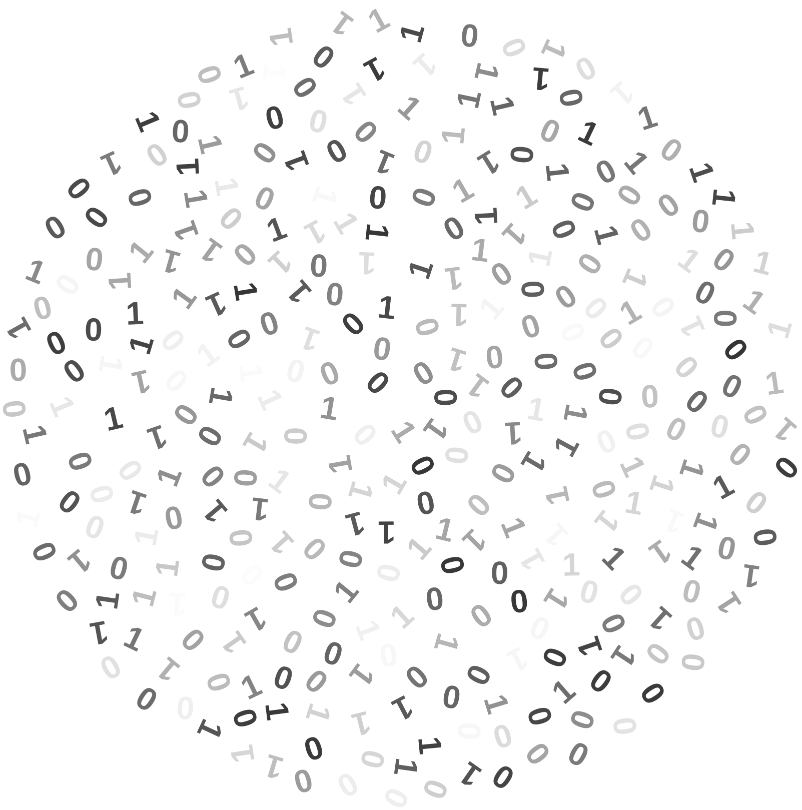
Základy pravdepodobnosti pre študentov informatiky a dátovej vedy



Radoslav Harman

Lenka Filová

Autori: Radoslav Harman, Lenka Filová
Názov: Základy pravdepodobnosti pre študentov informatiky a dátovej vedy
Vydavateľ: Knižničné a edičné centrum FMFI UK
Rok vydania: 2022
Miesto vydania: Bratislava
Vydanie: prvé
Počet strán: 109
ISBN 978-80-8147-126-1

Obsah

Prehľad značenia

$\mathbb{N}$	množina prirodzených čísel
$\mathbb{R}$	množina reálnych čísel
$\mathbb{R}^m$	množina m-rozmerných reálnych vektorov
I_n	jednotková matica s rozmermi $n \times n$
Ω	priestor elementárnych výsledkov, výberový priestor
$\mathcal{S}$	σ-algebra, obvykle σ-algebra udalostí
P	pravdepodobnostná miera, pravdepodobnosť
$(\Omega, \mathcal{S}, P)$	pravdepodobnostný priestor
$\mathcal{B}$	σ-algebra borelovských podmnožín množiny $\mathbb{R}$
$\mathcal{B}_m$	σ-algebra borelovských podmnožín množiny $\mathbb{R}^m$
$A, B, \ldots$	náhodné udalosti
$P(A)$	pravdepodobnosť udalosti A
$P(A\|B)$	podmienená pravdepodobnosť udalosti A za podmienky B
$\lambda(C)$	dĺžka, plocha alebo objem množiny C
$X, Y, \ldots$	náhodné premenné
$E(X)$	stredná hodnota náhodnej premennej X
$D(X)$	rozptyl náhodnej premennej X
$X \sim \ldots$	náhodná premenná X má rozdelenie ...
$R(M)$	diskrétne rovnomerné rozdelenie na množine M
$Alt(p)$	alternatívne rozdelenie s parametrom p
$Bin(n, p)$	binomické rozdelenie s parametrami n, p
$Po(\lambda)$	Poissonovo rozdelenie s parametrom λ
$Geo(p)$	geometrické rozdelenie s parametrom p
$Hyp(M, N, n)$	hypergeometrické rozdelenie s parametrami M, N, n
$Emp(x_1, \ldots, x_n)$	empirické rozdelenie určené dátami $x_1, \ldots, x_n$
$R(a, b)$	rovnomerné rozdelenie na intervale (a, b)
$Exp(\lambda)$	exponenciálne rozdelenie s parametrom λ
$N(\mu, \sigma^2)$	normálne rozdelenie s parametrami μ, σ^2
Φ	distribučná funkcia rozdelenia $N(0, 1)$
$cov(X_1, X_2)$	kovariancia náhodných premenných X_1, X_2
$\rho(X_1, X_2)$	korelačný koeficient náhodných premenných X_1, X_2
$\mathbf{X}, \mathbf{Y}, \ldots$	náhodné vektory
$Cov(\mathbf{X})$	kovariančná matica náhodného vektora $\mathbf{X}$
$\mathbf{X} \sim \ldots$	náhodný vektor $\mathbf{X}$ má rozdelenie ...
$Mult(n, p_1, \ldots, p_m)$	multinomické rozdelenie s parametrami $n, p_1, \ldots, p_m$
$N_m(\mu, \Sigma)$	m-rozmerné normálne rozdelenie s parametrami μ, Σ

Slovensko-anglický slovník

Pre lepšiu orientáciu v anglickej terminológii uvádzame slovník základných pojmov používaných v teórii pravdepodobnosti. Poznamenajme, že v anglickej odbornej literatúre sa môžeme stretnúť aj s odlišnými prekladmi niektorých uvedených pojmov.

Alternatívne rozdelenie	Bernoulli distribution
Aposteriórna pravdepodobnosť	Posterior probability
Apriórna pravdepodobnosť	Prior probability
Bayesova formula	Bayes formula
Bayesova veta	Bayes theorem
Binomická formula	Binomial formula
Binomické rozdelenie	Binomial distribution
Borelovská funkcia	Borel function
Borelovská množina	Borel set
Cauchyho-Schwarzova nerovnosť	Cauchy-Schwarz inequality
Centrálna limitná veta	Central limit theorem
Čebyševova nerovnosť	Chebyshev's inequality
Černovova nerovnosť	Chernoff's inequality
De Moivreova-Laplaceova veta	De Moivre-Laplace theorem
Diskrétna náhodná premenná	Discrete random variable
Distribučná funkcia	Cumulative distribution function
Elementárny výsledok	Elementary outcome
Empirická distribučná funkcia	Empirical distribution function
Empirické rozdelenie	Empirical distribution
Entropia	Entropy
Exponenciálne rozdelenie	Exponential distribution
Falošná kocka (minca)	Biased die (coin), Unfair die (coin)
Funkcia rozdelenia pravdepodobnosti	Probability mass function
Gaussovské rozdelenie	Gaussian distribution
Geometrická pravdepodobnosť	Geometric probability
Geometrické rozdelenie	Geometric distribution
Hlava (na minci)	Heads
Hodiť (mincou)	Toss (a coin)
Hustota	Density
Hypergeometrické rozdelenie	Hypergeometric distribution
Jensenova nerovnosť	Jensen's inequality
Kocka (hracia)	Die
Kombinatorická pravdepodobnosť	Combinatorial probability

Korelačný koeficient	Correlation coefficient
Korelačná matica	Correlation matrix
Kovariancia	Covariance
Kovariančná matica	Covariance matrix
Kvantilová funkcia	Quantile function
Medián	Median
Meranie	Measurement
Merateľná množina	Measurable set
Miera	Measure
Mnohorozmerné rozdelenie	Multivariate distribution
Multinomická formula	Multinomial formula
Multinomické rozdelenie	Multinomial distribution
Náhodná prechádzka	Random walk
Náhodná premenná	Random variable
Náhodná udalosť	Random event
Náhodný vektor	Random vector
Náhodný výber	Random sample
Nezávislé/závislé náhodné premenné	Independent/dependent random variables
Normálne rozdelenie	Normal distribution
Paretovo rozdelenie	Pareto distribution
Podmienená pravdepodobnosť	Conditional probability
Podmienka merateľnosti	Measurability condition
Poissonova aproximácia	Poisson approximation
Poissonova veta	Poisson theorem
Poissonovo rozdelenie	Poisson distribution
Pozorovanie	Observation
Pravdepodobnosť	Probability
Pravdepodobnostný model	Probability model
Pravdepodobnostný priestor	Probability space
Priemer	Mean
Priestor elementárnych výsledkov	Elementary outcomes space
Princíp zapojenia-vypojenia	Inclusion-exclusion principle
Rovnomerné rozdelenie	Uniform distribution
Rozdelenie pravdepodobnosti	Probability distribution
Rozklad výberového priestoru	Sample space partitioning
Rozptyl	Variance
Smerodajná odchýlka	Standard deviation
Spojitá náhodná premenná	Continuous random variable
Stredná hodnota	Expected value, Expectation
Teória miery	Measure theory
Udalosť	Event
Výberový priestor	Sample space
Vyvážená kocka (minca)	Fair die (coin), Unbiased die (coin)
Zákon veľkých čísel	Law of large numbers
Združene nezávislé udalosti	Mutually independent events
Združené rozdelenie	Joint distribution
Znak (na minci)	Tails

Predslov

Teória pravdepodobnosti poskytuje exaktné metódy, ktoré umožňujú zakomponovať do procesu spracovania dát apriórne vedomosti a kvantifikovať neistotu, chybovosť či riziko. Priamo alebo sprostredkovane sa používa vo všetkých oblastiach ľudskej činnosti, ktorých cieľom je získať z dát užitočnú informáciu a na jej základe sa efektívne rozhodovať.

Tieto vysokoškolské skriptá sú záznamom z prednášok z teórie pravdepodobnosti, určených pre študentov informatiky a dátovej vedy. Cieľom je vysvetliť matematické princípy, ktoré budú predstavovať základ pre štúdium špecializovaných predmetov z oblasti náhodných procesov, matematickej štatistiky, strojového učenia a niektorých predmetov teoretickej informatiky.

Skriptá pozostávajú zo sekvenčne číslovaných položiek typu Definícia, Lema, Veta, Poznámka a Príklad, čoho zmyslom je uľahčiť komunikáciu a odkazovanie sa. Najdôležitejšie pojmy sú zvýraznené hrubým písmom, najdôležitejšie vety sú pomenované. Pre ľahšiu orientáciu sú skriptá doplnené zoznamom symbolov a zoznamom základných pojmov s prekladom do anglického jazyka.

Cieľom týchto skrípt nie je nahradiť komunikáciu medzi vyučujúcim a študentom, čiže nejde o učebnicu určenú na samoštúdium. Záujemcov o samoštúdium teórie pravdepodobnosti odkazujeme na kvalitné knihy v anglickom jazyku, napríklad [2], [3], [5], [6] a [8].

Znalosti, ktoré sa vyžadujú na úspešné zvládnutie predmetu: diskrétna matematika (základy logiky, teórie množín, kombinatoriky), matematická analýza (limity, nekonečné rady, reálne funkcie, základy diferenciálneho a integrálneho počtu), lineárna algebra (základy teórie matíc).

doc. Mgr. Radoslav Harman, PhD. a
doc. Mgr. Lenka Filová, PhD.

Katedra aplikovanej matematiky a štatistiky
FMFI UK Bratislava

3. december 2022

Kapitola 1

Axiomatická definícia pravdepodobnosti

1.1 Priestor udalostí

Základom pre vybudovanie modernej teórie pravdepodobnosti je matematický objekt, ktorý nazývame σ-algebra.[1] Prvky σ-algebry, čiže množiny „elementárnych výsledkov“ sledovaného deja, budú reprezentovať „udalosti“, ktorým je možné priradiť „pravdepodobnosť“.

Definícia 1.1. Nech Ω je neprázdna množina a nech $\mathcal{S} \subseteq 2^{\Omega}$. Systém $\mathcal{S}$ podmnožín množiny Ω nazývame **σ-algebra** na množine Ω, ak platí 1) $\Omega \in \mathcal{S}$; 2) Ak $A \in \mathcal{S}$, tak aj $\Omega \setminus A \in \mathcal{S}$; 3) Ak $(A_i)_{i \in I}$ je postupnosť množín patriacich do $\mathcal{S}$, tak aj $\cup_{i \in I} A_i \in \mathcal{S}$. Vlastnosť 2) nazývame **uzavretosť** σ-algebry **na doplnky**[2] a vlastnosť 3) nazývame **uzavretosť** σ-algebry **na spočítateľné zjednotenia**.

Poznámka 1.2. a) Symbolom 2^{Ω} označujeme množinu všetkých podmnožín množiny Ω, takzvanú „potenčnú“ množinu. Čiže v definícii 1.1 zodpovedá zápis $\mathcal{S} \subseteq 2^{\Omega}$ tomu, že $\mathcal{S}$ je systém (nie nutne všetkých) podmnožín množiny Ω. b) Ak použijeme v texte pojem „postupnosť“, myslíme tým konečnú alebo aj nekonečnú spočítateľnú postupnosť. To znamená, že v definícii 1.1 je I konečná alebo nekonečná spočítateľná množina. c) Niekedy sa pod pojmom σ-algebra myslí celá usporiadaná dvojica $(\Omega, \mathcal{S})$ z definície 1.1. V týchto skriptách budeme pod pojmom σ-algebra rozumieť len systém množín $\mathcal{S}$.

Príklad 1.3. Nech Ω je ľubovoľná neprázdna množina a nech $\emptyset \neq A \subsetneq \Omega$. Potom $\{\emptyset, \Omega\}$, $\{\emptyset, A, \Omega/A, \Omega\}$ a 2^{Ω} sú σ-algebry na množine Ω, pretože spĺňajú všetky vlastnosti definície 1.1, ako sa dá ľahko presvedčiť. Ak $\Omega = \{1, 2, 3\}$, tak systém množín $\{\emptyset, \{1\}, \{2\}, \{3\}, \Omega\}$ nie je σ-algebra na množine Ω, pretože nesplĺňa vlastnosť 2) a dokonca ani vlastnosť 3) z definície 1.1. Uvedený systém podmnožín množiny Ω totiž nie je uzavretý ani vzhľadom na doplnky, ani vzhľadom na spočítateľné zjednotenia.

Veta 1.4 (Uzavretosť σ-algebry vzhľadom na spočítateľné prieniky)**.** Ak $\mathcal{S}$ je σ-algebra podmnožín množiny Ω a $(A_i)_{i \in I}$ je postupnosť prvkov systému $\mathcal{S}$, tak $\cap_{i \in I} A_i \in \mathcal{S}$.

[1] Hneď na úvod však dodajme, že pravdepodobnostné modely rôznych reálnych dejov môžu byť vybudované na rôznych σ-algebrách, čiže nepoužívame len jednu „univerzálnu“ σ-algebru ako model všetkých náhodných dejov súčasne.

[2] Doplnky sa často nazývajú aj „komplementy“; v týchto skriptách používame prvý uvedený pojem.

Dôkaz. Ak $A_i \in \mathcal{S}$ pre všetky $i \in I$, tak podľa vlastnosti 2) z definície 1.1 platí $\Omega \setminus A_i \in \mathcal{S}$ pre všetky $i \in I$; podľa vlastnosti 3) teda máme $\cup_{i \in I}(\Omega \setminus A_i) \in \mathcal{S}$ a opäť podľa vlastnosti 2) dostávame $\Omega \setminus (\cup_{i \in I}(\Omega \setminus A_i)) \in \mathcal{S}$. Lenže $\Omega \setminus (\cup_{i \in I}(\Omega \setminus A_i)) = \cap_{i \in I} A_i$ na základe De Morganových pravidiel. □

Poznámka 1.5. „Axiómy" definície 1.1 zaručujú, že systém $\mathcal{S}$ je uzavretý nielen vzhľadom na spočítateľné prieniky a zjednotenia, ale aj vzhľadom na akékoľvek množinové operácie, ktoré sú vyjadriteľné pomocou spočítateľných zjednotení, prienikov alebo doplnkov. Ukážme napríklad, že systém $\mathcal{S}$ je uzavretý vzhľadom na množinový rozdiel.

Veta 1.6 (Uzavretosť σ-algebry vzhľadom na množinový rozdiel)**.** Ak $\mathcal{S}$ je σ-algebra podmnožín množiny Ω a $A_1, A_2 \in \mathcal{S}$, tak aj $A_2 \setminus A_1 \in \mathcal{S}$.

Dôkaz. Ak $A_1, A_2 \in \mathcal{S}$, tak podľa vlastnosti 2) definície 1.1 platí $\Omega \setminus A_1 \in \mathcal{S}$. Teda podľa vety 1.4 obsahuje systém $\mathcal{S}$ aj množinu $A_2 \cap (\Omega \setminus A_1) = A_2 \setminus A_1$. □

Poznámka 1.7. Pojem σ-algebry je možné definovať viacerými ekvivalentnými spôsobmi. V literatúre sa najčastejšie vyskytuje definícia 1.1.[3] Takáto definícia je vhodná na následnú matematickú analýzu, avšak menej vhodná na pochopenie *významu* σ-algebry. Význam σ-algebry je totiž vymedziť systém všetkých tých podmnožín množiny Ω, ktorým je možné konzistentne priradiť „mieru", čiže niečo ako „dĺžku", „plochu" alebo „objem" a v prípade pravdepodobnostných modelov „pravdepodobnosť".[4] Intuitívne je zrejmé, že miera ako číselná charakteristika by mala mať nasledovnú vlastnosť: ak množinám $A_1, A_2, \dots$ vieme priradiť mieru a tieto množiny sa neprekrývajú, čiže sú navzájom disjunktné, tak vieme priradiť mieru aj množine $A_1 \cup A_2 \cup \dots$ a táto miera musí byť súčtom mier množín $A_1, A_2, \dots$. Ak by sme vlastnosť 3) v definícii 1.1 nahradili takouto interpretačne prirodzenejšou, hoci technicky o niečo zložitejšou požiadavkou uzavretosti vzhľadom na *disjunktné* spočítateľné zjednotenia, dostali by sme matematicky ekvivalentnú definíciu pojmu σ-algebry.

Poznámka 1.8. V pravdepodobnostných modeloch predstavujú prvky množiny Ω atomické, ďalej nerozložiteľné potenciálne výsledky modelovaného deja, čiže realizácie pokusu, pozorovania, merania a podobne. Tie podmnožiny množiny Ω, ktoré patria do systému $\mathcal{S}$, reprezentujú „udalosti", ktorým je možné pripísať pravdepodobnosť realizácie.

Definícia 1.9. Nech $\mathcal{S}$ je σ-algebra udalostí na množine Ω. V pravdepodobnostných aplikáciách nazývame množinu Ω **výberový priestor**, prvky množiny Ω nazveme **elementárne výsledky**[5] a podmnožiny Ω patriace do $\mathcal{S}$ nazývame **náhodné udalosti**.[6]

Príklad 1.10. Ak je výberový priestor Ω konečná množina, tak v našom modeli obvykle volíme $\mathcal{S} = 2^{\Omega}$.[7] Napríklad, nech náš experiment pozostáva z jedného hodu hracou kockou.

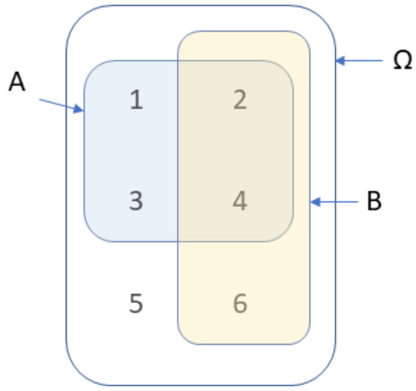
[3] Niekedy sa môžeme stretnúť s analogickou definíciou, v ktorej sa namiesto spočítateľných zjednotení používajú spočítateľné prieniky.

[4] Pojmom pravdepodobnosť sa budeme detailne zaoberať v ďalšej časti.

[5] V literatúre sa niekedy používa namiesto pojmu elementárny výsledok pojem „elementárna udalosť". To je ale formálne mätúce, lebo pri takejto terminológii elementárne udalosti (čoby *prvky* Ω) formálne nie sú náhodné udalosti.

[6] Skrátene budeme náhodné udalosti často nazývať len „udalosti".

[7] Ak by sme vždy pracovali len s konečnými výberovými priestormi, tak by sme pojem σ-algebry prakticky vôbec nepotrebovali. Ak však chceme vybudovať formálne matematicky korektnú definíciu pravdepodobnosti v situáciách *ne*konečnej množiny všetkých potenciálnych výsledkov, pojem σ-algebry už využijeme netriviálnym spôsobom.



Obr. 1.1: Ilustračný obrázok k príkladu 1.10

Ak nás na tomto experimente zaujíma jedine to, na ktorú stranu padne kocka (a ak z modelu vylúčime možnosť, že kocka zostane stáť na hrane a iné anomálne situácie), potom je zmysluplným výberovým priestorom $\Omega = \{1, 2, \ldots, 6\}$. Systém udalostí nech je $\mathcal{S} = 2^{\Omega}$. Udalosť $A = \{1, 2, 3, 4\}$ zodpovedá výroku „padne číslo menšie než 5“, udalosť $B = \{2, 4, 6\}$ zodpovedá výroku „padne párne číslo“ a podobne. Pozri obrázok 1.1.

Poznámka 1.11. Uvedomme si, že v modeli danom nejakým pravdepodobnostným priestorom existuje prirodzená korešpondencia medzi udalosťami a výrokmi týkajúcimi sa elementárnych výsledkov, ako je naznačené v predchádzajúcom príklade. Vo všeobecnosti udalosti $A \in \mathcal{S}$ zodpovedá výrok „nastane niektorý elementárny výsledok z A“. V tejto korešpondencii zodpovedajú množinové operácie medzi udalosťami logickým operáciám medzi príslušnými výrokmi. Množinový doplnok takto zodpovedá negácii výroku, zjednotenie logickému „alebo“, prienik logickému „a“ alebo napríklad symetrická diferencia množín ($A\Delta B = (A\backslash B)\cup(B\backslash A)$) zodpovedá logickému „výlučné alebo“ (známemu aj pod označením `xor`).

Príklad 1.12 (Jednoduchý náhodný výber bez návratu)**.** Uvažujme nasledovný experiment. Zo základného súboru $M > n$ objektov indexovaných $1, 2, \ldots, M$ náhodne vyberieme n-ticu (rôznych) objektov. V tomto prípade je možné formalizovať výberový priestor Ω ako množinu všetkých n-prvkových podmnožín množiny $\{1, 2, \ldots, M\}$. Keďže ide o konečnú množinu, je prirodzené zvoliť σ-algebru udalostí $\mathcal{S} = 2^{\Omega}$. Slovne vyjadrenú udalosť „Nevyberieme objekt s indexom 1“ reprezentuje množina A všetkých n-prvkových podmnožín množiny $\{2, 3, \ldots, M\}$. Alebo ak sú k a $N < M$ prirodzené čísla, tak slovne vyjadrenú udalosť „Práve k z vybratých objektov bude mať index nanajvýš N“ reprezentuje množina B tých podmnožín výberového priestoru, ktoré majú k-prvkový prienik s množinou $\{1, 2, \ldots, N\}$.

Poznámka 1.13. Ukazuje sa, že ak je množina Ω nekonečná a zvolili by sme $\mathcal{S} = 2^{\Omega}$, potom by mohlo byť obtiažne priradiť prvkom takto bohatého systému pravdepodobnosť spĺňajúcu prirodzené vlastnosti, ktoré od pravdepodobnosti očakávame. Ďalej stručne vysvetlíme, aké množiny udalostí je možné používať v takýchto situáciách. Potrebujeme na to ale ešte pojem σ-algebry „generovanej“ systémom množín.

Lema 1.14. Nech J je neprázdna indexová množina (nemusí byť spočítateľná) a nech $\mathcal{S}_j$ je σ-algebra na množine Ω, pre každé $j \in J$. Potom $\cap_{j\in J}\mathcal{S}_j$ je tiež σ-algebra na množine Ω.

Dôkaz. Lemu je možné dokázať priamočiarym overením toho, že $\mathcal{S} = \cap_{j\in J}\mathcal{S}_j$ spĺňa vlastnosti 1), 2), 3) definície 1.1. □

Definícia 1.15. Nech $\Omega \neq \emptyset$ a nech $\mathcal{F}$ je nejaký systém podmnožín množiny Ω. Nech $\sigma(\mathcal{F})$ je prienik všetkých takých systémov $\mathcal{S}$ podmnožín množiny Ω, pre ktoré platí: 1) $\mathcal{F} \subseteq \mathcal{S}$ a 2) $\mathcal{S}$ je σ-algebra na Ω. Potom σ-algebru[8] $\sigma(\mathcal{F})$ nazývame **σ-algebra podmnožín množiny Ω generovaná systémom množín** $\mathcal{F}$ alebo tiež **minimálna σ-algebra** podmnožín množiny Ω obsahujúca systém $\mathcal{F}$.

Príklad 1.16. V prípade, že je systém $\mathcal{F}$ konečný, je jednoduché nájsť $\sigma(\mathcal{F})$ postupným pridávaním prienikov, zjednotení a doplnkov. Nech napríklad $\Omega = \{1,2,3,4,5\}$ a $\mathcal{F} = \{\{1,2\},\{3,4\}\}$. Potom $\sigma(\mathcal{F}) = \{\emptyset, \Omega, \{1,2\}, \{3,4\}, \{1,2,3,4\}, \{5\}, \{1,2,5\}, \{3,4,5\}\}$. V prípade nekonečne veľkého systému $\mathcal{F}$ môže byť konštrukcia $\sigma(\mathcal{F})$ zložitá; takým situáciám sa nevyhneme ani v týchto skriptách, hoci nepôjdeme do detailov.

Príklad 1.17 (Nekonečné hádzanie mincou)**.** Hádžeme mincou (nekonečne dlho), pričom akákoľvek pre nás zaujímavá informácia je obsiahnutá v zázname o tom, v ktorých hodoch padla na minci hlava a v ktorých hodoch padol znak. Túto situáciu je prirodzené formalizovať tým spôsobom, že výberový priestor Ω bude množina všetkých nula-jednotkových nekonečných postupností, kde nula reprezentuje padnutie hlavy a jednotka reprezentuje padnutie znaku. Definovať na tejto množine vhodnú σ-algebru už však nie je celkom jednoduché. Ukazuje sa, že v tomto prípade je vhodnou σ-algebrou udalostí $\mathcal{S} = \sigma(\mathcal{F})$, kde $\mathcal{F} = \{A_k^{(b)} : k \in \mathbb{N} \wedge b \in \{0,1\}\}$ a pre každé $k \in \mathbb{N}$, $b \in \{0,1\}$ je $A_k^{(b)}$ množinou všetkých tých postupností z množiny Ω, ktorých k-ty člen je rovný hodnote b. Tento systém udalostí je dostatočne bohatý, ale nie *príliš*[9] bohatý. Napríklad o všetkých nasledovných udalostiach je možné ľahko ukázať, že patria do $\mathcal{S}$: $A_1 = \{(0,0,\ldots)\}$ („vo všetkých hodoch padne hlava“), $A_2 = \{(i_1, i_2, \ldots) \in \Omega : i_k \neq i_{k+1} \text{ pre všetky } k \in \mathbb{N}\}$ („budú sa striedať hlavy a znaky“), $A_3 = \{(i_1, i_2, \ldots) \in \Omega : i_1 = \cdots = i_{m-1} = 0 \wedge i_m = 1\}$ pre nejaké zadané $m \in \mathbb{N}$ („znak padne prvýkrát v m-tom hode“).

Poznámka 1.18. Na formálne exaktné pravdepodobnostné modelovanie je ešte komplikovanejšia taká (častá) situácia, v ktorej je vhodným výberovým priestorom $\Omega = \mathbb{R}$ alebo $\Omega = \mathbb{R}^m$, čiže ak je prirodzené uvažovať ako elementárne výsledky reálne čísla, prípadne vektory reálnych čísel. Pripomeňme, že $O \subseteq \mathbb{R}^m$ nazývame otvorenou množinou, ak pre všetky $x \in O$ existuje také $\epsilon > 0$, že každé y, ktorého Euklidovská vzdialenosť od x je menšia ako ϵ, tiež patrí do O. Voľne vyjadrené, otvorenou je taká množina, ktorej každý bod je „vo vnútri“ tento množiny, čiže množina, ktorá „nemá hraničné body“. V $\mathbb{R}$ sú typickými otvorenými množinami otvorené intervaly a ich zjednotenia, v $\mathbb{R}^2$ obdĺžniky (bez strán) a ich zjednotenia, $\mathbb{R}^3$ kvádre (bez stien) a ich zjednotenia a tak ďalej.

Definícia 1.19. Nech $\mathcal{O}$ je systém všetkých otvorených podmnožín množiny $\mathbb{R}^m$, $m \geq 1$. Potom σ-algebru $\mathcal{B}_m$ podmnožín $\mathbb{R}^m$ generovanú systémom $\mathcal{O}$ nazývame **σ-algebra borelovských podmnožín** $\mathbb{R}^m$. Systém $\mathcal{B}_1$ budeme zjednodušene značiť symbolom $\mathcal{B}$.

[8] To, že ide o σ-algebru, zaručuje lema 1.14.

[9] Je to prekvapivé, ale systém 2^Ω nemôžeme v tomto prípade zobrať ako $\mathcal{S}$. Dá sa totiž ukázať, že pre Ω rovné všetkým nula-jednotkovým postupnostiam neexistuje ani jedno zobrazenie $P : 2^\Omega \to \mathbb{R}$, ktoré by spĺňalo prirodzené vlastnosti pravdepodobnostnej miery, o ktorých píšeme v nasledujúcej podkapitole.

Veta 1.20 (Základné borelovské podmnožiny množiny $\mathbb{R}$)**.** Nech $a, b \in \mathbb{R}$, pričom $a < b$. Potom všetky nasledovné množiny patria do $\mathcal{B}$: $(-\infty, a)$, (a, ∞), $(-\infty, a]$, $[a, \infty)$ $(a, b]$, $[a, b)$, $[a, b]$, $\{a\}$. Do systému $\mathcal{B}$ tiež patrí akákoľvek spočítateľná podmnožina množiny $\mathbb{R}$.

Dôkaz. To, že otvorené intervaly sú borelovské množiny, plynie priamo z definície. Avšak interval akéhokoľvek typu (ako aj jednoprvková množina) musí byť borelovskou množinou, pretože ho je možné zapísať ako prienik spočítateľnej postupnosti otvorených intervalov a σ-algebra je uzavretá vzhľadom na spočítateľné prieniky. Spočítateľná množina je borelovská preto, lebo je spočítateľným zjednotením množín typu $\{a\}$, $a \in \mathbb{R}$, ktoré patria do systému $\mathcal{B}$ podľa prvej časti vety a systém $\mathcal{B}$ je uzavretý vzhľadom na spočítateľné zjednotenia. □

Poznámka 1.21. Vo vete 1.20 sme ukázali, že mnohé bežne používané podmnožiny množiny $\mathbb{R}$ patria do $\mathcal{B}$. V skutočnosti akákoľvek „slušná“ podmnožina množiny $\mathbb{R}$ patrí do $\mathcal{B}$. Neborelovské množiny existujú, dokonca ich je „veľmi veľa“, sú ale extrémne komplikované a v reálnych aplikáciách sa nevyskytujú. Neborelovskými množinami sa ďalej nebudeme zaoberať, hoci kvôli úplnosti sme museli spomenúť ich existenciu. Podobne ako v príklade 1.17, všetkých podmnožín množiny $\mathbb{R}$ je *príliš* veľa na to, aby sme na takto bohatom systéme vedeli definovať pravdepodobnosť vhodnú na modelovanie reálnych dejov.[10] Preto sa v teórii pravdepodobnosti sústreďujeme takmer výlučne na borelovské podmnožiny množiny $\mathbb{R}$. Podobná poznámka platí pre $\mathbb{R}^m$, resp. pre $\mathcal{B}_m$, ak $m \geq 2$.

Lema 1.22 (Zúženie σ-algebry)**.** Nech $\mathcal{S}$ je σ-algebra na množine Ω a nech $B \in \mathcal{S}$. Potom systém $\mathcal{S}_B = \{C \in \mathcal{S} : C \subseteq B\}$ je σ-algebra na množine B.

Dôkaz. Dôkaz je možné vykonať priamym overením vlastností 1), 2), 3) definície 1.1. □

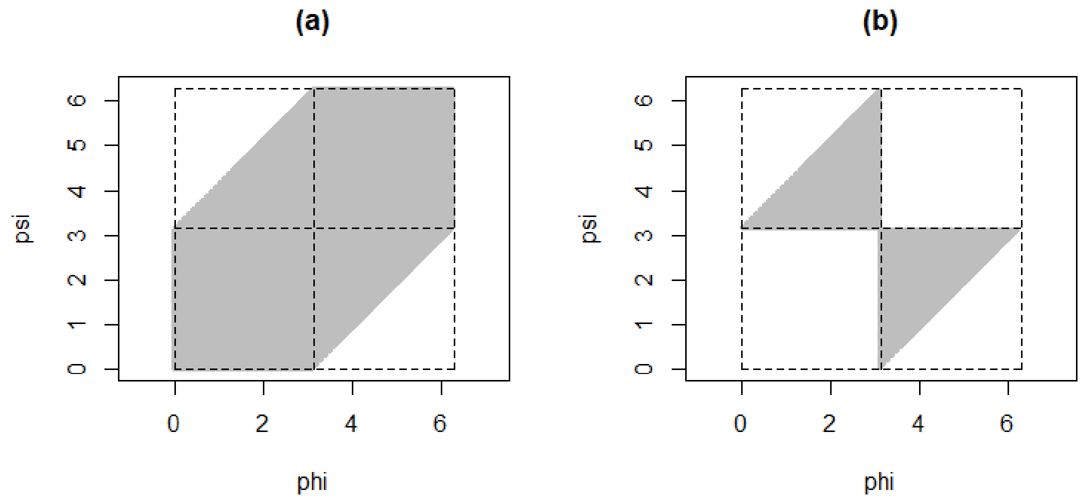
Poznámka 1.23. Systém $\mathcal{S}_B$ z lemy 1.22 môžeme nazvať „zúžením“ σ-algebry $\mathcal{S}$ na množinu $B \in \mathcal{S}$. Takúto zúženú σ-algebru používame ako model pre náhodné udalosti v situáciách, v ktorých vieme, že elementárny výsledok môže byť len z množiny B. Zúžený priestor udalostí úzko súvisí so zmenou pravdepodobnostného modelu po „podmienení“ udalosťou B, s čím sa viac oboznámime v časti 2.1. Lemu 1.22 však používame aj v nasledujúcich príkladoch, ktoré sú primárne určené na ilustrovanie situácií, v ktorých používame na reprezentovanie náhodných udalostí borelovské množiny.

Príklad 1.24. Sledujeme, o koľko sekúnd zaznamenáme novú požiadavku na určitý systém hromadnej obsluhy, napríklad na server. Ak predpokladáme, že čas vieme odmerať úplne presne, potom je možným výberovým priestorom $[0, \infty)$ a príslušným systémom udalostí je $\mathcal{B}_{[0,\infty)}$, čiže borelovské podmnožiny intervalu $[0, \infty)$.[11] Udalosť $[0, 60)$ zodpovedá výroku „novú požiadavku zaznamenáme skôr ako uplynie minúta“, udalosť $[3600, \infty)$ zodpovedá výroku „uplynie aspoň hodina, kým zaznamenáme novú požiadavku“ a tak ďalej.

Príklad 1.25. Na jednotkovej kružnici v rovine náhodne zvolíme dva body K, L. Body K, L je možné jednoznačne reprezentovať veľkosťami uhlov $\angle MOK$ a $\angle MOL$ v intervale $[0, 2\pi)$, meranými proti smeru hodinových ručičiek, kde $M = (1, 0)$ a $O = (0, 0)$. Takže za formálny

[10] Existujú rôzne monštrózne neborelovské množiny, ktoré by nám spôsobovali problémy. Záujemca sa môže zoznámiť s príbuznou problematikou takzvaného Banachovho-Tarského paradoxu.

[11] Opäť by nás mohlo zvádzať vziať za systém udalostí všetky podmnožiny výberového priestoru $[0, \infty)$. Ukazuje sa však, že, podobne ako v príklade 1.17, by sme sa dostali do problémov pri definícii pravdepodobnostnej miery na všetkých prvkoch takto bohatého systému množín.



