$, $y = r \sin \varphi$ in $dx dy = r dr d\varphi$:

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2/2} dx \right)^2 \\ &= \left(\int_{-\infty}^{\infty} e^{-x^2/2} dx \right) \cdot \left(\int_{-\infty}^{\infty} e^{-y^2/2} dy \right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy \\ &= \int_0^{\infty} \left(\int_0^{2\pi} e^{-r^2/2} r d\varphi \right) dr \\ &= 2\pi \int_0^{\infty} e^{-r^2/2} r dr \\ &= 2\pi. \end{aligned}$$

Torej je $I = \sqrt{2\pi}$ in zdaj s substitucijo $y = (x - \mu)/\sigma$, $dy = dx/\sigma$ hitro dobimo

$$\int_{-\infty}^{\infty} f(x) dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy = 1.$$

□

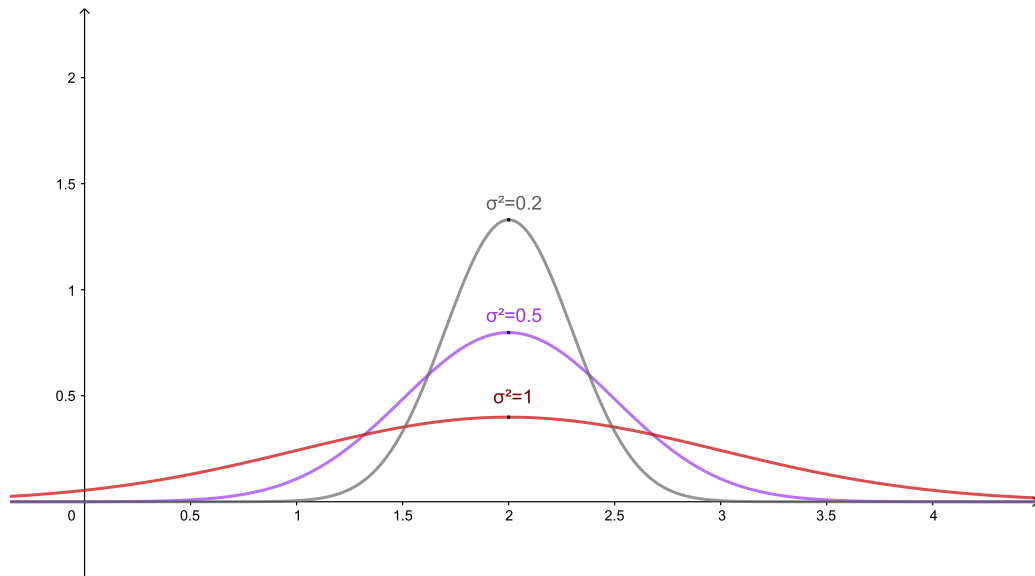
Ker je $f(x) > 0$ in je ploščina območja pod krivuljo enaka 1, je Gaussova funkcija gostota izjemno pomembne zvezne slučajne spremenljivke, ki jo imenujemo normalna (ali Gaussova).

Definicija. *Normalna slučajna spremenljivka* $X \sim N(\mu, \sigma^2)$ s parametroma μ, σ^2 , kjer je $\sigma > 0$, je zvezna slučajna spremenljivka, katere gostota je Gaussova funkcija

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

Normalno slučajno spremenljivko $X \sim N(0, 1)$ s parametroma $\mu = 0$ in $\sigma = 1$ imenujemo *standardna normalna slučajna spremenljivka*.

Zgled. Gostote normalne slučajne spremenljivke $X \sim N(\mu, \sigma^2)$ za $\mu = 2$ in različne vrednosti σ^2 .



Trditev. Če je X normalna, je tudi $aX + b$ normalna za vse $a, b \in \mathbb{R}$, $a \neq 0$. Natančneje, če je $X \sim N(\mu, \sigma^2)$, potem za transformiranko $Y = aX + b$ velja $Y \sim N(a\mu + b, (a\sigma)^2)$.

Dokaz. Naj bo $X \sim N(\mu, \sigma^2)$ in $Y = aX + b$. Če je $a > 0$, potem za porazdelitveno funkcijo $F_Y(x)$ velja

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

Ker za gostoto velja zveza $f_Y(x) = F_Y'(x)$, sledi

$$\begin{aligned} f_Y(x) &= F_X\left(\frac{x-b}{a}\right)' = \frac{1}{a} f_X\left(\frac{x-b}{a}\right) \\ &= \frac{1}{a} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-(\frac{x-b}{a}-\mu)^2/(2\sigma^2)} \\ &= \frac{1}{\sqrt{2\pi(a\sigma)^2}} \cdot e^{-\frac{(x-(a\mu+b))^2}{2(a\sigma)^2}}, \end{aligned}$$

kar je enako gostoti normalne slučajne spremenljivke s parametroma $a\mu + b$ in $(a\sigma)^2$, torej

$$Y \sim N(a\mu + b, (a\sigma)^2).$$

Dokaz za $a < 0$ poteka na podoben način. □

Trditev. Če je $X \sim N(\mu, \sigma^2)$, potem je $E(X) = \mu$ in $\text{var}(X) = \sigma^2$.

Dokaz. Najprej dokažimo, da trditev velja za standardno normalno spremenljivko $X \sim N(0, 1)$. Tedaj je $f_X(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}$ in integral

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \dots = 0$$

izračunamo s substitucijo $-x^2/2 = t$, $dx = -dt$ v nedoločeni integral in nato izračunom ustreznih limit. Za varianco potem sledi pa najprej izračunamo

$$\text{var}(X) = E(X^2) - E(X)^2 = E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \dots = 1,$$

ki ga izračunamo z metodo integracije po delih za substitucijo $u = x$ in $dv = x e^{-x^2/2} dx$. Podrobnosti prepustimo bralcu.

Po prejšnjem izreku potem za transformiranko $Y = \mu X + \sigma$ sledi $Y \sim N(\mu, \sigma^2)$ in torej $E(Y) = \mu$, $\text{var}(Y) = \sigma^2$. \square

Posledica (Standardizacija normalne spremenljivke). Če je $X \sim N(\mu, \sigma^2)$ nestandardna normalna, potem je $Z = (X - \mu)/\sigma$ standardna normalna, $Z \sim N(0, 1)$.

Dokaz. Sledi neposredno iz že dokazanega. \square

Skicirajmo še porazdelitveno funkcijo $F_X(x)$ normalne slučajne spremenljivke $X \sim N(\mu, \sigma^2)$. Velja

$$F_X(x) = \int_{-\infty}^x f_X(u) du,$$

vendar tega integrala ni mogoče zapisati z elementarnimi funkcijami. Velja tudi $\lim_{x \rightarrow -\infty} F_X(x) = 0$ in $\lim_{x \rightarrow \infty} F_X(x) = 1$, ter $F_X(\mu) = 1/2$. Levo od te točke je funkcija konveksna, desno pa konkavna.

V primeru, ko je $Z \sim N(0, 1)$ standardna, za funkcijo $F_Z(x)$ uporabljamo oznako Φ .

Definicija. Funkcija $\Phi(x)$ je porazdelitvena funkcija standardne normalne spremenljivke $Z \sim N(0, 1)$, torej

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

Tako definirana funkcija predstavlja verjetnost $P(Z \leq x)$, oziroma ploščino pod standardno normalno krivuljo na ustreznem območju. (SKICA) Zaradi izjemnega pomena te funkcije za verjetnost in statistiko se v različnih učbenikih običajno najde tabela približkov vrednosti $\Phi(x)$ za $x \geq 0$, za $x < 0$ pa lahko vrednosti izračunamo s pomočjo uporabe zveze

$$\Phi(-x) = P(Z \leq -x) = P(Z > x) = 1 - \Phi(x).$$

Pozor, včasih tabela namesto vrednosti $\Phi(x)$ podaja vrednosti $\Phi(x) - 1/2$, kar predstavlja verjetnost $P(0 \leq Z \leq x)$.

x	0.0	1.0	2.0	3.0
$\Phi(x)$	0.500	0.841	0.977	0.999

Zgled. S pomočjo skic in tabele določimo naslednje verjetnosti:

- $P(Z \leq 2)$, $P(Z \leq -1)$, $P(Z \leq (-2))$, $P(Z \geq 2)$, $P(-1 \leq Z \leq 1)$, kjer je $Z \sim N(0, 1)$.
- $P(|Z| \leq 1)$, $P(|Z| \leq 2)$, $P(|Z| \leq 3)$, kjer je $Z \sim N(0, 1)$.

Naslednja trditev je očitna.

Trditev. Če je $X \sim N(\mu, \sigma^2)$, potem je

$$F_X(x) = P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

Zgled. S pomočjo skic in tabele določimo naslednje verjetnosti:

1. $P(|X - \mu| \leq \sigma)$, $P(|X - \mu| \leq 2\sigma)$, $P(|X - \mu| \leq 3\sigma)$, kjer je $X \sim N(\mu, \sigma^2)$.
2. $P(X \leq 15)$, $P(11 \leq X \leq 15)$, $P(X \geq 17)$. kjer je $X \sim N(13, 2^2)$.

Zgled (Šolske ocene po krivulji). Pogost način ocenjevanja pri množičnih preizkusih znanja je ocenjevanje po krivulji. Če je preizkus ustrezno sestavljen, so rezultati preizkusa normalno porazdeljeni, kar pomeni, da jih lahko aproksimiramo z normalno krivuljo.