Obr. 1.2: Obrázok k príkladu 1.25. Panel (a): sivá oblasť znázorňuje množinovú reprezentáciu udalosti A. Panel (b): sivá oblasť znázorňuje množinovú reprezentáciu udalosti C.

výberový priestor môžeme zvoliť $[0, 2\pi) \times [0, 2\pi)$ a za σ-algebru udalostí systém $\mathcal{B}_2$ zúžený na množinu $[0, 2\pi) \times [0, 2\pi)$. Slovne vyjadrenej udalosti „Uhol $\angle KML$ je menší ako $\pi/2$" zodpovedá množina $A = \{(\psi, \phi) : 0 \leq \psi, \phi < 2\pi, |\psi - \phi| < \pi\}$, čo plynie z vety o vzťahu medzi obvodovým a stredovým uhlom. Udalosti „Trojuholník KML je ostrouhlý" zodpovedá množina $C = \{(\psi, \phi) : 0 \leq \psi, \phi < 2\pi, |\psi - \phi| < \pi, \min(\psi, \phi) < \pi, \max(\psi, \phi) > \pi\}$, čo sa dá odvodiť zo základných vlastností uhlov trojuholníka. Množiny A a C sú na obrázku 1.2.

Poznámka 1.26. Niekedy potrebujeme modelovať výskyt, generovanie či pozorovanie aj iných náhodných objektov než prvkov konečnej množiny (ako v príklade 1.10), postupností prvkov konečnej množiny (analogicky k príkladu 1.17), reálnych čísel ($m = 1$ z pohľadu definície 1.19, príklad 1.24) a náhodných konečno-rozmerných vektorov reálnych čísel ($m \geq 2$ z pohľadu definície 1.19, príklad 1.25). Napríklad by sme mohli pozorovať vývoj nejakej náhodnej kvantity v „spojitom čase" a v takej situácii by našimi elementárnymi výsledkami mohli byť celé funkcie definované na intervale $[0, \infty)$.[12] Pre potreby základnej štatistiky a analýzy dát si však obvykle vystačíme s modelmi podobnými tým, ktoré sú spomenuté vo vyššie uvedených príkladoch.

1.2 Pravdepodobnostná miera

Samotná σ-algebra udalostí je len, voľne vyjadrené, vymedzením toho, čo sa môže *v princípe* udiať. Na presnú špecifikáciu modelu potrebujeme ešte dodať jednotlivým potenciálnym udalostiam rôznu „mieru očakávania", resp. „asymptotickú frekvenciu výskytu", takzvanú

[12] Štúdiom takýchto náhodných objektov sa zaoberá „teória náhodných procesov" s aplikáciami napríklad vo fyzike a finančnej matematike.

„pravdepodobnostnú mieru“.[13]

Definícia 1.27. Pravdepodobnostná miera na σ-algebre $\mathcal{S}$ podmnožín množiny Ω, skrátene **pravdepodobnosť**, je zobrazenie $P : \mathcal{S} \to \mathbb{R}$ s nasledovnými vlastnosťami: 1) Pre všetky $A \in \mathcal{S}$ platí $0 \leq P(A) \leq 1$; 2) $P(\Omega) = 1$, $P(\emptyset) = 0$; 3) Ak $(A_i)_{i \in I}$ je postupnosť disjunktných udalostí, tak $P\left(\cup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)$. Vlastnosť 3) nazývame **aditivita** pravdepodobnosti, ak je I konečná množina a **σ-aditivita** pravdepodobnosti, ak je I nekonečná spočítateľná množina.

Poznámka 1.28. Vlastnosti z definície 1.27 nazývame „axiómy pravdepodobnosti“. Prvá axióma je dohodou, že pravdepodobnosť budeme merať číslami v rozsahu od 0 do 1. Druhá axióma požaduje, aby udalosť $\emptyset$ (nazývaná „nemožná udalosť“) mala pravdepodobnosť 0 a udalosť Ω (nazývaná „istá udalosť“) mala pravdepodobnosť 1.[14] Vlastnosť 3) je už netriviálna. Formalizuje naše intuitívne očakávanie, že ak máme udalosti, z ktorých môže nastať maximálne jedna, tak to, že nastane niektorá z týchto udalostí (v zmysle „akákoľvek“), má pravdepodobnosť rovnú súčtu pravdepodobností jednotlivých udalostí.[15]

Definícia 1.29. Nech $\mathcal{S}$ je σ-algebra udalostí na výberovom priestore Ω a nech P je pravdepodobnostná miera na $\mathcal{S}$. Potom trojicu $(\Omega, \mathcal{S}, P)$ nazývame **pravdepodobnostný priestor**.[16]

Príklad 1.30. Nech $\Omega = \{1, \ldots, 6\}$ a $\mathcal{S} = 2^{\Omega}$. Nech $P : \mathcal{S} \to \mathbb{R}$ je zobrazenie definované nasledovne: Pre každé $A \in \mathcal{S}$ platí $P(A) = |A|/6$, kde $|A|$ znamená počet prvkov množiny A. Potom $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor. Tento pravdepodobnostný priestor môžeme považovať za model hádzania vyváženou hracou kockou, ak nás zaujíma výlučne číselný výsledok hodu. (Pozri príklad 1.10.)

Poznámka 1.31. Dôležitou otázkou, ktorú nie je možné vyriešiť mechanickým, vždy použiteľným návodom, je *ako* určíme pre náš pravdepodobnostný model mieru P. V niektorých situáciách nás k vhodnej voľbe pravdepodobnosti P dovedú úvahy vychádzajúce zo symetrií modelovaného deja. Napríklad v príklade 1.30 je prirodzené predpokladať, že každý z výsledkov $1, \ldots, 6$ (čiže každá z udalostí $\{1\}, \ldots, \{6\}$) má rovnakú pravdepodobnosť p, z dôvodu symetrie a vyváženosti samotnej kocky, ktorú hádžeme. Aditivita pravdepodobnosti potom implikuje, že $p = |A|/6$ pre každé $A \in \mathcal{S}$. Pochopiteľne, v prípade falošnej kocky, by tento pravdepodobnostný model nebol vhodný. V takom prípade môže byť určenie dokonalého modelu, čiže presnej pravdepodobnosti P, nerealistické. Jedna z možností stanovenia približného modelu je hodiť danou kockou mnohokrát a pravdepodobnosti udalostí $\{1\}, \ldots, \{6\}$ určiť „empiricky“ ako podiely jednotlivých pozorovaných výsledkov.

[13] Zvrat „miera očakávania“ má blízko k takzvanej „bayesovskej“ interpretácii pravdepodobnosti a zvrat „asymptotická miera výskytu“ k takzvanej „frekventistickej“ interpretácii pravdepodobnosti. V tomto texte sa však témou interpretácie pojmu pravdepodobnosť nebudeme zaoberať; vystačíme si s intuitívnou predstavou a matematickým aparátom, ktorý je rovnaký pre ktorúkoľvek z uvedených dvoch interpretácií.

[14] V mnohých pravdepodobnostných modeloch však existujú aj neprázdne udalosti, ktoré majú nulovú pravdepodobnosť a udalosti, ktoré sú vlastnou podmnožinou množiny Ω, no ich pravdepodobnosť je 1. Povedzme v modeli z príkladu 1.33 s $p_i = 1/2$ (pre každé i) môžeme uvažovať udalosť $A = \{(1, 1, 1, \ldots)\}$, teda že nám v sérii hodov mincou budú donekonečna padať znaky. Udalosť A je očividne neprázdna, lebo obsahuje jeden elementárny výsledok, jej pravdepodobnosť je však 0. Udalosti nulovej pravdepodobnosti niekedy nazývame „nulové udalosti“ a udalosti pravdepodobnosti 1 nazývame „skoro isté udalosti“.

[15] Všimnime si, že niečo podobné požadujeme aj od pojmov „úhrnná dĺžka“, „celková plocha“ a „celkový objem“. Ako sme už naznačili, pojmy pravdepodobnosti, dĺžky, plochy a objemu sa dajú zastrešiť spoločnou matematickou teóriou, takzvanou „teóriou miery“.

[16] Niekedy sa pravdepodobnostný priestor nazýva aj „pravdepodobnostný model“.

Príklad 1.32 (Jednoduchý náhodný výber bez návratu)**.** Ako definujeme pravdepodobnostnú mieru pre situáciu z príkladu 1.12? Ak predpokladáme náhodný výber objektov, ktorý „neuprednostňuje“ žiadny objekt pred žiadnym iným objektom, tak je intuitívne zrejmé, že každá n-prvková podmnožina by mala mať rovnakú pravdepodobnosť, že bude výsledkom tohto výberu. Podobne ako v predchádzajúcom prípade je teda prirodzené predpokladať, že pre každú udalosť $\{\omega\} \in \mathcal{S}$, kde ω je n-prvková podmnožina množiny $\{1, 2, \ldots, M\}$, platí $P(\{\omega\}) = 1/|\Omega| = 1/\binom{M}{n}$. Preto ak A je akákoľvek udalosť, čiže akákoľvek množina n-prvkových podmnožín množiny $\{1, 2, \ldots, M\}$, tak $P(A) = |A|/\binom{M}{n}$. Najzložitejšie je v takýchto situáciách nie to, ako definovať pravdepodobnosť, ale ako spočítať počet prvkov $|A|$ (nejako, napríklad slovne) zadanej množiny A. To je obvykle kombinatorický problém a úlohy výpočtu pravdepodobností, ktoré sú ekvivalentné výpočtu počtu prvkov nejakej množiny, sa nazývajú problémy „kombinatorickej pravdepodobnosti“. Napríklad uvažujme množinu A z príkladu 1.12. Jej počet prvkov je rovný počtu n-prvkových podmnožín $(M-1)$-prvkovej množiny, čiže $|A| = \binom{M-1}{n}$, a teda $P(A) = \binom{M-1}{n}/\binom{M}{n} = 1 - \frac{n}{M}$. Pre udalosť B z príkladu 1.12 platí, že jej počet prvkov je počet všetkých k-prvkových podmnožín množiny $\{1, 2, \ldots, N\}$ krát počet všetkých $n-k$ prvkových podmnožín množiny $\{N+1, N+2, \ldots, M\}$. Takže pre netriviálny prípad $n + N - M \leq k \leq \min(n, N)$ platí $P(B) = \binom{N}{k}\binom{M-N}{n-k}/\binom{M}{n}$.[17] Rôzne schémy náhodného výberu z veľkej populácie objektov študuje takzvaná „teória náhodného výberu“.

Príklad 1.33 (Bernoulliho pokusy)**.** Nech $\Omega = \{0, 1\}^\infty$, t.j. Ω je množina všetkých (nekonečných spočítateľných) postupností čísel 0 a 1. Nech $p_1, p_2, \ldots$ je postupnosť čísel v intervale $(0, 1)$. Uvažujme σ-algebru $\mathcal{S}$ definovanú v príklade 1.17. Možno ukázať,[18] že na $\mathcal{S}$ existuje pravdepodobnosť P, ktorá spĺňa nasledovnú podmienku: Nech $b_1, \ldots, b_m \in \{0, 1\}$ sú pevné binárne čísla a nech $1 \leq k_1 < \cdots < k_m$ sú pevné prirodzené čísla. Nech $A \in \mathcal{S}$ je množina všetkých tých postupností hodnôt $\{0, 1\}$, ktoré majú na pozícii k_j hodnotu b_j pre všetky $j = 1, \ldots, m$ a na ostatných pozíciách akúkoľvek binárnu hodnotu. Pre takéto A platí $P(A) = r_1 r_2 \ldots r_m$, kde $r_j = 1 - p_{k_j}$, ak $b_j = 0$ a $r_j = p_{k_j}$, ak $b_j = 1$, pre všetky $j = 1, \ldots, m$. Pravdepodobnostný priestor $(\Omega, \mathcal{S}, P)$ môžeme považovať za model opakovaného hádzania mincou, pričom „hlava“ (formálne 0) padne v i-tom hode s pravdepodobnosťou $1 - p_i$, „znak“ (formálne 1) padne v i-tom hode s pravdepodobnosťou p_i a výsledky hodov sú navzájom „nezávislé“. Množiny A opísané vyššie zodpovedajú udalosti, že v hodoch s poradovými číslami $k_1, \ldots, k_m$ dostaneme výsledky $b_1, \ldots, b_m$ (a na tom, čo dostaneme v hodoch s inými poradovými číslami, nezáleží). Špeciálne, ak $p_i = p$ pre všetky i, ide o model nekonečného hádzania rovnakou mincou, takzvané „homogénne“ Bernoulliho pokusy. V tomto prípade $P(A) = p^s(1-p)^{m-s}$, kde s je počet jednotiek v postupnosti $b_1, \ldots, b_m$. Ak $p_j \neq p_i$ pre nejaké indexy i, j, hovoríme o „nehomogénnych“ Bernoulliho pokusoch.

Príklad 1.34 (Existencia rovnomernej pravdepodobnosti)**.** Uvažujme kváder $Q = [l_1, u_1] \times \cdots \times [l_m, u_m]$, $l_i < u_i, i = 1, \ldots, m$. Možno ukázať,[19] že na $\mathcal{B}_m$ (pozri definíciu 1.19) existuje

[17] Túto situáciu modeluje „hypergeometrické rozdelenie pravdepodobnosti“, s ktorým sa zoznámime v časti 4.2.6.

[18] Dôkaz spadá do oblasti teórie náhodných procesov a presahuje to, čo je rozumné uvádzať v týchto skriptách. Záujemca si môže vyhľadať pojem „Kolmogorovova veta o rozšírení“.

[19] Dôkaz spadá do oblasti teórie miery a presahuje to, čo je rozumné uvádzať v týchto skriptách; záujemca si môže vyhľadať pojem „Lebesgueova miera“.

pravdepodobnosť P_m, ktorá spĺňa nasledovnú podmienku: Pre každý kváder $A = [a_1, b_1] \times \cdots \times [a_m, b_m] \subseteq Q$, $a_i < b_i, i = 1, \dots, m$, platí

$$P_m(A) = \frac{\prod_{i=1}^m (b_i - a_i)}{\prod_{i=1}^m (u_i - l_i)}.$$

Pravdepodobnostný priestor $(\mathbb{R}^m, \mathcal{B}_m, P_m)$ je model rovnomerného náhodného výberu bodu (alebo m-rozmerného vektora) v kvádri Q. Špeciálne, pre $m = 1$ ide o model rovnomerného náhodného výberu reálneho čísla na úsečke $[l_1, u_1]$ a pre $m = 2$ o model rovnomerného náhodného výberu bodu (alebo dvojrozmerného vektora) na obdĺžniku $[l_1, u_1] \times [l_2, u_2]$. Dokonca ak by Q bola akákoľvek m-rozmerná borelovská množina, existuje na $\mathcal{B}_m$ pravdepodobnosť P, že pre každé $A \in \mathcal{B}_m$ platí

$$P(A) = \frac{\lambda(A \cap Q)}{\lambda(Q)},$$

kde $\lambda(B)$ označuje m-rozmerný objem (celkovú dĺžku pre $m = 1$ alebo plochu pre $m = 2$) množiny B. Pre jednoduché množiny B zodpovedá $\lambda(B)$ našim intuitívnym predstavám a znalostiam zo základných kurzov matematiky. Presná definícia hodnoty $\lambda(B)$, ktorá zahŕňa všetky borelovské množiny B, je veľmi zložitá.

Príklad 1.35 (Existencia gaussovskej miery na $\mathbb{R}$)**.** Nech $\mathcal{S} = \mathcal{B}$, $\mu \in \mathbb{R}$ a $\sigma > 0$. Možno ukázať,[20] že na σ-algebre $(\mathbb{R}, \mathcal{S})$ existuje pravdepodobnosť P, ktorá spĺňa nasledovnú podmienku:

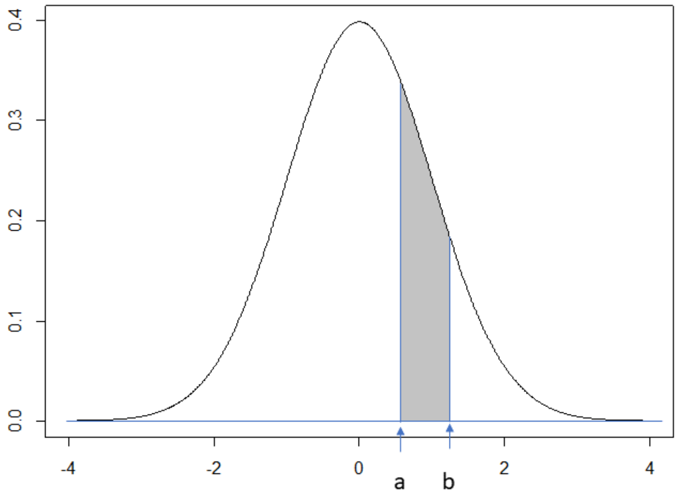
$$P([a, b]) = \int_a^b \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \mathrm{d}x. \tag{1.1}$$

pre každé $-\infty < a < b < \infty$. Pozri obrázok 1.3. Pravdepodobnosť P je takzvaná „gaussovská“ pravdepodobnosť. Viac sa s ňou zoznámime v časti 5.2.4 o „normálnom“ (resp. gaussovskom) rozdelení pravdepodobnosti. Okrem mnohých iných modelov, normálne rozdelenie pravdepodobnosti (pre $\mu = 100$ a $\sigma = 15$) približne zodpovedá výsledku IQ testu náhodne vybratého človeka.[21] Ako vidíme na tomto príklade, pravdepodobnostné modely sú niekedy len výsledkom zjednodušujúcich úvah, empirických pozorovaní a štandardizácií; ich platnosť nie je dokonalá. Aj nedokonalé pravdepodobnostné modely však môžu poskytovať užitočné podklady pre rozhodovanie.

Príklad 1.36. Stanovme pravdepodobnostnú mieru pre príklad 1.25 a vypočítajme pravdepodobnosti udalostí A a B. Keďže body K, L volíme „rovnomerne a nezávisle“, je prirodzené, že uhly ϕ a ψ by sa mali správať tak, že bod (ψ, ϕ) bude vygenerovaný „rovnomerne náhodne“ v štvorci $Q = [0, 2\pi) \times [0, 2\pi)$. Na základe príkladu 1.34 je preto $P(A)$ pomer plochy $\lambda(A \cap Q) = \lambda(A)$ k ploche $\lambda(Q)$. Z obrázku 1.2 teda vidíme, že $P(A) = 3/4$. Podobne $P(B) = \lambda(B \cap Q)/\lambda(Q) = \lambda(B)/\lambda(Q) = 1/4$. Úlohy takéhoto typu a prístup výpočtu pravdepodobnosti pomocou geometrických úvah, rovnomerného rozdelenia a výpočtu veľkostí plôch, nazývame „geometrická pravdepodobnosť“.

[20] Dôkaz je jednoduchý, pokiaľ uveríme tvrdeniu z predošlého príkladu. Stručne: stačí overiť, že P môžeme definovať ako $P(B) = P^R(\Phi^{-1}(B))$, kde P^R je rovnomerná pravdepodobnosť na intervale $[0, 1]$, ktorej existenciu predpokladáme, a Φ^{-1} je inverzná funkcia k takzvanej „distribučnej funkcii“ normálneho rozdelenia $N(\mu, \sigma^2)$.

[21] Jednoduché výpočty týkajúce sa IQ demonštrujeme v príklade 5.48.



Obr. 1.3: Ilustračný obrázok k príkladu 1.35. Na všetkých borelovských podmnožinách množiny $\mathbb{R}$ vieme definovať zobrazenie P tak, že spĺňa axiómy z definície 1.27 a jeho hodnoty na špeciálnych borelovských množinách – intervaloch – sú určené rovnosťou (1.1). Krivka na obrázku zodpovedá funkcii $x \to e^{-x^2/2}/\sqrt{2\pi}$, čo je podintegrálna funkcia v (1.1) pre $\mu = 0$ a $\sigma = 1$. Veľkosť plochy vyznačenej sivou farbou je rovná hodnote $P([a, b])$.

Príklad 1.37 (Bertrandov paradox)**.** Nech B je jednotkový kruh v rovine. Náhodne zvolíme tetivu T kruhu B. Aká je pravdepodobnosť, že dĺžka tejto tetivy bude väčšia ako $\sqrt{3}$? (Všimnime si, že $\sqrt{3}$ je dĺžka strany rovnostranného trojuholníka vpísaného do B.)

Riešenie: Problémom tohto zadania je nejednoznačnosť vety „Náhodne zvolíme tetivu T kruhu B." Ukážeme si, že rôzne matematicky poctivé spresnenia tejto vety vedú k rôznym výsledkom. Tento príklad slúži predovšetkým na zdôraznenie, že pri formulácii pravdepodobnostných problémov musíme byť opatrní, najmä pre situácie s nekonečným výberovým priestorom. Voľne vyjadrené, nie je „náhodne" ako „náhodne".

Spresnenie číslo 1: Na obdĺžniku $Q = [0, 2\pi) \times [0, 1)$ zvolíme rovnomerne náhodne bod a jeho súradnice si označíme (α, L). Tetivu T skonštruujeme tak, aby bola kolmá na úsečku spájajúcu body $(0, 0)$ a $(\cos(\alpha), \sin(\alpha))$, a aby vzdialenosť stredu tetivy T od bodu $(0, 0)$ bola L. Udalosť, že dĺžka tetivy T bude väčšia ako $\sqrt{3}$ je rovná udalosti, že $L < 1/2$, čo je formálne udalosť $A = [0, 2\pi) \times [0, 1/2)$. Zjavne $P(A) = \lambda(A)/\lambda(Q) = 1/2$.[22]

Spresnenie číslo 2: Na trojuholníku $Q = \{(\bar{\alpha}, \bar{\beta}) \in \mathbb{R}^2 : 0 \leq \bar{\alpha} < \bar{\beta} < 2\pi\}$ zvolíme rovnomerne náhodne bod so súradnicami (α, β). Tetivu T skonštruujeme tak, aby jej koncové body boli $(\cos(\alpha), \sin(\alpha))$ a $(\cos(\beta), \sin(\beta))$. Udalosť, že dĺžka tetivy T bude väčšia ako $\sqrt{3}$, je pri tomto spresnení zadania formálne rovná udalosti $A = \{(\alpha, \beta) \in Q : 2\pi/3 < \beta - \alpha < 4\pi/3\}$. V tomto prípade $P(A) = \lambda(A)/\lambda(Q) = 1/3$.

Spresnenie číslo 3: Na kruhu $Q = \{(x, y) \in \mathbb{R}^2 : \sqrt{x^2 + y^2} < 1\}$ zvolíme rovnomerne náhodne bod, ktorého súradnice si označíme (X, Y). Ak $(X, Y) \neq (0, 0)$, tak tetivu T skonštruujeme tak, aby bol bod (X, Y) stredom T a aby T bola kolmá na úsečku spájajúcu body $(0, 0)$ a (X, Y). Ak $(X, Y) = (0, 0)$, zvolíme za T úsečku spájajúcu body $(-1, 0)$ a $(1, 0)$. Udalosť,

[22] Všimnime si, že opísané zobrazenie z Q do množiny tetív nie je bijektívne, no napriek tomu je náhodný mechanizmus výberu tetivy definovaný jednoznačne. To isté platí aj o „Spresnení číslo 3".

že dĺžka tetivy T bude väčšia ako $\sqrt{3}$ je rovná udalosti, že vzdialenosť bodu (X, Y) od stredu kruhu Q bude menej než $1/2$, čo je formálne udalosť $A = \{(x, y) \in Q : \sqrt{x^2 + y^2} < 1/2\}$. V tomto prípade $P(A) = \lambda(A)/\lambda(Q) = 1/4$.

Pokúste sa vymyslieť ešte jeden „náhodný“ spôsob konštrukcie tetivy T a k nemu prislúchajúce formálne spresnenie. Aká je pre tento spôsob pravdepodobnosť, že dĺžka tetivy T bude väčšia ako $\sqrt{3}$?

1.3 Vlastnosti pravdepodobnosti

Axiómy pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$ z definícií 1.1 a 1.27 sú zvolené unikátne vhodným spôsobom: implikujú konzistentný a bohatý systém vlastností, na ktorých sa dá vybudovať aplikovateľná teória.[23] Niektoré z týchto vlastností si preberieme v tejto časti.

Veta 1.38 (Rozdielovosť a monotónnosť)**.** Nech A,B sú udalosti pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$, pričom $A \subseteq B$. Potom $P(B \setminus A) = P(B) - P(A)$ a $P(A) \leq P(B)$.

Dôkaz. Keďže udalosti A a $B \setminus A$ sú disjunktné, máme $P(A) + P(B \setminus A) = P(A \cup (B \setminus A)) = P(B)$ (použili sme aditivitu pravdepodobnosti), čiže $P(B \setminus A) = P(B) - P(A)$. Nerovnosť $P(A) \leq P(B)$ plynie z predchádzajúcej rovnosti a vlastnosti $P(B \setminus A) \geq 0$. □

Veta 1.39 (Pravdepodobnosť doplnku)**.** Nech $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor a nech A je udalosť. Potom

$$P(\Omega \setminus A) = 1 - P(A).$$

Dôkaz. Veta je špeciálny prípad predchádzajúcej vety pre $B = \Omega$. □

Poznámka 1.40 (Šanca udalosti)**.** Nech A je udalosť, $P(A) \neq 1$. Potom číslo[24]

$$o(A) = \frac{P(A)}{P(\Omega \setminus A)} = \frac{P(A)}{1 - P(A)}$$

by sme mohli nazvať „šanca“ udalosti A. Napríklad šanca, že na minci padne znak, je „jedna k jednej“, čiže 1. Podobne, šanca, že na kocke padne šestka, je „jedna k piatim“, čiže 0.2. Ľahko sa presvedčíme, že nielen z $P(A)$ vieme odvodiť $o(A)$, ale aj naopak, z $o(A)$ priamo dostaneme $P(A) = o(A)/(1 + o(A))$. Moderná teória pravdepodobnosti využíva pojem „šanca udalosti“ len ojedinele a my ho ďalej nebudeme používať.

Veta 1.41 (Pravdepodobnosť zjednotenia dvoch udalostí)**.** Nech A, B sú udalosti pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$. Potom

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

Dôkaz. Zrejme udalosti A a $B \backslash A$ sú disjunktné a ich zjednotenie je $A \cup B$, takže $P(A \cup B) = P(A) + P(B \backslash A)$. Súčasne sú $A \cap B$ a $B \backslash A$ disjunktné, pričom ich zjednotenie je udalosť B, takže máme $P(B \backslash A) = P(B) - P(A \cap B)$. Spojením týchto dvoch rovností dostávame rovnosť zo znenia vety. □

[23] Systém axióm pravdepodobnosti na σ-algebre vyvinul fenomenálny ruský matematik A. N. Kolmogorov na začiatku 30. rokov 20. storočia. Ten istý matematik sa zaslúžil aj o rozvoj informatiky; fascinujúca je napríklad takzvaná „kolmogorovská“ zložitosť, ktorá sa študuje v teoretickej informatike.

[24] Symbol o pochádza z anglického slova „odds“.

Veta 1.42 (Princíp inklúzie-exklúzie (zapojenia-vypojenia))**.** Nech $A_1, A_2, \ldots, A_n$ sú udalosti pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$. Potom

$$P\left(\cup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} P(A_{i_1} \cap \ldots \cap A_{i_k}).$$

Dôkaz. Pre $n = 2$ je tvrdenie tejto vety také isté ako vo vete 1.41. Pre všeobecné n je možné použiť matematickú indukciu. Odlišnou metódou dokážeme princíp inklúzie-exklúzie pre všeobecné n v príklade 4.74. □

Príklad 1.43. Za okrúhlym stolom je rozostavených $n \geq 3$ stoličiek. Náhodne za tento stôl rozsadíme troch ľudí. Aká je pravdepodobnosť p_n, že niektorí dvaja ľudia budú sedieť vedľa seba?

Riešenie: Zrejme $p_3 = 1$; ďalej budeme predpokladať, že $n \geq 4$. Označme ľudí ako 1,2,3 a definujme udalosť A (B, C) ako udalosť, že budú vedľa seba sedieť človek 1 a 2 (2 a 3, resp. 1 a 3). Zaujíma nás $P(A \cup B \cup C)$. Podľa princípu zapojenia-vypojenia máme:

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

Zrejme však

$$P(A) = P(B) = P(C) = \frac{2}{n-1} \text{ a } P(A \cap B) = P(B \cap C) = P(A \cap C) = \frac{2}{(n-1)(n-2)}.$$

Pre $n \geq 4$ udalosť $A \cap B \cap C$ nemôže nastať, preto $P(A \cap B \cap C) = 0$. Takže

$$P(A \cup B \cup C) = 3 \cdot \frac{2}{n-1} - 3 \cdot \frac{2}{(n-1)(n-2)} = \frac{6(n-3)}{(n-1)(n-2)}.$$

Špeciálne, $p_4 = p_5 = 1$ a $p_6 = 9/10$.

Príklad možno riešiť aj tak, že si definujme m udalostí $B_1, B_2, \ldots, B_m$, že budú obsadené susedné stoličky $1, 2$; $2, 3$; $\ldots$; resp. $m, 1$. Hľadaná pravdepodobnosť je $P(\cup_i B_i)$, čo sa dá opäť vypočítať pomocou princípu zapojenia-vypojenia.

Poznámka 1.44. V predchádzajúcom príklade sme formálne nedefinovali model na úrovni komponentov Ω, $\mathcal{S}$ a P, hoci by to bolo možné. (Vedeli by ste to?) To však ani nebolo potrebné, pretože sme používali úvahy platné pri *akejkoľvek* zmysluplnej formalizácii. Takto budeme riešiť príklady často.

Príklad 1.45. Postupnosť čísel $(1, 2, \ldots, n)$ dokonale náhodne premiešame. (Čiže každá spomedzi $n!$ permutácií má rovnakú pravdepodobnosť.) Nájdime pravdepodobnosť p_n, že aspoň jedno z čísel $1, 2, \ldots, n$ bude po premiešaní na svojom pôvodnom mieste. Určme $\lim_{n\to\infty} p_n$.

Riešenie: Nech n je pevné. Nech A je udalosť, že aspoň jedno z čísel $1, 2, \ldots, n$ bude po premiešaní na svojom pôvodnom mieste. Potom zrejme $A = \cup_{i=1}^{n} A_i$, kde A_i označuje udalosť, že číslo i zostane na svojom pôvodnom mieste. Keďže možností výberu rôznych indexov $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ je $\binom{n}{k}$ a pre každý takýto výber je $P(A_{i_1} \cap \ldots \cap A_{i_k}) = \frac{(n-k)!}{n!}$, tak podľa vety 1.42 dostávame

$$p_n = P(A) = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \frac{(n-k)!}{n!} = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!}.$$

Zrejme tiež

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

čo plynie zo známeho vzorca $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ pre všetky reálne čísla x.

Poznámka 1.46. Nasledovné lemy sa používajú pri dokazovaní niektorých limitných pravdepodobnostných tvrdení. Najprominentnejšia aplikácia uvedených „polospojitostí“ je pri dôkaze limitných vlastnosti „distribučnej funkcie“, s ktorou sa zoznámime v časti 3.2.

Lema 1.47 (Spojitosť pravdepodobnosti zdola)**.** Nech $(A_i)_{i=1}^{\infty}$ je nekonečná postupnosť udalostí pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$, neklesajúca v zmysle $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$. Potom

$$P\left(\cup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P(A_i).$$

Dôkaz. Položme $A_0 = \emptyset$ a $B_i = A_i \setminus A_{i-1}$ pre každé $i \in \mathbb{N}$. Všimneme si, že $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} B_i$, ďalej že udalosti B_i sú disjunktné a $P(B_i) = P(A_i) - P(A_{i-1})$. Taktiež ľahko overíme, že $\sum_{i=1}^{n} (P(A_i) - P(A_{i-1})) = P(A_n)$. Postupne dostávame $P\left(\cup_{i=1}^{\infty} A_i\right) = P\left(\cup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} (P(A_i) - P(A_{i-1})) = \lim_{n\to\infty} P(A_n)$. □

Lema 1.48 (Spojitosť pravdepodobnosti zhora)**.** Nech $(A_i)_{i=1}^{\infty}$ je nekonečná postupnosť udalostí pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$, nerastúca v zmysle $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$. Potom

$$P\left(\cap_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P(A_i).$$

Dôkaz. Lemu je možné dokázať podobne ako predchádzajúcu lemu alebo *aplikáciou* predchádzajúcej lemy na neklesajúcu postupnosť udalostí $(\Omega \setminus A_i)_{i=1}^{\infty}$. □

Príklad 1.49. Hádžeme vyváženou mincou (nekonečne dlho; pozri 1.17, resp. „homogénne“ Bernoulliho pokusy s $p_i = 1/2$ pre všetky i v príklade 1.33). Aká je pravdepodobnosť, že nám aspoň v jednom hode padne znak? Aká je pravdepodobnosť, že nám vo všetkých hodoch padne znak?

Riešenie: Odpoveď je intuitívne jasná, ukážme si však formálne presné riešenie. Nech C je udalosť, že aspoň v jednom hode padne znak. Zrejme $C = \cup_{i=1}^{\infty} C_i$, kde C_i je udalosť, že aspoň v jednom z prvých i hodov padne znak. Nech by bola formalizácia pravdepodobnostného priestoru pre tento príklad akákoľvek, musíme mať $P(C_i) = 1 - 2^{-i}$. Tiež si všimnime, že $C_1 \subseteq C_2 \subseteq C_3 \subseteq \cdots$. Podľa lemy 1.47 o spojitosti pravdepodobnosti zdola dostávame $P(C) = P(\cup_{i=1}^{\infty} C_i) = \lim_{i\to\infty} P(C_i) = \lim_{i\to\infty}(1-2^{-i}) = 1$. Analogicky môžeme postupovať pre druhú časť príkladu. Nech D je udalosť, že vo všetkých hodoch padne znak. Platí $D = \cap_{i=1}^{\infty} D_i$, kde D_i je udalosť, že v každom z prvých i hodov padne znak. Zrejme $P(D_i) = 2^{-i}$ a $D_1 \supseteq D_2 \supseteq D_3 \supseteq \cdots$. Podľa spojitosti pravdepodobnosti zhora 1.48 máme $P(D) = P(\cap_{i=1}^{\infty} D_i) = \lim_{i\to\infty} P(D_i) = \lim_{i\to\infty} 2^{-i} = 0$.

Veta 1.50 (Booleova nerovnosť)**.** Ak $(A_i)_{i\in I}$ je postupnosť udalostí pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$, tak

$$P\left(\cup_{i\in I} A_i\right) \leq \sum_{i\in I} P(A_i).$$

Dôkaz. Dokážme vetu pre prípad $I = \{1, \dots, n\}$. Položme $B_1 = A_1$ a $B_n = A_n \setminus \cup_{i=1}^{n-1} A_i$ pre $n \geq 2$. Udalosti B_n sú očividne disjunktné. Súčasne pre každé n platí $\cup_{i=1}^{n} A_i = \cup_{i=1}^{n} B_i$, ale aj $P(B_n) \leq P(A_n)$, pretože $B_n \subseteq A_n$. Dostávame: $P(\cup_{i=1}^{n} A_i) = P(\cup_{i=1}^{n} B_i) = \sum_{i=1}^{n} P(B_i) \leq \sum_{i=1}^{n} P(A_i)$. Pre prípad nekonečnej (spočítateľnej) množiny I je možné vetu dokázať pomocou verzie tejto vety pre konečné I, s využitím lemy 1.47. □

Poznámka 1.51. Vlastnosť pravdepodobnosti z vety 1.50 nazývame „subaditivita", ak je I konečná množina, resp. „σ-subaditivita", ak je I nekonečná spočítateľná množina. Analogická vlastnosť pre „klasickú mieru" množín v $\mathbb{R}^m$ (čiže pre úhrnnú dĺžku, plochu, objem, $m = 1, 2, 3$) je veľmi prirodzená: miera zjednotenia množín je nanajvýš rovná súčtu mier týchto množín.

Príklad 1.52. Presne 30 percent (nie nutne súvislých) dĺžky jednotkovej kružnice v rovine je zafarbených na zeleno a zvyšných 70 percent je zafarbených na modro. Formálne presnejšie: máme danú funkciu[25] $f : S \to \{Z, M\}$, pričom

$$P\{x \in [0, 1) : f(\cos(2\pi x), \sin(2\pi x)) = Z\} = 0.3,$$
$$P\{x \in [0, 1) : f(\cos(2\pi x), \sin(2\pi x)) = M\} = 0.7,$$

kde P je rovnomerná pravdepodobnosť na intervale $[0, 1]$ z príkladu 1.34. Dokážeme, že do tejto kružnice je možné vpísať *rovnostranný* trojuholník tak, aby všetky jeho vrcholy ležali na modrej farbe.

Riešenie: Rovnomerne náhodne zvolíme bod $\mathbf{X}$ na jednotkovej kružnici S. Symbolom A označme udalosť, že $\mathbf{X}$ padne do zelenej časti kružnice, symbolom B označme udalosť, že $\mathbf{X}$ po otočení o uhol $2\pi/3$ v smere hodinových ručičiek padne do zelenej farby a symbolom C označme udalosť, že $\mathbf{X}$ po otočení o uhol $4\pi/3$ v smere hodinových ručičiek padne do zelenej farby. Zrejme $P(A) = P(B) = P(C) = 0.3$, teda z Booleovej nerovnosti máme $P(A \cup B \cup C) \leq P(A) + P(B) + P(C) = 0.9$. Takže pravdepodobnosť, že ani jeden z trojice vytvorených bodov nepadne do zelenej farby, je aspoň 0.1. To znamená, že nutne musí *existovať* aspoň jedna príslušným spôsobom vytvorená trojica bodov ($\mathbf{X}$, $\mathbf{X}$ pootočené o $2\pi/3$ a $\mathbf{X}$ pootočené o $4\pi/3$), čiže rovnostranný trojuholník, ktorého všetky tri vrcholy ležia na modrej časti kružnice.

[25] Striktne vzaté, mali by sme ešte predpokladať vlastnosti „merateľnosti" zelenej a modrej časti kružnice, resp. funkcie f. V tomto príklade nám však ide o intuitívne presvedčivú demonštráciu využitia Booleovej nerovnosti, na čo nepotrebujeme byť dokonale formálne presní.

Kapitola 2

Podmieňovanie a nezávislosť udalostí

2.1 Podmienená pravdepodobnosť

V tejto kapitole budeme riešiť otázku, ako modifikovať pravdepodobnosť udalosti A, resp. celý pravdepodobnostný priestor (reprezentujúci náš pravdepodobnostný model) po získaní informácie, že nastala udalosť B.

Definícia 2.1. Nech A a B sú udalosti, pričom $P(B) > 0$. Potom hodnotu $P(A|B) = P(A \cap B)/P(B)$ budeme nazývať **pravdepodobnosť udalosti** A **za podmienky** B.

Poznámka 2.2. V súvislosti so zmenou $P(A)$ na $P(A|B)$ pri získaní informácie, že nastala udalosť B, nazývame niekedy $P(A)$ „apriórnou pravdepodobnosťou“ udalosti A a $P(A|B)$ „aposteriórnou pravdepodobnosťou“ udalosti A.

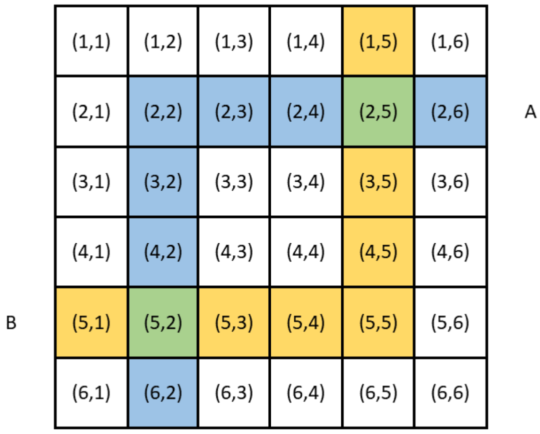
Definícia 2.3. Nech $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor a nech $B \in \mathcal{S}$, $P(B) > 0$. Definujme $P_B : \mathcal{S} \to \mathbb{R}$ nasledovne: $P_B(A) = P(A|B)$ pre všetky $A \in \mathcal{S}$. Zobrazenie P_B nazývame **podmienená pravdepodobnostná miera**[1] a pravdepodobnostný priestor $(\Omega, \mathcal{S}, P_B)$ nazývame **podmienený pravdepodobnostný priestor**.

Poznámka 2.4. Nahradenie $(\Omega, \mathcal{S}, P)$ priestorom $(\Omega, \mathcal{S}, P_B)$ predstavuje prirodzenú zmenu modelu po získaní (len) informácie obsiahnutej v tvrdení „nastane/nastala udalosť B“. Takýmto spôsobom si však ponechávame aj také elementárne výsledky, ktoré už určite nenastanú/nenastali. Model by sme ale mohli zmeniť aj tak, že elementárne výsledky mimo B z modelu vylúčime. Formálne by sme ako nový pravdepodobnostný priestor použili $(B, \mathcal{S}_B, P_B)$, kde $\mathcal{S}_B = \{C \in \mathcal{S} : C \subseteq B\}$[2] a $P_B(A) = P(A|B)$ pre všetky $A \in \mathcal{S}_B$.

Príklad 2.5. Vyššie uvedené definície ilustrujeme na príklade; pozri obrázok 2.1. Hádžeme naraz dvomi hracími kockami, pričom kocky považujeme za nezávislé a vyvážené, takže každá prípustná dvojica výsledkov padá s pravdepodobnosťou $1/36$. Model $(\Omega, \mathcal{S}, P)$ si môžeme formálne stanoviť nasledovne: $\Omega = \{(i,j) : i,j \in \{1,\dots,6\}\}$, $\mathcal{S} = 2^\Omega$ a $P(A) = |A|/36$ pre všetky $A \subseteq \Omega$. Ak označíme A udalosť, že menšie z padnutých čísel je 2 a B udalosť, že väčšie z padnutých čísel je 5, potom $P(A|B) = P(A \cap B)/P(B) = 2/9$. Podmienenú pravdepodobnostnú

[1] Dôkaz, že P_B je skutočne pravdepodobnostná miera, je jednoduchý; môžete si ho spraviť ako cvičenie. Niekedy používame pre pôvodnú pravdepodobnosť P zvrat „apriórna pravdepodobnostná miera“ a pre mieru P_B zvrat „aposteriórna pravdepodobnostná miera“.

[2] Pozrite lemu 1.22.



Obr. 2.1: Ilustrácia k príkladu 2.5. Modrá množina zodpovedá udalosti A, žltá udalosti B a zelená udalosti $A \cap B$.

mieru P_B dostaneme z pôvodnej pravdepodobnostnej miery P tak, že mimo B „skolabujeme“ P do nuly a na B „zdvihneme“ P „normalizačným“ faktorom $1/P(B)$.

Príklad 2.6. Máme dve nepriehľadné vrecká, pričom v jednom vrecku sú dve biele guľôčky a v druhom vrecku je jedna guľôčka biela a druhá čierna. Náhodne sme zvolili jedno vrecko (každé s pravdepodobnosťou 1/2) a z tohto vrecka sme náhodne vybrali jednu guľôčku, o ktorej sme sa presvedčili, že je biela. Aká je pravdepodobnosť, že aj druhá guľôčka vo vybratom vrecku je biela?

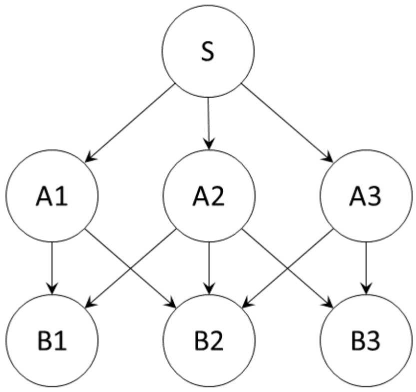
Riešenie: Takéto zadania chápeme nasledovne: Z hľadiska pred výberom vrecka a guľôčky, aká je pravdepodobnosť udalosti A = „vyberieme vrecko s dvomi bielymi guľôčkami“ za podmienky udalosti B = „vybratá guľôčka bude biela“? Alebo iná interpretácia zadania: Predpokladajme, že priestor $(\Omega, \mathcal{S}, P)$ zodpovedá pravdepodobnostnému modelu danej situácie z hľadiska pred začiatkom celého experimentu. Dodatočná informácia, že nastala udalosť B, mení náš pôvodný model na pravdepodobnostný priestor $(\Omega, \mathcal{S}, P_B)$ z definície 2.3 (alebo na priestor $(B, \mathcal{S}_B, P_B)$ z poznámky 2.4). Aká je pravdepodobnosť udalosti A (resp. udalosti $B \cap A$) v tomto novom priestore?

Takže riešenie (nezávisle od toho, ktorú interpretáciu úlohy prijmeme) je nasledovné: Máme $P(A|B) = P(A \cap B)/P(B)$. Zrejme však $P(A \cap B) = 1/2$ a $P(B) = 1/2.1/2 + 1/2 = 3/4$. Preto $P(A|B) = 2/3$.

Definícia 2.7. Nech $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor. Budeme hovoriť, že udalosti $A_1, \ldots, A_n$ tvoria **rozklad výberového priestoru** Ω, ak sú tieto udalosti disjunktné, každá z nich má nenulovú pravdepodobnosť a súčasne $\cup_{i=1}^n A_i = \Omega$.

Veta 2.8 (Veta o úplnej pravdepodobnosti)**.** Nech $A_1, \ldots, A_n$ tvoria rozklad výberového priestoru z definície 2.7. Nech B je akákoľvek udalosť. Potom

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



Obr. 2.2: Ilustračný obrázok k príkladu 2.10.

Dôkaz. $P(B) = P(B \cap (A_1 \cup \ldots \cup A_n)) = P((B \cap A_1) \cup \ldots \cup (B \cap A_n)) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i)P(A_i)$. Využili sme postupne fakt, že $\Omega = A_1 \cup \ldots \cup A_n$, základné vlastnosti množinových operácií, disjunktnosť množín $A_1, \ldots, A_n$ (a teda aj disjunktnosť množín $B \cap A_1, \ldots, B \cap A_n$) spolu s aditivitou pravdepodobnosti a definíciu podmienenej pravdepodobnosti. □

Veta 2.9 (Bayesov vzorec)**.** Nech udalosti $A_1, A_2, \ldots, A_n$ tvoria rozklad výberového priestoru. Nech B je akákoľvek udalosť nenulovej pravdepodobnosti a nech $k \in \{1, \ldots, n\}$. Potom

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^n P(B|A_i)P(A_i)}.$$

Dôkaz. Pre každé $k \in \{1, \ldots, n\}$ máme $P(B \cap A_k) = P(B|A_k)P(A_k)$, preto $P(A_k|B) = P(B \cap A_k)/P(B) = P(B|A_k)P(A_k)/P(B)$. Použitím vzorca pre úplnú pravdepodobnosť z vety 2.8 dostávame dokazovanú rovnosť. □

Príklad 2.10. Majme systém, ktorý sa môže nachádzať v stavoch $S, A1, A2, A3, B1, B2, B3$, pričom sú možné nasledovné prechody medzi stavmi: $S \to A1$, $S \to A2$, $S \to A3$, $A1 \to B1$, $A1 \to B2$, $A2 \to B1$, $A2 \to B2$, $A2 \to B3$, $A3 \to B2$, $A3 \to B3$ (pozri obrázok 2.2). Ak má systém viacero možností prechodu, tak prejde do každého nového prípustného stavu s rovnakou pravdepodobnosťou. Vieme len to, že na začiatku sa systém nachádzal v stave S a na konci sa nachádzal v stave $B2$. Určme pravdepodobnosť (v zmysle nášho subjektívneho hodnotenia), že systém prešiel stavom $A2$.

Riešenie: Pozrieme sa na situáciu z hľadiska pred experimentom (t.j. z pohľadu pozorovateľa, ktorý vidí, že systém je v počiatočnom stave S; tým si implicitne definujeme apriórny pravdepodobnostný priestor $(\Omega, \mathcal{S}, P)$). Označme ako A_1 (A_2, A_3,…,B_3) udalosť, že sa systém vyskytne v stave $A1$ ($A2$, $A3$, …, $B3$). Zaujíma nás podmienená pravdepodobnosť $P(A_2|B_2)$. Podľa Bayesovej vety je

$$P(A_2|B_2) = \frac{P(B_2|A_2)P(A_2)}{\sum_{i=1}^3 P(B_2|A_i)P(A_i)}.$$

Avšak $P(A_1) = P(A_2) = P(A_3)$ a $P(B_2|A_2) = 1/3$, $P(B_2|A_1) = P(B_2|A_3) = 1/2$. Dosadením dostávame riešenie $P(A_2|B_2) = 1/4$. Takže po po získaní informácie, že systém skončil v stave

$B2$, klesne pre udalosť A_2 naša apriórna pravdepodobnosť 1/3 na aposteriórnu pravdepodobnosť 1/4.

Príklad 2.11. Na vstupe do laboratória sú 2 nezávisle pracujúce biometrické autorizačné zariadenia (typ C1: snímanie prstu a typ C2: snímanie očnej dúhovky.) Typ C1 vykazuje 0.6 % mylných odmietnutí a 0.1 % mylných prijatí.[3] Typ C2 vykazuje 2 % mylných odmietnutí a 0.01 % mylných prijatí. Systém nám ohlási pokus o neautorizovaný vstup, ak aspoň jeden typ autorizácie ohlási odmietnutie. (Systém ale neoznámi, ktorý typ testu zlyhal). Určme pravdepodobnosť, že ak sa o vstup pokúsi neautorizovaná osoba, tak systém zareaguje správne (čiže ohlási pokus o neautorizovaný vstup). Z dlhodobých skúseností vieme, že v 99.9 percentách prípadov sa o vstup pokúša osoba, ktorá má autorizáciu na vstup a len v 0.1 percentách prípadov ide o pokus o neautorizovaný vstup. Aká je pravdepodobnosť, že naozaj ide o neautorizovaný vstup, ak vieme (len to), že systém *ohlásil* pokus o neautorizovaný vstup?

Riešenie: Z pohľadu apriórneho modelu si ako N označme udalosť, že dôjde k neautorizovanému pokusu o vstup a ako H udalosť, že systém *ohlási* neautorizovaný vstup. Označme si tiež ako H_1 a H_2 udalosti, že zariadenie C1, resp. zariadenie C2 ohlási pokus o neautorizovaný vstup. Tiež si uvedomme, že pravdepodobnosti 0.6 %, 0.1 %, 2 %, 0.01 %, 99.9 % a 0.1 % v zadaní sú postupne $P(H_1|N^c)$, $P(H_1^c|N)$, $P(H_2|N^c)$, $P(H_2^c|N)$, $P(N^c)$ a $P(N)$.[4]

V prvej otázke nás zaujíma $P(H|N)$. Keďže $H = H_1 \cup H_2$ a zariadenia sa správajú navzájom nezávisle, máme $P(H|N) = P(H_1 \cup H_2|N) = 1 - P(H_1^c \cap H_2^c|N) = 1 - P(H_1^c|N)P(H_2^c|N) = 1 - 10^{-3} \cdot 10^{-4} = 1 - 10^{-7}$.

V druhej otázke potrebujeme vypočítať $P(N|H)$. Podľa Bayesovho vzorca máme

$$P(N|H) = \frac{P(H|N)P(N)}{P(H|N)P(N) + P(H|N^c)P(N^c)}.$$

V tomto vzorci musíme dopočítať už len $P(H|N^c)$, čo je podobné ako $P(H|N)$. Máme $P(H|N^c) = P(H_1 \cup H_2|N^c) = P(H_1|N^c) + P(H_2|N^c) - P(H_1|N^c)P(H_2|N^c) = 0.006 + 0.02 - 0.006 \cdot 0.02 = 0.02588$.

Celkovo teda máme

$$P(N|H) = \frac{(1 - 10^{-7})0.001}{(1 - 10^{-7})0.001 + 0.02588 \cdot 0.999} \approx 0.037.$$

Takýto systém by asi nebol praktický; mali by sme príliš veľké percento falošných poplachov.

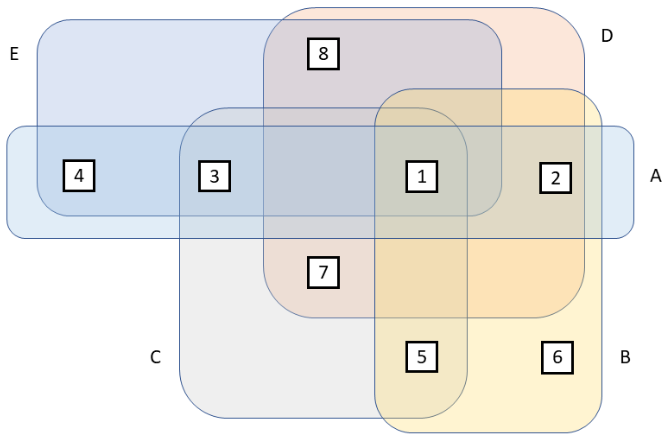
Uvedomme si, že analogické výpočty aposteriórnej pravdepodobnosti je možné vykonať pre akékoľvek nesekvenčné aj sekvenčné testovanie, v ktorom je výsledok testu jedna z možností „výsledok je negatívny“ a „výsledok je pozitívny“. Takto je možné skonštruovať celý „rozhodovací strom“.

2.2 Nezávislosť a závislosť udalostí

Pojem nezávislosti, resp. závislosti náhodných udalostí je jedným zo základných konštruktov, o ktoré sa opiera teória a aplikácie pravdepodobnosti. Nezávislosť udalostí A a B je vyjadrením toho, že z pravdepodobnostného hľadiska A „nenesie informáciu“ o B (a B nenesie informáciu o A).

[3] Tieto dva typy omylov sa niekedy nazývajú falošne negatívny výsledok, resp. falošne pozitívny výsledok.

[4] Kvôli stručnosti označujeme v tomto príklade doplnkovú udalosť symbolom c, čiže napríklad N^c je označenie udalosti $\Omega \setminus N$.



Obr. 2.3: Ilustračný obrázok k príkladu 2.15.

Definícia 2.12. Nech pre udalosti A a B platí $P(A \cap B) = P(A)P(B)$. Potom hovoríme, že **udalosti A a B sú nezávislé**. Ak $P(A \cap B) \neq P(A)P(B)$, tak hovoríme, že **udalosti A a B sú závislé**.

Poznámka 2.13. Všimnime si, že pre nezávislé udalosti A a B ($P(B) > 0$) platí $P(A|B) = P(A)$, čiže, voľne vyjadrené, znalosť toho, že nastala udalosť B, neovplyvní našu mieru očakávania, že nastane (alebo subjektívnu mieru očakávania, že nasta*la*) aj udalosť A. Naopak, ak $P(A|B) = P(A)$ (a $P(B) > 0$), tak sú udalosti A a B nezávislé. Inými slovami: ak vieme, že znalosť výsledku udalosti B nijak nemení naše pravdepodobnostné očakávanie, že nastane A, potom sú udalosti A a B nezávislé.

Definícia 2.14. Nech $A_1, \ldots, A_n$ sú udalosti, pričom pre každú množinu indexov $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ platí

$$P\left(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}\right) = P(A_{i_1})P(A_{i_2}) \ldots P(A_{i_k}).$$

Potom hovoríme, že **udalosti $A_1, \ldots, A_n$ sú združene nezávislé**. V opačnom prípade hovoríme, že **udalosti $A_1, \ldots, A_n$ sú združene závislé**.

Príklad 2.15. Majme priestor $(\Omega, \mathcal{S}, P)$, kde $\Omega = \{1, 2, \ldots, 8\}$, $\mathcal{S} = 2^{\Omega}$, $P(A) = |A|/8$ pre každé $A \subseteq \Omega$. (Tento model zodpovedá rovnomernému náhodnému výberu jedného z čísel $1, 2, \ldots, 8$, alebo modelu hádzania tromi vyváženými mincami, ak každú z ôsmich rôznych kombinácií výsledkov na jednotlivých minciach označíme jedným z čísel $1, 2, \ldots, 8$.)

Definujme nasledovné udalosti: $A = \{1, 2, 3, 4\}$, $B = \{1, 2, 5, 6\}$, $C = \{1, 3, 5, 7\}$, $D = \{1, 2, 7, 8\}$ a $E = \{1, 3, 4, 8\}$ (pozri obrázok 2.3). Ľahko sa presvedčíme, že a) Udalosti A, B, C sú združene nezávislé. b) Udalosti A, B, D sú po dvojiciach nezávislé, ale združene nezávislé nie sú, pretože $P(A \cap B \cap D) \neq P(A)P(B)P(D)$. c) Udalosti A, B, E nie sú všetky po dvojiciach nezávislé, napriek tomu, že $P(A \cap B \cap E) = P(A)P(B)P(E)$.

Príklad 2.16 (Voľne prevzaté z [6])**.** Napísali sme (klasický, „deterministický“) algoritmus na overovanie nasledovnej rovnosti dvoch polynómov:

$$\prod_{i=1}^{d} (a_i x - b_i) = c_d x^d + \ldots + c_1 x + c_0. \tag{2.1}$$

Vstupom algoritmu sú celé čísla $a_1, \ldots, a_d, b_1, \ldots, b_d, c_0, \ldots, c_d$ a výstupom je logická hodnota `true`/`false` podľa toho, či (2.1) platí (pre všetky $x \in \mathbb{R}$) alebo nie. Tento algoritmus overí, či platí

$$\begin{aligned} a_1 a_2 \cdots a_{d-1} a_d &= c_d, \\ b_1 a_2 \cdots a_{d-1} a_d + \cdots + a_1 a_2 \cdots a_{d-1} b_d &= -c_{d-1}, \\ &\cdots \\ b_1 b_2 \cdots b_{d-1} b_d &= (-1)^d c_0. \end{aligned}$$

(Predpokladáme, že s celými číslami vieme počítať presne.) Všimnime si, že priamočiare overenie týchto rovností, tak ako sú vyjadrené vyššie, si vyžaduje až $(d-1)2^d$ násobení.[5] Pre väčšie d, povedzme $d = 1000$, by bol tento počet násobení v rozumnom čase nevykonateľný.

Efektívnejší postup by bol ten, že by sme vyčíslili hodnotu polynómu $R - Q$ v $d + 1$ rôznych celých číslach, kde Q je polynóm na ľavej strane rovnosti (2.1) a R je polynóm na pravej strane rovnosti (2.1). Ak by boli niektoré výsledky tohto vyčíslenia nenulové, tak s istotou $R \neq Q$. Súčasne platí, že ak by všetky tieto výsledky boli nulové, tak $R - Q = 0$,[6] a teda $R = Q$. Prakticky si ale vieme vystačiť s menej ako $d + 1$ vyčísleniami polynómu $R - Q$, ako ukážeme.

Uvažujme nasledovný „znáhodnený“ algoritmus: rovnomerne náhodne nezávisle vygenerujeme prirodzené čísla $X_1, \ldots, X_k$ z množiny $\{1, 2, \ldots, N\}$. Formálne, všetky k-tice udalostí $[X_1 \in M_1], [X_2 \in M_2], \ldots, [X_k \in M_k]$, kde $M_1, M_2, \ldots, M_k \subseteq \{1, \ldots, N\}$, sú združene nezávislé a $P([X_i \in M_i]) = m_i/N$, kde m_i je počet prvkov množiny M_i, $i = 1, \ldots, k$.[7] Pokiaľ bude platiť $Q(X_i) = R(X_i)$ pre každú hodnotu $X_1, \ldots, X_k$, znáhodnený algoritmus vráti hodnotu `true`, v opačnom prípade (t.j. ak $Q(X_i) \neq R(X_i)$ pre čo i len jednu hodnotu $X_1, \ldots, X_k$) znáhodnený algoritmus vráti hodnotu `false`. Aká je pravdepodobnosť, že sa znáhodnený algoritmus „pomýli“?

Riešenie: Ak (2.1) platí, tak sa znáhodnený algoritmus s istotou nepomýli. V prípade, že rovnosť (2.1) neplatí, tak už nie je úplne vylúčené, že sa znáhodnený algoritmus pomýli, teda že vráti hodnotu `true`. Ohraničme túto pravdepodobnosť omylu zhora. Nech $Q \neq R$. Znáhodnený algoritmus sa pomýli, ak bude platiť $Q(X_i) - R(X_i) = 0$ pre všetky $i = 1, \ldots, k$. Množina M reálnych koreňov polynómu $R - Q$ má maximálne d prvkov, pretože $R - Q$ je nenulový polynóm stupňa nanajvýš d. Algoritmus sa teda pomýli, len ak súčasne nastanú udalosti $[X_1 \in M]$, $[X_2 \in M]$, $\ldots$, $[X_k \in M]$. Keďže tieto udalosti sú združene nezávislé a každá z nich má pravdepodobnosť menšiu alebo rovnú hodnote d/N, dostávame

$$P(\cap_{i=1}^k [X_i \in M]) = \prod_{i=1}^k P([X_i \in M]) \leq \left(\frac{d}{N}\right)^k.$$

Ak zvolíme, povedzme, $N = 10d$ a $k = 10$ evaluácií polynómu $R - Q$, tak pravdepodobnosť mylného výsledku znáhodneného algoritmu je nulová alebo prakticky zanedbateľná.

[5] Dá sa napísať aj (menej priamočiary) deterministický algoritmus s menším počtom násobení, ktorý overí tieto rovnosti.

[6] Všimnime si, že $R - Q$ je polynóm, ktorého stupeň je maximálne d, a teda ak je nulový v $d + 1$ rôznych číslach, je nulový na celom $\mathbb{R}$.

[7] $X_1, X_2, \ldots, X_k$ sú takzvané „nezávislé náhodné premenné“, čo je pojem, s ktorými sa viac oboznámime v ďalších kapitolách. V tomto príklade ale nepotrebujeme presne rozumieť pojmu náhodná premenná; stačí len brať udalosť $[X_i \in M_i]$ ako zápis udalosti „i-te vygenerované číslo bude patriť do množiny M_i“.

Podobná situácia sa vyskytuje pomerne často. Problém vieme vyriešiť použitím znáhodnených algoritmov nie síce úplne dokonale z teoretického hľadiska, z praktického hľadiska však úplne dostatočne a omnoho rýchlejšie či jednoduchšie.

Lema 2.17 (Nezávislosť doplnkov)**.** Ak sú udalosti A a B nezávislé, tak sú nezávislé aj udalosti A a $\Omega \setminus B$, ako aj udalosti $\Omega \setminus A$ a $\Omega \setminus B$. Všeobecne, nech $A_1, \dots, A_n$ sú združene nezávislé udalosti. Pre každé $i = 1, \dots, n$ zvoľme za A'_i buď A_i, alebo $\Omega \setminus A_i$. Potom aj udalosti $A'_1, \dots, A'_n$ sú združene nezávislé.

Dôkaz. Urobme dôkaz pre dvojicu udalostí; pre všeobecný počet udalostí je dôkaz analogický.

Ak sú udalosti A a B nezávislé, tak s využitím základných vlastností pravdepodobnosti dostávame $P(A \cap (\Omega \setminus B)) = P(A \setminus (A \cap B)) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\Omega \setminus B)$. Tým sme ukázali, že A a $\Omega \setminus B$ sú nezávislé. Nezávislosť udalostí $\Omega \setminus A$ a $\Omega \setminus B$ plynie opakovaným použitím už dokázanej časti lemy. □

Predchádzajúcu vetu vieme zovšeobecniť do nasledujúcej podoby; pri jej čítaní je dobré si uvedomiť, že *každú* množinovú operáciu vieme vyjadriť pomocou operácií doplnku, zjednotenia a prieniku.

Veta 2.18 (Nezávislosť udalostí po skupinách)**.** Majme združene nezávislé udalosti $A_1, \dots, A_n$. Nech $n_0 = 0 < n_1 < n_2 < \dots < n_{k-1} < n_k = n$. Pre každé $j = 1, \dots, k$ vytvorme B_j z $A_{n_{j-1}+1}, \dots, A_{n_j}$ akokoľvek, pomocou operácií doplnku, zjednotenia alebo prieniku. Potom sú aj udalosti $B_1, \dots, B_k$ združene nezávislé.

Dôkaz. Nech sú udalosti $A_1, A_2, \dots, A_n$ združene nezávislé, $n \geq 3$. Potom a) $A_1 \cap A_2, A_3, \dots, A_n$ sú združene nezávislé a b) $A_1 \cup A_2, A_3, \dots, A_n$ sú združene nezávislé. Tvrdenie a) plynie priamo z definície združenej nezávislosti a tvrdenie b) ukážeme nasledovne: Ak sú $A_1, A_2, \dots, A_n$ združene nezávislé, tak sú podľa lemy 2.17 aj $\Omega \setminus A_1, \Omega \setminus A_2, A_3, \dots, A_n$ združene nezávislé, takže podľa tvrdenia a) sú aj $(\Omega \setminus A_1) \cap (\Omega \setminus A_2), A_3, \dots, A_n$ združene nezávislé, a teda opäť podľa lemy 2.17 sú aj $\Omega \setminus ((\Omega \setminus A_1) \cap (\Omega \setminus A_2)), A_3, \dots, A_n$ združene nezávislé. Avšak prvá z týchto udalostí je práve $A_1 \cup A_2$. Celé tvrdenie vety o nezávislosti po skupinách plynie opakovaným použitím lemy 2.17 a tvrdení a), b). □

Poznámka 2.19. Intuitívne teda môžeme združenú nezávislosť udalostí $A_1, \dots, A_n$ chápať tak, že informácia o nastaní alebo nenastaní udalosti založenej na akejkoľvek podmnožine $A_{i_1}, \dots, A_{i_k}$ týchto udalostí neovplyvní pravdepodobnostné očakávanie ohľadom udalosti založenej na podmnožine iných udalostí $A_{j_1}, \dots, A_{j_l}$ (slovo „iných“ chápeme v zmysle $\{i_1, \dots, i_k\} \cap \{j_1, \dots, j_l\} = \emptyset$).

Poznámka 2.20. Niekedy sa stretávame s takým prípadom združene nezávislých udalostí $A_1, A_2, \dots, A_n$, ktoré reprezentujú sériu nezávislých pokusov rovnakého charakteru, čiže všetky tieto udalosti majú rovnakú pravdepodobnosť nastania. Vtedy nás často zaujíma, s akou pravdepodobnosťou sa realizuje práve (akýchkoľvek) k udalostí spomedzi $A_1, A_2, \dots, A_n$. Typickým príkladom je n-násobné hádzanie hracou kockou, pričom udalosť A_i znamená padnutie šestky v i-tom hode, $i = 1, \dots, n$. V takom prípade by nás mohlo zaujímať, s akou pravdepodobnosťou padne šestka práve v k hodoch, $0 \leq k \leq n$. Touto situáciou sa formálnejšie zaoberáme v nasledovnej vete.

Veta 2.21 (Binomická formula)**.** Nech $A_1, A_2, \ldots, A_n$ sú združene nezávislé udalosti, pričom každá má pravdepodobnosť p. Nech A je udalosť, že nastane práve k spomedzi udalostí $A_1, A_2, \ldots, A_n$, kde $k \in \{0, \ldots, n\}$, t.j.

$$A = \{\omega \in \Omega : |\{i \in \{1, \ldots, n\} : \omega \in A_i\}| = k\}.$$

Potom

$$P(A) = \binom{n}{k} p^k (1-p)^{n-k}. \tag{2.2}$$

Dôkaz. Pre každú k-prvkovú množinu indexov $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ označme ako $B_{i_1,\ldots,i_k}$ udalosť, že nastanú všetky A_i pre $i \in \{i_1, \ldots, i_k\}$ a súčasne nenastane žiadna udalosť A_j pre $j \notin \{i_1, \ldots, i_k\}$, t.j.

$$B_{i_1,\ldots,i_k} = \left[\bigcap_{i \in \{i_1,\ldots,i_k\}} A_i\right] \bigcap \left[\bigcap_{j \notin \{i_1,\ldots,i_k\}} (\Omega \setminus A_j)\right].$$

Všimnime si, že zjednotením $\binom{n}{k}$ udalostí $B_{i_1,\ldots,i_k}$ je udalosť A, tieto udalosti sú navzájom disjunktné a každá z nich má pravdepodobnosť $p^k(1-p)^{n-k}$. Preto

$$P(A) = P\left(\cup_{\{i_1,\ldots,i_k\}} B_{i_1,\ldots,i_k}\right) = \sum_{\{i_1,\ldots,i_k\}} P(B_{i_1,\ldots,i_k}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

□

Príklad 2.22. Vykonáme 10 nezávislých hodov kockou. S akou pravdepodobnosťou padne šestka 7-krát? S akou pravdepodobnosťou padne šestka viac ako 7-krát?

Riešenie: Označme symbolmi $A_1, \ldots, A_{10}$ udalosti, že v prvom hode padne šestka, ..., v desiatom hode padne šestka. Tieto udalosti sú združene nezávislé a každá z nich má pravdepodobnosť $1/6$. Pýtame sa na pravdepodobnosť udalosti, že nastane práve 7 z udalostí $A_1, \ldots, A_{10}$. Podľa binomickej formuly (2.2) je táto pravdepodobnosť

$$\binom{10}{7} \left(\frac{1}{6}\right)^7 \left(\frac{5}{6}\right)^3 \approx 0.00025.$$

Analogicky dostaneme odpoveď na druhú otázku:

$$\binom{10}{8} \left(\frac{1}{6}\right)^8 \left(\frac{5}{6}\right)^2 + \binom{10}{9} \left(\frac{1}{6}\right)^9 \left(\frac{5}{6}\right)^1 + \binom{10}{10} \left(\frac{1}{6}\right)^{10} \left(\frac{5}{6}\right)^0 \approx 1.95 \times 10^{-5}.$$

Príklad 2.23. Komunikačný kanál sa skladá zo série uzlov, pričom vždy i-ty uzol predáva jedno-bitovú informáciu na vstup $i+1$-vému uzlu. Na každom uzle však s pravdepodobnosťou p dochádza k chybe, ktorá sa prejaví tým, že na výstupe tohto uzla bude opačný bit ako na jeho vstupe. Navyše, chyby na jednotlivých uzloch sa vyskytujú navzájom nezávisle. Napíšme vzorec udávajúci pravdepodobnosť, že bit na vstupe prvého uzla bude rovnaký ako bit na výstupe n-tého uzla.

Riešenie: Je zrejmé, že bit na vstupe prvého uzla bude rovnaký ako bit na výstupe n-tého uzla práve vtedy, ak dôjde k chybe prenosu na párnom počte uzlov. Pre jednoduchosť zápisu predpokladajme, že n je párne. Ak označíme $A^{(k)}$ udalosť, že dôjde k chybe práve na k uzloch ($k = 0, \ldots, n$), tak podľa binomickej formuly je hľadaná pravdepodobnosť

$$P(A^{(0)} \cup A^{(2)} \cup \ldots \cup A^{(n)}) = \sum_{0 \leq j \leq n/2} \binom{n}{2j} p^{2j}(1-p)^{n-2j} = \frac{(1-2p)^n}{2} + \frac{1}{2}.$$

Všimnite si, ako sa výsledná hodnota správa v prípadoch $p = 0$, $p = 1/2$ a $p = 1$.

Príklad 2.24. Vo vrecku máme dve na pohľad nerozlíšiteľné mince; vieme však, že sú obe falošné. Dokonca vieme, že na jednej z týchto mincí padá znak s pravdepodobnosťou 1/3 a hlava s pravdepodobnosťou 2/3, a na druhej minci presne naopak, t.j. znak na nej padá s pravdepodobnosťou 2/3 a hlava s pravdepodobnosťou 1/3. Náhodne sme zvolili z tejto dvojice jednu mincu a hodili sme ňou šesťkrát, z čoho nám štyrikrát padol znak a dvakrát hlava. S akou pravdepodobnosťou nám padne znak pri siedmom hode zvolenou mincou?

Riešenie: Pozrieme sa na situáciu z hľadiska pred začatím celého experimentu a vypočítajme pravdepodobnosť, že nám v siedmom hode náhodne zvolenou mincou padne znak (udalosť A) za podmienky, že z prvých šiestich hodov touto mincou padne znak štyrikrát (udalosť B). Pre oba indexy $i = 1, 2$ definujme ešte udalosť C_i, ktorá znamená, že na hádzanie náhodne vyberieme mincu i. Keďže udalosti C_1, C_2 tvoria rozklad výberového priestoru a pravdepodobnosť každej z nich je 1/2, dostávame podľa vety o úplnej pravdepodobnosti a binomickej formuly nasledovné rovnosti:

$$P(B) = \sum_{i=1}^{2} P(B|C_i)P(C_i) = \frac{1}{2}\left(\binom{6}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2 + \binom{6}{4}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^2\right),$$

$$P(A \cap B) = \sum_{i=1}^{2} P(A \cap B|C_i)P(C_i) = \frac{1}{2}\left(\binom{6}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2\frac{1}{3} + \binom{6}{4}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^2\frac{2}{3}\right).$$

Po mechanických úpravách zisťujeme, že

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3}{5} = 0.6.$$

Poznámka 2.25. Niekedy sa zaoberáme sériou nezávislých pokusov rovnakého charakteru, pričom každý pokus môže skončiť nie len binárnym výsledkom typu „úspech/neúspech“, ale viacerými rôznymi výsledkami, napríklad tromi – „biela/modrá/červená“. Takejto situácie sa týka nasledovná veta.