Na podlagi rezultatov najprej določimo parametra μ (ar. sredina vzorca) in σ (standardni odklon), nato razdelimo ocene: A za tiste, ki dosežejo več kot $\mu + \sigma$ točk, B za rezultat med μ in $\mu + \sigma$, C za rezultat med $\mu - \sigma$ in μ , D za rezultat med $\mu - 2\sigma$ in $\mu - \sigma$ ter F (fail) za rezultat pod $\mu - 2\sigma$. Tako bo porazdelitev ocen približno naslednja:

A	B	C	D	E
15.87%	34.13%	34.13%	13.59%	2.28%

Za konkreten zgled si predstavljajmo, da so bili rezultati izpita iz matematike na veliki univerzi leta 2015 normalno porazdeljeni s povprečjem $\mu = 134$ točk in standardnim odklonom $\sigma = 25$ točk, leta 2016 pa normalno porazdeljeni s povprečjem $\mu = 139$ točk in standardnim odklonom $\sigma = 14$ točk. Kakšno oceno bi dobil študent z rezultatom 110 točk leta 2015 in kakšno leta 2016? Katerega leta je bil rezultat 150 točk boljši glede na celotno populacijo?

Normalna porazdelitev je naravno povezana z binomsko, kar je prvi opisal De Moivre okoli leta 1733, ko je določal porazdelitev števila grbov pri 3600 metih kovanca.

Izrek (De Moivre-Laplaceov limitni izrek ali normalna aproksimacija binomske slučajne spremenljivke). Če je $S_n \sim \text{Bin}(n, p)$, potem je

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{S_n - \mu}{\sigma} \leq b\right) = \Phi(b) - \Phi(a),$$

kjer je $\mu = np$ in $\sigma^2 = npq$.

Izreka ne bomo dokazali.

Zgled. Denimo, da 40-krat vržemo pošten kovanec. Zanima nas verjetnost, da pade natanko $X = 20$ grbov, kjer je $X \sim \text{Bin}(40, 0.5)$. Po Bernoullijevem obrazcu dobimo

$$P(X = 20) = \binom{40}{20} (0.5)^{40} \doteq 0.1254.$$

Z uporabo normalne aproksimacije pa dobimo

$$\begin{aligned} P(X = 20) &\doteq P(19.5 \leq X \leq 20.5) \\ &\doteq P\left(\frac{19.5 - 20}{\sqrt{10}} \leq \frac{X - 20}{\sqrt{10}} \leq \frac{20.5 - 20}{\sqrt{10}}\right) \\ &\doteq P(-0.16 \leq (X - 20)/\sqrt{10} \leq 0.16) \\ &= \Phi(0.16) - \Phi(-0.16) \\ &= \Phi(0.16) - (1 - \Phi(0.16)) \\ &= 2\Phi(0.16) - 1 \\ &= 0.1272. \end{aligned}$$

Zgled. Galtonova plošča.

Zgled. Gauss in zgodba o Ceresu.

X	$f_X(x)$	$F_X(x)$	$E(X)$	$\text{var}(X)$
$U(a, b), a < b$	$\begin{cases} \frac{1}{b-a}, & a < x < b; \\ 0, & \text{sicer.} \end{cases}$	$\begin{cases} 0, & x \leq a; \\ \frac{x-a}{b-a}, & a < x < b; \\ 1, & x \geq b. \end{cases}$	$\frac{a+b}{2}$	$\frac{(a-b)^2}{12}$
$\text{Exp}(\lambda), \lambda > 0$	$\begin{cases} \lambda e^{-\lambda x}, & x > 0; \\ 0, & x \leq 0. \end{cases}$	$\begin{cases} 0, & x \leq 0; \\ 1 - e^{-\lambda x}, & x \geq 0. \end{cases}$	λ^{-1}	λ^{-2}
$N(\mu, \sigma^2), \sigma^2 > 0$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	Ni zaklj. formule.	μ	σ^2
$\text{Par}(\alpha), \alpha > 0$	$\begin{cases} 0, & x < 1; \\ \alpha x^{-(\alpha+1)}, & x \geq 1. \end{cases}$	$\begin{cases} 0, & x < 1; \\ 1 - x^{-\alpha}, & x \geq 1. \end{cases}$	$\begin{cases} \infty, & \alpha \leq 1; \\ \frac{\alpha}{\alpha-1}, & \alpha > 1. \end{cases}$	$\begin{cases} \infty, & \alpha \leq 2; \\ \frac{\alpha}{(\alpha-1)^2(\alpha-2)}, & \alpha > 2. \end{cases}$
$\text{Cauchy}(x_0, \gamma)$	$\frac{1}{\pi\gamma\left(1 + \left(\frac{x-x_0}{\gamma}\right)^2\right)}$	$\frac{1}{\pi} \arctg\left(\frac{x-x_0}{\gamma}\right) + \frac{1}{2}$	Ne obstaja.	Ne obstaja.

Tabela 2: Nekaj osnovnih tipov zveznih slučajnih spremenljivk in njihove lastnosti

4.1.5 Paretova porazdelitev

Glej tabelo.

4.1.6 Cauchyjeva porazdelitev

Cauchyjeva porazdelitev se pojavlja pri različnih fizikalnih problemih. Njena gostota ima krivuljo, ki je simetrična in zvonaste oblike, zato spominja na normalno porazdelitev. Posebnost Cauchyjeve porazdelitve je, da nima pričakovane vrednosti, saj ustrezeni integral ni definiran.

Standardno Cauchyjevo porazdelitev srečamo pri naslednji nalogi. Izvir seva delce proti neskončnemu ravnemu zaslonu, ki je oddaljen za 1. Posamezni delec je izstreljen pod naključnim kotom $\Phi \in (-\pi/2, \pi/2)$ glede na navpičnico in leti premočrtno do točke X , v kateri zadane. Torej je Φ slučajna spremenljivka, ki je porazdeljena zvezno enakomerno na intervalu $(-\pi/2, \pi/2)$,

Opomba. Poleg tukaj predstavljenih zvezno porazdeljenih slučajnih spremenljivk poznamo še vrsto drugih, na primer χ^2 , študentova, Weibullova, Beta, Gama in podobne porazdelitve, ki so za statistiko izjemno pomembne.

4.2 Srednje vrednosti: Pričakovana vrednost, mediana in modus

V statistiki in verjetnosti poznamo različne srednje vrednosti. Pri tem je potrebno najprej jasno ločiti, da se srednje vrednosti v verjetnosti nanašajo na teoretični model, srednje vrednosti v statistiki pa na izmerjeni vzorec podatkov. Mi bomo govorili predvsem o prvih.

Definicija. Naj bo X zvezna slučajna spremenljivka. *Pričakovana vrednost* oziroma *matematično upanje* $E(X)$ je definirano z izrazom

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx,$$

(če ta integral obstaja). *Mediana* $\text{Med}(X)$ je tako realno število $a \in \mathbb{R}$, za katerega velja $P(X \leq a) = 1/2$, torej razpolovi ploščino ustreznega lika na grafu. *Modus* $\text{Mod}(X)$ je vrednost $a \in \mathbb{R}$, v kateri ima gostota $f(x)$ globalni maksimum.

Opomba. Slučajna spremenljivka ima lahko več različnih median ali modulusov. Če je modus en sam, rečemo, da je spremenljivka *unimodalna*.

Vrednost $E(X)$ si geometrijsko predstavljamo kot težišče lika, ki ga določata os x in krivulja f_X , vrednost $\text{Med}(X)$ pa kot točko, v kateri pravokotnica razdeli lik na 2 ploščinsko enaka dela. Vrednost $\text{Mod}(X)$ je seveda točka, v kateri gostota f_X zavzame globalni maksimum, če taka točka obstaja. Pri nekaterih slučajnih spremenljivkah zato modus izračunamo s pomočjo odvoda in določanja stacionarnih točk.

Zgled. Bralec bo sam razmislil, zakaj sta pri zvezni enakomerni porazdelitvi pričakovana vrednost in mediana enaki $\frac{a+b}{2}$, modus pa je vsaka točka na intervalu (a, b) .

Zgled. Izračunajmo mediano, modus in pričakovano vrednost za Paretovo slučajno spremenljivko $X \sim \text{Par}(\alpha)$. Očitno je modus enak $\text{Mod}(X) = 1$ za vsak α . Za izračun mediane moramo rešiti enačbo $P(X \leq x) = 1/2$, torej

$$\frac{1}{2} = \int_{-\infty}^x f_X(t) dt = \int_1^x \alpha t^{-\alpha-1} dt = -t^{-\alpha} \Big|_1^x = -x^{-\alpha} + 1.$$

Sledi $x^{-\alpha} = 1/2$ in od tod $x = 2^{1/\alpha}$. Torej je

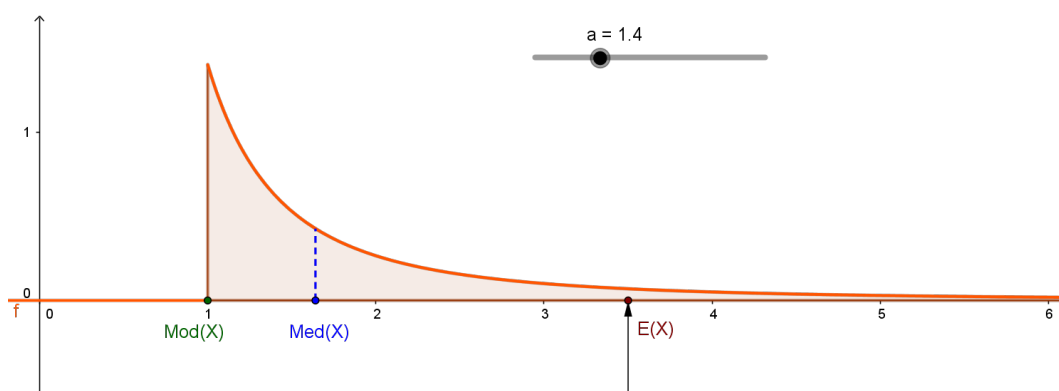
$$\text{Med}(X) = 2^{1/\alpha}.$$

Za pričakovano vrednost pa izračunamo integral

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_1^{\infty} \alpha x^{-\alpha} dx = \frac{\alpha}{\alpha-1},$$

če je $\alpha > 1$, sicer pa integral ne konvergira. Za primere $\alpha = 1, 2, 3, 4$ tako dobimo približne vrednosti

α	$E(X)$	$\text{Mod}(X)$	$\text{Med}(X)$
1	∞	1	2
2	2	1	1.41
3	1.5	1	1.26
4	1.33	1	1.19



Pojma modus in mediana na podoben način definiramo tudi za dss. Mediana diskretne slučajne spremenljivke X je vrednost $a \in \mathbb{R}$, za katero je $P(X \leq a) = 1/2$ in $P(X \geq a) = 1/2$, modus pa (vsak) izid z največjo verjetnostjo.