Veta 2.26 (Multinomická formula)**.** Uvažujme udalosti $A_i^{(j)}$, $i = 1, \ldots, n$ a $j = 1, \ldots, m$. Nech pre každé i tvoria udalosti $A_i^{(1)}, \ldots, A_i^{(m)}$ rozklad výberového priestoru, pričom $P(A_i^{(1)}) = p_1, \ldots, P(A_i^{(m)}) = p_m$. (Pravdepodobnosti $p_1, \ldots, p_m$ nezávisia na i.) Ďalej nech pre každý výber indexov $j_1, j_2, \ldots, j_n \in \{1, \ldots, m\}$ sú udalosti $A_1^{(j_1)}, \ldots, A_n^{(j_n)}$ združene nezávislé. Nech $k_1, \ldots, k_m$ sú čísla z množiny $\{0, \ldots, n\}$, ktorých súčet je n. Nech $A_{k_1,\ldots,k_m}$ je udalosť, že spomedzi udalostí $A_1^{(j)}, \ldots, A_n^{(j)}$ nastane práve k_j, a to pre každé j. Potom

$$P(A_{k_1,\ldots,k_m}) = \frac{n!}{k_1!k_2!\ldots k_m!} p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m}. \tag{2.3}$$

Dôkaz. Dôkaz nie je ťažký, ale je technicky zdĺhavý, takže ho nebudeme uvádzať. □

Poznámka 2.27. Presná formulácia vety 2.26 je neprehľadná, jej myšlienku si ľahšie zapamätáme napríklad nasledovne. Predpokladajme, že máme tabuľku s n riadkami a m stĺpcami a predpísané kladné pravdepodobnosti $p_1, \ldots, p_m$, ktorých súčet je jedna. V prvom riadku tabuľky zvolíme a zaškrtneme práve jedno políčko, a to tak, že prvé políčko volíme s pravdepodobnosťou p_1, druhé s pravdepodobnosťou p_2, ... ,m-te políčko volíme s pravdepodobnosťou p_m.[8] Nezávisle, avšak rovnakým spôsobom, zvolíme a zaškrtneme jedno políčko v každom ďalšom riadku. Vzorec (2.3) udáva pravdepodobnosť, že na konci bude v prvom stĺpci k_1 zaškrtnutých políčok, v druhom stĺpci k_2 zaškrtnutých políčok, ..., v m-tom stĺpci bude k_m zaškrtnutých políčok.

Príklad 2.28. V krabici máme 10 loptičiek, z ktorých je 5 bielych, 3 sú modré a 2 sú červené. Z krabice 7-krát vyberieme loptičku, pričom každú vybranú loptičku vrátime naspäť do krabice ešte pred výberom ďalšej loptičky (takýto systém voláme „náhodný výber s návratom“). Vypočítajme pravdepodobnosť, že takto vyberieme spolu 3 biele loptičky, 2 modré loptičky a 2 červené loptičky (nezávisle od toho, v ktorom ťahu).

Riešenie: Ak udalosť $A_i^{(j)}$, $i = 1, \ldots, 7$, $j = 1, 2, 3$ zodpovedá tomu, že v i-tom ťahu vyberieme loptičku j-tej farby (farba 1 je biela, farba 2 modrá a farba 3 červená), tak máme situáciu z predchádzajúcej vety. Dostávame

$$P(A_{3,2,2}) = \frac{7!}{3!2!2!} 0.5^3 0.3^2 0.2^2 = 0.0945.$$

[8] Ako konkrétne túto náhodnú voľbu s rôznymi pravdepodobnosťami realizujeme je už iná otázka. Existujú napríklad metódy, ako urobiť počítačový program, ktorý nám bude túto náhodnú voľbu realizovať; pozri vetu 3.26.

Kapitola 3

Všeobecné náhodné premenné

3.1 Vlastnosti náhodných premenných

V mnohých situáciách nás na elementárnych výsledkoch ω náhodného pokusu (pozorovania, merania, simulácie a podobne) zaujíma jedna alebo viac *číselných* charakteristík. Napríklad ak náhodne generujeme graf, čiže elementárnym výsledkom ω bude graf, môže nás zaujímať počet hrán $X(\omega)$, počet komponentov súvislosti $Y(\omega)$ a počet izolovaných vrcholov $Z(\omega)$ tohto grafu. Alebo ak náhodne generujeme bod v rovine, čiže ω je bod z $\mathbb{R}^2$, môže nás zaujímať horizontálna súradnica $X(\omega)$, vertikálna súradnica $Y(\omega)$ alebo napríklad vzdialenosť $Z(\omega)$ tohto bodu od počiatku súradnicovej sústavy.

Číselné charakteristiky elementárnych výsledkov, ktoré spĺňajú istú technickú podmienku, voláme v teórii pravdepodobnosti náhodné premenné. Čiže náhodná premenná sa dá intuitívne stotožniť s budúcim pozorovaním, ktorého konkrétnu hodnotu nie je možné vopred vedieť.

Definícia 3.1. Nech $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor. Budeme hovoriť, že funkcia $X : \Omega \to \mathbb{R}$ je **náhodná premenná**,[1] ak spĺňa **podmienku merateľnosti**

$$\{\omega \in \Omega : X(\omega) < x\} \in \mathcal{S} \text{ pre všetky } x \in \mathbb{R}.$$

Poznámka 3.2. Všimnime si, že podmienka merateľnosti nezávisí od P a že v často sa vyskytujúcom prípade $\mathcal{S} = 2^\Omega$ je vždy splnená.

Veta 3.3 (Vzory borelovských množín sú náhodné udalosti)**.** Nech $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor. Potom funkcia $X : \Omega \to \mathbb{R}$ je náhodná premenná vtedy a len vtedy, ak pre každé $B \in \mathcal{B}$ platí

$$\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{S}.$$

Dôkaz. Dôkaz implikácie $\Rightarrow$: Nech $X : \Omega \to \mathbb{R}$ je náhodná premenná. Pre akúkoľvek množinu $A \subseteq \mathbb{R}$ označíme symbolom $X^{-1}(A)$ vzor množiny A v zobrazení X. Uvažujme systém $\mathcal{H}$ tých podmnožín $H \subseteq \mathbb{R}$, pre ktoré $X^{-1}(H) \in \mathcal{S}$. Ľahko overíme, že $\mathcal{H}$ je σ-algebra na $\mathbb{R}$, ktorá obsahuje všetky otvorené intervaly. (Uvedomíme si, že pre každú množinu $A \subseteq \mathbb{R}$ platí $X^{-1}(\mathbb{R} \setminus A) = \Omega \setminus X^{-1}(A)$ a tiež pre akýkoľvek systém množín $A_i \subseteq \mathbb{R}$, $i \in I$ platí $X^{-1}(\cup_{i \in I} A_i) = \cup_{i \in I} X^{-1}(A_i)$.) Keďže $\mathcal{B}$ je najmenšia σ-algebra na $\mathbb{R}$ obsahujúca všetky otvorené intervaly, tak musí platiť $\mathcal{B} \subseteq \mathcal{H}$, čím je dôkaz priamej implikácie ukončený. Implikácia $\Leftarrow$ je zrejmá, pretože každý interval typu $(-\infty, a)$ patrí do systému $\mathcal{B}$. $\square$

[1] V slovenčine sa ako synonymum používa aj pojem „náhodná veličina“.

Poznámka 3.4. Pre $B \in \mathcal{B}$ budeme udalosť $\{\omega \in \Omega : X(\omega) \in B\}$ zapisovať skrátene $[X \in B]$. Teda $\{\omega \in \Omega : X(\omega) \in (a,b)\} = [X \in (a,b)] = [a < X < b]$ alebo $\{\omega \in \Omega : X(\omega) = a\} = [X = a]$ a podobne.

Poznámka 3.5. Intuitívne zdôvodnenie podmienky merateľnosti v definícii 3.1 je nasledovné. Pre akékoľvek $x \in \mathbb{R}$ vyžadujeme, aby sme vedeli priradiť pravdepodobnosť[2] výroku „Výsledkom pokusu bude nejaké také ω, ktorého číselná charakteristika X je menšia ako x“. Všimnime si, že podľa vety 3.3 táto jednoduchá podmienka znamená, že vieme priradiť pravdepodobnosť aj všetkým, potenciálne veľmi komplikovaným výrokom typu „Výsledkom pokusu bude nejaké také ω, ktorého číselná charakteristika X patrí do borelovskej množiny B“.

Poznámka 3.6. Náhodná premenná X na pravdepodobnostnom priestore $(\Omega, \mathcal{S}, P)$ definuje „indukovanú“ pravdepodobnostnú mieru P_X na σ-algebre $\mathcal{B}$ borelovských množín, a to nasledovne: $P_X(B) = P[X \in B]$. Pravdepodobnosť P_X sa niekedy formálne stotožňuje s „pravdepodobnostným rozdelením“ náhodnej premennej X. Pod pojmom pravdepodobnostné rozdelenie náhodnej premennej sa však niekedy myslí aj „distribučná funkcia“, „pravdepodobnostná funkcia rozdelenia“ alebo „hustota“, čo sú pojmy, s ktorými sa zoznámime neskôr.

Príklad 3.7 (Príklady náhodných premenných)**.**

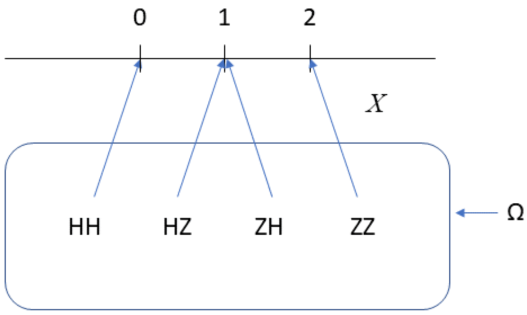
1. Nech $\mathcal{S}$ je σ-algebra na množine $\Omega \neq \emptyset$ a nech $A \subseteq \Omega$. Definujme zobrazenie $I : \Omega \to \mathbb{R}$ nasledovne: $I(\omega) = 1$ pre všetky $\omega \in A$ a $I(\omega) = 0$ pre všetky $\omega \in \Omega/A$. (Pozri tiež definíciu 4.73.) Potom I je náhodná premenná na tomto priestore vtedy a len vtedy, keď $A \in \mathcal{S}$.

2. Hodíme naraz dvomi mincami. Ako pravdepodobnostný priestor zvolíme $(\Omega, 2^\Omega, P)$, kde $\Omega = \{\mathrm{HH}, \mathrm{HZ}, \mathrm{ZH}, \mathrm{ZZ}\}$ a $P(A) = |A|/4$ pre každé $A \subseteq \Omega$. Počet X mincí, na ktorých padne znak, je náhodná premenná; pozri obrázok 3.1. (Keďže každá podmnožina množiny Ω je v našom modeli udalosťou, podmienka merateľnosti funkcie X je automaticky splnená.)

3. Generujeme tri náhodné čísla z množiny $\{0, \dots, 9\}$. Súčet X všetkých troch vygenerovaných čísel, najväčšie vygenerované číslo Y aj počet Z vygenerovaných núl sú náhodné premenné. Formálne, ak by sme uvažovali pravdepodobnostný model $(\Omega, \mathcal{S}, P)$, kde $\Omega = \{(i_1, i_2, i_3) : i_1, i_2, i_3 \in \{0, \dots, 9\}\}$ a $\mathcal{S} = 2^\Omega$, tak spomenuté náhodné premenné by boli funkcie na Ω definované $X(i_1, i_2, i_3) = i_1 + i_2 + i_3$, $Y(i_1, i_2, i_3) = \max(i_1, i_2, i_3)$ a $Z(i_1, i_2, i_3) = |\{k \in \{1, 2, 3\} : i_k = 0\}|$. (Podobne ako v predošlom príklade, keďže každá podmnožina množiny Ω je v našom modeli udalosťou, podmienka merateľnosti funkcií X, Y, Z je automaticky splnená.)

4. Pozorujeme nekonečnú postupnosť náhodne generovaných núl a jednotiek. Použime formálny model reprezentovaný σ-algebrou udalostí z príkladu 1.17, resp. pravdepodobnostným priestorom $(\Omega, \mathcal{S}, P)$ z príkladu 1.33. Náhodnou premennou na tomto priestore môže byť napríklad prvá hodnota X vo vygenerovanej postupnosti, súčet Y hodnôt

[2] Ak uvažujeme pravdepodobnostný model $(\Omega, \mathcal{S}, P)$, tak to, že vieme priradiť pravdepodobnosť nejakej množine $A \subseteq \Omega$, je ekvivalentné tomu, že A je udalosť, teda že $A \in \mathcal{S}$.



Obr. 3.1: Ilustračný obrázok k príkladu 3.7

prvých 10 prvkov postupnosti, počet Z pozorovaných hodnôt kým prvýkrát zaznamenáme jednotku a tak ďalej. Formálne, pre elementárny výsledok – nekonečnú binárnu postupnosť, ktorú si označíme $\omega = (b_i)_{i=1}^{\infty}$ – definujeme spomenuté náhodné premenné nasledovne: $X(\omega) = b_1$, $Y(\omega) = \sum_{i=1}^{10} b_i$ a $Z(\omega) = \min\{i \in \mathbb{N} : b_i = 1\}$.[3] Pre všetky tri spomenuté funkcie je možné dokázať, že spĺňajú podmienku merateľnosti. Dôkaz merateľnosti X je jednoduchý, dôkaz merateľnosti zvyšných dvoch náhodných premenných je už zložitejší a v týchto skriptách je lepšie ho vynechať.

Poznámka 3.8. Neskôr budeme často analyzovať súčasne viacero náhodných premenných, podobne ako v častiach 3 a 4 z príkladu 3.7. V takom prípade automaticky predpokladáme, že všetky tieto náhodné premenné sú definované na spoločnom pravdepodobnostnom priestore.

Definícia 3.9. Funkcia $g : \mathbb{R}^n \to \mathbb{R}$ sa nazýva **borelovská**, ak pre každé $B \in \mathcal{B}$ platí $g^{-1}(B) \in \mathcal{B}_n$, čiže ak vzor každej (jednorozmernej) borelovskej množiny je (n-rozmerná) borelovská množina.

Nasledovné lemy majú predovšetkým teoretický význam a ich dôkazy sú technicky pomerne zdĺhavé, preto ich neuvedieme. Je ale dobré vedieť, že takéto lemy existujú, aby sa ani opatrný študent nebál náhodné premenné transformovať a kombinovať.

Lema 3.10. Nech $B_1, B_2, \ldots \in \mathcal{B}_n$ sú navzájom disjunktné a také, že $\cup_{i=1}^{\infty} B_i = \mathbb{R}^n$. Nech $g_1, g_2, \ldots$ sú spojité funkcie na $\mathbb{R}^n$. Potom funkcia $g : \mathbb{R}^n \to \mathbb{R}$ spĺňajúca $g(x) = g_i(x)$ pre všetky $x \in B_i$ je borelovská.

Lema 3.11. Nech $X_1, \ldots, X_n$ sú náhodné premenné a nech g je borelovská funkcia. Potom je náhodnou premennou aj zobrazenie $g(X_1, \ldots, X_n) : \Omega \to \mathbb{R}$ definované $g(X_1, \ldots, X_n)(\omega) = g(X_1(\omega), \ldots, X_n(\omega))$ pre všetky $\omega \in \Omega$.

Poznámka 3.12. Dôležité je uvedomiť si to, že prakticky každá „slušná“ funkcia z $\mathbb{R}^n$ do $\mathbb{R}$ sa dá napísať v tvare borelovskej funkcie g, a preto prakticky akákoľvek funkcia jednej alebo viacerých náhodných premenných je tiež náhodnou premennou. Povedzme, ak je X náhodná premenná, tak aj X^2 je náhodná premenná (lebo transformačná funkcia $g(x) = x^2$ je spojitá), ak X_1, X_2 sú náhodné premenné, tak aj $X_1 + X_2$ je náhodná premenná (lebo transformačná

[3] Uvedená definícia náhodnej premennej Z má drobný nedostatok: nemá priradenú hodnotu pre elementárny výsledok ω_0 zodpovedajúci nekonečnej postupnosti núl. Pre tento výsledok môžeme definovať $Z(\omega_0) = 0$.

funkcia $g(x_1, x_2) = x_1 + x_2$ je spojitá na $\mathbb{R}^2$), ale aj napríklad $X_1^{\lceil |X_2| \rceil}$, kde $\lceil \cdot \rceil$ označuje hornú celú časť, je náhodná premenná (pretože príslušná transformačná funkcia $g(x_1, x_2) = x_1^{\lceil |x_2| \rceil}$ sa dá napísať v tvare zo znenia lemy 3.10 ako funkcia po častiach spojitá na spočítateľnom systéme dvojrozmerných borelovských množín).

Poznámka 3.13. Zmysluplný je aj pojem náhodnej premennej s oborom hodnôt iným než $\mathbb{R}$.[4] Napríklad by takýmto oborom hodnôt mohla byť akákoľvek konečná množina, na ktorej nie je definované žiadne usporiadanie;[5] vtedy by išlo o takzvanú „kategorickú“ náhodnú premennú.[6] Analogicky by sme vedeli definovať napríklad náhodnú premennú, ktorej hodnoty sú matice, postupnosti, funkcie a podobne. V týchto skriptách sa budeme primárne zaoberať klasickými náhodnými premennými v zmysle definície 3.1 a takzvanými „náhodnými vektormi“, čo sú náhodné premenné, ale s oborom hodnôt $\mathbb{R}^m$ pre všeobecné $m \in \mathbb{N}$.

3.2 Distribučná a kvantilová funkcia

„Distribučná funkcia“ náhodnej premennej X je funkcia $F : \mathbb{R} \to \mathbb{R}$ s jednoduchými vlastnosťami, ktorá reprezentuje pravdepodobnosti typu $P[X \in B]$ pre všetky borelovské množiny B.[7] „Kvantilová funkcia“ $G : (0, 1) \to \mathbb{R}$ a „kvantily“ sú dôležité napríklad pre počítačové simulovanie realizácií náhodných premenných a pre testovanie štatistických hypotéz. Pre veľa náhodných premenných sú distribučná a kvantilová funkcia navzájom inverzné funkcie.

Definícia 3.14. Distribučná funkcia náhodnej premennej X je funkcia $F : \mathbb{R} \to \mathbb{R}$, ktorá je v bode $x \in \mathbb{R}$ definovaná[8]

$$F(x) = P[X < x].$$

Príklad 3.15. Hádžeme dvakrát mincou. Nech náhodná premenná X je počet znakov v týchto dvoch hodoch. Potom distribučná funkcia náhodnej premennej X je (pozri panel (a) na obrázku 3.2)

$$F(x) = \begin{cases} 0 & \text{pre } x \le 0, \\ 1/4 & \text{pre } 0 < x \le 1, \\ 3/4 & \text{pre } 1 < x \le 2, \\ 1 & \text{pre } x > 2. \end{cases}$$

Veta 3.16 (Základné vlastnosti distribučnej funkcie)**.** Nech F je distribučná funkcia akejkoľvek náhodnej premennej X. Potom 1) $0 \le F(x) \le 1$ pre všetky $x \in \mathbb{R}$, 2) F je neklesajúca a spojitá zľava, 3) $\lim_{x \to \infty} F(x) = 1$ a $\lim_{x \to -\infty} F(x) = 0$.

Dôkaz. Vlastnosť $0 \le F(x) \le 1$ pre každé $x \in \mathbb{R}$ je zrejmá. Neklesajúcosť F je tiež jednoduchá: ak $x \le y$ sú dve reálne čísla, tak $F(x) = P[X < x] \le P[X < y] = F(y)$, pretože $[X < x] \subseteq [X < y]$.

[4] Pokiaľ by sme mali na tomto obore hodnôt definovanú σ-algebru.
[5] Príslušnou σ-algebrou by bola potenčná množina.
[6] Niekedy označovanú aj ako „kvalitatívna“ náhodná premenná.
[7] Hoci my si ukážeme ako reprezentuje F len pravdepodobnosti typu $P[X \in I]$, kde I je interval.
[8] Časť literatúry definuje distribučnú funkciu predpisom $F(x) = P[X \le x]$, $x \in \mathbb{R}$. Vlastnosti takto definovanej distribučnej funkcie sú odlišné, ale ide len o málo podstatné technické odlišnosti.

Dokážeme spojitosť zľava. Nech $a \in \mathbb{R}$. Pre každé prirodzené číslo n platí $[X < a-1/n] \subseteq [X < a-1/(n+1)]$ a navyše $\cup_{n=1}^{\infty}[X < a-1/n] = [X < a]$, takže z neklesajúcosti distribučnej funkcie a lemy 1.47 o spojitosti pravdepodobnosti zdola: $\lim_{x\to a^-} F(x) = \lim_{n\to\infty} F(a-1/n) = \lim_{n\to\infty} P[X < a-1/n] = P(\cup_{n=1}^{\infty}[X < a-1/n]) = P[X < a] = F(a)$.

Podobne odvodíme: $\lim_{x\to\infty} F(x) = \lim_{n\to\infty} F(n) = \lim_{n\to\infty} P[X < n] = P(\cup_{n=1}^{\infty}[X < n]) = P(\Omega) = 1$. Rovnosť $\lim_{x\to-\infty} F(x) = 0$ môžeme dokázať analogicky. □

Veta 3.17 (Vyjadrenie pravdepodobnosti intervalov pomocou distribučnej funkcie)**.** Nech F je distribučná funkcia náhodnej premennej X. Nech $a, b \in \mathbb{R}$, pričom $a < b$. Potom $P[a \leq X < b] = F(b) - F(a)$, $P[a \leq X] = 1 - F(a)$, $P[a \leq X \leq b] = \lim\limits_{x\to b^+} F(x) - F(a)$, $P[X = a] = \lim\limits_{x\to a^+} F(x) - F(a)$, $P[a < X < b] = F(b) - \lim\limits_{x\to a^+} F(x)$, $P[a < X] = 1 - \lim\limits_{x\to a^+} F(x)$, $P[a < X \leq b] = \lim\limits_{x\to b^+} F(x) - \lim\limits_{x\to a^+} F(x)$, $P[X \leq b] = \lim\limits_{x\to b^+} F(x)$.

Dôkaz. Priamo z definície máme $P[a \leq X < b] = P[X < b] - P[X < a] = F(b) - F(a)$. Ukážeme ešte napríklad $P[X = a] = \lim_{x\to a^+} F(x) - F(a)$; ostatné rovnosti je možné dokázať buď z tejto rovnosti, alebo analogicky. Platí $P[X = a] = P(\cap_{n=1}^{\infty}[a-1/n \leq X < a+1/n]) = \lim_{n\to\infty} P[a-1/n \leq X < a+1/n] = \lim_{n\to\infty}(F(a+1/n) - F(a-1/n)) = \lim_{x\to a^+} F(x) - F(a)$. Druhá rovnosť plynie zo spojitosti pravdepodobnosti zhora 1.48, posledná rovnosť plynie z toho, že F je spojitá zľava. □

Definícia 3.18. Nech $F : \mathbb{R} \to [0, 1]$ je distribučná funkcia náhodnej premennej X. **Kvantilová funkcia** náhodnej premennej X alebo tiež kvantilová funkcia distribučnej funkcie F je funkcia $G : (0, 1) \to \mathbb{R}$ definovaná

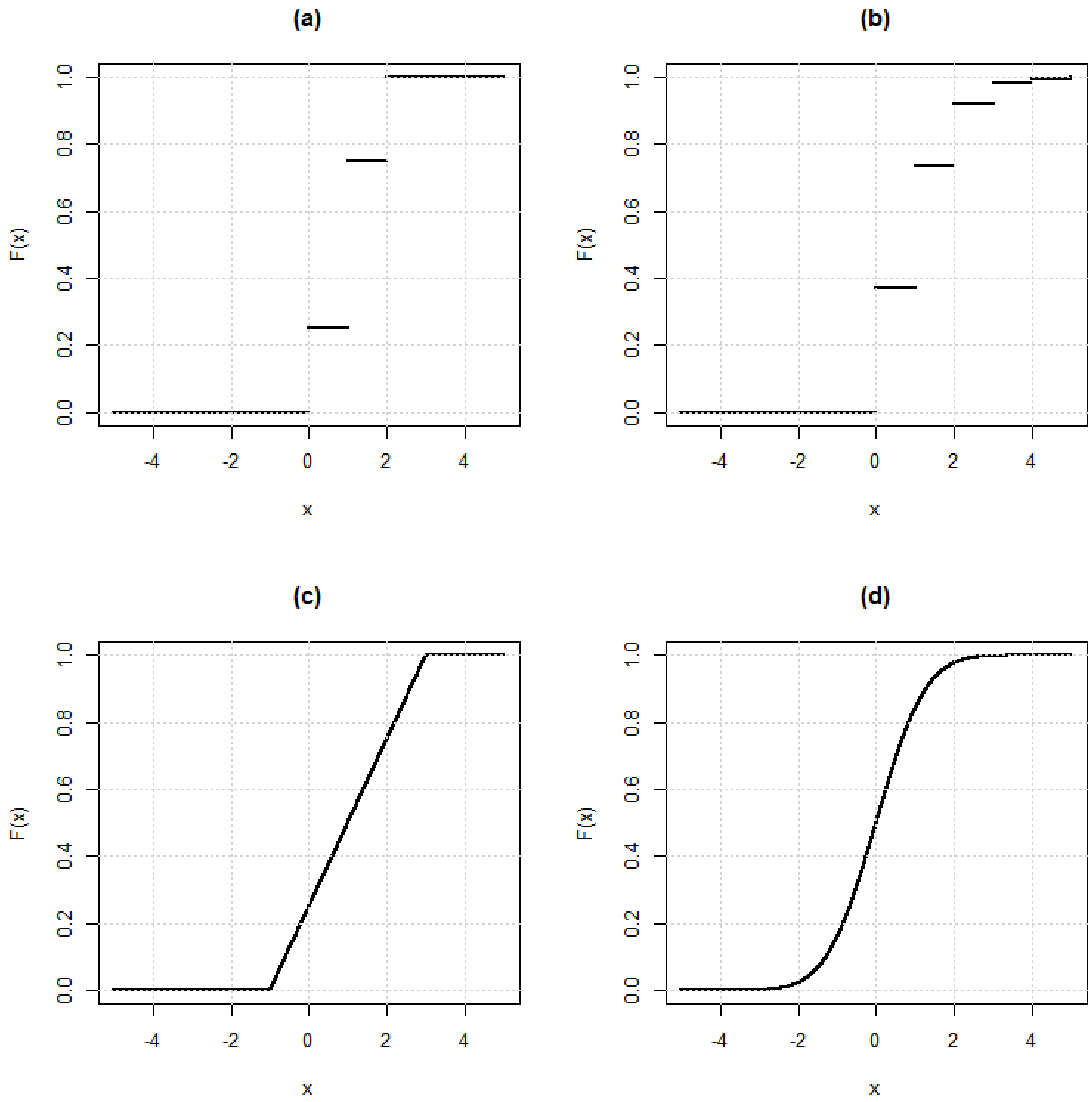
$$G(u) = \sup\{x \in \mathbb{R} : F(x) \leq u\} \text{ pre všetky } u \in (0, 1). \tag{3.1}$$

Poznámka 3.19. Všimnime si, že kvantilová funkcia je dobre definovaná, pretože množina na pravej strane rovnosti (3.1) je vždy neprázdna, keďže $\lim_{x\to-\infty} F(x) = 0$ a u je ostro kladné. Ľahko nahliadneme, že kvantilová funkcia je neklesajúca. S trochou úsilia sa dá ukázať, že kvantilová funkcia je spojitá sprava.

Poznámka 3.20. Nech F je distribučná funkcia a nech $I = \{x \in \mathbb{R} : 0 < F(x) < 1\}$. Distribučné funkcie používané v aplikáciách majú často tú vlastnosť, že sú spojité na celom $\mathbb{R}$ a striktne rastúce na I. V takom prípade je kvantilová funkcia G spojitá a rastúca na celom intervale $(0, 1)$, a je zároveň inverzná funkcia k funkcii, ktorá je zúžením F na interval I.

Definícia 3.21. Nech $G : (0, 1) \to \mathbb{R}$ je kvantilová funkcia náhodnej premennej X. Hodnotu $G(u)$ pre $u \in (0, 1)$ nazývame **$100u$-percentným kvantilom** rozdelenia náhodnej premennej X. Hodnotu $G(1/2)$, čiže 50-percentný kvantil, nazývame **medián** náhodnej premennej X, ktorý budeme značiť $\mathrm{m}(X)$.

Poznámka 3.22. V typickom prípade spomenutom v poznámke 3.20 zjavne platí tvrdenie: $P[X < \mathrm{m}(X)] = P[X > \mathrm{m}(X)] = 1/2$ (a navyše $P[X = \mathrm{m}(X)] = 0$). Medián je teda miera „stredu rozdelenia" náhodnej premennej X, taká, že „polovica pravdepodobnosti je naľavo a polovica je napravo" od mediánu.



Obr. 3.2: Typické distribučné funkcie náhodných premenných. Panel (a) zobrazuje distribučnú funkciu náhodnej premennej z príkladu 3.15, ktorá zodpovedá takzvanému „binomickému rozdeleniu s parametrami $n = 2$ a $p = 1/2$“; pozri časť 4.2.3. Ostatné panely taktiež zodpovedajú náhodným premenným s rozdeleniami základných typov, ktoré si vysvetlíme v neskorších kapitolách. Poznamenajme ešte, že distribučné funkcie na paneloch (a), (b) síce nie sú spojité, ale sú spojité zľava, čo obrázok nezachytáva.

Poznámka 3.23. Veľmi užitočným nástrojom je počítačové simulovanie reálnych dejov. Pri simulovaní máme typicky k dispozícii procedúru, ktorá dokáže generovať realizáciu náhodnej premennej U s „rovnomerným rozdelením na intervale $(0,1)$“, čo zapisujeme $U \sim R(0,1)$. Toto rozdelenie bližšie opisujeme v časti 5.2.1.[9] Nasledovná veta sa využíva na prevod medzi náhodnou premennou U a takou náhodnou premennou, ktorá má iné rozdelenie ako rovnomerné na intervale $(0,1)$.[10]

Veta 3.24 (Veta o inverznej transformácii)**.** Nech F je distribučná funkcia akejkoľvek náhodnej premennej a nech G je kvantilová funkcia funkcie F. Nech $U \sim R(0,1)$. Potom náhodná premenná $G(U)$ má distribučnú funkciu F.

Dôkaz. Vetu je možné veľmi jednoducho ukázať pre prípad, že F je funkcia, ktorá je rastúca a spojitá na celom $\mathbb{R}$, čo je inštruktívne cvičenie pre čitateľa. Uvedieme však dôkaz vety o inverznej transformácii v jej plnej všeobecnosti.

Najprv ukážeme, že pre každé $z \in \mathbb{R}$ a pre každé $u \in (0,1)$ platí $G(u) < z \Leftrightarrow u < F(z)$, alebo ekvivalentne, že pre každé $z \in \mathbb{R}$ a pre každé $u \in (0,1)$ platí $F(z) \leq u \Leftrightarrow G(u) \geq z$.

Označme $M_u = \{z \in \mathbb{R} : F(z) \leq u\}$. Ak $F(z) \leq u$, potom $z \in M_u$, a teda $G(u) = \sup M_u \geq z$. Tým sme ukázali implikáciu „$\Rightarrow$“. Opačná implikácia plynie nasledovne: Nech $G(u) = \sup M_u \geq z$. Z definície supréma dostávame, že pre každé $n \in \mathbb{N}$ existuje $z_n \in M_u$, t.j. z_n spĺňajúce $F(z_n) \leq u$, pre ktoré $z - 1/n \leq z_n$. Z neklesajúcosti F máme $F(z - 1/n) \leq F(z_n) \leq u$ a zo spojitosti zľava dostávame $F(z) \leq u$. Tým sme ukázali aj opačnú implikáciu.

Teraz už môžeme vyjadriť distribučnú funkciu náhodnej premennej $G(U)$, kde $U \sim R(0,1)$ v bode $z \in \mathbb{R}$: $F_{G(U)}(z) = P[G(U) < z] = P[U < F(z)] = F(z)$. □

Poznámka 3.25. Ak chceme použiť vetu 3.24 na generovanie realizácií náhodnej premennej s distribučnou funkciou F, potrebujeme počítať hodnoty príslušnej kvantilovej funkcie G. Tento výpočet môže, ale nemusí byť numericky obtiažny. Typické situácie, keď je tento výpočet relatívne jednoduchý, sú, keď F zodpovedá „diskrétnemu“ rozdeleniu pravdepodobnosti[11], exponenciálnemu rozdeleniu (veta 5.34) alebo Paretovmu rozdeleniu (veta 5.41).[12]

Príklad 3.26 (Generovanie realizácií diskrétnej náhodnej premennej)**.** Nech $p_1, p_2, \ldots, p_m$ sú kladné čísla také, že $\sum_{i=1}^m p_i = 1$ a nech $x_1 < x_2 < \ldots < x_m$. Uvažujme distribučnú funkciu $F(x) = 0$ pre $x \leq x_1$ a $F(x) = \sum_{i:x_i<x} p_i$ pre $x > x_1$.[13] Kvantilová funkcia takejto distribučnej funkcie je $G(y) = x_1$, ak $0 < y < p_1$ a $G(y) = x_k$, ak $\sum_{i=1}^{k-1} p_i \leq y < \sum_{i=1}^{k} p_i$ pre $k = 2, \ldots, m$. Realizácie príslušnej náhodnej premennej preto môžeme generovať ako $G(U)$, kde $U \sim R(0,1)$.

Konkrétnejšie, interval $(0,1)$ rozbijeme na intervaly $I_1 = (0, p_1)$, $I_2 = [p_1, p_1 + p_2)$, $I_3 = [p_1 + p_2, p_1 + p_2 + p_3), \ldots, I_m = [1 - p_m, 1)$. Ak realizácia $u \in (0,1)$ náhodnej premennej $U \sim R(0,1)$ padne do intervalu I_k, tak x_k môžeme považovať za realizáciu náhodnej premennej s distribučnou funkciou F. Efektivita generovania pomocou tohto princípu veľmi závisí od toho, ako rýchlo dokážeme identifikovať interval, do ktorého padne číslo u.

[9] Pri prvom čítaní môžeme na tomto mieste chápať pojem rovnomerného rozdelenia skôr intuitívne. K zostatku tejto časti je dobré sa vrátiť po vyjasnení si pojmu rovnomernej spojitej náhodnej premennej.

[10] Počítačové generovanie realizácií náhodných premenných je zaujímavý, avšak netriviálny proces, ktorému sa v týchto skriptách nebudeme podrobnejšie venovať.

[11] Pozri nasledujúci príklad; podrobnejšie o diskrétnych náhodných premenných v ďalšej kapitole.

[12] V prípade exponenciálneho a Paretovho rozdelenia má totiž kvantilová funkcia G veľmi jednoduchý analytický predpis.

[13] F je distribučná funkcia takzvanej „diskrétnej náhodnej premennej, ktorá nadobúda hodnoty $x_1, \ldots, x_m$ s pravdepodobnosťami $p_1, \ldots, p_m$“.

Kapitola 4

Diskrétne náhodné premenné

4.1 Vlastnosti diskrétnych náhodných premenných

„Diskrétnu“ náhodnú premennú používame ako model neurčitosti pre taký výsledok pozorovania, merania či experimentu, o ktorom vopred vieme, že bude patriť do nejakej spočítateľnej podmnožiny reálnych čísel.[1] Diskrétna náhodná premenná môže reprezentovať napríklad výsledok hodu kockou, počet priateľov náhodne zvoleného člena sociálnej siete, počet iterácií znáhodneného algoritmu a podobne. Základnými číselnými charakteristikami diskrétnej náhodnej premennej sú „stredná hodnota“ a „rozptyl“, ktoré zodpovedajú, voľne vyjadrené, „centru“, resp. „šírke“ rozdelenia pravdepodobnosti príslušnej náhodnej premennej.[2]

Definícia 4.1. Náhodnú premennú X na pravdepodobnostnom priestore $(\Omega, \mathcal{S}, P)$ nazývame **diskrétna náhodná premenná**, ak jej obor hodnôt $X(\Omega)$ je spočítateľná množina. Funkciu, ktorá priraďuje pravdepodobnosť $P[X = x]$ číslam $x \in X(\Omega)$, budeme nazývať **pravdepodobnostná funkcia** diskrétnej náhodnej premennej X.

Poznámka 4.2. Hovoríme, že diskrétna náhodná premenná X „nadobúda“ spočítateľne veľa hodnôt a že X nadobúda číslo x „s pravdepodobnosťou“ $P[X = x]$.

Príklad 4.3. Hodíme dvoma (vyváženými klasickými hracími) kockami. Nech náhodná premenná X znamená súčet bodov na týchto kockách. Ľahko sa presvedčíme, že X nadobúda hodnoty $2, 3, \dots, 12$ (obor hodnôt náhodnej premennej X je $X(\Omega) = \{2, 3, \dots, 12\}$, čiže ide o diskrétnu náhodnú premennú). Jednotlivé hodnoty nadobúda X s nasledovnými pravdepodobnosťami: $P[X = x] = (x - 1)/36$ pre $x = 2, \dots, 7$ a $P[X = x] = (13 - x)/36$ pre $x = 8, \dots, 12$.