Pri statističnih podatkih z besedo "povpreče" običajno mislimo aritmetično sredino vzorca podatkov. Modus vzorca je (vsaka) vrednost, ki se največkrat ponovi, mediana pa je vrednost v sredini takoimenovane ranžirne vrste, v kateri podatke razvrstimo od najmanjšega do največjega.

Zgled. V majhnem podjetju ima direktor plačo 5000 EUR, 2 vodstvena delavca vsak 1500 EUR, 10 proizvodnih delavcev pa vsak 500 EUR. Koliko je povprečna plača, kaj je modus in kaj mediana?

Povprečna plača je enaka $\frac{5000+2 \cdot 1500+10 \cdot 500}{13} = 1000$ EUR. Modus je plača, ki jo prejema največ delavcev, torej 500 EUR. Mediana pa je vrednost na sredi ranžirne vrste:

500, 500, 500, 500, 500, 500, **500**, 500, 500, 500, 1500, 1500, 5000,

torej v našem primeru prav tako 500 EUR.

Omenimo še kvartile in percentile. Vrednost x , pri kateri je $P(X \leq x) = 1/4$, se imenuje prvi kvartil, oziroma $3/4$ tretji kvartil. Podobno je deseti percentil vrednost, pri kateri je $P(X \leq x) = 0.1$. Definicije so smiselne tako za zvezne kot za diskretne slučajne spremenljivke, najpogosteje pa se omenjajo v zvezi z različnimi statističnimi podatki.

4.3 Funkcije zveznih slučajnih spremenljivk

Če je X zvezna slučajna spremenljivka in $g: \mathbb{R} \rightarrow \mathbb{R}$ poljubna realna funkcija, za razliko od diskretnega primera kompozitum $g \circ X: S \rightarrow \mathbb{R}$ ni nujno zvezna slučajna spremenljivka. Kljub temu se najpogosteje srečujemo s funkcijami, za katere to velja. Oglejmo si najprej primer, ko je g strogo naraščajoča linearna funkcija.

Trditev. Naj bo X zvezna slučajna spremenljivka in naj bosta $a, b \in \mathbb{R}$ poljubni konstanti, tako da je $a > 0$. Potem je $Y = aX + b$ zvezna slučajna spremenljivka s porazdelitveno funkcijo

$$F_{aX+b}(x) = F_X\left(\frac{x-b}{a}\right)$$

in gostoto

$$f_{aX+b}(x) = \frac{1}{a} f_X\left(\frac{x-b}{a}\right).$$

Dokaz. Velja

$$F_{aX+b}(x) = P(aX + b \leq x) = P\left(X \leq \frac{x-b}{a}\right) = F_X\left(\frac{x-b}{a}\right).$$

Gostoto nato dobimo z odvajanjem:

$$f_{aX+b} = (F_{aX+b}(x))' = \left(F_X\left(\frac{x-b}{a}\right)\right)' = \frac{1}{a} f_X\left(\frac{x-b}{a}\right).$$

□

Bralec lahko sam razmisli, zakaj bi v primeru $a < 0$ v zgornjem izreku dobili nekoliko drugačni enakosti $F_{aX+b}(x) = 1 - F_X\left(\frac{x-b}{a}\right)$ in $f_{aX+b}(x) = -f_X\left(\frac{x-b}{a}\right)$, kljub temu pa je $aX + b$ zss tudi v takem primeru. Oglejmo si še zgled, v katerem je g kvadratna funkcija.

Zgled. Naj bo $X \sim \text{Exp}(\lambda)$ in $g(x) = x^2$. Pokazali bomo, da je $g(X) = X^2$ zvezna slučajna spremenljivka, in določili njeno gostoto. Označimo $X^2 = Y$. Porazdelitvena funkcija za eksponentno porazdeljeno X je enaka

$$F_X(x) = P(X \leq x) = 1 - e^{-\lambda x} \quad \text{za } x \geq 0.$$

Zato velja

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = \\ &= P(X \leq \sqrt{y}) = 1 - e^{-\lambda y^{1/2}} \quad \text{za } y \geq 0. \end{aligned}$$

Gostoto za Y dobimo z odvodom:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{\lambda y^{-1/2}}{2} e^{-\lambda y^{1/2}} \quad \text{za } y \geq 0,$$

in $f_Y(y) = 0$ sicer.

Za prejšnjo trditev in zadnji zgled je bilo ključno, da sta bili funkciji $g(x) = ax + b$ in $g(x) = x^2$ strogo naraščajoči na območju, na katerem je bila gostota f_X neničelna, torej sta imeli na tem območju definiran obrat. Zato lahko na podoben način dokažemo splošnejšo trditev.

Trditev. Naj bo X takšna zss, da velja $P(a < X < b) = 1$. Naj bo $g: (a, b) \rightarrow \mathbb{R}$ zvezna in strogo naraščajoča in naj bo $h = g^{-1}$ inverzna funkcija, ki interval $(g(a), g(b))$ preslika nazaj na (a, b) . Potem je $Y = g(X)$ zvezna slučajna spremenljivka z gostoto

$$f_Y(x) = f_X(h(x))h'(x) \quad \text{za vse } x \in (a, b),$$

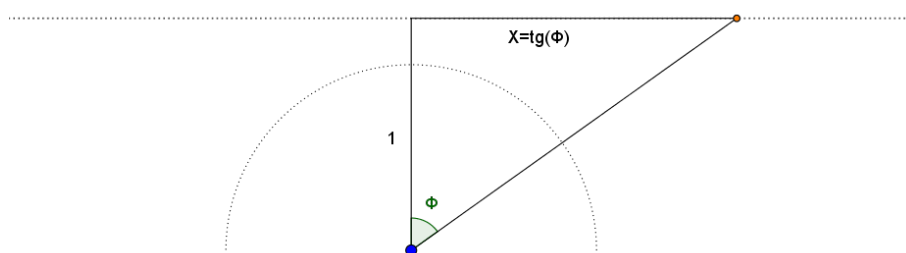
in $f_Y(x) = 0$ sicer.

Dokaz. Ključni korak predstavlja verižno pravilo za odvod:

$$\begin{aligned} f_Y(x) &= (F_Y(x))' = (P(Y \leq x))' = (P(g(X) \leq x))' = \\ &= (P(X \leq g^{-1}(x)))' = (F_X(h(x)))' = F_X'(h(x))h'(x) = \\ &= f_X(h(x))h'(x), \end{aligned}$$

kot smo želeli. □

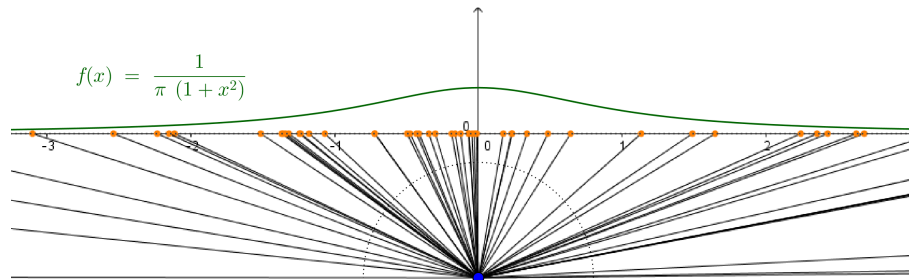
Zgled (Cauchyjeva porazdelitev). Denimo, da stojimo 1 enoto stran od ravnega zidu in z lučjo naključno posvetimo v vodoravni smeri proti zidu pod kotom Φ glede na premico, ki je pravokotna na zid. Denimo, da so vsi koti enako verjetni, torej je $\Phi \sim U(-\pi/2, \pi/2)$ zvezno enakomerno porazdeljena, njena gostota je enaka $\frac{1}{\pi}$ na intervalu $(-\pi/2, \pi/2)$. Naj bo X zvezna slučajna spremenljivka, ki opisuje položaj žarka na zidu. Določimo njeno gostoto.



Iz skice sledi, da je $X = \operatorname{tg}\Phi$. Ker je $\operatorname{tg}\Phi$ zvezna in naraščajoča na intervalu $(-\pi/2, \pi/2)$, je po prejšnji trditvi gostota za X enaka

$$f_X(x) = f_\Phi(\arctg(x))(\arctg(x))' = \frac{1}{\pi(1+x^2)}.$$

To pomeni, da X z največjo verjetnostjo zavzame vrednosti blizu $x = 0$, verjetnost, da je $|X| \geq x$ pa hitro pada. To je razvidno tudi iz naslednje skice, pri kateri smo narisali več kot 50 ponovitev poskusa.



Dobljeno porazdelitev slučajne spremenljivke X imenujemo (standardizirana) **Cauchyjeva porazdelitev**, $X \sim \text{Cauchy}(0, 1)$. Njena posebnost je, da pričakovana vrednost $E(X)$ ne obstaja, mediana in modus za X pa sta enaka 0. Našteto naj bralec preveri sam.

Če je funkcija g dovolj lepa, je mogoče tudi za zss dokazati naslednjo trditev, ki smo jo v podobni obliki srečali že pri dss. Dokaz je v zveznem primeru zahtevnejši in ga bomo tukaj izpustili.

Trditev. Naj bo X zss z gostoto f_X in naj bo $g: \mathbb{R} \rightarrow \mathbb{R}$ taka funkcija, da je $g(X) = g \circ X$ spet zss. Potem velja:

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f_X(x) dx.$$

Dokaz. □

Zdaj lahko za funkcijo $g(x)$ vzamemo $g(x) = x^n$ oziroma $g(x) = (x - \mu)^n$ in definiramo **n -ti moment** zvezne slučajne spremenljivke kot

$$E(X^n) = \int_{-\infty}^{\infty} x^n f_X(x) dx,$$

če ustrezní integral obstaja, in **n -ti centralni moment** kot $E((X - E(X))^n)$. **Varianca zss** je torej definirana kot $\operatorname{var}(X) = E((X - E(X))^2)$.

Trditev. Naj bo X zvezna slučajna spremenljivka, za katero obstajata $E(X)$ in $\operatorname{var}(X)$.

- (i) $\operatorname{var}(X) = E(X^2) - E(X)^2 \geq 0$.
- (ii) $E(aX + b) = aE(X) + b$ za poljubna $a, b \in \mathbb{R}$.
- (iii) $\operatorname{var}(aX + b) = a^2 \operatorname{var}(X)$ za poljubna $a, b \in \mathbb{R}$.

Dokaz. Dokazi so podobni kot pri dss. Da je $\operatorname{var}(X) \geq 0$, pa sledi iz Jensenove neenakosti, ki pravi, da za konveksno funkcijo g velja $g(E(X)) \leq E(g(X))$, torej za $g(x) = x^2$ dobimo $E(X)^2 \leq E(X^2)$. □

Zgled. Računanje variance za zvezno enakomerno, eksponentno, normalno in Paretovo slučajno spremenljivko je odlična vaja iz integralov, ki jo prepuščamo bralcu.

4.4 Neenakost Čebiševa

Za grobo oceno, kakšna je verjetnost, da vrednost poljubne (zvezne ali diskretne) slučajne spremenljivke zelo odstopa od pričakovane vrednosti, je izjemno uporabna neenakost Čebiševa, ki jo pogosto potrebujemo tudi pri dokazih limitnih izrekov.

Izrek. Naj bo X poljubna slučajna spremenljivka, za katero sta definirani $E(X)$ in $\text{var}(X)$. Potem za poljubno realno število $a > 0$ velja:

$$P(|X - E(X)| \geq a) \leq \frac{1}{a^2} \text{var}(X).$$

Dokaz. Zapišimo dokaz za zss, za dss je dokaz podoben. Označimo $\mu = E(X)$. Potem velja

$$\begin{aligned} \text{var}(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x) dx \geq \int_{|x - \mu| \geq a} (x - \mu)^2 f_X(x) dx \geq \\ &\geq \int_{|x - \mu| \geq a} a^2 f_X(x) dx = a^2 \int_{|x - \mu| \geq a} f_X(x) dx = a^2 P(|X - \mu| \geq a). \end{aligned}$$

Dobljeno neenakost delimo z a^2 in iskani rezultat sledi. □

Posledica. Za poljubno slučajno spremenljivko X s pričakovano vrednostjo $\mu = E(X)$ in varianco $\text{var}(X) = \sigma^2$ za vsako število $k \in \mathbb{N}$ veljata neenakosti

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \quad \text{in} \quad P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

Dokaz. Če vstavimo $a = k\sigma$ v neenakost Čebiševa, dobimo

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2 \sigma^2} \sigma^2 = \frac{1}{k^2}.$$

□

Zgled. Za vsako slučajno spremenljivko X , ki ima definirano pričakovano vrednost in varianco, torej veljajo ocene:

$$\begin{aligned} P(|X - \mu| < \sigma) &\geq 0 \doteq 0\% \\ P(|X - \mu| < 2\sigma) &\geq 3/4 = 75\% \\ P(|X - \mu| < 3\sigma) &\geq 8/9 \doteq 88\% \\ P(|X - \mu| < 4\sigma) &\geq 15/16 \doteq 94\% \end{aligned}$$

To je sicer precej manj natančno kot pravilo 68 – 95 – 99.7 za normalno porazdeljeno X , toda ocene po Čebiševu veljajo za poljubno slučajno spremenljivko z definiranim μ in σ .

Zgled. Pošten kovanec vržemo 1600-krat. Zanima nas verjetnost, da je padlo vsaj 1200 grbov. Število grbov X ima binomsko porazdelitev $\text{Bin}(1600, 1/2)$, torej je $\mu = E(X) = np = 800$ in $\sigma^2 = npq = 1600/4 = 400$ oziroma $\sigma = 20$. Točno verjetnost bi torej dobili z vsoto

$$P(X \geq 1200) = \sum_{x=1200}^{1600} \binom{1600}{x} \left(\frac{1}{2}\right)^{1600},$$

kar je praktično nemogoče izračunati na roke, pa tudi računalnik bi znal imeti probleme. Po neenačbi Čebiševa pa lahko ocenimo

$$P(|X - 800| \geq 20k) \leq \frac{1}{k^2}$$

oziroma za $k = 20$

$$P(|X - 800| \geq 400) \leq \frac{1}{400}.$$

Zaradi simetričnosti velja

$$P(|X - 800| \geq 400) = P(X \geq 1200) + P(X \leq 400) = 2P(X \geq 1200).$$

Od tod dobimo oceno

$$P(X \geq 1200) = \frac{1}{2}P(|X - 800| \geq 400) \leq \frac{1}{800} \doteq 0.125\%.$$

Za natančnejši rezultat pa bi morali slučajno spremenljivko X aproksimirati z normalno porazdeljeno $X' \sim N(800, 400)$ in izračunati ustrezno verjetnost s pomočjo tabel ali računalnika.

5 Slučajni vektorji

Pri uporabi verjetnostnega računa hitro naletimo na situacije, v katerih je smiselno obravnavati več različnih slučajnih spremenljivk. Tedaj nas zanima, ali so te spremenljivke enako porazdeljene, ali so njihove vrednosti med seboj neodvisne, kaj lahko povemo o njihovi vsoti in podobno. Zaradi enostavnosti bomo ločeno obravnavali primera, ko so slučajne spremenljivke X_1, X_2, \dots, X_n bodisi vse diskretne bodisi vse zvezne, n -terico (X_1, \dots, X_n) pa bomo glede na to imenovali **slučajni vektor** diskretnih oziroma zveznih slučajnih spremenljivk. Zaradi krajšega zapisa bomo večino trditev formulirali za vektor z dvema slučajnima spremenljivkama (X, Y) in bralcu prepustili, da sam razmisli, kako bi jih posplošil na 3 ali več slučajnih spremenljivk.

5.1 Diskretni slučajni vektorji

5.1.1 Verjetnostne in porazdelitvene funkcije za dsv

Naj bo (S, P) verjetnostni prostor in naj bosta $X, Y: S \rightarrow \mathbb{R}$ diskretni slučajni spremenljivki. Par (X, Y) imenujemo **diskretni slučajni vektor** ali **slučajni vektor diskretnih slučajnih spremenljivk**. Njegova **verjetnostna funkcija** $p_{X,Y}: \mathbb{R}^2 \rightarrow [0, 1]$ je definirana s predpisom

$$p_{X,Y}(x, y) = P(X = x, Y = y).$$

Pri tem zapis $(X = x, Y = y)$ pomeni dogodek $\{s \in S \mid X(s) = x \wedge Y(s) = y\} \subseteq S$, torej presek dogodkov $(X = x)$ in $(Y = y)$. Vrednosti verjetnostne funkcije $p_{x,y}$ običajno podamo v tabeli, katere vrstice označimo z možnimi vrednostmi $x \in X(S)$, stolpce pa z možnimi vrednostmi $y \in Y(S)$.

Zgled. Iz vreče z 1 rdečo, 2 belima in 3 zelenimi kroglicami naključno in brez vračanja izvlečemo tri kroglice. Naj bo X število rdečih in Y število zelenih kroglic med izvlečenimi. Potem verjetnostno funkcijo $p_{X,Y}$ predstavimo z naslednjo tabelo vrednosti:

$p_{X,Y}$	$y = 0$	1	2	3
$x = 0$	0	3/20	6/20	1/20
1	1/20	6/20	3/20	0

Bralec bo hitro opazil, da je vsota vseh vrednosti v tabeli enaka 1. Prav tako je iz dane tabele za $p_{X,Y}$ zelo enostavno izračunati verjetnostni funkciji p_X in p_Y , ki ju v tem primeru imenujemo **robni verjetnostni funkciji** za dsv. Dovolj je, da seštejemo ustrezne vrednosti v vsaki vrstici oziroma stolpcu.

Trditev. Naj bo (X, Y) dsv z verjetnostno funkcijo $p_{X,Y}$. Potem velja

$$(a) \sum_{x,y} p_{X,Y}(x, y) = 1, \text{ kjer seštevamo po vseh } x \in X(S) \text{ in vseh } y \in Y(S).$$

$$(b) p_X(x) = \sum_{y \in Y(S)} p_{X,Y}(x, y) \text{ in } p_Y(y) = \sum_{x \in X(S)} p_{X,Y}(x, y).$$

Dokaz. Prva trditev sledi iz dejstva, da je $S = \bigcup_{x,y} (X = x \wedge Y = y)$ unija ločenih dogodkov, zato enakost sledi po aksiomu o števeni aditivnosti in aksiomu $P(S) = 1$. Za drugo trditev pa zapišemo $(X = x) = \bigcup_{y \in Y(S)} (X = x \wedge Y = y)$ in spet uporabimo števno aditivnost. \square

Zgled. Za X in Y kot v prejšnjem zgledu dobimo robni verjetnostni funkciji p_X in p_Y s seštevanjem po vrsticah oziroma stolpcih:

$p_{X,Y}$	$y = 0$	1	2	3	p_X
$x = 0$	0	3/20	6/20	1/20	1/2
1	1/20	6/20	3/20	0	1/2
p_Y	1/20	9/20	9/20	1/20	

Drugače zapisano, velja

$$X \sim \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix} \text{ in } Y \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1/20 & 9/20 & 9/20 & 1/20 \end{pmatrix}.$$

Poleg verjetnostne funkcije za dsv lahko definiramo tudi (skupno) porazdelitveno funkcijo (angl. *joint cumulative distribution function*) $F_{X,Y}: \mathbb{R}^2 \rightarrow [0, 1]$ s predpisom

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y).$$

Verjetnostna in porazdelitvena funkcija ter ustrezne robne porazdelitve so seveda neposredno povezane.