Lema 4.4. Ak X je diskrétna náhodná premenná a B je akákoľvek borelovská množina,[3] tak

$$P[X \in B] = \sum_{t \in B \cap X(\Omega)} P[X = t].$$

[1] Najčastejšie je touto množinou nejaká podmnožina nezáporných celých čísel.

[2] S číselnými charakteristikami, ktoré nazývame kvantily, sme sa už stretli v časti 3.2. Poznamenajme, že náhodným premenným môžeme priradiť nielen číselné hodnoty, ale aj viacero typov charakterizačných funkcií. Tieto funkcie sú pre teóriu pravdepodobnostných rozdelení veľmi užitočné nástroje, ale v tomto základnom texte sa im nevenujeme.

[3] Typicky je $B = (a, b)$, kde $-\infty \le a < b \le \infty$.

Dôkaz. Stačí si najprv uvedomiť, že $\{[X \in t] : t \in B \cap X(\Omega)\}$ je spočítateľný systém disjunktných udalostí, ktorých zjednotenie je udalosť $[X \in B]$ a následne využiť aditivitu alebo σ-aditivitu pravdepodobnosti. □

Príklad 4.5. Presvedčme sa pomocou predchádzajúcej lemy, že ak je X diskrétna náhodná premenná, tak pre jej distribučnú funkciu F platí

$$F(x) = \sum_{t<x,\, t\in X(\Omega)} P[X = t] \text{ pre všetky } x \in \mathbb{R}.$$

Slovne, $F(x)$ je súčtom atomických pravdepodobností typu $P[X = t]$ pre tie čísla t, ktoré X nadobúda a ktoré sú menšie než x. Z toho je zrejmé, že distribučná funkcia je schodovitá, so „skokmi" v bodoch, ktoré X nadobúda. Formálne $F(x) = P[X < x] = \sum_{t<x,\, t\in X(\Omega)} P[X = t]$, pretože X nadobúda spočítateľne veľa hodnôt. Viaceré príklady distribučných funkcií diskrétnych náhodných premenných sú na obrázkoch v časti 4.2.

Definícia 4.6. Nech X je diskrétna náhodná premenná[4] a nech rad $\sum_{x\in X(\Omega)} xP[X = x]$ absolútne konverguje. Potom hovoríme, že náhodná premenná X má konečnú strednú hodnotu a číslo

$$E(X) = \sum_{x\in X(\Omega)} xP[X = x]$$

nazývame **stredná hodnota** náhodnej premennej X. Ak rad $\sum_{x\in X(\Omega)} xP[X = x]$ nekonverguje alebo nekonverguje absolútne, hovoríme, že náhodná premenná X nemá konečnú strednú hodnotu.

Poznámka 4.7. Pod absolútnou konvergenciou radu $\sum_{x\in\mathbb{X}} r(x)$, kde $\mathbb{X}$ je spočítateľná množina, myslíme to, že $\sum_i |r(x_i)| < \infty$, pričom $x_1, x_2, \ldots$ je (akékoľvek) očíslovanie prvkov množiny $\mathbb{X}$. Každý konečný rad očividne konverguje absolútne,[5] ale nekonečný rad absolútne konvergovať nemusí. Pripomeňme, že absolútna konvergencia okrem iného zaručuje, že hodnota $\sum_{x\in\mathbb{X}} r(x)$ je konečná a rovnaká nezávisle od toho, v akom poradí sčitujeme členy tohto radu.

Poznámka 4.8. Fyzikálne sa dá na strednú hodnotu pozerať ako na „ťažisko rozdelenia" diskrétnej náhodnej premennej: Majme náhodnú premennú X s konečnou strednou hodnotou a nech $X(\Omega) = \{x_1, x_2, \ldots\}$. Uvažujme systém hmotných bodov s hmotnosťami $P[X = x_1]$, $P[X = x_2], \ldots$ ležiacich na orientovanej priamke $\overrightarrow{p}$ vo vzdialenostiach $x_1, x_2, \ldots$ od pevne zvoleného bodu O v smere orientácie tejto priamky. Potom ťažisko tohto systému bodov bude ležať na $\overrightarrow{p}$, presne vo vzdialenosti $E(X)$ od bodu O v smere orientácie priamky $\overrightarrow{p}$.

Príklad 4.9. Stredná hodnota náhodnej premennej z príkladu 4.3 je

$$E(X) = 2\frac{1}{36} + \cdots + 6\frac{5}{36} + 7\frac{6}{36} + 8\frac{5}{36} \cdots + 12\frac{1}{36} = 7.$$

Ak umiestnime body s hmotnosťami $\frac{1}{36}, \ldots, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \ldots, \frac{1}{36}$ do intervalu $[0, \infty)$ tak, aby ich súradnice boli $2, 3, \ldots, 12$, potom ťažisko tejto sústavy bodov bude mať súradnicu 7, v súlade s poznámkou 4.8.

[4] Strednú hodnotu je možné definovať aj pre všeobecnú náhodnú premennú, vyžaduje si to ale hlbšie technické základy. V týchto skriptách definujeme strednú hodnotu osobitne pre diskrétne náhodné premenné, pre takzvané „spojité" náhodné premenné, a nedefinujeme vôbec pre náhodné premenné, ktoré nie sú ani diskrétne, ani spojité. Analogicky postupujeme v prípade rozptylu náhodných premenných.

[5] Čiže každá náhodná premenná, ktorá nadobúda len konečne veľa hodnôt, má konečnú strednú hodnotu.

Poznámka 4.10. Všimnime si, že ak $X_1, \dots, X_n$ sú diskrétne náhodné premenné a $g : \mathbb{R}^n \to \mathbb{R}$ je akákoľvek funkcia, tak $g(X_1, \dots, X_n)$ je tiež diskrétna náhodná premenná, pretože jej obor hodnôt musí byť spočítateľný.

Veta 4.11 (Stredná hodnota funkcie diskrétnej náhodnej premennej)**.** Nech X je diskrétna náhodná premenná a nech $g : \mathbb{R} \to \mathbb{R}$ je funkcia. Potom stredná hodnota diskrétnej náhodnej premennej[6] $g(X)$ je konečná, ak rad $\sum_{x \in X(\Omega)} g(x)P[X = x]$ absolútne konverguje. V takom prípade platí:

$$E(g(X)) = \sum_{x \in X(\Omega)} g(x)P[X = x].$$

Dôkaz. Označme $Y = g(X)$. Potom $P[Y = y] = \sum_{x:g(x)=y} P[X = x]$, kde suma prebieha cez tie $x \in X(\Omega)$, pre ktoré $g(x) = y$. Takže[7]

$$\begin{aligned} E(g(X)) &= E(Y) = \sum_{y \in Y(\Omega)} yP[Y = y] = \sum_{y \in Y(\Omega)} y \sum_{x:g(x)=y} P[X = x] \\ &= \sum_{y \in Y(\Omega)} \sum_{x:g(x)=y} yP[X = x] = \sum_{y \in Y(\Omega)} \sum_{x:g(x)=y} g(x)P[X = x] \\ &= \sum_{x \in X(\Omega)} g(x)P[X = x]. \end{aligned}$$

□

Veta 4.12 (Linearita strednej hodnoty diskrétnej náhodnej premennej)**.** Nech X a Y sú diskrétne náhodné premenné, ktoré majú konečnú strednú hodnotu. Nech a, b sú reálne čísla. Potom aj diskrétna náhodná premenná[8] $aX + bY$ má konečnú strednú hodnotu a platí

$$E(aX + bY) = aE(X) + bE(Y).$$

Špeciálne, $E(aX) = aE(X)$ a $E(X + Y) = E(X) + E(Y)$.

Dôkaz. Z definície strednej hodnoty a aditivity (prípadne σ-aditivity) pravdepodobnosti dostávame:[9]

$$\begin{aligned} E(aX + bY) &= \sum_{z \in (aX+bY)(\Omega)} zP[aX + bY = z] \\ &= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} (ax + by)P[X = x, Y = y] \\ &= a \sum_{x \in X(\Omega)} x \sum_{y \in Y(\Omega)} P[X = x, Y = y] + b \sum_{y \in Y(\Omega)} y \sum_{x \in X(\Omega)} P[X = x, Y = y] \\ &= a \sum_{x \in X(\Omega)} xP[X = x] + b \sum_{y \in Y(\Omega)} yP[Y = y] = aE(X) + bE(Y). \end{aligned}$$

[6] Pozri poznámku 4.10.

[7] Samotnú existenciu konečnej $E(Y)$, čiže absolútnu konvergenciu radu $\sum_{y \in Y(\Omega)} yP[Y = y]$, vieme dokázať z predokladu vety 4.11 o absolútnej konvergencii radu $\sum_{x \in X(\Omega)} g(x)P[X = x]$. Tento dôkaz existencie nechávame na čitateľa.

[8] Pozri poznámku 4.10.

[9] Podobne ako v dôkaze vety 4.11, ani tu nebudeme podrobne dokazovať samotnú existenciu konečnej $E(aX + bY)$, čiže absolútnu konvergenciu radu $\sum_{z \in (aX+bY)(\Omega)} zP[aX + bY = z]$. Táto existencia plynie z predpokladu existencie konečných stredných hodnôt náhodných premenných X a Y. Podrobnosti nechávame na usilovného čitateľa.

(V prípade, že by predchádzajúce rovnosti súm boli nejasné, je inštruktívne si ich platnosť premyslieť pre špeciálny prípad, napríklad taký, že $X(\Omega) = Y(\Omega) = \{0, 1\}$, $a = b = 1$.) □

Príklad 4.13. Náhodná premenná X z príkladu 4.3 sa dá napísať ako súčet náhodných premenných X_1 a X_2, ktoré zodpovedajú výsledkom na prvej, resp. na druhej kocke. Zrejme $E(X_1) = E(X_2) = 1\frac{1}{6} + \cdots + 6\frac{1}{6} = 3.5$, takže podľa vety 4.12 platí $E(X) = E(X_1) + E(X_2) = 7$. To je podstatne kratší výpočet ako v príklade 4.9. Ešte užitočnejšie môže byť využitie vety 4.12, keď sčítavané náhodné premenné nie sú nezávislé; pozri časť 4.3.

Poznámka 4.14. Z vety 4.12 plynie, že pre diskrétne náhodné premenné $X_1, \ldots, X_n$, ktoré majú konečnú strednú hodnotu a pre akékoľvek reálne konštanty $a_1, \ldots, a_n$ platí, že diskrétna náhodná premenná[10] $\sum_{i=1}^{n} a_i X_i$ má konečnú strednú hodnotu a $E(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i E(X_i)$. Špeciálne, $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$.

Definícia 4.15. Nech X je diskrétna náhodná premenná a nech náhodná premenná X^2 má konečnú strednú hodnotu. Potom hovoríme, že náhodná premenná X má konečný rozptyl[11] a číslo

$$D(X) = E\left((X - E(X))^2\right)$$

nazývame **rozptyl** náhodnej premennej X. Ak náhodná premenná X^2 nemá konečnú strednú hodnotu, tak hovoríme, že X má nekonečný rozptyl.

Poznámka 4.16. Predchádzajúca definícia implicitne využíva nasledovné tvrdenie, ktoré nebudeme dokazovať: ak má X^2 konečnú strednú hodnotu, tak samotná náhodná premenná X má tiež konečnú strednú hodnotu a náhodná premenná $(X - E(X))^2$ má taktiež konečnú strednú hodnotu.[12]

Poznámka 4.17. Z definície 4.15 okamžite plynie, že $D(X) \geq 0$ pre akúkoľvek diskrétnu náhodnú premennú X, pretože samotná náhodná premenná $(X - E(X))^2$ je nezáporná.

Poznámka 4.18. Podobne ako stredná hodnota náhodnej premennej, aj rozptyl sa dá intuitívne pochopiť pomocou fyzikálnej predstavy. Ak uvažujeme systém hmotných bodov ako v poznámke 4.8, tak rozptyl náhodnej premennej X je úmerný kinetickej energii takéhoto systému, ak by sa točil okolo ťažiska (ktoré má polohu určenú strednou hodnotou) konštantnou uhlovou rýchlosťou.

Príklad 4.19. Vypočítajme rozptyl náhodnej premennej X z príkladu 4.3. Stredná hodnota náhodnej premennej X je 7, ako vieme z príkladov 4.9 a 4.13. Náhodná premenná $(X - E(X))^2$ teda nadobúda hodnoty $(2-7)^2, (3-7)^2, \ldots, (12-7)^2$ s pravdepodobnosťami $\frac{1}{36}, \ldots, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \ldots, \frac{1}{36}$, takže $D(X) = E((X - E(X))^2) = (2-7)^2\frac{1}{36} + \ldots + (6-7)^2\frac{5}{36} + (7-7)^2\frac{6}{36} + (8-7)^2\frac{5}{36} + \cdots + (12-7)^2\frac{1}{36} = \frac{35}{6}$.

Veta 4.20 (Základné vlastnosti rozptylu diskrétnej náhodnej premennej)**.** Nech X je diskrétna náhodná premenná, ktorá má konečný rozptyl. Potom $D(X) = E(X^2) - E(X)^2$. Nech a, b sú reálne čísla. Potom aj diskrétna náhodná premenná $aX + b$ má konečný rozptyl a platí $D(aX + b) = a^2 D(X)$.

[10] Pozri poznámku 4.10.

[11] Často sa v slovenskom jazyku používa na pomenovanie rozptylu aj pojem „disperzia“, niekedy aj „variancia“.

[12] Všimnime si ešte to, že ak X nadobúda konečne veľa hodnôt, tak aj X^2 nadobúda konečne veľa hodnôt, čiže existuje konečná $E(X^2)$, takže existuje aj konečný rozptyl $D(X)$.

Dôkaz. Z definície rozptylu a linearity strednej hodnoty máme: $D(X) = E((X - E(X))^2) = E(X^2 - 2E(X)X + (E(X))^2) = E(X^2) - 2E(X)E(X) + (E(X))^2 = E(X^2) - (E(X))^2$. Dôkaz druhej časti vety je podobne jednoduché cvičenie. □

Poznámka 4.21. Nech X je diskrétna náhodná premenná a nech $(x_i)_{i \in I}$ sú všetky hodnoty, ktoré X nadobúda s nenulovou pravdepodobnosťou. Ak má X konečnú strednú hodnotu, tak z vety 4.11 a predchádzajúcej vety plynie, že rozptyl X môžeme vypočítať podľa ktoréhokoľvek z nasledujúcich dvoch vzorcov

1. $D(X) = \sum_{i \in I} (x_i - E(X))^2 P[X = x_i]$,
2. $D(X) = \sum_{i \in I} x_i^2 P[X = x_i] - (E(X))^2$,

ak sú príslušné sumy konečné.

Príklad 4.22. Výpočet rozptylu v príklade 4.19 využíva prvý vzorec z poznámky 4.21. Výpočet podľa druhého uvedeného vzorca je $D(X) = (2^2 \frac{1}{36} + \ldots + 6^2 \frac{5}{36} + 7^2 \frac{6}{36} + 8^2 \frac{5}{36} + \cdots + 12^2 \frac{1}{36}) - 7^2 = \frac{35}{6}$.

Definícia 4.23. Ak má diskrétna náhodná premenná X konečný rozptyl $D(X)$, tak číslo $\sigma(X) = \sqrt{D(X)}$ nazývame **smerodajná odchýlka** náhodnej premennej X.

Poznámka 4.24. Všimnime si, že pokým rozptyl reaguje na lineárne transformácie kvadratickou zmenou, smerodajná odchýlka len lineárnou zmenou (formálne $\sigma(aX + b) = a\sigma(X)$ pre $a, b \in \mathbb{R}$, $a \geq 0$). Inými slovami, smerodajnú odchýlku môžeme „vyjadriť v rovnakých jednotkách“ ako samotnú premennú X. Smerodajná odchýlka je preto v niektorých situáciách interpretačne prirodzenejšia charakteristika „šírky rozdelenia“ náhodnej premennej X.

Poznámka 4.25. Stredná hodnota a rozptyl náhodných premenných sú špeciálne prípady číselných charakteristík nazývaných „momenty“. Presnejšie, pre $k \in \mathbb{N}$ definujeme „počiatočný“ moment k-teho rádu ako $E(X^k)$ a, v prípade že existuje $\mu = E(X)$, „centrálny“ moment k-teho rádu ako $E((X - \mu)^k)$. Stredná hodnota náhodnej premennej je teda jej počiatočný moment prvého rádu a rozptyl je jej centrálny moment druhého rádu. Dá sa ukázať, že existencia a konečnosť momentu k-teho rádu implikuje existenciu a konečnosť momentov nižších rádov. Momenty vyšších rádov ako druhého sa v pravdepodobnosti používajú napríklad na definíciu pojmov „šikmosti“ a „špicatosti“ náhodnej premennej alebo tým spôsobom, že ich existencia a konečnosť zaručuje platnosť niektorých pokročilejších teoretických tvrdení. V tomto texte sa nebudeme zaoberať momentami rádov vyšších ako 2.

Definícia 4.26. Nech X je diskrétna náhodná premenná a nech $\sum_{x \in X(\Omega)} P[X = x] \log_2(P[X = x])$ konverguje.[13] Potom hovoríme, že náhodná premenná X má konečnú entropiu $H(X)$ a číslo

$$H(X) = -\sum_{x \in X(\Omega)} P[X = x] \log_2(P[X = x]) \tag{4.1}$$

nazývame **entropia** náhodnej premennej X. Entropiu meriame v bitoch.

[13] Symbolom $\log_2$ označujeme binárny logaritmus a pre $p = 0$ kladieme $p \log_2(p) = 0$. Všimnime si tiež, že všetky sčítance uvedeného radu sú nekladné, takže konvergencia tohto radu je ekvivalentná jeho absolútnej konvergencii. Ak tento rad nekonverguje (čiže „konverguje do $-\infty$“), hovoríme, že X má nekonečnú entropiu.

Poznámka 4.27. Všimnime si, že entropia nezávisí od toho, aké hodnoty X nadobúda, len od toho, aké sú pravdepodobnosti nadobudnutia jednotlivých hodnôt. To znamená, že entropiu by sme mohli rovnako definovať aj pre „náhodnú premennú“, ktorá nenadobúda reálne čísla, ale hodnoty z nejakej konečnej (alebo spočítateľne nekonečnej) „abecedy“; pozri tiež poznámku 3.13.

Príklad 4.28 (Rozšírený príklad 1.1.2 z [4])**.** V dostihových pretekoch bude súťažiť 8 koní. Pravdepodobnosti výhry jednotlivých koní sú $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}$. a) Nájdime entropiu náhodnej premennej X, ktorá reprezentuje číslo víťazného koňa (každý kôň má iné číslo). b) Čísla koní sme sa rozhodli zakódovať nasledovnými postupnosťami: 0, 10, 110, 1110, 111100, 111101, 111110 a 111111, v uvedenom poradí podľa najvyššej pravdepodobnosti výhry. Nájdime strednú hodnotu náhodnej premennej Y, ktorá reprezentuje počet bitov (čiže dĺžku) kódu víťazného koňa.

Riešenie: a) $H(X) = -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{4}\log_2(\frac{1}{4}) - \frac{1}{8}\log_2(\frac{1}{8}) - \frac{1}{16}\log_2(\frac{1}{16}) - 4\frac{1}{64}\log_2(\frac{1}{64}) = 2$ bity. b) $E(Y) = 1\frac{1}{2} + 2\frac{1}{4} + 3\frac{1}{8} + 4\frac{1}{16} + 6\frac{1}{16} = 2$.

Veta 4.29 (Základná vlastnosť entropie)**.** Nech X je diskrétna náhodná premenná, ktorá nadobúda n rôznych hodnôt. Potom má X konečnú entropiu a platí $0 \leq H(X) \leq \log_2(n)$. Ak X nadobúda len jedinú hodnotu, tak $H(X) = 0$. Ak X nadobúda každú z n hodnôt s rovnakou pravdepodobnosťou, tak $H(X) = \log_2(n)$.

Dôkaz. Nerovnosť $0 \leq H(X)$ plynie priamo z definície entropie a toho, že $-p\log_2(p) \geq 0$ pre všetky $p \in [0, 1]$. Stručne ukážme, že $H(X) \leq \log_2(n)$.

V matematickej analýze sa dokazuje takzvaná „Jensenova“ nerovnosť,[14] ktorá sa dá v reči diskrétnych náhodných premenných formulovať nasledovne: Nech je g reálna funkcia, konkávna na nejakom intervale $[a, b]$. Nech diskrétna náhodná premenná Z nadobúda hodnoty v intervale $[a, b]$. Potom $E(g(Z)) \leq g(E(Z))$.

Nech $X(\Omega) = \{x_1, \ldots, x_n\}$. Označme $p_i = P[X = x_i]$ pre všetky $i = 1, \ldots, n$. Uvažujme náhodnú premennú Z, ktorá nadobúda hodnoty $\frac{1}{np_1}, \ldots, \frac{1}{np_n}$ s pravdepodobnosťami $p_1, \ldots, p_n$.[15] Máme

$$H(X) - \log_2(n) = \sum_{i=1}^{n} p_i \log_2\left(\frac{1}{np_i}\right) = E(\log_2(Z)) \leq \log_2(E(Z)) = \log_2(1) = 0.$$

Prvá rovnosť plynie z definície entropie a zo základných vlastností logaritmu, druhá rovnosť je dôsledkom vety 4.11, nerovnosť nám zaručuje Jensen (pretože $\log_2$ je konkávna funkcia na nejakom konečnom intervale obsahujúcom obor hodnôt náhodnej premennej Z) a predposledná rovnosť plynie z toho, že $E(Z) = 1$.

Zostávajúce časti vety plynú priamo z definície entropie. □

Poznámka 4.30. Entropia diskrétnej náhodnej premennej X vyjadruje „neurčitosť“ budúcej realizácie premennej X. Túto predstavu podporujú základné vlastnosti z vety 4.29. Skutočne; neurčitosť by mala byť nezáporná charakteristika a nulovú hodnotu by mala nadobúdať pre

[14] Jensenovu nerovnosť tu nebudeme dokazovať; jej dôkaz je bežne dostupný, principiálne jednoduchý, no technický, pozri napríklad časť 2.6 v [4]. Ide ale o veľmi zaujímavé a užitočné tvrdenie pre pokročilejšiu teóriu pravdepodobnosti.

[15] Mlčky tu predpokladáme, že $\frac{1}{np_1}, \ldots, \frac{1}{np_n}$ sú rôzne čísla; modifikovať dôkaz pre prípad, že niektoré z týchto hodnôt sú rovnaké, je jednoduché, ale formálne zdĺhavejšie.

každú degenerovanú náhodnú premennú, čiže takú, ktorej výsledok je vopred istý. Intuitívne je tiež zrejmé, že najväčšiu neurčitosť spomedzi náhodných premenných nadobúdajúcich n rôznych hodnôt by mala mať taká náhodná premenná, ktorá nadobúda každú z týchto hodnôt s rovnakou pravdepodobnosťou.[16] Entropia je kľúčovou charakteristikou v takzvanej „teórii informácie" a v príbuznej „teórii kódovania". Pomocou výsledkov z týchto oblastí je napríklad možné ukázať, že vlastnosť $H(X) = E(Y)$ v príklade 4.28 znamená, že použitý kód je najúspornejší možný rozumný kód. Uvedený kód je tiež najefektívnejší z pohľadu opakovaného generovania realizácie náhodnej premennej X pomocou nezávislých hodov vyváženou mincou.[17] Pomocou entropie je taktiež možné definovať prirodzenú číselnú mieru informácie, ktorú nesie jedna náhodná premenná o druhej náhodnej premennej; to už však presahuje rámec týchto skrípt.

4.2 Typy diskrétnych náhodných premenných

4.2.1 Diskrétne rovnomerné rozdelenie

Definícia 4.31. Nech a, b sú celé čísla, $a \leq b$. Označme $M = \{a, a+1, \dots, b\}$. Hovoríme, že diskrétna náhodná premenná X má **diskrétne rovnomerné rozdelenie** na množine M, ak $P[X = x] = \frac{1}{n}$ pre všetky $x \in M$, kde $n = b - a + 1$ je počet prvkov množiny M. Túto skutočnosť značíme $X \sim R(M)$.

Veta 4.32 (Základné číselné charakteristiky diskrétneho rovnomerného rozdelenia[18])**.** Nech $X \sim R\{a, a+1, \dots, b\}$,[19] $a \leq b$, $n = b - a + 1$. Potom $E(X) = \frac{a+b}{2}$, $D(X) = \frac{n^2-1}{12}$ a $H(X) = \log_2(n)$.

Dôkaz. Dôkaz je jednoduché cvičenie,[20] resp. plynie z vety 4.29. □

Príklad 4.33. Náhodná premenná $X \sim R\{0, 1\}$ zodpovedá výsledku hodu vyváženou mincou alebo rovnomernému náhodnému výberu bitu. Náhodná premenná $X \sim R\{1, 2, \dots, 6\}$ zodpovedá výsledku hodu vyváženou hracou kockou. Náhodná premenná $X \sim R\{0, 1, \dots, 9\}$ zodpovedá rovnomernému náhodnému výberu dekadickej cifry (pozri obrázok 4.1).

Poznámka 4.34. Náhodná premenná $X \sim R\{a\}$ (čiže $a = b$ v definícii 4.31) má „singulárne" rozdelenie v bode a. Hoci ide formálne o náhodnú premennú, jej budúca realizácia je vopred istá. (S takouto náhodnou premennou sme sa už stretli vo vete 4.29.)

Poznámka 4.35. Diskrétne rovnomerné rozdelenie by sme mohli prirodzeným spôsobom definovať aj na všeobecnej konečnej množine M. Takéto všeobecnejšie diskrétne rozdelenia sú špeciálnym prípadom „empirického" rozdelenia dát (pozri časť 4.2.7).

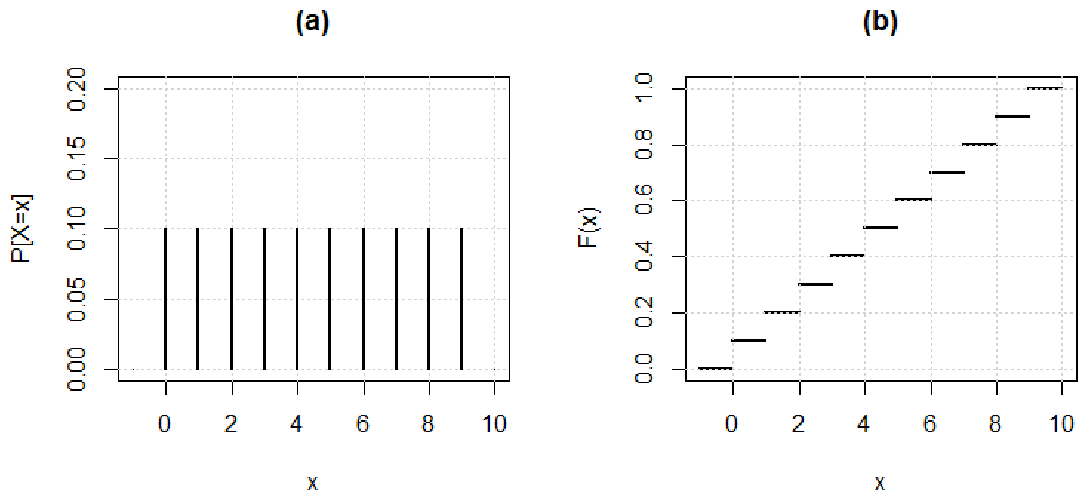
[16] Takéto náhodné premenné podrobnejšie študujeme v časti 4.2.1.

[17] Hádžeme mincou dovtedy, kým postupnosť výsledkov na minci nebude zodpovedať niektorému z daných ôsmych kódov.

[18] Presnejší zvrat by bol „číselné charakteristiky náhodnej premennej s diskrétnym rovnomerným rozdelením". Uvedomme si však, že naozaj ide v podstate o charakteristiky rozdelenia ako takého, pretože dve náhodné premenné s rovnakým rozdelením – pozri poznámku 3.6 – majú rovnakú strednú hodnotu, rozptyl aj entropiu.

[19] Namiesto formálne presného zápisu $R(\{\dots\})$ budeme používať stručnejší $R\{\dots\}$.

[20] Pripomeňme, že $\sum_{i=0}^{n-1} i^2 = n(n-1)(2n-1)/6$.



Obr. 4.1: Pravdepodobnosti $P[X = x]$ a distribučná funkcia náhodnej premennej s rovnomerným rozdelením na množine $\{0, 1, \ldots, 9\}$.

4.2.2 Alternatívne rozdelenie

Definícia 4.36. Hovoríme, že diskrétna náhodná premenná X má **alternatívne rozdelenie** s parametrom $p \in (0, 1)$, ak $P[X = 0] = 1 - p$ a $P[X = 1] = p$. Túto skutočnosť značíme $X \sim Alt(p)$.[21]

Príklad 4.37. Hádžeme falošnou mincou, na ktorej padá znak s pravdepodobnosťou $p \neq \frac{1}{2}$. Potom náhodná premenná

$$X = \begin{cases} 1 & \text{ak padne znak,} \\ 0 & \text{ak padne hlava} \end{cases}$$

má alternatívne rozdelenie $X \sim Alt(p)$.

Veta 4.38 (Základné číselné charakteristiky alternatívneho rozdelenia)**.** Nech $X \sim Alt(p)$. Potom $E(X) = p$, $D(X) = p(1 - p)$, $H(X) = -p \log_2(p) - (1 - p) \log_2(1 - p)$.

Dôkaz. Dôkaz je jednoduché cvičenie. □

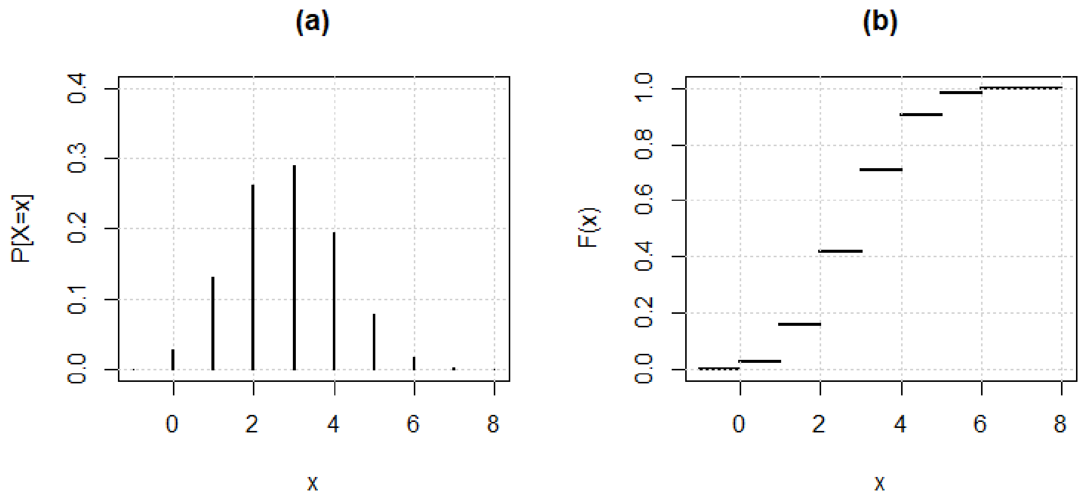
4.2.3 Binomické rozdelenie

Definícia 4.39. Hovoríme, že diskrétna náhodná premenná X má **binomické rozdelenie** s parametrami $p \in (0, 1)$ a $n \in \mathbb{N}$, ak pre každé $k = 0, 1, \ldots, n$ platí

$$P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Túto skutočnosť značíme $X \sim Bin(n, p)$.

[21] Alternatívne rozdelenie sa niekedy nazýva aj „Bernoulliho“ rozdelenie alebo „nula-jednotkové“ rozdelenie. Všimnime si, že rozdelenie $Alt(\frac{1}{2})$ je rovnaké ako rozdelenie $R\{0, 1\}$ z časti 4.2.1.



Obr. 4.2: Pravdepodobnosti $P[X = x]$ a distribučná funkcia náhodnej premennej s rozdelením $Bin(7, 0.4)$.

Poznámka 4.40. Z binomickej formuly máme $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1$, takže $X \sim Bin(n, p)$ nadobúda s nenulovou pravdepodobnosťou *len* hodnoty $\{0, 1, \ldots, n\}$.

Poznámka 4.41. Uvažujme sériu n nezávislých pokusov, z ktorých každý môže skončiť buď úspechom (s pravdepodobnosťou $p \in (0, 1)$), alebo neúspechom (s pravdepodobnosťou $q = 1 - p$). Náhodná premenná $X \sim Bin(n, p)$ potom zodpovedá celkovému počtu úspešných pokusov.[22]

Príklad 4.42. Hádžeme mincou z príkladu 4.37, no nie raz, ale n-krát. Definujme náhodnú premennú X ako počet znakov, ktorý padne v týchto n hodoch. Potom X má binomické rozdelenie $X \sim Bin(n, p)$.

Veta 4.43 (Základné číselné charakteristiky binomického rozdelenia)**.** Nech $X \sim Bin(n, p)$. Potom $E(X) = np$ a $D(X) = np(1 - p)$.[23]

Dôkaz. Nech $X \sim Bin(n, p)$, $q = 1 - p$. Platí

$$E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^{n} k \frac{n!}{(n-k)!k!} p^k q^{n-k} =$$

$$np \sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} q^{n-k} = np \sum_{i=0}^{n-1} \frac{(n-1)!}{((n-1)-i)!i!} p^i q^{(n-1)-i} =$$

[22] S binomickým rozdelením sme sa implicitne stretli už vo vete 2.21 a v nadväzujúcich príkladoch.

[23] Entropiu náhodnej premennej s binomickým rozdelením nie je možné vyjadriť oveľa jednoduchším zápisom, ako je ten, ktorý je priamočiarym prepisom definície entropie. Podobné konštatovanie platí pre entropiu náhodnej premennej s Poissonovým a hypergeometrickým rozdelením z ďalších častí tejto kapitoly, preto príslušné vzorčeky nebudeme uvádzať.

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

Posledná rovnosť plynie z binomického rozvoja súčtu $(p+q)^{n-1}$. Podobne odvodíme[24] $E(X(X-1)) = \sum_{k=2}^{n} k(k-1)\binom{n}{k}p^k q^{n-k} = n(n-1)p^2$, a preto $D(X) = E(X^2) - (E(X))^2 = E(X(X-1)) + E(X) - (E(X))^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)$. □

Poznámka 4.44. Všimnime si, že $X \sim Alt(p)$ vtedy a len vtedy, keď $X \sim Bin(1,p)$.

Poznámka 4.45. Z interpretácie v poznámke 4.41 plynie, že súčet „nezávislých“[25] náhodných premenných $X_i \sim Alt(p)$, $i = 1,\dots,n$, má rozdelenie $Bin(n,p)$. Táto úvaha (spolu s vetou 4.38 a poznámkou 4.14) nám poskytuje alternatívny dôkaz, že $E(X) = np$ pre $X \sim Bin(n,p)$.

Príklad 4.46. Pre potreby „genetického algoritmu“ modelujeme „chromozóm dĺžky n“ postupnosťou n binárnych hodnôt 0 alebo 1. Nech x je chromozóm pozostávajúci z k jednotiek a $n-k$ núl. Chromozóm y vytvoríme z chromozómu x náhodnou „mutáciou“, t.j. tak, že každý bit preklopíme na opačný s pravdepodobnosťou p. Nájdime strednú hodnotu počtu jednotiek, ktoré bude obsahovať chromozóm y.

Riešenie: Z interpretácie binomického rozdelenia je zrejmé, že počet $N_{0\to1}$ nulových bitov chromozómu x, ktoré sa zmenia na jednotku, má rozdelenie $Bin(n-k,p)$ a počet $N_{1\to1}$ jednotkových bitov chromozómu x, ktoré sa nezmenia, má rozdelenie $Bin(k,1-p)$. Vidíme, že pre počet N jednotkových bitov chromozómu y platí $N = N_{0\to1} + N_{1\to1}$, a teda na základe linearity strednej hodnoty a vety 4.43 dostávame:

$$E(N) = E(N_{0\to1}) + E(N_{1\to1}) = (n-k)p + k(1-p).$$

Príklad 4.47. Hráme hru, v ktorej sa ťahajú 4 čísla zo 40. Vyhráme, ak uhádneme všetky štyri čísla. Určme pravdepodobnosť, že ak sa hry zúčastníme 500 krát, vyhráme raz alebo dvakrát.

Riešenie: Pravdepodobnosť výhry v jednej hre je $p = 1/\binom{40}{4} \approx 1.1\times10^{-5}$. Potom náhodná premenná X, ktorá označuje počet výhier v 500 hrách, má rozdelenie $Bin(500,p)$ a

$$P[X=1] + P[X=2] = \binom{500}{1}p(1-p)^{499} + \binom{500}{2}p^2(1-p)^{498} \approx 0.00546.$$

V takomto prípade, čiže keď je n „veľké“ a p „veľmi malé“, môže byť výhodnejšie aproximovať binomické rozdelenie nasledovným spôsobom.