Trditev. Naj bosta X, Y dss. Potem velja:

$$(a) F_{X,Y}(x, y) = \sum_{\substack{a \in X(S), a \leq x \\ b \in Y(S), b \leq y}} p_{X,Y}(a, b).$$

$$(b) F_X(x) = \lim_{y \rightarrow \infty} F_{X,Y}(x, y) \text{ za vsak } x \in X \text{ in podobno za } F_Y.$$

Dokaz. Dokaz prve trditve je očiten, saj seštejemo verjetnosti ločenih dogodkov $(X = a \wedge Y = b)$ po ustreznih vrednostih $a \leq x$ in $b \leq y$. Za drugo trditev pa je potrebno zapisati dogodek $(X \leq x)$ kot unijo naraščajočih dogodkov in uporabiti limitni proces. To zahteva nekaj tehnične spretnosti – detajle naj bralec poišče v katerem od standardnih virov. \square

5.1.2 Neodvisnost in pogojna porazdelitev dss

Spomnimo se, da sta dogodka A in B neodvisna, če velja $P(A \cap B) = P(A)P(B)$. Rekli bomo, da sta dss X in Y neodvisni, če sta dogodka $(X = x)$ in $(Y = y)$ neodvisna za vse možne vrednosti x, y , torej, če velja

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

za vse $x \in X(S), y \in Y(S)$.

Zgled. Slučajni spremenljivki X in Y iz zгледа s kroglicami nista neodvisni, saj velja denimo $P(X = 0, Y = 0) = 0$, kar ni enako produktu verjetnosti $P(X = 0) = 1/2$ in $P(Y = 0) = 1/20$. Iz tega zгледа lahko razberemo preprost zadostni pogoj za odvisnost. Če v tabeli za $p_{X,Y}$ nastopi vrednost 0, čeprav ustrezni robni verjetnosti $p_X(x)$ in $p_Y(y)$ nista enaki 0, potem sta X in Y gotovo odvisni.

Zgled. Naj bo za neki dss X, Y ustrezna verjetnostna funkcija $p_{X,Y}$ podana s tabelo:

$p_{X,Y}$	$y = 0$	1	p_X
$x = 0$	0.4	0.3	0.7
1	0.2	0.1	0.3
p_Y	0.6	0.4	

Potem X in Y nista neodvisni, saj je na primer vrednost $P(X = 1, Y = 0)$ različna od produkta $P(X = 1)P(Y = 0)$. Izkaže se, da lahko tabelo za $p_{X,Y}$, v kateri so vnaprej predpisane robne verjetnosti, na en sam način dopolnimo tako, da bosta X in Y neodvisni. Bralec naj to preizkusi na spodnji tabeli.

$p_{X,Y}$	$y = 0$	1	p_X
$x = 0$			0.7
1			0.3
p_Y	0.6	0.4	

Trditve. Naj bosta X, Y dss. Potem so naslednje trditve enakovredne:

- (i) X in Y sta neodvisni.
- (ii) $p_{X,Y}(x, y) = p_X(x)p_Y(y)$ za vse $x \in X(S)$, $y \in Y(S)$.
- (iii) $F_{X,Y}(x, y) = F_X(x)F_Y(y)$ za vse $x, y \in \mathbb{R}$.

Dokaz. Enakovrednost prvih dveh trditve je očitna, prav tako ni težko preveriti, da iz druge trditve sledi tretja. Nekoliko zahtevnejši je dokaz v obratno smer, ki ga prepuščamo bralcu. \square

Če sta X in Y odvisni dss, je lahko verjetnost dogodka ($X = x$) odvisna od tega, ali se je zgodil dogodek ($Y = y$). Zato lahko definiramo **pogojno verjetnostno funkcijo** za X pri pogoju $Y = y$ na podoben način, kot smo definirali pogojno verjetnost odvisnih dogodkov:

$$p_X(x | Y = y) = P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

Podobno definiramo tudi **pogojno pričakovano vrednost** oziroma **pogojno matematično upanje** za X pri pogoju $Y = y$ kot vsoto

$$E(X | Y = y) = \sum_{x \in X(S)} x \cdot P(X = x | Y = y).$$

Zgled (Vir: R. Durrett). Na neki ameriški fakulteti analizirajo študijski uspeh študentov in njihove srednješolske dosežke. Označimo z X oceno na izpitu iz Analize, z zalogo vrednosti $X(S) = \{1, 2, 3, 4, 5\}$, in z Y oceno iz matematike v srednji šoli, z zalogo vrednosti $Y(S) = \{2, 3, 4\}$. Na podlagi podatkov so dobili naslednjo tabelo:

$p_{X,Y}$	$y = 4$	3	2	p_X
$x = 5$	0.10	0.05	0	0.15
4	0.15	0.15	0	0.30
3	0.10	0.15	0.10	0.35
2	0	0.05	0.10	0.15
1	0	0	0.05	0.05
p_Y	0.35	0.40	0.25	

Naravno vprašanje je, kako uspešni so pri Analizi študentje, ki so v SŠ dosegli najvišjo možno oceno. Iz podatkov zlahka izračunamo pogojno porazdelitev $p_X(x | Y = 4)$:

$$\begin{aligned} p_X(5 | Y = 4) &= P(X = 5 | Y = 4) = \frac{0.10}{0.35} = 2/7, \\ p_X(4 | Y = 4) &= P(X = 4 | Y = 4) = \frac{0.15}{0.35} = 3/7, \\ p_X(3 | Y = 4) &= P(X = 3 | Y = 4) = \frac{0.10}{0.35} = 2/7, \\ p_X(2 | Y = 4) &= P(X = 2 | Y = 4) = 0, \\ p_X(1 | Y = 4) &= P(X = 1 | Y = 4) = 0. \end{aligned}$$

Od tod dobimo tudi pogojno pričakovano vrednost

$$E(X | Y = 4) = \frac{5 \cdot 2 + 4 \cdot 3 + 3 \cdot 2}{7} = 4,$$

ki predstavlja povprečno oceno študenta, ki je kot dijak dosegel najvišjo oceno. Podobno bi izračunali še $E(X | Y = 3) = 3.5$ in $E(X | Y = 2) = 2.2$.

Zadnji zgled namiguje, da so študentje, ki imajo višje ocene v SŠ, v povprečju uspešnejši tudi pri študiju. Naravno vprašanje v statistiki je seveda ugotoviti, s kakšno gotovostjo lahko iz vrednosti ene slučajne spremenljivke napovedujemo vrednosti druge. Nekaj o tem bomo povedali v naslednjih razdelkih.

5.1.3 Vsote, produkti in kovarianca dss

Naj bosta X in Y poljubni dss, in naj bo $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ poljubna funkcija. Potem predpis

$$g(X, Y)(s) = g(X(s), Y(s))$$

določa neko funkcijo $g(X, Y): S \rightarrow \mathbb{R}$, katere zaloga vrednosti ima kardinalnost največ $|X(S) \times Y(S)|$, torej je števna. Zato je $g(X, Y)$ spet dss. V posebnem primeru za $g(x, y) = x + y$ ali $g(x, y) = x y$ torej velja, da sta vsota $X + Y$ in produkt $X Y$ spet dss.

Trditev (Linearnost pričakovane vrednosti). Naj bosta X in Y poljubni dss, za kateri obstajata pričakovani vrednosti $E(X)$ in $E(Y)$. Potem velja

$$E(X + Y) = E(X) + E(Y).$$

Splošneje, za poljubne dss X_1, \dots, X_n s pričakovanimi vrednostmi $E(X_1), \dots, E(X_n)$ velja

$$E(\alpha_1 X_1 + \dots + \alpha_n X_n) = \alpha_1 E(X_1) + \dots + \alpha_n E(X_n).$$

Dokaz. Za prvo enakost zapišemo

$$\begin{aligned} E(X + Y) &= \sum_{x \in X(S), y \in Y(S)} (x + y) P(X = x, Y = y) = \\ &= \sum_{x \in X(S)} \sum_{y \in Y(S)} x P(X = x, Y = y) + \sum_{x \in X(S)} \sum_{y \in Y(S)} y P(X = x, Y = y) = \\ &= \sum_{x \in X(S)} x \sum_{y \in Y(S)} P(X = x, Y = y) + \sum_{y \in Y(S)} y \sum_{x \in X(S)} P(X = x, Y = y) = \\ &= \sum_{x \in X(S)} x P(X = x) + \sum_{y \in Y(S)} y P(Y = y) = \\ &= E(X) + E(Y), \end{aligned}$$

kjer smo upoštevali zvezo med $p_{X,Y}$ in p_X oziroma p_Y . Enakost za n členov dokažemo z indukcijo. Pri tem uporabimo od prej znano zvezo $E(\alpha X) = \alpha E(X)$. \square

Dokazana trditev je izjemno uporabna tako pri konkretnih izračunah kot tudi pri teoretičnih izpeljavah.

Zgled. Denimo, da n kroglic naključno porazdelimo v m škatel. Koliko praznih škatel lahko pričakujemo?

Oštevilčimo škatle in označimo z X_i dss, ki ima vrednost 1, če je i -ta škatla prazna, in 0 sicer. Potem je $X = X_1 + \dots + X_m$ število vseh praznih škatel, zato je

$$E(X) = E(X_1) + \dots + E(X_m) = mE(X_1),$$

saj so X_i enako porazdeljene (ne pa tudi neodvisne!). Ker velja $P(X_1 = 1) = \left(\frac{m-1}{m}\right)^n$, sledi $E(X_1) = \left(\frac{m-1}{m}\right)^n$, torej $E(X) = m\left(\frac{m-1}{m}\right)^n$.

Za konkreten primer si predstavljajmo, da poskušamo napolniti album, v katerem je prostora za $m = 20$ različnih smrkcev, zato smo kupili $n = 30$ naključnih figuric. V tem primeru lahko v albumu pričakujemo približno $20(0.95)^{30} \doteq 4.3$ praznih mest.

Zgled. Spomnimo se, da smo s precej spretnim premetavanjem vsot in binomskih simbolov dokazali, da za $X \sim \text{Bin}(n, p)$ velja $E(X) = np$. Ker je X število uspehov v n ponovitvah Bernoullijevega poskusa, bi lahko z X_i označili dss z vrednostjo 1, če je i -ti poskus uspel. Potem je $X = X_1 + \dots + X_n$, kjer so X_i enako porazdeljene in velja $E(X_1) = 1 \cdot p + 0 \cdot q = p$. Zato je $E(X) = nE(X_1) = np$.

Podobna enakost za varianco vsote ne velja nujno. Zaradi linearnosti E velja

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

Zadnji člen na desni označimo s $\text{cov}(X, Y)$ in imenujemo **kovarianca** diskretnih slučajnih spremenljivk X in Y . Bralec bo sam preveri, da lahko definicijo kovariance preoblikujemo tudi v

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

ki jo običajno uporabljamo pri konkretnih računih. Pri statističnih raziskavah pa se namesto kovariance pogosto uporablja tako imenovani (Pearsonov) **korelacijski koeficient**

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)},$$

ki kovarianco normira glede na velikost standardnih odklonov varianc, zato ima vrednost na intervalu $[-1, 1]$.

Zgled. Naj bosta X in Y dss iz zglada z vlečenjem kroglic. Vemo že, da X in Y nista neodvisni. Za izračun njune kovariance najprej izračunajmo pričakovane vrednosti $E(X)$, $E(Y)$, $E(XY)$. Velja $E(X) = \frac{0 \cdot 1 + 1 \cdot 1}{2} = 1/2$ in $E(Y) = \frac{0 \cdot 1 + 1 \cdot 9 + 2 \cdot 9 + 3 \cdot 1}{20} = \frac{30}{20} = 3/2$. Za XY pa najprej določimo zalogo vrednosti in opišimo porazdelitev s tabelo

$$XY \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ 11/20 & 6/20 & 3/20 & 0/20 \end{pmatrix}.$$

Sledi $E(XY) = \frac{0 \cdot 11 + 1 \cdot 6 + 2 \cdot 3 + 3 \cdot 0}{20} = 12/20 = 3/5$ in od tod

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y) = 3/5 - 3/4 = -3/20 = -0.15.$$

Da bi izračunali korelacijski koeficient, je potrebno izračunati še $E(X^2)$ in $E(Y^2)$. Velja $X^2 \sim X$ in zato $E(X^2) = E(X) = \frac{1}{2}$ ter

$$Y^2 \sim \begin{pmatrix} 0 & 1 & 4 & 9 \\ 1/20 & 9/20 & 9/20 & 1/20 \end{pmatrix},$$

zato je $E(Y^2) = 27/10$. Sledi $\sigma(X) = \sqrt{E(X^2) - E(X)^2} = 1/2$ in $\sigma(Y) = \sqrt{E(Y^2) - E(Y)^2} = 3\sqrt{5}/10$. Sledi

$$\rho = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)} = -\frac{\sqrt{5}}{5} \doteq -0.44.$$

Dobljena negativna korelacija pomeni, da se v primeru, ko ima ena spremenljivka vrednost nad povprečjem, precej poveča verjetnost, da bo vrednost druge pod povprečjem.

Rečemo, da sta spremenljivki **pozitivno/negativno korelirani**, če je ρ pozitiven oziroma negativen. Pozitivna korelacija pomeni veliko verjetnost, da sta obe spremenljivki bodisi obe hkrati nad svojim povprečjem bodisi obe hkrati pod svojim povprečjem. Negativna korelacija pa pomeni, da je ena nad svojim povprečjem takrat, ko je druga pod svojim povprečjem in obratno. Če je $\rho = 0$, pa sta spremenljivki **nekorelirani**. Geometrijsko si lahko kovarianco predstavljamo kot skalarni produkt, korelacijo pa kot kosinus kota med "vektorjema" $X = \vec{u}$ in $Y = \vec{v}$:

$$\cos(\vec{u}, \vec{v}) = \frac{\vec{u} \cdot \vec{v}}{|\vec{u}| |\vec{v}|}.$$

Dokazati je mogoče, da je $|\rho(X, Y)| = 1$ natanko tedaj, ko je $Y = aX + b$, torej, ko je Y linearna funkcija X .

Opomba. Koreliranost predstavlja neko mero odvisnosti med slučajnima spremenljivkama, vendar ne pomeni vzročnosti (kavzalnosti). Recimo, da slučajni spremenljivki X in Y opisujeta mesečno število prodanih sladoledov oziroma smrti zaradi utopitve. Statistični podatki bodo pokazali, da sta X in Y pozitivno korelirani z ρ blizu 1, saj obe vrednosti narasteta v poletnih mesecih in upadeta v zimskih. Seveda pa bi bilo povsem zgrešeno iz tega sklepati, da uživanje sladoleda poveča verjetnost utopitve. Slabo razumevanje tovrstnih pojmov povzroča številne napake pri interpretaciji statističnih podatkov.

Zdaj lahko končno dokažemo še naslednjo trditev.

Trditev. Če sta dss X in Y neodvisni in ustrezne variance in pričakovane vrednosti obstajajo, potem velja:

(i) $E(XY) = E(X)E(Y)$.

(ii) $\text{cov}(X, Y) = 0$.

(iii) $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$.

Dokaz. Vemo že, da je $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2 \text{cov}(X, Y)$. Iz enakosti $E(XY) = E(X)E(Y)$ neposredno sledi $\text{cov}(X, Y) = 0$ in zato $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$. Dokazati moramo torej le

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

saj je $P(X = x, Y = y) = P(X = x)P(Y = y)$ zaradi neodvisnosti. □

Posledica. Naj bodo X_1, \dots, X_n paroma neodvisne dss. Potem je

$$\text{var}(\alpha_1 X_1 + \dots + \alpha_n X_n) = \alpha_1^2 \text{var}(X_1) + \dots + \alpha_n^2 \text{var}(X_n),$$

če ustrezne variance obstajajo.

Dokaz. Uporabimo indukcijo in že znano zvezo $\text{var}(\alpha X) = \alpha^2 \text{var}(X)$. □

Zgled. Če binomsko slučajno spremenljivko $X \sim \text{Bin}(n, p)$ zapišemo kot vsoto neodvisnih Bernoullijevih $X_i \sim \text{Ber}(p)$, lahko izračunamo varianco z uporabo zadnje trditve veliko hitreje kot neposredno po definiciji:

$$\text{var}(X) = \text{var}(X_1) + \dots + \text{var}(X_n) = n \text{var}(X_1) = npq.$$

Tudi varianco za hipergeometrijsko dss bi s pomočjo zadnje trditve izračunali veliko lažje kot neposredno po definiciji.

5.1.4 Vsota neodvisnih dss

Razdelek zaključimo še z razmislekom o vsoti neodvisnih spremenljivk. Če vemo, da sta dss X, Y tudi neodvisni, potem lahko razmeroma preprosto opišemo verjetnostno funkcijo njune vsote na naslednji način.

Trditev. Naj bosta X, Y neodvisni dss. Potem za vsoto $X + Y$ velja

$$p_{X+Y}(z) = P(X + Y = z) = \sum_{x \in X(S)} P(X = x)P(Y = z - x).$$

Dokaz. Preveri sam. □

Zgled. Dokažimo, da če sta $X \sim \text{Bin}(n, p)$ in $Y \sim \text{Bin}(m, p)$ neodvisni pri istem parametru p , potem za vsoto $X + Y$ velja, da je $X + Y \sim \text{Bin}(n + m, p)$.

Po prejšnji trditvi velja

$$\begin{aligned} p_{X+Y}(z) &= P(X + Y = z) = \\ &= \sum_{x=0}^n P(X = x)P(Y = z - x) = \\ &= \sum_{x=0}^n \binom{n}{x} p^x q^{n-x} \binom{m}{z-x} p^{z-x} q^{m-(z-x)} = \\ &= p^z q^{n+m-z} \sum_{x=0}^n \binom{n}{x} \binom{m}{z-x} = \\ &= p^z q^{n+m-z} \binom{n+m}{z}. \end{aligned}$$

Pri tem enakost $\binom{n+m}{z} = \sum_{x=0}^n \binom{n}{x} \binom{m}{z-x}$ najlažje utemeljimo kombinatorično: število izbir z elementov izmed $n + m$ je enako številu izbir x elementov izmed n in izbir $z - x$ elementov izmed m , za vse smiselne x .

Še krajši dokaz iste trditve pa dobimo z naslednjim razmislekom: če X predstavlja število uspehov v prvih n ponovitvah Bernoullijevega poskusa in Y število uspehov v naslednjih m ponovitvah istega poskusa, potem je $X + Y$ ravno število uspehov v $n + m$ ponovitvah, torej je $X + Y \sim \text{Bin}(n + m, p)$.

Zgled. Naj bosta $X \sim \text{Poiss}(\lambda)$ in $Y \sim \text{Poiss}(\mu)$ neodvisni Poissonovi. Z nekaj računanja lahko s pomočjo prejšnje trditve dokažemo, da velja $X + Y \sim \text{Poiss}(\lambda + \mu)$.

Zgled. Bralec naj sam preveri, da je za neodvisni dss $X, Y \sim \text{Geo}(\lambda)$ verjetnostna funkcija enaka $P(X + Y = z) = (z - 1)p^2(1 - p)^{z-2}$. V tem primeru vsota $X + Y$ nima kakšne od nam znanih standardnih porazdelitev.