Veta 4.48 (Poissonova limitná veta)**.** Majme postupnosť náhodných premenných $X_{n_1}, X_{n_1+1},\dots$ pričom $X_n \sim Bin(n,\lambda/n)$, kde $0 < \lambda < n_1$. Potom pre každé $k \in \{0,1,2,\dots\}$ platí

$$\lim_{n\to\infty} P[X_n = k] = e^{-\lambda}\frac{\lambda^k}{k!}. \tag{4.2}$$

Dôkaz.

$$\lim_{n\to\infty} P[X_n = k] = \lim_{n\to\infty}\binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} =$$

[24] Technicky použijeme vetu 4.11 s $g(x) = x(x-1)$.

[25] S pojmom nezávislosti náhodných premenných sa viac zoznámime v kapitole 6.

$$\frac{\lambda^k}{k!}\lim_{n\to\infty}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{-k}\left(1-\frac{\lambda}{n}\right)^{n}.$$

Avšak zrejme platí

$$\lim_{n\to\infty}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{-k}=1,$$

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n}=e^{-\lambda}.$$

Z predchádzajúcich rovností dostávame požadované tvrdenie. □

Poznámka 4.49. Akej veľkej chyby sa dopustíme, ak na základe vzťahu (4.2) nahradíme hodnotu $P[X_n = k]$ číslom $e^{-\lambda}\lambda^k/k!$ pre konečné n? Dá sa ukázať (pozri napríklad [8], strana 64), že

$$\sum_{k=0}^{\infty}\left|P[X_n=k]-e^{-\lambda}\frac{\lambda^k}{k!}\right|\leq\frac{2\lambda}{n}\min(2,\lambda).$$

.

Príklad 4.50. Predpokladajme, že pri prenose binárnych dát cez určitý komunikačný kanál dochádza k preklopeniu bitu na opačný s pravdepodobnosťou $p = 10^{-6}$. Udalosti preklopenia bitu na opačný sa vyskytujú navzájom nezávisle a k iným typom chyby prenosu dát nedochádza. Odhadnime pravdepodobnosť, že pri prenose 10^6 bitov dôjde k viac ako dvom chybám prenosu.

Riešenie: Počet X preklopení bitov na opačný má rozdelenie $Bin(n,p)$, kde $n = 10^6$ a $p = 10^{-6}$. Podľa Poissonovej limitnej vety platí $P[X \in \{0,1,2\}] \approx \sum_{i=0}^{2} e^{-\lambda}\lambda^k/k!$, kde $\lambda = np = 1$, numericky $P[X \in \{0,1,2\}] \approx 0.9196986$. Hľadaná pravdepodobnosť je teda $P[X > 2] = 1 - P[X \in \{0,1,2\}] \approx 0.0803014$. Podľa poznámky 4.49 je chyba, ktorej sme sa dopustili „Poissonovou aproximáciou“, rádovo maximálne na úrovni 10^{-6}.

4.2.4 Poissonovo rozdelenie

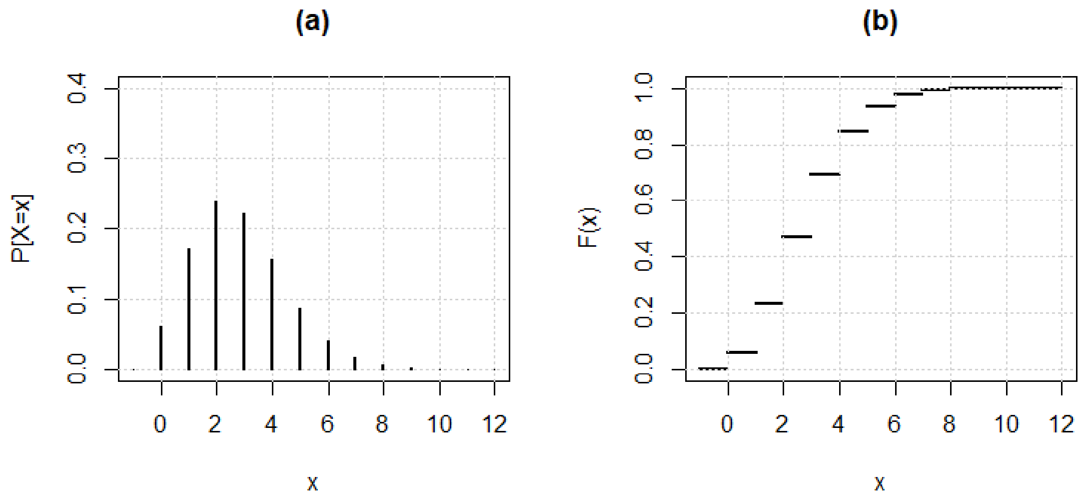
Definícia 4.51. Hovoríme, že diskrétna náhodná premenná X má **Poissonovo rozdelenie** s parametrom $\lambda > 0$, ak pre každé $k = 0, 1, 2, \ldots$ platí

$$P[X=k]=e^{-\lambda}\frac{\lambda^k}{k!}.$$

Túto skutočnosť značíme $X \sim Po(\lambda)$.

Poznámka 4.52. Všimnime si, že $\sum_{k=0}^{\infty} e^{-\lambda}(\lambda^k/k!) = 1$, čo plynie z už spomínaného slávneho vzťahu $\sum_{k=0}^{\infty}\lambda^k/k! = e^{\lambda}$. Takže $X \sim Po(\lambda)$ nadobúda s nenulovou pravdepodobnosťou *len* celé nezáporné čísla.

Veta 4.53 (Základné číselné charakteristiky Poissonovho rozdelenia)**.** Nech $X \sim Po(\lambda)$. Potom $E(X) = \lambda$ a $D(X) = \lambda$.



Obr. 4.3: Pravdepodobnosti $P[X = x]$ a distribučná funkcia náhodnej premennej s rozdelením $Po(2.8)$.

Dôkaz.

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

Podobne máme

$$E(X(X-1)) = \sum_{k=0}^{\infty} k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = e^{-\lambda} \lambda^2 e^{\lambda} = \lambda^2.$$

Preto $D(X) = E(X^2) - (E(X))^2 = E(X(X-1)) + E(X) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$ □

Poznámka 4.54. Poissonovo rozdelenie sa používa napríklad na modelovanie počtu rozpadov atómov rádioaktívnej látky za určitý čas, volaní na telefónnu ústredňu, impulzov prichádzajúcich na neurónovú bunku počas určitého časového úseku a podobne. Poissonovo rozdelenie je tiež „limitným rozdelením“ niektorých postupností náhodných premenných; iný príklad aproximácie, než predstavuje veta 4.48, je uvedený v nasledujúcom príklade.

Príklad 4.55. Nech $n \in \mathbb{N}$ a nech X_n znamená počet prvkov, ktoré zostanú na svojom pôvodnom mieste po dokonalej náhodnej permutácii postupnosti $(1, \ldots, n)$. Dokážeme, že limitné rozdelenie náhodných premenných X_n je $Po(1)$, t.j. že $\lim_{n\to\infty} P[X_n = k] = e^{-1}/k!$ pre každé nezáporné celé číslo k.

Riešenie: Najprv si uvedomme, že $p_n = P[X_n > 0]$ sme už určili v príklade 1.45. Položme $q_n = P[X_n = 0] = 1 - p_n$, t.j. q_n je pravdepodobnosť, že po náhodnej permutácii n čísel nezostane ani jedno z nich na svojom pôvodnom mieste. Dodefinujme tiež $q_0 = 1$. Pre všeobecné $k \geq 1$ máme

$$P[X_n = k] = P\left(\bigcup_{1 \leq i_1 < \ldots < i_k \leq n} B_{i_1,\ldots,i_k} \right),$$

kde $B_{i_1,\dots,i_k}$ znamená udalosť, že čísla $i_1, \dots, i_k$ zostanú na svojom pôvodnom mieste a žiadne iné číslo nezostane na svojom pôvodnom mieste. Je zrejmé, že takýchto udalostí je $\binom{n}{k}$, že sú navzájom disjunktné a pravdepodobnosť každej takejto udalosti je

$$P(B_{i_1,\dots,i_k}) = \frac{q_{n-k}}{n(n-1)\cdots(n-k+1)}.$$

Spojením týchto výsledkov dostávame

$$P[X_n = k] = \binom{n}{k} \frac{q_{n-k}}{n(n-1)\dots(n-k+1)} = \frac{q_{n-k}}{k!}.$$

Použitím riešenia príkladu 1.45 máme

$$\lim_{n\to\infty} P[X_n = k] = \lim_{n\to\infty} \frac{q_{n-k}}{k!} = \lim_{n\to\infty} \frac{1 - p_{n-k}}{k!} = e^{-1}/k!.$$

4.2.5 Geometrické rozdelenie

Definícia 4.56. Hovoríme, že diskrétna náhodná premenná X má **geometrické rozdelenie** s parametrom $p \in (0,1)$, ak pre každé $k = 0, 1, 2, \dots$ platí

$$P[X = k] = p(1-p)^k.$$

Túto skutočnosť značíme $X \sim Geo(p)$.

Poznámka 4.57. Všimnime si, že $\sum_{k=0}^{\infty} p(1-p)^k = 1$, čo plynie zo známeho vzťahu pre súčet členov nekonečnej geometrickej postupnosti. Takže $X \sim Geo(p)$ nadobúda s nenulovou pravdepodobnosťou *len* celé nezáporné čísla.

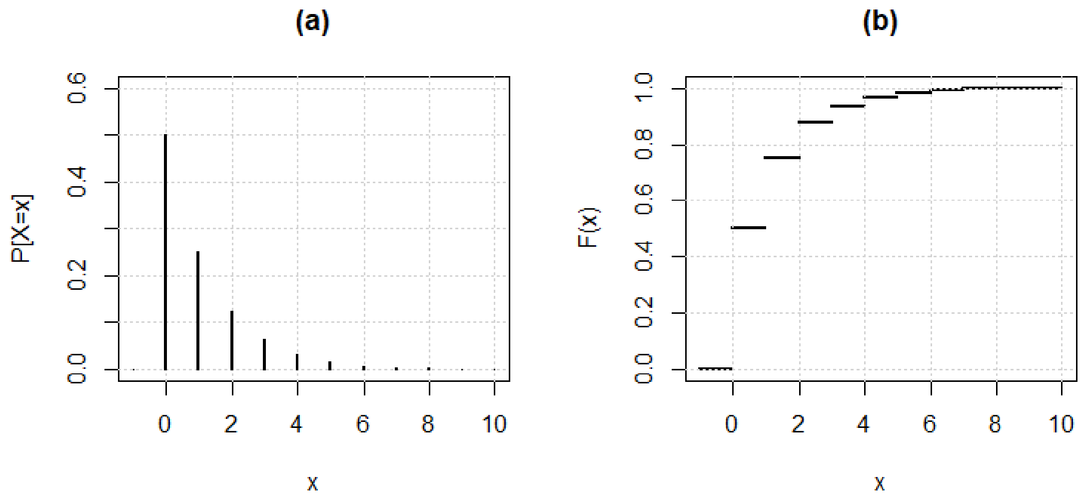
Poznámka 4.58. Uvažujme (nekonečne dlhú) sériu nezávislých pokusov, pričom každý jednotlivý pokus končí „úspechom“ s pravdepodobnosťou p a neúspechom s pravdepodobnosťou $q = 1 - p$. Náhodná premenná X s rozdelením $Geo(p)$ potom zodpovedá počtu neúspešných pokusov predchádzajúcich prvý úspešný pokus.

Príklad 4.59. Hádžeme falošnou mincou, na ktorej padá znak s pravdepodobnosťou p. Touto mincou budeme hádzať až dovtedy, kým nám prvýkrát nepadne znak. Potom náhodná premenná, ktorá znamená počet padnutí hlavy (alebo počet všetkých hodov mínus 1) má geometrické rozdelenie s parametrom p.

Veta 4.60 (Základné číselné charakteristiky geometrického rozdelenia)**.** Nech $X \sim Geo(p)$. Potom $E(X) = (1-p)/p$, $D(X) = (1-p)/p^2$ a $H(X) = -\frac{1}{p}(p\log_2(p) + (1-p)\log_2(1-p))$.

Dôkaz. Pre každé $q \in (0,1)$ platí $f(q) = \sum_{k=0}^{\infty} q^k = (1-q)^{-1}$. Derivovaním funkcie f dostávame jednak $f'(q) = \sum_{k=0}^{\infty} kq^{k-1}$, ako aj $f'(q) = (1-q)^{-2}$. Preto $\sum_{k=0}^{\infty} kq^{k-1} = (1-q)^{-2}$. Takže

$$E(X) = \sum_{k=0}^{\infty} kp(1-p)^k = p(1-p)\sum_{k=0}^{\infty} k(1-p)^{k-1} = p(1-p)p^{-2} = (1-p)/p.$$



Obr. 4.4: Pravdepodobnosti $P[X = x]$ a distribučná funkcia náhodnej premennej s rozdelením $Geo(0.5)$.

Dvojnásobným derivovaním funkcie f dostaneme vzorec pre súčet nekonečného radu, pomocou ktorého odvodíme rozptyl veľmi podobne, ako sme vypočítali strednú hodnotu. Ukážme ešte, ako je možné vypočítať entropiu. Označme $q = 1 - p$.

$$H(X) = -\sum_{k=0}^{\infty} pq^k \log_2(pq^k) = -\log_2(p)\sum_{k=0}^{\infty} pq^k - \log_2(q)\sum_{k=0}^{\infty} kpq^k = -\log_2(p) - \log_2(q)E(X),$$

pretože $\sum_{k=0}^{\infty} pq^k = 1$, keďže je to súčet všetkých pravdepodobností typu $P[X = k]$. Výsledný vzťah pre entropiu dostaneme použitím už odvodeného vzťahu pre $E(X)$. □

Príklad 4.61. Kódom PIN môže byť akákoľvek postupnosť štyroch cifier od 0000 po 9999. Predpokladajme, že systém akceptuje kód PIN, ak správne zadáme aspoň 3 zo štyroch cifier (na zodpovedajúcich miestach). Budeme náhodne voliť kódy PIN, až kým náš kód nebude akceptovaný. (Voľbu vykonávame tak, že každú cifru volíme s pravdepodobnosťou 1/10 bez ohľadu na to, čo sme volili predtým.) Aká je stredná hodnota počtu pokusov?

Riešenie: Pravdepodobnosť, že zo štyroch náhodne zadaných cifier uhádneme aspoň tri, je podľa binomickej formuly

$$p = \binom{4}{3}(0.1)^3(0.9)^1 + \binom{4}{4}(0.1)^4(0.9)^0 = 0.0037.$$

Počet „neúspešných pokusov“ X, t.j. odmietnutých kódov PIN, má rozdelenie $Geo(p)$. Preto stredná hodnota počtu $N = X+1$ zadaní kódu PIN, so započítaním aj posledného, úspešného, je podľa vety 4.60:

$$E(N) = E(X) + 1 = \frac{1-p}{p} + 1 = \frac{1}{p} = 10000/37 \approx 270.$$

Príklad 4.62. Predpokladajme, že veľkosť N vstupu algoritmu má rozdelenie $Geo(p) + 1$ a čas T výpočtu algoritmu závisí od veľkosti vstupu podľa vzťahu $T = \alpha N^2 + \beta N + \gamma$, kde $p \in (0, 1)$, $\alpha, \beta, \gamma > 0$ sú známe konštanty. Určme strednú hodnotu času výpočtu T.

Riešenie: Z linearity strednej hodnoty, vety 4.60 a vety 4.20 dostávame $E(T) = \alpha E(N^2) + \beta E(N) + \gamma$, kde $E(N) = 1/p$ a $E(N^2) = D(N) + (E(N))^2 = (1-p)/p^2 + (1/p)^2 = (2-p)/p^2$.

4.2.6 Hypergeometrické rozdelenie

Definícia 4.63. Hovoríme, že diskrétna náhodná premenná X má **hypergeometrické rozdelenie** s parametrami $M, N, n \in \mathbb{N}$, spĺňajúcimi $N < M$ a $n < M$, ak

$$P[X = k] = \frac{\binom{N}{k}\binom{M-N}{n-k}}{\binom{M}{n}} \tag{4.3}$$

pre každé $k = \max(0, n+N-M), \dots, \min(n, N)$. Túto skutočnosť značíme $X \sim Hyp(M, N, n)$.

Poznámka 4.64. Dá sa ukázať, že súčet členov na pravej strane (4.3) pre celé čísla k rozsahu od $\max(0, n + N - M)$ do $\min(n, N)$ je 1, čo plynie z takzvanej „Vandermondeovej" identity. Preto $X \sim Hyp(M, N, n)$ nadobúda s nenulovou pravdepodobnosťou *len* celé čísla od $\max(0, n + N - M)$ do $\min(n, N)$.

Poznámka 4.65. S hypergeometrickým rozdelením sme sa už stretli v príkladoch 1.12, 1.32 a 4.47, hoci sme ešte nepoužívali tento pojem. Predpokladajme, že máme krabicu s M loptičkami, z ktorých je N čiernych. Z krabice (súčasne alebo, ekvivalentne, bez vrátenia) vyberieme n loptičiek. Ak X znamená počet čiernych loptičiek v tomto výbere, tak X má rozdelenie $Hyp(M, N, n)$. Hypergeometrické rozdelenie sa používa napríklad pri pravdepodobnostnej analýze lotérií, v štatistickej kontrole kvality, ale napríklad aj pri subselekcii dátového súboru, ako ukážeme v nasledujúcom príklade.

Príklad 4.66. Dátový súbor obsahuje 90 objektov typu A a 10 objektov typu B. Z tohto súboru vyberieme rovnomerne náhodne 20 objektov ako malú reprezentatívnu vzorku. Vypočítajme pravdepodobnosť, že tento súbor bude obsahovať aspoň 2 objekty typu B.

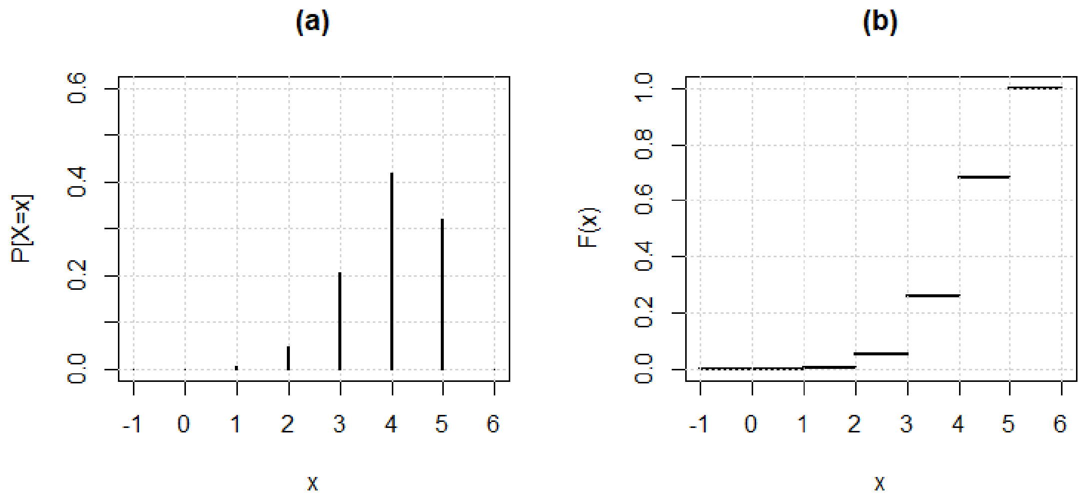
Riešenie: Počet objektov typu B, ktoré sa budú nachádzať v selektovanej podmnožine dát, má hypergeometrické rozdelenie $Hyp(100, 10, 20)$. Pravdepodobnosť, že vo vzorke sa budú nachádzať aspoň 2 objekty typu B, je teda

$$\sum_{k=2}^{10} \frac{\binom{10}{k}\binom{90}{20-k}}{\binom{100}{20}} \approx 0.637.$$

Numericky jednoduchšie je najprv vypočítať hodnotu doplnkovej udalosti, čiže udalosti, že vzorka bude obsahovať maximálne jeden objekt typu B:

$$\sum_{k=0}^{1} \frac{\binom{10}{k}\binom{90}{20-k}}{\binom{100}{20}} \approx 0.363,$$

z čoho tiež dostaneme požadovaný výsledok.



Obr. 4.5: Pravdepodobnosti $P[X = x]$ a distribučná funkcia náhodnej premennej s rozdelením $Hyp(100, 80, 5)$.

Veta 4.67 (Základné číselné charakteristiky hypergeometrického rozdelenia)**.** Nech $X \sim Hyp(M, N, n)$. Potom $E(X) = n\frac{N}{M}$ a $D(X) = n\frac{N}{M}\left(1 - \frac{N}{M}\right)\frac{M-n}{M-1}$.

Dôkaz. Vzťah pre $E(X)$ dokážeme v príklade 4.75. Odvodenie vzťahu pre $D(X)$ nechávame ako cvičenie pre veľmi usilovného študenta (je možné použiť viacero prístupov, napríklad aj prístup využívajúci linearitu strednej hodnoty; pozri techniky v časti 4.3). □

Poznámka 4.68. Zaujímavý postreh je ten, že ak si vo vete 4.67 označíme $p = N/M$, čo je pri klasickej interpretácii pravdepodobnosť výberu čiernej loptičky v prvom výbere, tak $E(X) = np$, rovnako ako pre binomické rozdelenie. Rozptyl je $D(X) = np(1 - p)c$, kde c je číslo blízke jednotke v prípade veľkého M, čo je opäť podobné vzťahu pre rozptyl binomického rozdelenia. V skutočnosti je binomické rozdelenie istým limitným prípadom hypergeometrického rozdelenia, čím sa ale nebudeme podrobnejšie zaoberať.

4.2.7 Empirické rozdelenie

Definícia 4.69. Nech $x_1, \ldots, x_n \in \mathbb{R}$.[26] Hovoríme, že náhodná premenná X má **empirické rozdelenie** určené dátami $x_1, \ldots, x_n$, ak je to diskrétna náhodná premenná a

$$P[X = x] = \frac{N(x)}{n}$$

pre všetky čísla x, ktoré sa vyskytujú v postupnosti $x_1, \ldots, x_n$, pričom $N(x)$ označuje počet výskytov čísla x v postupnosti $x_1, \ldots, x_n$. Túto skutočnosť budeme značiť $X \sim Emp(x_1, \ldots, x_n)$.

[26] Čísla $x_1, \ldots, x_n$ tu nemusia byť všetky rôzne. Všimnime si, že ak tieto čísla sú rôzne a $\{x_1, \ldots, x_n\} = \{a, a + 1, \ldots, b\}$ pre nejaké celé čísla $a \le b$, tak práve definované rozdelenie je rovnomerné rozdelenie z časti 4.2.1.

Veta 4.70 (Základné číselné charakteristiky empirického rozdelenia)**.** Predpokladajme, že $X \sim Emp(x_1, \ldots, x_n)$. Potom

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i,\ D(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - E(X))^2,\ H(X) = -\frac{1}{n}\sum_{i=1}^{n}\log_2 \frac{N(x_i)}{n}.$$

Dôkaz. Dôkaz vzťahu pre $E(X)$ a $D(X)$ plynie priamo z definícií. Vzťah pre entropiu sa dá ukázať napríklad nasledovne. Nech M označuje množinu všetkých tých čísel, ktoré sa aspoň raz vyskytujú v postupnosti $x_1, \ldots, x_n$. Potom z definície entropie a empirického rozdelenia

$$H(X) = -\sum_{x \in M} \frac{N(x)}{n}\log_2 \frac{N(x)}{n} = -\frac{1}{n}\sum_{x\in M}\sum_{i:x_i=x}\log_2 \frac{N(x)}{n} = -\frac{1}{n}\sum_{i=1}^{n}\log_2 \frac{N(x)}{n}.$$

Druhá rovnosť plynie z faktu, že $N(x) = \sum_{i:x_i=x} 1$ pre každé $x \in M$. □

Poznámka 4.71. Predpokladajme, že máme dáta $x_1, \ldots, x_n$, ktoré modelujeme ako realizácie náhodných premenných $X_1, \ldots, X_n$ so spoločným, no neznámym rozdelením.[27] Potom rozdelenie $Emp(x_1, \ldots, x_n)$ môžeme chápať ako „empirický odhad“ tohto spoločného rozdelenia. Ukazuje sa, že ak $X \sim Emp(x_1, \ldots, x_n)$, potom distribučná funkcia náhodnej premennej X, $E(X)$, $D(X)$ a $H(X)$ sú za určitých predpokladov „dobré“ odhady zodpovedajúcich charakteristík daného spoločného neznámeho rozdelenia. Tieto číselné charakteristiky dát sa nazývajú „empirická distribučná funkcia“, „výberový priemer“, „výberový rozptyl“[28] a „empirická entropia“. Okrem mnohých iných otázok týkajúcich sa spracovania dát sa otázkami odhadu charakteristík rozdelenia, ktoré generuje dáta, zaoberá „matematická štatistika“.

Poznámka 4.72. Okrem uvedených diskrétnych rozdelení (presnejšie parametrických tried diskrétnych rozdelení) existuje množstvo iných, každé so špecifickou interpretáciou, aplikáciami, vlastnosťami a vzťahmi k iným rozdeleniam. Môžete si napríklad vyhľadať rozdelenie „negatívne binomické“, „negatívne hypergeometrické“, „Zipfovo“, prípadne Vás môže zaujať príbuzný „Benfordov zákon“.

4.3 Využitie linearity strednej hodnoty

Definícia 4.73. Nech $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor a nech $A \in \mathcal{S}$. Náhodnú premennú[29] $I : \Omega \to \mathbb{R}$ definovanú $I(\omega) = 1$ pre všetky $\omega \in A$ a $I(\omega) = 0$ pre všetky $\omega \in \Omega/A$ nazývame **indikátor udalosti** A.[30]

Príklad 4.74. Dokážme princíp zapojenia-vypojenia 1.42. Nech $A_1, A_2, \ldots, A_n$ sú udalosti na pravdepodobnostnom priestore $(\Omega, \mathcal{S}, P)$. Pre každé $i = 1, \ldots, n$ nech je Y_i indikátor udalosti A_i. Potom

$$P\left(\cup_{i=1}^n A_i\right) = 1 - P\left(\cap_{i=1}^n [Y_i = 0]\right) = 1 - P\left[\prod_{i=1}^{n}(1 - Y_i) = 1\right] = 1 - E\left(\prod_{i=1}^{n}(1 - Y_i)\right),$$

[27] Ide o jeden z najčastejších modelov pre pozorované dáta.

[28] Niekedy sa $\frac{1}{n}\sum_{i=1}^{n}(x_i - E(X))^2$ nazýva aj „vychýlený“ výberový rozptyl a za „nevychýlený“ výberový rozptyl sa považuje $\frac{1}{n-1}\sum_{i=1}^{n}(x_i - E(X))^2$. Táto terminológia sa vyjasní v základnom kurze matematickej štatistiky. Pre veľké n sú uvedené charakteristiky podobné.

[29] Dôkaz, že uvedené zobrazenie je náhodná premenná, je priamočiary.

[30] Všimnime si, že ak je I indikátor akejkoľvek udalosti, tak $I^2 = I$, čiže napríklad $E(I^2) = E(I)$. Tento jednoduchý postreh v ďalšom texte využívame bez upozornenia.

kde sme využili to, že $\prod_{i=1}^n (1-Y_i)$ má alternatívne rozdelenie a stredná hodnota takejto náhodnej premennej je pravdepodobnosť, že nadobudne hodnotu 1. Ak roznásobíme výraz $\prod_{i=1}^n (1-Y_i)$, využijeme linearitu strednej hodnoty a fakt, že pre akékoľvek $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ má $Y_{i_1} \cdots Y_{i_k}$ alternatívne rozdelenie so strednou hodnotou $P(A_{i_1} \cap \ldots \cap A_{i_k})$, máme

$$\begin{aligned} 1 - E\left(\prod_{i=1}^n (1-Y_i)\right) &= \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} E(Y_{i_1} \cdots Y_{i_k}) = \\ & \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} P(A_{i_1} \cap \ldots \cap A_{i_k}). \end{aligned}$$

Príklad 4.75. Nech $X \sim Hyp(M, N, n)$; pozri definíciu 4.63. Určme $E(X)$.

Riešenie: Ak sa chceme vyhnúť výpočtu komplikovaných súm, môžeme urobiť dôkaz s využitím linearity strednej hodnoty, a to nasledovne: Uvažujme schému z poznámky 4.65. Označme pre $i = 1, \ldots, n$ ako U_i indikátor udalosti „i-ta vybratá guľôčka je čierna". Zo symetrie plynie, že $U_i \sim Alt(N/M)$, preto $E(U_i) = N/M$. Keďže $X = \sum_{i=1}^n U_i$, tak máme:

$$E(X) = E\left(\sum_{i=1}^n U_i\right) = \sum_{i=1}^n E(U_i) = n(N/M).$$

Takto by bolo možné odvodiť aj rozptyl; výpočet je však zdĺhavejší.

Príklad 4.76. Nech $n \in \mathbb{N}$, $n \geq 2$ a nech X_n znamená počet prvkov, ktoré zostanú na svojom pôvodnom mieste po dokonalej náhodnej permutácii postupnosti $(1, \ldots, n)$, rovnako ako v príklade 4.55. Určme $E(X_n)$ a $D(X_n)$.

Riešenie: Nech $U_1, \ldots, U_n$ sú náhodné premenné, pričom U_i je indikátor udalosti, že číslo i zostane na svojom pôvodnom mieste. To znamená, že U_i je 1, ak číslo i zostane na svojom mieste a U_i je 0, ak číslo i nezostane na svojom pôvodnom mieste. Zrejme $U_i \sim Alt(1/n)$, t.j. $E(U_i) = 1/n$ podľa vety 4.38. Keďže $X = U_1 + \cdots + U_n$, dostávame z tvrdenia 4.12 $E(X) = \sum_{i=1}^n E(U_i) = 1$.

Podobne, použitím linearity strednej hodnoty máme

$$E(X_n^2) = E\left(\sum_{i=1}^n U_i \sum_{j=1}^n U_j\right) = \sum_{i=1}^n \sum_{j=1}^n E(U_i U_j).$$

Zrejme $U_i U_j \sim Alt(p)$, kde p je pravdepodobnosť, že zároveň číslo i aj číslo j zostane na svojom mieste. Ak $i = j$, tak $p = \frac{1}{n}$ a ak $i \neq j$, tak $p = \frac{1}{n}\frac{1}{n-1}$. Preto

$$E(X_n^2) = \sum_i E(U_i U_i) + \sum\sum_{i \neq j} E(U_i U_j) = n\frac{1}{n} + n(n-1)\frac{1}{n}\frac{1}{n-1} = 2.$$

Z odvodených rovností a vety 4.20 dostávame

$$D(X_n) = E(X_n^2) - (E(X_n))^2 = 2 - 1 = 1.$$

Z príkladu 4.55 (a vety 4.53) vieme, že limitné rozdelenie náhodných premenných X_n má strednú hodnotu aj rozptyl 1. Práve sme sa však presvedčili, že $E(X_n) = D(X_n) = 1$ pre *každé* $n \geq 2$.

Príklad 4.77. Nech $\sigma = (\sigma_1, \dots, \sigma_n)$ je permutácia postupnosti $(1, \dots, n)$. Index $i \in \{1, \dots, n\}$ nazveme „rekordom“ v permutácii σ, ak buď $i = 1$, alebo ak $i > 1$ a $\sigma_i > \sigma_j$ pre všetky $j \in \{1, \dots, i-1\}$. Nech X_n označuje počet rekordov v dokonalej náhodnej permutácii σ postupnosti $(1, \dots, n)$.[31] Nájdime $E(X_n)$ a $D(X_n)$.

Riešenie: Pre všetky $i = 1, \dots, n$ označme ako U_i indikátor udalosti „v náhodnej permutácii čísel $(1, \dots, n)$ bude i rekordom“. Zjavne $X_n = U_1 + \cdots + U_n$. Tiež je zrejmé, že pre každé $i \in \{1, \dots, n\}$ je udalosť [i je rekord] ekvivalentná udalosti $[\sigma_i = \max(\sigma_1, \dots, \sigma_i)]$, ktorej pravdepodobnosť je $1/i$, pretože z dôvodov symetrie má každá z i hodnôt $\sigma_1, \dots, \sigma_i$ rovnakú pravdepodobnosť, že bude najväčšia v množine $\{\sigma_1, \dots, \sigma_i\}$. Na základe linearity strednej hodnoty dostávame:

$$E(X_n) = \sum_{i=1}^{n} E(U_i) = \sum_{i=1}^{n} \frac{1}{i}.$$

Poznamenajme, že súčet prvých n členov harmonického radu je $\sum_{i=1}^{n} \frac{1}{i} \approx \ln(n) + \gamma$, kde $\gamma \approx 0,577$ je Eulerova-Mascheroniho konštanta. Vypočítajme rozptyl. Priamo z definície nezávislosti si ľahko kombinatoricky overíme, že pre všetky $1 \leq i < j \leq n$ sú udalosti [i je rekord] a [j je rekord] nezávislé. Takže $E(U_iU_j) = P[U_iU_j = 1] = P[U_i = 1, U_j = 1] = P[U_i = 1]P[U_j = 1] = \frac{1}{ij}$. Máme preto

$$E(X_n^2) = \sum_{i=1}^{n}\sum_{j=1}^{n} E(U_iU_j) = \sum_{i=1}^{n} E(U_i) + 2\sum_{i<j} E(U_iU_j) = \sum_{i=1}^{n} \frac{1}{i} + 2\sum_{i<j} \frac{1}{ij},$$

z čoho dostávame

$$D(X_n) = E(X_n^2) - (E(X_n))^2 = \sum_{i=1}^{n} \frac{1}{i} + 2\sum_{i<j} \frac{1}{ij} - \left(\sum_{i=1}^{n} \frac{1}{i}\right)^2 = \sum_{i=1}^{n} \frac{1}{i} - \sum_{i=1}^{n} \frac{1}{i^2}.$$

Pre veľké n je $\sum_{i=1}^{n} \frac{1}{i^2} \approx \frac{\pi^2}{6}$, čiže $D(X_n) \approx \ln(n) + \gamma - \frac{\pi^2}{6}$.

Príklad 4.78. Za okrúhlym stolom je rozostavených 17 stoličiek očíslovaných $0, \dots, 16$, na ktoré sa posadilo 12 mužov a 5 žien. Ukážeme, že existuje sedmica susedných stoličiek, na ktorých sedia aspoň tri ženy.

Riešenie: Najprv si definujeme (nenáhodné) konštanty R_k a V_k pre $k = 0, \dots, 16$, a to nasledovne: R_k označuje počet žien, ktoré sedia na stoličkách s číslami $k, k+1, \dots, k+6$ modulo 17. Nech V_k je indikátor udalosti „na stoličke s číslom k sedí žena“.

Ďalej náhodne zvoľme jednu stoličku (každú s pravdepodobnosťou 1/17). Nech K je číslo tejto stoličky, t.j. K je náhodná premenná s rovnomerným rozdelením na množine $\{0, \dots, 16\}$. Pre náhodnú premennú R_K platí

$$R_K = V_K + V_{K+1} + \cdots + V_{K+6},$$

kde indexy stále berieme modulo 17. Všimnime si, že $E(V_{K+i}) = 5/17$ pre všetky $i = 0, \dots, 6$, lebo $E(V_{K+i}) = P[V_{K+i} = 1]$, čo je pravdepodobnosť, že na stoličke $K + i$ (modulo 17) je žena. Máme:

$$E(R_K) = E(V_K) + E(V_{K+1}) + \cdots + E(V_{K+6}) = 7 \times (5/17) = 35/17 > 2.$$

[31] Kvôli jednoduchosti značíme v tomto príklade aj pevnú aj náhodnú permutáciu symbolom σ.

Aby mohlo platiť $E(R_K) > 2$, musí byť splnené $P[R_K > 2] > 0$, t.j. náhodná premenná R_K nadobúda s nenulovou pravdepodobnosťou jednu z hodnôt $3, 4, \ldots$. Z toho ale plynie, že nutne musí existovať aspoň jedna sedmica susedných stoličiek, na ktorých sedia aspoň 3 ženy.

Poznámka 4.79. Predchádzajúci príklad (ale aj príklad 1.52) ukazuje, že pravdepodobnostné metódy môžeme použiť aj na riešenie niektorých čisto deterministických úloh (napríklad geometrických, kombinatorických alebo grafových). Súhrnne sa takéto techniky nazývajú „pravdepodobnostná metóda“; pozri napríklad [1].