5.2 Zvezni slučajni vektorji

5.2.1 Osnovni pojmi

Večino pojmov, ki smo jih srečali pri vektorjih diskretnih slučajnih spremenljivk, lahko definiramo tudi za zvezne spremenljivke, vendar se pri tem srečamo z nekaj dodatnimi omejitvami, povezanih z obstojem ustreznih funkcij ali njihovih integralov.

Naj bosta X, Y zvezni slučajni spremenljivki. Potem je par (X, Y) **zvezni slučajni vektor**, če obstaja takšna funkcija $f_{X,Y}: \mathbb{R}^2 \rightarrow \mathbb{R}$, da velja:

$$(i) \quad f_{X,Y}(x, y) \geq 0 \text{ za vse } x, y \in \mathbb{R}.$$

$$(ii) \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1.$$

$$(iii) \quad P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f_{X,Y}(x, y) dx dy \text{ za vsa realna števila } a \leq b, c \leq d.$$

Vsako funkcijo $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, ki ustreza prvima dvema pogojema, imenujemo **gostota slučajnega vektorja**, saj taka funkcija natanko določa neki zsv (X, Y) . Obratno sicer ne velja: če sta X, Y zss, par (X, Y) še ni nujno zsv, ker ustrezna gostota morda ne obstaja.

Porazdelitvena funkcija za zsv (X, Y) je preslikava $F_{X,Y}: \mathbb{R}^2 \rightarrow [0, 1]$, določena s predpisom

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) dv du.$$

Zgled. Naj bo (X, Y) zsv z gostoto

$$f_{X,Y}(x, y) = \begin{cases} \frac{x^2 y}{12}, & 0 \leq x \leq 2, 0 \leq y \leq 3, \\ 0, & \text{sicer.} \end{cases}$$

Bralec naj sam preveri, da je ta funkcija res gostota, saj je nenegativna, njen dvojni integral po celotni ravnini pa je enak 1. Za zgled uporabe pa izračunajmo verjetnosti $P(X < 1, Y > 2)$, $P(X < 1)$ in $P(Y < X)$.

$$P(X < 1, Y > 2) = P(0 < X < 1, 2 < Y < 3) = \int_0^1 \left(\int_2^3 \frac{x^2 y}{12} dy \right) dx = \frac{5}{72},$$

$$P(X < 1) = P(0 < X < 1, 0 < Y < 3) = \int_0^1 \left(\int_0^3 \frac{x^2 y}{12} dy \right) dx = \frac{1}{8},$$

$$P(Y < X) = P(0 < X < 2, 0 < Y < X) = \int_0^2 \left(\int_0^x \frac{x^2 y}{12} dy \right) dx = \frac{4}{15}.$$

Iz definicije porazdelitvene funkcije lahko izračunamo tudi

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y \frac{u^2 v}{12} dv du = \frac{x^3 y^2}{72} \text{ za } 0 \leq x \leq 2, 0 \leq y \leq 3.$$

S pomočjo tega izraza lahko zdaj direktno računamo nekatere verjetnosti, na primer

$$P(X < 1, Y < 2) = F_{X,Y}(1, 2) = \frac{4}{72},$$

$$P(X < 1) = F_{X,Y}(1, 3) = \frac{9}{72},$$

in od tod potem še druge, na primer

$$P(X < 1, Y > 2) = P(X < 1) - P(X < 1, Y < 2) = \frac{5}{72}.$$

Zveza med gostoto in porazdelitveno funkcijo za zsv je podobna zvezi med verjetnostno in porazdelitveno funkcijo za dsv, le da vsote zamenja integral.

Trditev. Za zsv (X, Y) velja:

- (i) $F_X(x) = \lim_{y \rightarrow \infty} F_{X,Y}(x, y)$ in podobno za $F_Y(y)$.
- (ii) $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$ in podobno za $f_Y(y)$.
- (iii) $f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y}$.

Dokaz. Prepuščamo bralcu. □

5.2.2 Neodvisnost zss

Tudi pri zveznih slučajnih spremenljivkah bi želeli govoriti o neodvisnosti. Ker so verjetnosti $P(X = x)$, $P(Y = y)$ in $P(X = x, Y = y)$ vselej enake 0, definiramo, da sta zvezni X, Y **neodvisni**, če velja

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

za vse $x, y \in \mathbb{R}$.

Trditev. Naj bosta X, Y zvezni slučajni spremenljivki. Potem so naslednje trditve enakovredne:

- (i) X in Y sta neodvisni.
- (ii) $F_{X,Y}(x, y) = F_X(x)F_Y(y)$.
- (iii) Obstajata taki funkciji $g, h: \mathbb{R} \rightarrow \mathbb{R}$, da za vsak par $x, y \in \mathbb{R}$ velja $f_{X,Y}(x, y) = g(x)h(y)$. Z drugimi besedami, gostoto para (X, Y) lahko zapišemo kot produkt gostot za X in Y .

Dokaz. (dodati) □

Zgled. Spremenljivki X, Y iz zadnjega zгледа sta neodvisni, saj lahko njuno gostoto $f_{X,Y}(x, y) = \frac{x^2 y}{12}$ na ustreznem območju izrazimo kot produkt funkcij $g(x) = \frac{3}{8}x^2$ in $h(y) = \frac{2}{9}y$, ki predstavljata ravno gostoti f_X oziroma f_Y .

Če pa bi imel neki par (X, Y) gostoto $f_{X,Y}(x, y) = x + y$ na ustreznem območju, potem tega izraza ne bi mogli faktorizirati v produkt $g(x)h(y)$. Taki spremenljivki bi bili odvisni.

Zgled (Buffonov poskus). Na neskončno ravnino so narisane vzporedne premice z medsebojno razdaljo 1. Na ravnino vržemo iglo dolžine 1. Zanima nas verjetnost, da igla seka katero od premic.

Označimo z X ostri kot med iglo in vodoravno črto, z Y pa razdaljo središča igle do premice in. Potem sta X in Y funkciji $S \rightarrow \mathbb{R}$, katerih zalogi vrednosti sta enaki $X(S) = [0, \pi/2]$ in $Y(S) =$

$[0, 1/2]$. Ker je met naključen, lahko predpostavimo, da so vse možne vrednosti za X in Y enako verjetne. Torej sta slučajni spremenljivki porazdeljeni zvezno enakomerno, $X \sim U(0, \pi/2)$ in $Y \sim U(0, 1/2)$. Njuni gostoti sta torej enaki

$$f_X(x) = \begin{cases} 2/\pi, & 0 \leq x \leq \pi/2, \\ 0, & \text{sicer,} \end{cases}$$

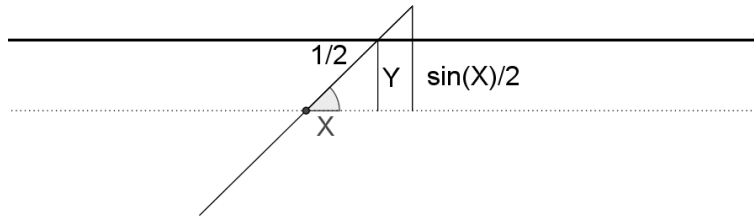
in

$$f_Y(y) = \begin{cases} 2, & 0 \leq y \leq 1/2, \\ 0, & \text{sicer.} \end{cases}$$

Ker je met naključen, je smiselno predpostaviti, da sta X in Y tudi neodvisni, zato je gostota para (X, Y) enaka produktu gostot:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) = \begin{cases} 4/\pi, & 0 \leq x \leq \pi/2 \text{ in } 0 \leq y \leq 1/2, \\ 0, & \text{sicer.} \end{cases}$$

Iz skice lahko razberemo, da igla seka kakšno izmed premic natanko tedaj, ko velja $Y < \frac{\sin X}{2}$.



Zanima nas torej verjetnost $P(Y < \frac{\sin(X)}{2})$, ki jo dobimo z integralom gostote $f_{X,Y}$ po ustreznem območju $D = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq \pi/2, 0 \leq y \leq \sin(x)/2\}$ v ravnini:

$$P\left(Y < \frac{\sin(X)}{2}\right) = \iint_D f_{X,Y}(x, y) dx dy = \int_0^{\pi/2} \left(\int_0^{\sin(x)/2} \frac{4}{\pi} dy \right) dx = \frac{2}{\pi} \doteq 0.64.$$

Bolj zanimiva od dejanske vrednosti je ugotovitev, da lahko torej približek za π določimo eksperimentalno tako, da iglo v resnici mečemo na ravnino in beležimo razmerje med številom uspešnih metov u in vseh metov m . Potem je $\pi \approx \frac{2m}{u}$.

Izkaže pa se, da je potrebno za približek na dve decimalni mesti poskus ponoviti prevečkrat, da bi bila ta metoda uporabna za praktične namene določanja vrednosti π z dejanskim metom igle, lahko pa jo simuliramo s pomočjo računalnika. Radovedni bralec naj si ogleda katerega od apletov na to temo na svetovnem spletu (išči Buffon Needle). S podobnimi simulacijami se v fiziki včasih izračuna približne vrednosti za nekatere integrale, za katere niso znane točne vrednosti (Monte Carlo simulacije).

V zaključku razdelka brez dokazov omenimo še, da je tudi za zss X in Y izraz $g(X, Y)$ spet zvezna slučajna spremenljivka, če je funkcija $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ "dovolj lepa". Posledično sta vsota $X + Y$ in produkt XY spet zvezni slučajni spremenljivki. Zato lahko tudi za zvezne slučajne spremenljivke definiramo pričakovano vrednost

$$E(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy,$$

če ustrezni integral obstaja. Tako lahko torej definiramo pričakovani vrednosti vsote $E(X + Y)$ in produkta $E(XY)$ in posledično tudi kovarianco $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$ ter korelacijo. Izkaže se, da lahko tudi za zss izpeljemo vse bistvene zveze, kot jih poznamo pri dss.

Trditev. Za poljubni zvezni slučajni spremenljivki X in Y velja

$$E(X + Y) = E(X) + E(Y),$$

če ustrezne vrednosti obstajajo. Če sta X in Y neodvisni, pa velja še

$$E(XY) = E(X)E(Y), \quad \text{cov}(X, Y) = 0, \quad \text{var}(X + Y) = \text{var}(X) + \text{var}(Y),$$

če ustrezne vrednosti obstajajo, in za gostoto vsote velja

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y) dy.$$

Dokaz. Posnemamo dokaze za dss, pri čemer vsote nadomestimo z integrali. □

6 Limitni izreki

6.1 Zakon velikih števil

Jacob Bernoulli se je kar 20 let ukvarjal z vprašanjem, kolikokrat je treba ponoviti neki eksperiment, da se bo povprečni rezultat meritev dovolj natančno ujema s pričakovano vrednostjo. V delu *Ars conjectandi* (Umetnost napovedovanja), ki je izšlo šele leta 1713, je zapisal izrek, ki ga je poimenoval kar Zlati izrek, danes pa ga imenujemo *Šibki oziroma Bernoullijev zakon velikih števil*.

Denimo torej, da neko meritev pod enakimi pogoji ponovimo velikokrat. Če so ponovitve neodvisne, si lahko predstavljamo, da slučajna spremenljivka X_i zabeleži vrednost meritve pri i -ti ponovitvi. Predpostavimo lahko, da so vse slučajne spremenljivke X_i enako porazdeljene, torej imajo enako pričakovano vrednost in varianco. Potem zlahka vidimo, da za njihovo povprečje

$$Z_n = \frac{X_1 + \dots + X_n}{n}$$

velja naslednje.

Trditev. Če so X_1, \dots, X_n neodvisne in enako porazdeljene s pričakovano vrednostjo $\mu = E(X_i)$ in varianco $\sigma^2 = \text{var}(X_i)$, potem za povprečje $Z_n = \frac{X_1 + \dots + X_n}{n}$ velja $E(Z_n) = \mu$ in $\text{var}(Z_n) = \frac{\sigma^2}{n}$.

Dokaz. $E(Z_n) = \frac{n\mu}{n} = \mu$ in $\text{var}(Z_n) = \frac{1}{n^2} \text{var}(X_1 + \dots + X_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$. □

Posledica. S ponavljanjem ponovitev meritev se razpršenost povprečja zmanjšuje proti 0, pričakovana vrednost pa se ne spreminja.

Z uporabo neenakosti Čebiševa lahko zdaj dokažemo Zlati izrek veliko hitreje, kot je to uspelo samemu Bernoulliju.

Izrek ((Bernoullijev) Šibki zakon velikih števil). Naj bo $Z_n = \frac{X_1 + \dots + X_n}{n}$ povprečje enako porazdeljenih neodvisnih slučajnih spremenljivk in naj bo $\mu = E(X_i)$. Potem za vsak $\epsilon > 0$ velja:

$$\lim_{n \rightarrow \infty} P(|Z_n - \mu| \geq \epsilon) = 0.$$

Dokaz. Naj bo $\epsilon > 0$. Po prejšnji trditvi je $E(Z_n) = \mu$ in $\text{var}(Z_n) = \sigma^2/n$, zato po neenakosti Čebiševa za Z_n sledi

$$P(|Z_n - \mu| \geq \epsilon) \leq \frac{1}{\epsilon^2} \frac{\sigma^2}{n}.$$

Izraz na desni seveda konvergira proti 0, ko gre $n \rightarrow \infty$. □

Denimo, da večkrat ponavljamo neko meritev, ki zaradi slučajnih vzrokov ni povsem točna. Potem Šibki zakon velikih števil pomeni, da lahko z dovolj velikim številom meritev dosežemo, da bo aritmetična sredina izmerkov z veliko verjetnostjo poljubno blizu točni vrednosti. Koliko meritev je potrebnih, pa iz tega izreka ni mogoče neposredno izračunati.

Opomba. Bernoulli je v svojem delu obravnaval posebne primere z vlečenjem kroglic iz žare. Za primer žare s 30 belimi in 20 črnimi kroglicami je poskusil izračunati, koliko vlečenj z vračanjem je potrebnih, da bo razmerje med vlečenji belih in vsemi vlečenji odstopalo od teoretičnega povprečja $3/5$ za manj kot $1/50$ z verjetnostjo več kot $1000/1001$. Dobil je rezultat 25500 ponovitev, ki se mu je verjetno zdel povsem neuporaben za praktične namene, in morda je ravno zato do smrti zavlačeval z objavo svoje knjige.

Opomba. Močnejši od Šibkega je takoimenovani Krepki Zakon velikih števil, ki ima močnejše zaključke, a tudi strožje predpostavke in zahtevnejši dokaz. Tudi iz Krepkega zakona pa ni mogoče neposredno izračunati števila meritev, potrebnih za željeno natančnost.

6.2 Centralni limitni izrek

Centralni limitni izrek (CLI) je eden osrednjih izrekov verjetnosti in statistike. Zagotavlja namreč, da se vsota velikega števila neodvisnih, enako porazdeljenih slučajnih spremenljivk obnaša podobno kot normalno porazdeljena slučajna spremenljivka. Obstajajo številne različice CLI, ki se razlikujejo glede na splošnost predpostavk in formulacijo moč rezultata, seveda pa tudi glede na zahtevnost dokaza. V tem besedilu bomo eno od osnovnih različic CLI navedli kar brez dokaza, bralec ga lahko poišče denimo v knjigah [5] ali [2].

Izrek (Centralni limitni izrek). *Naj bodo X_1, X_2, \dots neodvisne in enako porazdeljene slučajne spremenljivke, tako da je $\mu = E(X_i)$ in $\sigma^2 = \text{var}(X_i) < \infty$. Označimo vsoto $S_n = X_1 + \dots + X_n$, torej je $E(S_n) = n\mu$, $\text{var}(S_n) = n\sigma^2$ in $\sigma(S_n) = \sigma\sqrt{n}$. Potem velja*

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = F_Y(b) - F_Y(a),$$

kjer je F_Y porazdelitvena funkcija standardizirane normalne slučajne spremenljivke $Y \sim N(0, 1)$.

Z drugimi besedami, če vsoti S_n velikega števila neodvisnih enako porazdeljenih slučajnih spremenljivk odštejemo njeno povprečje in rezultat delimo s standardnim odklonom, dobimo približno standardizirano normalno slučajno spremenljivko $N(0, 1)$.

V posebnem primeru, ko je S_n vsota Bernoullijevih slučajnih spremenljivk, dobimo iz CLI klasičen rezultat De-Moivreja in Laplacea, da lahko za velike n lahko binomsko slučajno spremenljivko $X \sim \text{Bin}(n, p)$ aproksimiramo z normalno $Y \sim N(np, npq)$. Ta rezultat se je v praktičnem računanju uporabljal vse do nedavnega, ko so tabele vrednosti normalne porazdelitve zamenjali računalniki.

Zgled. [1] Denimo, da kovanec vržemo 100x. Kakšna je verjetnost, da pade vsaj 56 grbov?

Zanima nas $P(X \geq 56)$, kjer je $X \sim \text{Bin}(100, 0.5)$, torej velja $E(X) = 50$ in $\sigma(X) = 5$. Iz pogoja $X \geq 56$ sledi $\frac{X-50}{5} \geq \frac{6}{5}$, zato po CLI izračunamo

$$\begin{aligned} P(X \geq 56) &= P\left(\frac{X-50}{5} \geq \frac{6}{5}\right) \\ &\approx P(Y \geq 1.2) = 1 - P(Y < 1.2) = \\ &= 1 - 0.8849 = 0.1151. \end{aligned}$$

Pri tem smo vrednost

$$P(Y < 1.2) = F_Y(1.2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1.2} e^{-\frac{1}{2}x^2} dx = 0.8849$$

odčitali iz ustrezne tabele.

Rezultat sicer ni zelo točen, saj nismo upoštevali takoimenovane korekcije histograma. Ker binomska slučajna spremenljivka zavzame le celoštevilске vrednosti, normalna pa je zvezna, se pri tovrstnih aproksimacijah namesto verjetnosti $P(X \geq 56)$ običajno računa verjetnost $P(X \geq 55.5) = 1 - P(Y < 1.1) = 0.1357$, kar da precej boljši približek.

Zgled. Statistični podatki pravijo, da so učenci pri NPZ dosegli v povprečju 61 točk, s standardnim odklonom 8.2 točke. Ker so rezultati posameznikov neodvisne in enako porazdeljene diskretne slučajne spremenljivke, lahko domnevamo, da je njihovo povprečje konvergira k normalni porazdelitvi. Če nas zanima, kakšen delež učencev je dosegel več kot 82 točk ali koliko točk je moral doseči učenec, da je bil uvrščen med 5% najuspešnejših, lahko izračunamo ustrezne rezultate s pomočjo tabel normalne porazdelitve ali računalnika, na primer v GeoGebri.

Centralni limitni izrek predstavlja tudi osnovo za preverjanje hipotez v statistiki. Več o tem bo bralec izvedel pri kakšnem specializiranem predmetu iz področja statistike.

Literatura

- [1] Durrett, Rick, Elementary Probability for Applications, Cambridge University Press, 2009.
- [2] Hladnik, Milan, Verjetnost in statistika, zapiski predavanj, Založba FRI in FE, 2002.
- [3] Prakash Gorroochurn, Classic Problems of Probability, Wiley, 2012.
- [4] Charles M. Grinstead, J. Laurie Snell, Introduction to Probability, American Mathematical Society, 1997.
- [5] Jamnik, Rajko, Verjetnostni račun in statistika, DMFA Slovenije, 1986.
- [6] Šparl, Primož, Verjetnostni račun in statistika, zapiski predavanj. Spletna učilnica UL PeF, 2015.
- [7] Širca, Simon, Verjetnost v fiziki, DMFA Slovenije, 2015.
- [8] Jeffrey S. Rosenthal, Ko strela udari, Skrivnostni svet verjetnosti, DMFA, 2006.