Kapitola 5

Spojité náhodné premenné

5.1 Vlastnosti spojitých náhodných premenných

„Spojitou“ náhodnou premennou reprezentujeme číselný výsledok „spojitého náhodného deja“. Presnejšie, model spojitou náhodnou premennou používame vtedy, ak je množina všetkých potenciálnych výsledkov nespočítateľná, na rozdiel od diskrétnej náhodnej premennej, pre ktorú je množina všetkých potenciálnych výsledkov spočítateľná.[1] Ide, povedzme, o meranie hmotnosti účinnej látky v tabletke nejakého lieku, pozorovanie výšky hladiny rieky v zadanom čase, generovanie „náhodného čísla“ v intervale $(0, 1)$ pre potreby stochastického optimalizačného algoritmu a tak ďalej. V aplikáciách sa spojité náhodné premenné používajú aj ako aproximácia pre také náhodné premenné, ktoré sú vo svojej podstate diskrétne, avšak potenciálne nadobúdajú veľmi veľký počet hodnôt. Príkladom by mohlo byť meranie hmotnosti, ktoré zaokrúhľujeme na miligramy, meranie výšky, ktoré zaokrúhľujeme na milimetre, alebo počítačové generovanie náhodného čísla v intervale $(0, 1)$, ktoré je v skutočnosti reprezentované vysokým, ale konečným počtom bitov. Špecifickou funkciou, charakterizujúcou rozdelenie spojitej náhodnej premennej je jej „hustota“. Základnými číselnými charakteristikami spojitej náhodnej premennej sú stredná hodnota a rozptyl, podobne ako pre diskrétne náhodné premenné.

Poznámka 5.1. Poznamenajme, že tvrdenia týkajúce sa spojitých náhodných premenných je možné dokázať pomocou základných vlastností integrálu, analogicky, ako sa tvrdenia pre diskrétne náhodné premenné dokazujú pomocou vlastností súm. Dôkazy pre spojité náhodné premenné však závisia na detailoch toho, ako mali študenti vybudovaný integrálny počet, čo nie je úplne jednotné, takže ich v tejto časti väčšinou nebudeme uvádzať.

Definícia 5.2. Nech X je náhodná premenná s distribučnou funkciou F a nech existuje integrovateľná funkcia $f : \mathbb{R} \to [0, \infty)$ taká, že pre každé $x \in \mathbb{R}$ platí

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t. \tag{5.1}$$

[1] Existujú však aj také situácie, v ktorých pozorovaná číselná charakteristika potenciálne nadobúda nespočítateľne veľa hodnôt, no zodpovedajúca náhodná premenná nie je spojitá a v zmysle našej definície 5.2 samozrejme ani diskrétna; taká náhodná premenná je „zmiešaného“ typu. Náhodné premenné, ktoré nie sú ani spojité ani diskrétne, sa v aplikáciách využívajú zriedka a ďalej sa im nebudeme venovať.

Potom hovoríme, že X je **spojitá**[2] **náhodná premenná**[3] s **hustotou** f.

Príklad 5.3. V kruhovom okne s polomerom 1 meter sa rovnomerne náhodne vyskytne častica. Zaujíma nás náhodná premenná X, ktorá reprezentuje vzdialenosť tejto častice od stredu okna. Distribučná funkcia náhodnej premennej X, definovaná $F(x) = P[X < x]$ pre všetky $x \in \mathbb{R}$, je $F(x) = 0$, ak $x < 0$, $F(x) = x^2$, ak $x \in [0, 1]$ a $F(x) = 1$, ak $x > 1$. Všimnime si, že platí $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$ pre každé $x \in \mathbb{R}$, kde $f(t) = 0$ pre $t < 0$, $f(t) = 2t$ pre $t \in [0, 1]$ a $f(t) = 0$ pre $t > 1$. To znamená, že X je spojitá náhodná premenná a uvedená funkcia f je jej hustota.[4]

Poznámka 5.4. Hustota spojitej náhodnej premennej nie je určená jednoznačne. Skutočne, ak f je hustota spojitej náhodnej premennej X a funkcia f° sa rovná funkcii f všade s výnimkou konečného počtu bodov, tak aj f° je hustota náhodnej premennej X, pretože ak platia rovnosti (5.1) pre f, tak platia aj pre f°. Dá sa však ukázať, že ak sú f a f° dve hustoty náhodnej premennej X, ktoré sú obe spojité na nejakom intervale (a, b), tak $f(x) = f^\circ(x)$ pre všetky $x \in (a, b)$.

Príklad 5.5. Pre spojitú náhodnú premennú X z príkladu 5.3 je okrem f hustotou aj funkcia f° definovaná $f^\circ(t) = 0$ pre $t < 0$, $f^\circ(t) = 2t$ pre $t \in [0, 1)$ a $f^\circ(t) = 0$ pre $t \geq 1$.

Veta 5.6 (Vyjadrenie pravdepodobnosti intervalu pomocou hustoty)**.** Nech X je spojitá náhodná premenná s hustotou f. Potom $\int_{-\infty}^{\infty} f(t)\mathrm{d}t = 1$. Navyše, pre akékoľvek $a, b \in \mathbb{R}$, $a < b$:

$$P[a \leq X < b] = \int_a^b f(t)\mathrm{d}t. \tag{5.2}$$

a $P[X = a] = 0$.[5]

Príklad 5.7. Pre náhodnú premennú X z príkladu 5.3 platí napríklad $P[1/3 \leq X < 2/3] = \int_{1/3}^{2/3} 2t\mathrm{d}t = 1/3$. Pochopiteľne, ten istý výsledok vieme dostať aj priamo z distribučnej funkcie náhodnej premennej X: $P[1/3 \leq X < 2/3] = F(2/3) - F(1/3) = 1/3$.

Poznámka 5.8. Rovnosť (5.2) je dôležitá pre interpretáciu pojmu hustoty spojitej náhodnej premennej: Pravdepodobnosť, že náhodná premenná nadobudne hodnotu medzi číslami a, b ($a < b$) je rovná ploche pod grafom hustoty na intervale (a, b).[6] Taktiež si všimnime nasledovné

[2] V literatúre sa pod „spojitou“ náhodnou premennou niekedy myslí každá taká náhodná premenná, ktorá má spojitú distribučnú funkciu F a náhodná premenná spojitá v našom (silnejšom) zmysle sa volá „*absolútne* spojitá“. Napríklad náhodná premenná s takzvaným „Cantorovým“ rozdelením nie je spojitá v zmysle definície z týchto skrípt, avšak je spojitá v zmysle uvedenej slabšej definície. Tento jemný rozdiel v definíciách však nie je aplikačne dôležitý a ďalej sa mu nebudeme venovať.

[3] Pojem „spojitej“ náhodnej premennej je zaužívaný, môže však byť mätúci. Totiž, ak je X spojitá náhodná premenná v zmysle definície 5.2, tak X nemusí byť spojité zobrazenie z Ω do $\mathbb{R}$ v zmysle definície spojitosti z matematickej analýzy, aj v prípade, ak by sme mali na Ω definovanú „topológiu“ a malo by zmysel hovoriť o spojitosti X v analytickom zmysle. Presnejšie by sme mohli hovoriť, že náhodná premenná X má „spojité rozdelenie“, nie že je X sama osebe spojitá.

[4] Ide o takzvané „beta rozdelenie s parametrami 2 a 1“.

[5] Čiže pre spojitú náhodnú premennú X platí $P[a \leq X < b] = P[a < X \leq b] = P[a \leq X \leq b] = P[a < X < b]$; všetky tieto pravdepodobnosti sú rovné integrálu na pravej strane rovnosti (5.2). Všimnime si tiež, že vlastnosť „$P[X = a] = 0$ pre všetky $a \in \mathbb{R}$“ implikuje, že distribučná funkcia každej spojitej náhodnej premennej je nutne spojitá funkcia.

[6] Čo je to isté ako plocha pod grafom hustoty na intervale $[a, b)$, $(a, b]$ alebo $[a, b]$.

pozorovanie: Nech hustota f náhodnej premennej X je spojitá na krátkom intervale $B = [a, a+\delta)$, $a \in \mathbb{R}$, $\delta > 0$. Nech $x \in B$. Potom $P[X \in B] \approx f(x)\delta$. Takže $f(x)$ vyjadruje relatívnu „koncentráciu“ (resp. „hustotu“) pravdepodobnosti v okolí bodu x.

Veta 5.9 (Výpočet hustoty z distribučnej funkcie)**.** Nech distribučná funkcia F náhodnej premennej X je spojitá na celom $\mathbb{R}$ a spojite diferencovateľná všade, maximálne s výnimkou konečnej množiny bodov H. Potom X je spojitá náhodná premenná a akákoľvek nezáporná funkcia spĺňajúca $f(x) = \mathrm{d}F(x)/\mathrm{d}x$ pre všetky $x \notin H$ je hustotou X.[7]

Príklad 5.10. Všimnime si, že pre premennú X z príkladu 5.3 platí $f(x) = \mathrm{d}F(x)/\mathrm{d}x$ pre všetky $x \neq 1$. Avšak aj hustota f° z príkladu 5.5 spĺňa $f^\circ(x) = \mathrm{d}F(x)/\mathrm{d}x$ pre všetky $x \neq 1$.

Poznámka 5.11. Vieme, že funkcia diskrétnej náhodnej premennej, alebo aj viacerých diskrétnych náhodných premenných, je opäť diskrétna náhodná premenná (poznámka 4.10). Avšak ak je X spojitá náhodná premenná a g borelovská funkcia, tak náhodná premenná $Y = g(X)$ môže byť ako spojitá, tak aj diskrétna (napr. ak $g(x) = \lfloor x \rfloor$, t.j. dolná celá časť x), ale aj taká náhodná premenná, ktorá nie je ani spojitá, ani diskrétna. Podobná „nepríjemná“ vlastnosť platí aj pre borelovskú funkciu viacerých spojitých náhodných premenných; napríklad ak sú X_1, X_2 spojité náhodné premenné, tak $X_1 + X_2$ už nemusí byť spojitá náhodná premenná. Často však transformácia spojitej náhodnej premennej (alebo viacerých náhodných premenných) je opäť spojitá náhodná premenná a dokonca vieme jednoducho odvodiť hustotu tejto transformovanej náhodnej premennej, ako ukážeme ďalej.

Veta 5.12 (Lineárna transformácia spojitej náhodnej premennej)**.** Nech X je spojitá náhodná premenná s distribučnou funkciou F_X a hustotou f_X. Nech $a, b \in \mathbb{R}$, $a \neq 0$. Potom náhodná premenná $Y = aX + b$ je tiež spojitá s distribučnou funkciou $F_Y(y) = F_X((y-b)/a)$ pre všetky $y \in \mathbb{R}$, ak $a > 0$ a $F_Y(y) = 1 - F_X((y-b)/a)$ pre všetky $y \in \mathbb{R}$, ak $a < 0$. Hustota náhodnej premennej Y je daná predpisom

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) \text{ pre každé } y \in \mathbb{R}.$$

Dôkaz. Nech $a > 0$ a $y \in \mathbb{R}$. Platí $F_Y(y) = P[aX + b < y] = P[X < (y-b)/a] = F_X((y-b)/a)$. Hustotu f_Y dostaneme derivovaním F_Y (všade tam, kde derivácia existuje). Podobne odvodíme vzťah pre F_Y a následne aj pre f_Y pre prípad $a < 0$. □

Poznámka 5.13. Všimnime si nasledovný špeciálny prípad predchádzajúcej vety: $Y = aX$, kde $a > 0$, čo vedie len k „roztiahnutiu a splošteniu“ (ak $a > 1$) alebo „stiahnutiu a zvýšeniu“ (ak $a < 1$) hustoty. Z interpretačného hľadiska je transformácia $Y = aX$ „zmenou mierky“ alebo „zmenou jednotiek“, v ktorých meriame výsledky náhodnej premennej X. Napríklad, ak by sme merali vzdialenosť X z príkladu 5.3 nie v metroch, ale v centimetroch, tak by sme museli použiť náhodnú premennú $Y = 100X$. Ak je f_X hustota náhodnej premennej X, tak hustota náhodnej premennej Y by bola $f_Y(y) = 0.01 f_X(0.01y)$ pre všetky $y \in \mathbb{R}$.

Príklad 5.14 (Kvadratická transformácia spojitej náhodnej premennej)**.** Nech X je spojitá náhodná premenná so spojitou[8] hustotou f_X. Ukážeme, že náhodná premenná $Y = X^2$ je tiež spojitá a nájdeme jej hustotu f_Y.

[7]Existujú aj silnejšie verzie tejto vety; uvedená formulácia však postačuje v skoro všetkých aplikáciách.
[8]Spojitosť hustoty na celom $\mathbb{R}$ predpokladáme len kvôli jednoduchosti riešenia.

Riešenie: Najprv pomocou distribučnej funkcie F_X náhodnej premennej X určíme distribučnú funkciu F_Y náhodnej premennej Y. Pre každé $y \geq 0$ platí: $F_Y(y) = P[Y < y] = P[X^2 < y] = P[-\sqrt{y} < X < \sqrt{y}] = P[-\sqrt{y} \leq X < \sqrt{y}] = F_X(\sqrt{y}) - F_X(-\sqrt{y})$. Pre $y < 0$ zrejme platí: $F_Y(y) = P[Y < y] = P[X^2 < y] = 0$.

Keďže f_X je spojitá, je F_X spojite diferencovateľná, a teda z odvodeného vzťahu medzi F_X a F_Y vidíme, že aj F_Y je spojite diferencovateľná všade, s prípadnou výnimkou bodu 0. To znamená, že Y je spojitá náhodná premenná s hustotou f_Y, ktorú získame derivovaním F_Y všade tam, kde derivácia existuje a inde (t.j. v bode 0) zvolíme hustotu ľubovoľne. Teda $f_Y(y) = 0$ pre $y \leq 0$ a, keďže f_X je deriváciou funkcie F_X, pre $y > 0$ dostávame

$$f_Y(y) = \frac{1}{2\sqrt{y}} \left(f_X(\sqrt{y}) + f_X(-\sqrt{y}) \right).$$

Presvedčte sa pomocou tohto všeobecného vzťahu, že funkcia $f_Y(y) = 1$ pre $y \in [0,1]$ a $f_Y(y) = 0$ pre $y \notin [0,1]$ je hustotou náhodnej premennej $Y = X^2$ pre náhodnú premennú X z príkladu 5.3.[9]

Poznámka 5.15. Okrem lineárnej (veta 5.12) a kvadratickej (príklad 5.14) transformácie náhodnej premennej sa často vyskytuje aj exponenciálna transformácia $Y = e^X$ alebo logaritmická transformácia $Y = \ln(X)$.[10] Hustotu takto transformovaných náhodných premenných vieme odvodiť analogicky.

Definícia 5.16. Nech X je spojitá náhodná premenná s hustotou f. Ak existuje konečný integrál $\int_{-\infty}^{\infty} |x| f(x) \mathrm{d}x$,[11] hovoríme, že náhodná premenná X má konečnú strednú hodnotu a číslo

$$E(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x$$

nazývame **stredná hodnota** náhodnej premennej X. V opačnom prípade hovoríme, že náhodná premenná X nemá konečnú strednú hodnotu.

Poznámka 5.17. Stredná hodnota spojitej náhodnej premennej má podobnú fyzikálnu interpretáciu ako v prípade diskrétnej náhodnej premennej (poznámka 4.8), ide však o ťažisko objektu, v ktorom je, voľne povedané, hmotnosť distribuovaná spojito, v súlade s hustotou danej náhodnej premennej.

Veta 5.18 (Stredná hodnota funkcie spojitej náhodnej premennej)**.** Nech X je spojitá náhodná premenná s hustotou f a nech $g : \mathbb{R} \to \mathbb{R}$ je borelovská funkcia. Nech navyše existuje konečný integrál $\int_{-\infty}^{\infty} |g(x)| f(x) \mathrm{d}x$. Potom náhodná premenná $g(X)$ má konečnú strednú hodnotu a platí[12]

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) \mathrm{d}x.$$

[9] Y má „rovnomerné rozdelenie na intervale $(0,1)$“; viac o tomto rozdelení budeme mať v časti 5.2.1.

[10] Logaritmickú transformáciu môžeme pochopiteľne použiť iba vtedy, ak X nadobúda výlučne kladné hodnoty. To sa ale stáva často; principiálne kladná náhodná premenná X môže reprezentovať napríklad náhodný čas, výšku finančného plnenia, vzdialenosť atď.

[11] Všimnite si, že uvedený integrál je konečný napríklad vtedy, keď je hustota f nenulová len na ohraničenom intervale.

[12] Uvedená rovnosť platí nezávisle of toho, akého typu je náhodná premenná $g(X)$; pozri poznámku 5.11.

Príklad 5.19. Vypočítame strednú hodnotu náhodnej premennej X z príkladu 5.3. Na základe definície 5.16 a hustoty f uvedenej v príklade 5.3 máme $E(X) = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x = \int_0^1 2x^2\mathrm{d}x = [2x^3/3]_0^1 = 2/3$. Vypočítajme tiež strednú hodnotu náhodnej premennej X^2. Pomocou vety 5.18 máme $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x = \int_0^1 2x^3\mathrm{d}x = [2x^4/4]_0^1 = 1/2$. Všimnime si, že rovnaký výsledok by sme dostali aj vtedy, ak by sme najprv odvodili hustotu f_Y náhodnej premennej $Y = X^2$ (pozri poslednú časť príkladu 5.14) a potom použili definíciu 5.16 na f_Y.

Veta 5.20 (Linearita strednej hodnoty spojitej náhodnej premennej)**.** Nech X a Y sú spojité náhodné premenné, ktoré majú konečnú strednú hodnotu. Nech a, b sú akékoľvek reálne čísla. Potom $aX + bY$ má konečnú strednú hodnotu a platí[13]

$$E(aX + bY) = aE(X) + bE(Y).$$

Špeciálne, $E(aX) = aE(X)$ a $E(X + Y) = E(X) + E(Y)$.

Poznámka 5.21. Z vety 5.20 plynie, že pre n-ticu $X_1, \dots, X_n$ spojitých náhodných premenných, ktoré majú konečnú strednú hodnotu, a pre akékoľvek reálne konštanty $a_1, \dots, a_n$ platí, že náhodná premenná $\sum_{i=1}^n a_i X_i$ má konečnú strednú hodnotu $E(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i E(X_i)$.[14] Špeciálne, $E(\sum_{i=1}^n X_i) = \sum_{i=1}^n E(X_i)$.

Príklad 5.22. Uvažujme obmenu príkladu 5.3, ktorá spočíva v tom, že v jednotkovom kruhu pozorujeme *dve* nezávislé, rovnomerne sa vyskytujúce častice. Aká je stredná hodnota kvadrátu vzdialenosti medzi týmito časticami?

Riešenie: Príklad je množné riešiť mnohými spôsobmi; uvedieme jeden, ktorý zásadným spôsobom využíva linearitu strednej hodnoty, hoci budeme potrebovať aj niektoré výsledky z nasledujúcej kapitoly.

Definujme si kartézsky súradnicový systém tak, aby jeho počiatok bol v strede jednotkového kruhu. Nech náhodné premenné U_1, U_2 reprezentujú súradnice prvej častice (pomenujme ju $\mathbf{U}$) a náhodné premenné V_1, V_2 nech reprezentujú súradnice druhej častice (tú pomenujeme $\mathbf{V}$). Ľahko nahliadneme, že stredná hodnota každej z náhodných premenných U_1, U_2, V_1, V_2 je 0. Nech D je náhodná premenná znamenajúca vzdialenosť častíc $\mathbf{U}$ a $\mathbf{V}$. Na základe Pytagorovej vety a vety 5.20, resp. poznámky 5.21 máme

$$\begin{aligned} E(D^2) &= E\left((U_1 - V_1)^2 + (U_2 - V_2)^2\right) = E\left((U_1^2 + U_2^2) + (V_1^2 + V_2^2) - 2U_1V_1 - 2U_2V_2\right) = \\ &= E(U_1^2 + U_2^2) + E(V_1^2 + V_2^2) - 2E(U_1V_1) - 2E(U_2V_2). \end{aligned}$$

Lenže $U_1^2 + U_2^2 = X_1^2$, kde X_1 je vzdialenosť častice $\mathbf{U}$ od stredu kruhu a v príklade 5.19 sme vypočítali, že $E(X_1^2) = 1/2$. Podobne $V_1^2 + V_2^2 = X_2^2$, kde X_2 je vzdialenosť častice $\mathbf{V}$ od stredu kruhu, takže $E(X_2^2) = 1/2$. Na druhej strane, náhodné premenné U_1 a V_1 sú „nezávislé“,[15] takže z vety 6.10 dostávame $E(U_1V_1) = E(U_1)E(V_1) = 0$. Podobne $E(U_2V_2) = 0$. Dostali sme teda

$$E(D^2) = 1/2 + 1/2 - 0 - 0 = 1.$$

[13] Uvedená rovnosť platí nezávisle od toho, akého typu je náhodná premenná $aX + bY$; pozri poznámku 5.11.

[14] Toto tvrdenie platí, nech by náhodná premenná $\sum_{i=1}^n a_i X_i$ bola akéhokoľvek typu.

[15] Nezávislosť náhodných premenných je kľúčový pojem v teórii pravdepodobnosti a podrobnejšie sa ním zaoberáme v kapitole 6.

Definícia 5.23. Pre spojitú náhodnú premennú X definujeme **rozptyl** $D(X)$ a **smerodajnú odchýlku** $\sigma(X)$) rovnako ako pre diskrétnu náhodnú premennú (definícia 4.15, resp. 4.23; pozri tiež poznámky 4.17, 4.24 a 4.25).

Poznámka 5.24. Pre rozptyl spojitej náhodnej premennej platia analogické poznámky, ako poznámky 4.16 a 4.18 týkajúce sa diskrétnej náhodnej premennej. Tiež preň platia rovnaké vzťahy ako vo vete 4.20. Ak spojitá náhodná premenná X má hustotu f a konečnú strednú hodnotu, tak aj náhodná premenná $(X - E(X))^2$ je spojitá a platí

1. $D(X) = \int_{-\infty}^{\infty}(x - E(X))^2 f(x)\mathrm{d}x$,
2. $D(X) = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x - (E(X))^2$,

ak sú príslušné integrály konečné.

Príklad 5.25. Vypočítajme ešte rozptyl náhodnej premennej X z príkladu 5.3. Ak poznáme len hustotu, tak najjednoduchšie je obvykle využiť vzťahy uvedené v poznámke vyššie. My sme však už v príklade 5.19 odvodili nielen $E(X)$, ale aj $E(X^2)$, takže môžeme použiť vzťah $D(X) = E(X^2) - (E(X))^2 = 1/2 - 4/9 = 1/18$.

5.2 Typy spojitých náhodných premenných

5.2.1 Rovnomerné rozdelenie na intervale

Definícia 5.26. Hovoríme, že náhodná premenná X má (spojité) **rovnomerné rozdelenie** na intervale (a, b), ak je X spojitá náhodná premenná s hustotou

$$f(x) = \frac{1}{b-a} \text{ pre } x \in (a, b)$$

a $f(x) = 0$ pre $x \notin (a, b)$. Túto skutočnosť značíme $X \sim R(a, b)$. Pre $X \sim R(0, 1)$ sa niekedy používa pojem **náhodné číslo**.

Veta 5.27 (Základné vlastnosti rovnomerného spojitého rozdelenia)**.** Nech $X \sim R(a, b)$. Potom pre distribučnú funkciu F náhodnej premennej X platí

$$F(x) = \begin{cases} 0 & \text{pre } x \leq a, \\ (x-a)/(b-a) & \text{pre } x \in (a, b), \\ 1 & \text{pre } x \geq b. \end{cases}$$

Navyše, ak $c, d \in \mathbb{R}$, $c > 0$, tak $cX + d \sim R(ca + d, cb + d)$. Špeciálne, $\frac{X-a}{b-a} \sim R(0, 1)$.

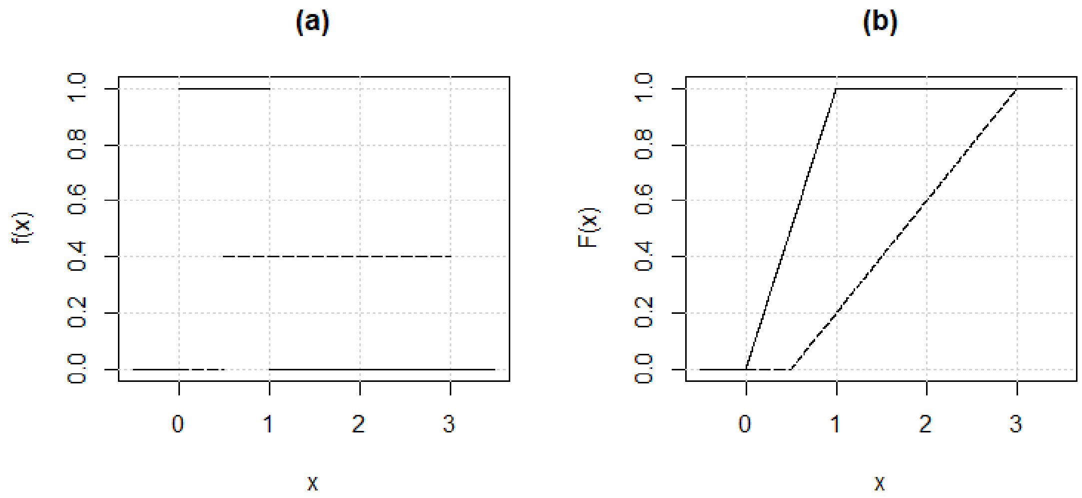
Dôkaz. Dôkaz je jednoduché cvičenie. □

Poznámka 5.28. Všimnime si nasledovný špeciálny prípad predchádzajúcej vety: Ak $U \sim R(0, 1)$ a $a < b$ sú reálne čísla, tak $(b - a)U + a \sim R(a, b)$.

Veta 5.29 (Základné číselné charakteristiky rovnomerného spojitého rozdelenia)**.** Nech $X \sim R(a, b)$. Potom $E(X) = (a + b)/2$ a $D(X) = (b - a)^2/12$.

Dôkaz. Dôkaz je jednoduché cvičenie. □

Poznámka 5.30. Rovnomerné rozdelenie sa používa napríklad na generovanie realizácií náhodných premenných z iných typov rozdelení pomocou vhodných transformácií; pozri vetu 3.24 a diskusiu za touto vetou.



Obr. 5.1: Hustota $f(x)$ a distribučná funkcia $F(x)$ náhodnej premennej s rozdelením $R(0,1)$ (plná čiara) a $R(0.5,3)$ (prerušovaná čiara).

5.2.2 Exponenciálne rozdelenie

Definícia 5.31. Hovoríme, že náhodná premenná X má **exponenciálne rozdelenie** s parametrom $\lambda > 0$, ak je X spojitá náhodná premenná s hustotou

$$f(x) = \lambda e^{-\lambda x} \text{ pre } x \geq 0$$

a $f(x) = 0$ pre $x < 0$. Túto skutočnosť značíme $X \sim Exp(\lambda)$.[16]

Veta 5.32 (Základné vlastnosti exponenciálneho rozdelenia)**.** Nech $X \sim Exp(\lambda)$. Potom pre distribučnú funkciu F náhodnej premennej X platí

$$F(x) = \begin{cases} 0 & \text{pre } x \leq 0, \\ 1 - e^{-\lambda x} & \text{pre } x > 0. \end{cases}$$

Navyše, ak $c > 0$, tak $cX \sim Exp(\lambda/c)$.

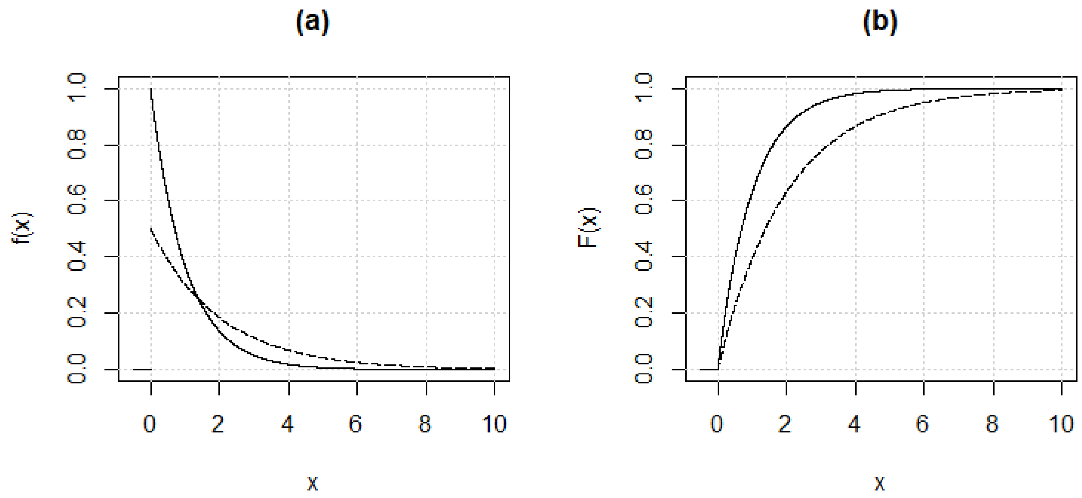
Dôkaz. Dôkaz je jednoduché cvičenie. □

Veta 5.33 (Základné číselné charakteristiky exponenciálneho rozdelenia)**.** Nech $X \sim Exp(\lambda)$. Potom $E(X) = 1/\lambda$ a $D(X) = 1/\lambda^2$.

Dôkaz. Najprv odvodíme strednú hodnotu a rozptyl pre $X_1 \sim Exp(1)$. Použijúc metódu per partes dostávame

$$E(X_1) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x = \int_0^{\infty} x e^{-x} \mathrm{d}x = \left[-xe^{-x}\right]_0^{\infty} + \int_0^{\infty} e^{-x} \mathrm{d}x = 1.$$

[16] Niekedy sa v literatúre používa aj parametrizácia triedy hustôt exponenciálneho rozdelenia v tvare $f(x) = \exp(-x/\mu)/\mu$ pre $x \geq 0$. V tejto parametrizácii reprezentuje μ strednú hodnotu, v parametrizácii, ktorú budeme používať my, reprezentuje λ prevrátenú hodnotu k strednej hodnote, ktorá sa pri exponenciálnom rozdelení niekedy nazýva „intenzita“.



Obr. 5.2: Hustota $f(x)$ a distribučná funkcia $F(x)$ náhodnej premennej s rozdelením $Exp(1)$ (plná čiara) a $Exp(1/2)$ (prerušovaná čiara).

$$E(X_1^2) = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x = \int_0^{\infty} x^2 e^{-x}\mathrm{d}x = \left[-x^2 e^{-x}\right]_0^{\infty} + \int_0^{\infty} 2xe^{-x}\mathrm{d}x = 2.$$

Takže $D(X_1) = E(X_1^2) - (E(X_1))^2 = 2 - 1 = 1$.

Ak $X_\lambda \sim Exp(\lambda)$ pre $\lambda > 0$, potom $\lambda X_\lambda \sim Exp(1)$, podľa vety 5.32. Zo základných vlastností strednej hodnoty a rozptylu, s použitím vzťahov $E(X_1) = 1$ a $D(X_1) = 1$ dostávame

$$E(X_\lambda) = E(\lambda X_\lambda)/\lambda = 1/\lambda \text{ a } D(X_\lambda) = D(\lambda X_\lambda)/\lambda^2 = 1/\lambda^2.$$

□

Veta 5.34 (Generovanie exponenciálneho rozdelenia rovnomerným)**.** Nech kladná náhodná premenná U má rozdelenie $R(0,1)$ a nech $\lambda > 0$. Potom $-\ln(U)/\lambda \sim Exp(\lambda)$.

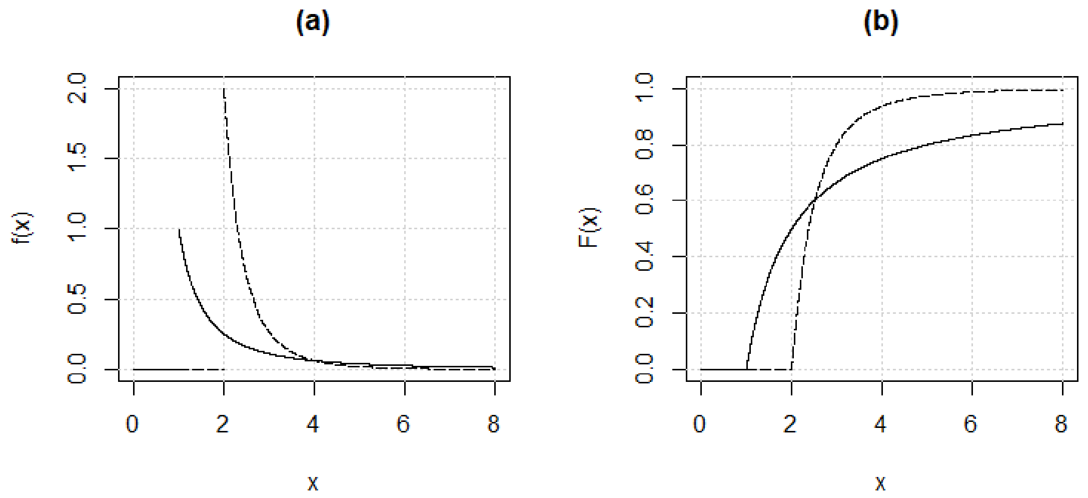
Dôkaz. Dôkaz plynie z vety o inverznej transformácii 3.24, z tvaru kvantilovej funkcie náhodnej premennej $X \sim Exp(\lambda)$, ktorý je $G(u) = -\frac{1}{\lambda}\ln(1-u)$ pre $u \in (0,1)$ a z faktu, že $U \sim R(0,1)$ práve vtedy, keď $1 - U \sim R(0,1)$. □

Poznámka 5.35. Exponenciálne rozdelenie sa používa v „teórii spoľahlivosti“ na modelovanie doby bezporuchovej činnosti zariadení, v „teórii hromadnej obsluhy“ na modelovanie doby medzi príchodmi zákazníkov do systému obsluhy, vo fyzike na modelovanie doby medzi špecifickými udalosťami týkajúcimi sa detekcie častíc a podobne.

Veta 5.36 („Bezpamäťovosť“ exponenciálneho rozdelenia)**.** Nech $X \sim Exp(\lambda)$, kde $\lambda > 0$. Nech $0 < t < s$. Potom $P[X > s | X > t] = P[X > s - t]$.

Dôkaz. Dôkaz je jednoduché cvičenie. □

Poznámka 5.37. Vlastnosť exponenciálneho rozdelenia z vety 5.36 znamená, voľne vyjadrené, toto: Ak doba X do nastania nejakého javu má exponenciálne rozdelenie a v nejakom



Obr. 5.3: Hustota $f(x)$ a distribučná funkcia $F(x)$ náhodnej premennej s rozdelením $Par(1,1)$ (plná čiara) a $Par(4,2)$ (prerušovaná čiara).

okamihu t dostaneme informáciu, že očakávaný jav ešte nenastal (čiže dostaneme informáciu, že $X > t$), tak to, koľko ešte budeme čakať do nastania daného javu, má stále to isté exponenciálne rozdelenie. Inými slovami, sme tam, kde sme boli na začiatku. Premyslite si, že analogickú vlastnosť má aj geometrické rozdelenie. Exponenciálne rozdelenie sa dá chápať ako „spojitá verzia“ geometrického rozdelenia.

5.2.3 Paretovo rozdelenie

Definícia 5.38. Hovoríme, že náhodná premenná X má **Paretovo rozdelenie** s parametrami $\alpha, k > 0$, ak je X spojitá náhodná premenná s hustotou

$$f(x) = \frac{\alpha}{k}\left(\frac{k}{x}\right)^{\alpha+1} \quad \text{pre } x \geq k$$

a $f(x) = 0$ pre $x < k$. Túto skutočnosť značíme $X \sim Par(\alpha, k)$.

Veta 5.39 (Distribučná funkcia Paretovho rozdelenia)**.** Nech $X \sim Par(\alpha, k)$. Potom pre distribučnú funkciu F náhodnej premennej X platí

$$F(x) = \begin{cases} 0 & \text{pre } x \leq k, \\ 1 - (k/x)^{\alpha} & \text{pre } x > k. \end{cases}$$

Dôkaz. Dôkaz je jednoduché cvičenie. □

Veta 5.40 (Stredná hodnota Paretovho rozdelenia)**.** Nech $X \sim Par(\alpha, k)$, pričom $\alpha > 1$. Potom $E(X) = \frac{\alpha \cdot k}{\alpha - 1}$. (Pre $\alpha \in (0,1]$ nemá náhodná premenná s rozdelením $Par(\alpha, k)$ konečnú strednú hodnotu.)

Dôkaz. Dôkaz je jednoduché cvičenie. □

Veta 5.41 (Generovanie Paretovho rozdelenia rovnomerným)**.** Nech náhodná premenná U má rozdelenie $R(0,1)$ a nech $\lambda > 0$. Potom $kU^{-1/\alpha} \sim Par(\alpha, k)$.

Dôkaz. Dôkaz plynie z vety o inverznej transformácii 3.24, z tvaru kvantilovej funkcie náhodnej premennej $X \sim Par(\alpha, k)$, ktorý je $G(u) = k(1-u)^{-1/\alpha}$ pre $u \in (0,1)$ a z faktu, že $U \sim R(0,1) \Leftrightarrow 1-U \sim R(0,1)$. □

Poznámka 5.42. Existuje viacero variantov Paretovho rozdelenia; tu sme sa stručne zoznámili len s jedným z nich. Zaujímavé pozorovanie je, že $X \sim Par(\alpha, k)$, tak $\ln(X/k) \sim Exp(\alpha)$, čo sa dá priamo nahliadnuť z viet 5.41 a 5.34. To znamená, že Paretovo rozdelenie je e na (posunuté) exponenciálne rozdelenie.

Poznámka 5.43. Paretovo rozdelenie sa niekedy používa na modelovanie doby, za ktorú vykoná CPU určitý proces, na modelovanie veľkosti súborov na internetových serveroch a podobne.

5.2.4 Normálne rozdelenie

Definícia 5.44. Nech $\mu \in \mathbb{R}$ a $\sigma > 0$. Hovoríme, že náhodná premenná X má **normálne rozdelenie**[17] s parametrami μ a σ^2, ak je X spojitá náhodná premenná s hustotou

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

pre všetky $x \in \mathbb{R}$. Túto skutočnosť značíme $X \sim N(\mu, \sigma^2)$. Ak $X \sim N(0,1)$, tak hovoríme, že X má **štandardizované normálne rozdelenie**, nazývané aj „normalizované“ normálne rozdelenie. Distribučnú funkciu náhodnej premennej X s rozdelením $N(0,1)$ budeme označovať symbolom Φ.

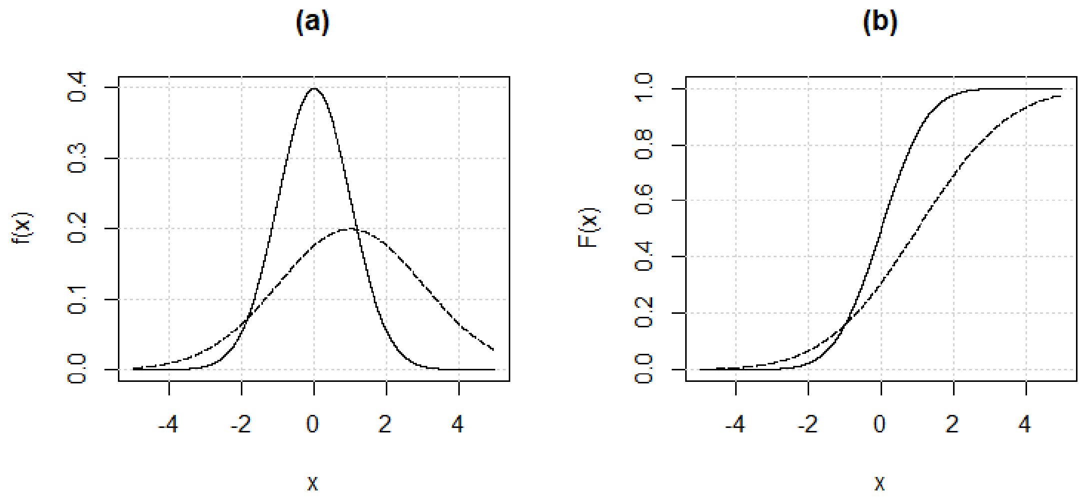
Poznámka 5.45. Normálne rozdelenie má ústredné postavenie v teoretickej aj aplikovanej teórii pravdepodobnosti a matematickej štatistike. Množstvo náhodných premenných používaných na modelovanie prirodzených dejov má približne normálne rozdelenie. Okrem toho má normálne rozdelenie niektoré výnimočné teoretické vlastnosti, napríklad je často limitným rozdelením prirodzene definovaných postupností náhodných premenných.

Poznámka 5.46. Distribučnú funkciu normálneho rozdelenia s akýmikoľvek parametrami vieme jednoducho vyjadriť pomocou Φ, ako ukazuje nasledovná veta. Samotná Φ je špeciálna funkcia, ktorá je „príjemná analyticky“ (je hladká, rastúca, symetrická v zmysle $\Phi(-x) = 1 - \Phi(x)$). Hodnoty $\Phi(x)$ je možné vypočítať s prakticky dostatočnou presnosťou rýchlymi numerickými algoritmami.

Veta 5.47 (Lineárna transformácia a štandardizácia normálneho rozdelenia)**.** Nech $X \sim N(\mu, \sigma^2)$ a nech $a, b \in \mathbb{R}$, $a \neq 0$. Potom $aX + b \sim N(a\mu + b, a^2\sigma^2)$. Špeciálne

$$\frac{X-\mu}{\sigma} \sim N(0,1).$$

[17] Normálne rozdelenie niekedy voláme gaussovské, na počesť Carla Friedricha Gaussa, ktorý toto rozdelenie spopularizoval.



Obr. 5.4: Hustota $f(x)$ a distribučná funkcia $F(x)$ náhodnej premennej s rozdelením $N(0,1)$ (plná čiara) a $N(1,2)$ (prerušovaná čiara).

Ak je F distribučná funkcia náhodnej premennej X, tak

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \text{ pre každé } x \in \mathbb{R}.$$

Dôkaz. Priamo z vety 5.12 dostávame, že hustota f_Y náhodnej premennej $Y = aX + b$ je v každom $y \in \mathbb{R}$:

$$f_Y(y) = \frac{1}{\sqrt{2\pi}|a|\sigma} \exp\left(-\frac{(y-b-a\mu)^2}{2a^2\sigma^2}\right).$$

Z definície normálneho rozdelenia teda máme $Y \sim N(a\mu + b, a^2\sigma^2)$. Druhú časť tvrdenia dostaneme voľbou $a = 1/\sigma$ a $b = -\mu/\sigma$ a keďže $(X-\mu)/\sigma \sim N(0,1)$, tak

$$F(x) = P[X < x] = P[(X-\mu)/\sigma < (x-\mu)/\sigma] = \Phi((x-\mu)/\sigma).$$

□

Príklad 5.48. Predpokladajme, že výsledok štandardizovaného testu IQ pre človeka náhodne zvoleného z populácie má rozdelenie $N(\mu, \sigma^2)$, kde $\mu = 100$ a $\sigma = 15$.[18] Predpokladajme tiež, že máme k dispozícii program, ktorý numericky počíta funkčné hodnoty funkcie Φ. Určme pravdepodobnosť, že výsledok IQ testu človeka náhodne vybratého z populácie bude nižší ako 130.

Riešenie: Ak $X \sim N(100, 15^2)$, tak podľa vety 5.47 platí $(X-100)/15 \sim N(0,1)$, preto

$$P[X < 130] = P\left[\frac{X-100}{15} < \frac{130-100}{15}\right] = \Phi(2) \approx 0.97725.$$

[18]Tento predpoklad je naozaj splnený pre štandardne používané IQ testy. Ide však len o približnú platnosť, pretože v skutočnosti je výsledok IQ testu celé číslo (takže, striktne vzaté, výsledok IQ testu je diskrétna náhodná premenná, ktorú normálnym rozdelením len aproximujeme). Navyše platnosť tohto predpokladu pre veľmi nízke a veľmi vysoké hodnoty IQ je z rôznych dôvodov problematická. Taktiež, môže záležať na populácii, z ktorej robíme náhodný výber. Podobnú kritickú úvahu by sme mali vykonať aj v iných situáciách, a to napriek tomu, že nakoniec model *presným* normálnym rozdelením môže byť úplne dostatočný z praktického hľadiska.

Príklad 5.49. Z dlhodobých štatistík sme zistili, že výška (v cm) mužov a žien v istej krajine má rozdelenie približne $N(180, 6^2)$, resp. $N(168, 5^2)$. O človeku z tejto krajiny vieme len to, že má výšku menšiu ako 170 centimetrov. Odhadneme pravdepodobnosť, že ide o ženu. (Naše apriórne očakávanie, čiže predtým ako sme sa dozvedeli výšku, je, že na 50 % pôjde o ženu a na 50 % pôjde o muža.) Predpokladáme, že máme k dispozícii numerickú procedúru na výpočet distribučnej funkcie Φ rozdelenia $N(0, 1)$.

Riešenie: Využijeme Bayesov vzorec a základné vlastnosti normálneho rozdelenia pravdepodobnosti. Označme si udalosť, že ide o ženu ako Z a že ide o muža ako M. Doplňujúcu informáciu, teda že výška daného človeka je menšia ako 170 centimetrov, si označme B. Hľadáme $P(Z|B)$. Keďže $P(Z) = P(M) = 0.5$, podľa Bayesovho vzorca dostávame

$$P(Z|B) = \frac{P(B|Z)P(Z)}{P(B|Z)P(Z) + P(B|M)P(M)} = \frac{P(B|Z)}{P(B|Z) + P(B|M)}.$$

Ak $X \sim N(168, 5^2)$ je náhodná premenná vyjadrujúca výšku ženy a $Y \sim N(180, 6^2)$ je náhodná premenná vyjadrujúca výšku muža, tak

$$\begin{aligned} P(B|Z) &= P[X < 170] = P[(X - 168)/5 < (170 - 168)/5] = \Phi((170 - 168)/5) \approx 0.6554, \\ P(B|M) &= P[Y < 170] = P[(Y - 180)/6 < (170 - 180)/6] = \Phi((170 - 180)/6) \approx 0.0478. \end{aligned}$$

Využili sme, že náhodné premenné $(X - 168)/5$ a $(Y - 180)/6$ majú rozdelenie $N(0, 1)$; pozri vetu 5.47. Dostávame teda $P(Z|B) \approx 0.6554/(0.6554 + 0.0478) \approx 0.932$.

Poznámka 5.50. Na základe rovnakých úvah aké sme použili v príklade 5.49 by sme vedeli zostaviť nasledovný algoritmus: Jeho vstupom je výška človeka a výstupom je „žena“ alebo „muž“, podľa toho, ktorá z týchto alternatív má vyššiu aposteriórnu pravdepodobnosť. Tým by sme dostali veľmi jednoduchý príklad „binárneho klasifikátora“. Na pravdepodobnostných a štatistických metódach, špeciálne na Bayesovom vzorci, je založených viacero metód, ktoré sa používajú v oblasti „strojového učenia“.[19]

Poznámka 5.51. Pre hustotu f normálneho rozdelenia je možné ukázať platnosť $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$ pomocou integrálu[20]

$$\int_{-\infty}^{\infty} e^{-x^2}\mathrm{d}x = \sqrt{\pi}.$$

Pomocou tohto výsledku a základných metód integrovania je možné ukázať, že pre akékoľvek párne $m \geq 2$ platí

$$\int_{-\infty}^{\infty} x^m e^{-x^2/2}\mathrm{d}x = \sqrt{2\pi}(m - 1)!!.$$

Pre nepárne m je funkcia $x^m e^{-x^2/2}$ nepárna, preto

$$\int_{-\infty}^{\infty} x^m e^{-x^2/2}\mathrm{d}x = 0.$$

Pomocou predchádzajúcich vzťahov určíme strednú hodnotu a rozptyl normálneho rozdelenia.

[19]Všimnite si však, že spomenutý algoritmus by ešte potreboval vopred odhadnúť pravdepodobnostné rozdelenie výšok žien a mužov, ktoré sme v príklade „zistili z dlhodobých štatistík“. Moderné algoritmy strojového učenia vedia na odhady zložitých pravdepodobnostných charakteristík a štatistických vzťahov využiť veľké objemy dát.

[20]Tento integrál by ste mohli poznať z pokročilejšieho kurzu matematickej analýzy.

Veta 5.52 (Základné číselné charakteristiky normálneho rozdelenia)**.** Nech $X \sim N(\mu, \sigma^2)$. Potom $E(X) = \mu$ a $D(X) = \sigma^2$.

Dôkaz. Najprv určíme $E(X_{0,1})$ a $D(X_{0,1})$ pre náhodnú premennú $X_{0,1} \sim N(0,1)$. Podľa výsledkov uvedených v poznámke 5.51 máme

$$E(X_{0,1}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-x^2/2} \mathrm{d}x = 0,$$

$$E(X_{0,1}^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} \mathrm{d}x = 1.$$

Preto $D(X_{0,1}) = E(X_{0,1}^2) - (E(X_{0,1}))^2 = 1$.

Ak $X \sim N(\mu, \sigma^2)$, tak $(X - \mu)/\sigma \sim N(0,1)$, podľa vety 5.47. Preto

$$0 = E((X - \mu)/\sigma) = (E(X) - \mu)/\sigma,$$

$$1 = D((X - \mu)/\sigma) = D(X)/\sigma^2.$$

Z týchto rovností dostávame požadované vzťahy pre $E(X)$ a $D(X)$. □

Veta 5.53 (Integrálna De Moivrova-Laplaceova veta)**.** Nech $p \in (0,1)$ a nech $X_1, X_2, \dots$ sú náhodné premenné, $X_n \sim Bin(n,p)$ pre každé $n \in \mathbb{N}$. Nech $x \in \mathbb{R}$. Potom

$$\lim_{n \to \infty} P\left[\frac{X_n - np}{\sqrt{np(1-p)}} < x\right] = \Phi(x). \tag{5.3}$$

Dôkaz. Táto veta je dôsledkom takzvanej „centrálnej limitnej" vety 6.24 a skutočnosti, že náhodná premenná s rozdelením $Bin(n,p)$ sa dá vyjadriť ako súčet n nezávislých náhodných premenných s rozdelením $Alt(p)$ (poznámka 4.45). □

Poznámka 5.54. Predchádzajúca veta sa používa na aproximáciu pravdepodobností týkajúcich sa binomického rozdelenia v tom zmysle, že rozdelenie $Bin(n,p)$ je možné aproximovať rozdelením $N(np, np(1-p))$. Podobne ako v prípade „Poissonovej" aproximácie, pozri poznámku 4.49, aj tu sa núka otázka, ako veľmi sa pomýlime, ak použijeme limitné tvrdenie (5.3) na aproximáciu pravdepodobností týkajúcich sa binomického rozdelenia pre konečné n. Zložitými technikami sa dá ukázať (pozri napríklad [8], strana 63) nasledovné jednoduché tvrdenie:

$$\sup_{x \in \mathbb{R}} \left| P\left[\frac{X_n - np}{\sqrt{np(1-p)}} < x\right] - \Phi(x) \right| \leq \frac{p^2 + (1-p)^2}{\sqrt{np(1-p)}}.$$

Táto nerovnosť ukazuje, že „gaussovská" aproximácia náhodnej premennej s rozdelením $Bin(n,p)$ je vhodná v prípade, že n je „dostatočne veľké" a p „nie je blízko nuly ani jednotky".

Príklad 5.55. Uvažujme matematický problém, ktorý má dva možné výsledky – V_1 a V_2 – pričom práve jeden z týchto výsledkov je „správny". Skonštruovali sme znáhodnený algoritmus, ktorý spustíme n-krát. Nech A_i je udalosť, že v i-tom spustení, $i = 1, \dots, n$, dostaneme správny výsledok. Náš algoritmus je síce extrémne rýchly, avšak vracia pri každom spustení správny výsledok s pravdepodobnosťou len mierne lepšou než hod mincou, konkrétne

$P(A_i) = 0.51$. Navyše, udalosti $A_1, \dots, A_n$, že vráti správny výsledok v jednotlivých spusteniach, sú nezávislé. Rozhodli sme sa, že tento algoritmus spustíme 9999-krát a ako finálny výsledok vezmeme ten, ktorý nám algoritmus vráti aspoň 5000-krát. Odhadnime pravdepodobnosť, že finálny výsledok bude správny. Pomocou tvrdenia uvedeného v poznámke 5.54 vypočítajme maximálnu chybu nášho odhadu.

Riešenie: Keďže ide o nezávislé spustenia algoritmu s rovnakou pravdepodobnosťou správneho výsledku, náhodná premenná označujúca počet správnych výsledkov má binomické rozdelenie $Bin(n, p)$, kde $n = 9999$ a $p = 0.51$. Toto rozdelenie môžeme aproximovať normálnym rozdelením $N(np, np(1-p)) = N(5099.49, 2498.75)$ a použitím vety 5.53 dostávame

$$P[X \geq 5000] = P\left[\frac{X - np}{\sqrt{np(1-p)}} \geq \frac{5000 - np}{\sqrt{np(1-p)}}\right] = 1 - \Phi\left(\frac{5000 - 5099.49}{\sqrt{2498.75}}\right) \approx 0.977.$$

Maximálna chyba nášho odhadu je

$$\frac{p^2 + (1-p)^2}{\sqrt{np(1-p)}} \approx 0.01.$$

Príklad 5.56. (Čiastočne motivované príkladom 3.9 z [2]) Prenášame binárny signál S, ktorý môže byť buď -1, alebo $+1$. Komunikačný kanál tento signál zašumí aditívnym normálnym šumom X so strednou hodnotou $\mu = 0$ a známym rozptylom σ^2. Predpokladáme, že šum X je nezávislý od signálu S. Prijímač usúdi, že správna hodnota signálu je $+1$ (resp. -1), ak po prenose zaznamená hodnotu aspoň 0 (resp. menšiu ako 0). Pomocou distribučnej funkcie Φ rozdelenia $N(0, 1)$ približne vyjadrime pravdepodobnosť, že pri n nezávislých prenosoch binárneho signálu dôjde k chybe menej ako k-krát.

Riešenie: K chybe dochádza, ak $S = -1$ a šum X je väčší alebo rovný 1, alebo ak $S = +1$ a šum je menší ako -1. Takže pravdepodobnosť chyby je pri jednotlivom prenose

$$P[\text{Chyba}] = P[X \geq 1 | S = -1]P[S = -1] + P[X < -1 | S = +1]P[S = +1].$$

Zvrat, že pôvodný signál nezávisí od pridaného šumu znamená, že udalosti $[X \geq 1]$ a $[S = -1]$ sú nezávislé, a taktiež udalosti $[X < -1]$ a $[S = +1]$ sú nezávislé. Preto $P[X \geq 1 | S = -1] = P[X \geq 1]$ a $P[X < -1 | S = +1] = P[X < -1]$. Avšak

$$\begin{aligned} P[X \geq 1] &= 1 - P[X < 1] = 1 - P\left[\frac{X - \mu}{\sigma} < \frac{1 - \mu}{\sigma}\right] = 1 - \Phi\left(\frac{1}{\sigma}\right) = p, \\ P[X < -1] &= P\left[\frac{X - \mu}{\sigma} < \frac{-1 - \mu}{\sigma}\right] = \Phi\left(\frac{-1}{\sigma}\right) = p, \end{aligned}$$

kde sme symbolom p označili číslo $1 - \Phi(1/\sigma)$. Využili sme pritom vetu 5.47 a vlastnosť $1 - \Phi(x) = \Phi(-x)$, ktorá platí pre všetky $x \in \mathbb{R}$. Dostávame teda

$$P[\text{Chyba}] = pP[S = -1] + pP[S = +1] = p,$$

pretože $P[S = -1] + P[S = +1] = 1$. Pri n nezávislých pokusoch o prenos binárneho signálu má počet chýb N_e rozdelenie $Bin(n, p)$, kde $p = 1 - \Phi(1/\sigma)$, preto podľa vety 5.53 dostávame

$$P[N_e < k] = P\left[\frac{N_e - np}{\sqrt{np(1-p)}} < \frac{k - np}{\sqrt{np(1-p)}}\right] \approx \Phi\left(\frac{k - np}{\sqrt{np(1-p)}}\right).$$

Poznámka 5.57. Okrem uvedených tried rozdelení spojitých náhodných premenných bolo opísaných a pomenovaných množstvo iných, so zaujímavými špecifickými vlastnosťami, vhodných na modelovanie rôznych typov dát. Spomeňme napríklad rozdelenia „beta“, „gama“, „Cauchyho“, „Laplaceovo“, „Paretovo“ a „lognormálne“ rozdelenie. Špeciálnu kategóriu rozdelení tvoria také, ktoré sa využívajú prevažne na konštrukciu intervalov spoľahlivosti a testovanie štatistických hypotéz, predovšetkým „chí-kvadrát rozdelenie“, „t-rozdelenie“ a „F-rozdelenie“.

Kapitola 6

Nezávislosť náhodných premenných

6.1 Nezávislosť a závislosť náhodných premenných

Najdôležitejšou charakteristikou vzájomného vzťahu viacerých náhodných premenných je ich nezávislosť, resp. závislosť.

Definícia 6.1. Náhodné premenné $X_1, \dots, X_n$ sa nazývajú **združene nezávislé**, ak pre akékoľvek množiny $B_1, \dots, B_n \in \mathcal{B}$ platí, že *udalosti* $[X_1 \in B_1], [X_2 \in B_2], \dots, [X_n \in B_n]$ sú združene nezávislé. Náhodné premenné $X_1, X_2, \dots$ nekonečnej postupnosti sa nazývajú združene nezávislé, ak sú združene nezávislé náhodné premenné $X_1, \dots, X_n$ pre každé n. Ak $X_1, \dots, X_n$, resp. $X_1, X_2, \dots$ nie sú združene nezávislé, hovoríme, že sú tieto náhodné premenné **združene závislé**.[1]

Poznámka 6.2. Združená nezávislosť udalostí $[X_1 \in B_1], [X_2 \in B_2], \dots, [X_n \in B_n]$ z definície 6.1 znamená,[2] že pre každú množinu indexov $\{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$ platí

$$P\Big([X_{i_1} \in B_{i_1}] \cap \dots \cap [X_{i_k} \in B_{i_k}]\Big) = \prod_{j=1}^{k} P[X_{i_j} \in B_{i_j}]. \tag{6.1}$$

Udalosť na ľavej strane v (6.1) často skrátene zapisujeme $P[X_{i_1} \in B_{i_1}, \dots, X_{i_k} \in B_{i_k}]$.

Poznámka 6.3. Voľne a intuitívne vyjadrené, náhodné premenné sú združene nezávislé, ak získanie akýchkoľvek poznatkov o výsledkoch realizácie niektorých z týchto náhodných premenných neovplyvní pravdepodobnostné očakávanie týkajúce sa výsledkov ostatných z týchto náhodných premenných. Ak sú náhodné premenné nezávislé, tak „nenesú o sebe informáciu“.

Poznámka 6.4. V pravdepodobnostných (a štatistických) modeloch obvykle *predpokladáme* nezávislosť nejakých náhodných premenných, či už priamo pozorovaných alebo latentných, a to na základe znalostí modelovaného deja. Vlastnosť nezávislosti náhodných premenných nám často umožňuje vypočítať charakteristiky, ktoré by bez nezávislosti bolo ťažké alebo nemožné vypočítať.

[1] Združene nezávislé a združene závislé náhodné premenné niekedy nazývame len „nezávislé“, resp. „závislé“; my budeme tiež používať toto skrátené vyjadrenie. Treba však mať na pamäti, že náhodné premenné môžu byť napríklad nezávislé po dvojiciach, (t.j. X_i a X_j sú nezávislé ako dvojica náhodných premenných pre akékoľvek indexy $i \neq j$, $i, j \in \{1, \dots, n\}$), ale pritom nebyť (združene) nezávislé.

[2] Pozri definíciu 2.14.

Príklad 6.5. Každý z troch ľudí (Adam, Eva a Daniel) hodí vyváženou kockou. Títo ľudia sú však ďaleko od seba a pri hodoch sa nijako neovplyvňujú. Náhodné premenné X_A, X_E, X_D, ktoré reprezentujú výsledky ich hodov, môžeme preto v našom pravdepodobnostnom modeli považovať za nezávislé.[3] Aká je pravdepodobnosť, že Adamov výsledok hodu bude menší ako Evin a súčasne Evin výsledok hodu bude menší ako Danielov?

Riešenie: Pre diskrétne náhodné premenné X_A, X_E, X_D potrebujeme vypočítať $P[X_A < X_E < X_D]$. Udalosť $[X_A < X_E < X_D]$ vieme napísať ako zjednotenie disjunktných udalostí $[X_A < X_E < X_D, X_E = k]$ pre $k = 2, 3, 4, 5$. Tiež si všimnime, že pre každé $k = 2, 3, 4, 5$ máme

$$\begin{aligned}[X_A < X_E < X_D, X_E = k] &= [X_A < k, k < X_D, X_E = k] \\ &= [X_A \in B_{A,k}, X_D \in B_{D,k}, X_E \in B_{E,k}]\end{aligned}$$

pre $B_{A,k} = \{1, \ldots, k-1\}$, $B_{D,k} = \{k+1, \ldots, 6\}$, $B_{E,k} = \{k\}$. Aditivita pravdepodobnosti, združená nezávislosť náhodných premenných X_A, X_E, X_D a elementárne metódy kombinatorickej pravdepodobnosti dávajú:

$$\begin{aligned}P[X_A < X_E < X_D] &= \sum_{k=2}^{5} P[X_A \in B_{A,k}, X_D \in B_{D,k}, X_E \in B_{E,k}] \\ &= \sum_{k=2}^{5} P[X_A \in B_{A,k}]P[X_D \in B_{D,k}]P[X_E \in B_{E,k}] \\ &= \sum_{k=2}^{5} \frac{k-1}{6} \cdot \frac{6-k}{6} \cdot \frac{1}{6} = \frac{5}{54}.\end{aligned}$$

Príklad 6.6. Máme dva identické počítače, ktorých činnosti sa navzájom nijako neovplyvňujú. Na každom z týchto počítačov naraz spustíme výpočet rovnakého znáhodneného algoritmu. Časy T_1, T_2 od spustenia po ukončenie výpočtu na prvom, resp. druhom počítači môžeme teda považovať za nezávislé náhodné premenné, navyše s rovnakým pravdepodobnostným rozdelením. Predpokladajme, že distribučná funkcia každej z náhodných premenných T_1 a T_2 je $F(t) = t^\delta$, kde $t \in (0,1)$ a δ je známa kladná konštanta. Nech Y je doba do ukončenie prvého z dvojice výpočtov (v zmysle „toho, ktorý skončí skôr“), čiže $Y = \min\{T_1, T_2\}$. Pre náhodnú premennú Y určme distribučnú funkciu, hustotu a strednú hodnotu.

Riešenie: Pre distribučnú funkciu F_Y náhodnej premennej Y a pre akékoľvek $y \in (0,1]$:

$$\begin{aligned}F_Y(y) &= P[\min\{T_1, T_2\} < y] = P([T_1 < y] \cup [T_2 < y]) = 1 - P[T_1 \geq y, T_2 \geq y] = \\ &\quad 1 - P[T_1 \geq y]P[T_2 \geq y] = 1 - (1 - F(y))^2 = 1 - (1 - y^\delta)^2.\end{aligned}$$

Kľúčovou je štvrtá rovnosť, kde sme využili nezávislosť náhodných premenných T_1 a T_2.[4] Očividne platí $F_Y(y) = 0$ pre $y \leq 0$ a $F_Y(y) = 1$ pre $y > 1$. Vidíme (pozri vetu 5.9), že Y je spojitá náhodná premenná a za jej hustotu môžeme vziať

$$f_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} = 2\delta(1 - y^\delta)y^{\delta - 1}$$

[3] Predpoklad nezávislosti by bol v tejto situácii správny aj vtedy, ak by kocky, ktoré títo ľudia používajú, boli falošné.

[4] Udalosti „B_1 a B_2“ tu boli intervaly $[y, \infty)$, ale to sme už nerozpisovali.

pre $y \in (0,1)$, inde $f_Y(y) = 0$. Strednú hodnotu dostaneme pomocou hustoty integrovaním:

$$E(Y) = \int_0^1 y f_Y(y)\mathrm{d}y = 2\delta \int_0^1 y^\delta - y^{2\delta}\mathrm{d}y = 2\delta((\delta+1)^{-1} - (2\delta+1)^{-1}).$$

Jeden z možných zápisov je

$$E(Y) = \frac{2\delta^2}{(2\delta+1)(\delta+1)}.$$

Napríklad ak je $\delta = 1$, čiže doba do ukončenia jednotlivého výpočtu má pre oba prípady rozdelenie $R(0,1)$, tak hustota náhodnej premennej Y má na intervale $(0,1)$ tvar klesajúcej lineárnej funkcie a jej stredná hodnota je $1/3$.[5]

Veta 6.7 (Teoretické charakterizácie nezávislosti náhodných premenných)**.** Nech $X_1,\dots,X_n$ sú náhodné premenné. Potom sú všetky štyri nasledovné výroky ekvivalentné. a) $X_1,\dots,X_n$ sú nezávislé, b) Náhodné premenné $g_1(X_1,\dots,X_{k_1}),\dots,g_r(X_{k_{r-1}+1},\dots,X_n)$ sú nezávislé pre akékoľvek $1 \le k_1 < \dots < k_{r-1} < n$ a akékoľvek borelovské funkcie $g_1,\dots,g_r$, c) Nech $\mathcal{F}$ je nejaký systém podmnožín množiny $\mathbb{R}$, taký, že $\mathcal{B} = \sigma(\mathcal{F})$. Potom $P[X_1 \in B_1,\dots,X_n \in B_n] = P[X_1 \in B_1]\dots P[X_n \in B_n]$ pre akýkoľvek výber množín $B_1,\dots,B_n \in \mathcal{F}$, d) $P[X_1 < x_1,\dots,X_n < x_n] = P[X_1 < x_1]\dots P[X_n < x_n]$ pre všetky $x_1,\dots,x_n \in \mathbb{R}$.

Dôkaz. Niektoré časti vety 6.7 je jednoduché dokázať (pozrite nasledujúcu poznámku), avšak iné sú náročnejšie a ich dôkazy presahujú ciele týchto skrípt. □

Poznámka 6.8. Všimnime si, že vo vete 6.7 je implikácia a)⇒b)[6] analogická tvrdeniu vety 2.18; voľne vyjadrené, ak máme nezávislé náhodné premenné a vyrobíme z „disjunktných skupiniek“ týchto náhodných premenných nové náhodné premenné, tak tieto nové náhodné premenné budú tiež nezávislé. Implikácia a)⇒c)[7] je taktiež zaujímavá: Hovorí, že na to, aby sme overili nezávislosť náhodných premenných, stačí overiť rovnosti $P[X_1 \in B_1,\dots,X_n \in B_n] = P[X_1 \in B_1]\dots P[X_n \in B_n]$ pre množiny $B_1,\dots,B_n$ zo systému množín $\mathcal{F}$, ktorý môže byť oveľa chudobnejší ako systém $\mathcal{B}$ borelovských množín; stačí, ak $\mathcal{F}$ *generuje* systém $\mathcal{B}$. Implikácia d)⇒a)[8] je dôležitá pre charakterizáciu nezávislosti náhodných premenných pomocou „združenej“ distribučnej funkcie náhodných premenných $X_1,\dots,X_n$, s ktorou sa zoznámime v časti 7.1.1. Implikáciu d)⇒a) využijeme aj pri dôkaze nasledovnej vety.

Veta 6.9 (Charakterizácia nezávislosti diskrétnych náhodných premenných)**.** Diskrétne náhodné premenné $X_1,\dots,X_n$ sú nezávislé vtedy a len vtedy, keď pre akékoľvek reálne čísla $x_1,\dots,x_n$ platí

$$P[X_1 = x_1,\dots,X_n = x_n] = \prod_{i=1}^n P[X_i = x_i]. \tag{6.2}$$

Dôkaz. Implikácia „⇒“ je zrejmá z definície 6.1, pretože jednoprvkové podmnožiny množiny $\mathbb{R}$ sú borelovské. Dokážme opačnú implikáciu. Pre každé $i = 1,\dots,n$ nech $x_i \in \mathbb{R}$

[5] V tomto príklade sme vypočítali hustotu náhodnej premennej, ktorá je funkciou iných náhodných premenných, konkrétne ich minimom. To je v teórii pravdepodobnosti častá úloha, ktorej sa budeme viac venovať v časti 7.2, kde využijeme aparát náhodných vektorov.

[6] Implikácia b)⇒a) je triviálna.

[7] Implikácia c)⇒a) je zrejmá.

[8] Implikácia a)⇒d) je očividná.

a $B_i = (-\infty, x_i)$. Označme $C_i = X_i(\Omega) \cap B_i$ množinu tých čísel z oboru hodnôt náhodnej premennej X_i, ktoré patria do B_i. Keďže náhodné premenné $X_1, \dots, X_n$ sú podľa predpokladu diskrétne, tak množiny C_i sú spočítateľné. Všimnime si, že $[X_1 \in B_1, \dots, X_n \in B_n] = \cup_{x_1 \in C_1, \dots, x_n \in C_n} [X_1 = x_1, \dots, X_n = x_n]$ je rozklad na disjunktné udalosti a $[X_i \in B_i] = \cup_{x_i \in C_i} [X_i = x_i]$ je tiež rozklad na disjunktné udalosti pre všetky $i = 1, \dots, n$. Ak okrem týchto rozkladov použijeme aditivitu (alebo σ-aditivitu) pravdepodobnosti, predpoklad (6.2) a základné vlastnosti konečných alebo nekonečných absolútne konvergentných súm, dostávame:

$$\begin{aligned} P[X_1 \in B_1, \dots, X_n \in B_n] &= P\left(\bigcup_{x_1 \in C_1, \dots, x_n \in C_n} [X_1 = x_1, \dots, X_n = x_n]\right) = \\ \sum_{x_1 \in C_1, \dots, x_n \in C_n} P[X_1 = x_1, \dots, X_n = x_n] &= \sum_{x_1 \in C_1, \dots, x_n \in C_n} P[X_1 = x_1] \dots P[X_n = x_n] = \\ \sum_{x_1 \in C_1} P[X_1 = x_1] &\cdots \sum_{x_n \in C_n} P[X_n = x_n] = \\ P\left(\bigcup_{x_1 \in C_1} [X_1 = x_1]\right) \cdots P\left(\bigcup_{x_n \in C_n} [X_n = x_n]\right) &= P[X_1 \in B_1] \dots P[X_n \in B_n]. \end{aligned}$$

Takže pre akékoľvek $x_1, \dots, x_n \in \mathbb{R}$ platí

$$\begin{aligned} P[X_1 < x_1, \dots, X_n < x_n] &= P[X_1 \in B_1, \dots, X_n \in B_n] \\ &= P[X_1 \in B_1] \dots P[X_n \in B_n] = P[X_1 < x_1] \dots P[X_n < x_n]. \end{aligned}$$

Podľa implikácie d)⇒a) z vety 6.7 sú teda náhodné premenné $X_1, \dots, X_n$ nezávislé. □

Veta 6.10 (Stredná hodnota súčinu nezávislých náhodných premenných)**.** Nech X_1 a X_2 sú nezávislé, obe diskrétne alebo obe spojité náhodné premenné s konečnou strednou hodnotou. Potom $X_1 X_2$ je diskrétna, resp. spojitá náhodná premenná s konečnou strednou hodnotou a platí

$$E(X_1 X_2) = E(X_1) E(X_2).$$

Dôkaz. Dôkaz urobíme len pre diskrétne náhodné premenné X_1, X_2. Pre jednoduchosť označíme ako A_i obor hodnôt X_i, t.j. $A_i = X_i(\Omega)$ pre $i = 1, 2$. Náhodná premenná $Z = X_1 X_2$ je zrejme diskrétna, lebo jej obor hodnôt $Z(\Omega)$ je spočítateľná množina súčinov $z = x_1 x_2$, kde $x_1 \in A_1$ a $x_2 \in A_2$. Uvedomme si, že pre každé $z \in Z(\Omega)$ máme $P[Z = z] = \sum P[X_1 = x_1, X_2 = x_2]$, kde súčet berieme pre všetky dvojice $x_1 \in A_1$ a $x_2 \in A_2$, pre ktoré platí $x_1 x_2 = z$. Navyše, z nezávislosti X_1 a X_2 vieme, že $P[X_1 = x_1, X_2 = x_2] = P[X_1 = x_1] P[X_2 = x_2]$; pozri vetu 6.9. Máme teda:[9]

$$E(X_1 X_2) = E(Z) = \sum_{z \in Z(\Omega)} z P[Z = z] = \sum_{x_1 \in A_1} \sum_{x_2 \in A_2} x_1 x_2 P[X_1 = x_1, X_2 = x_2] =$$

$$\sum_{x_1 \in A_1} \sum_{x_2 \in A_2} x_1 x_2 P[X_1 = x_1] P[X_2 = x_2] = \sum_{x_1 \in A_1} x_1 P[X_1 = x_1] \sum_{x_2 \in A_2} x_2 P[X_2 = x_2] =$$

$$E(X_1) E(X_2).$$

□

[9] Nebudeme dokazovať to, že *existuje* konečná stredná hodnota náhodnej premennej $Z = X_1 X_2$, čiže že rad $\sum_{z \in Z(\Omega)} z P[Z = z]$ absolútne konverguje; nechávame to na čitateľa.

Poznámka 6.11. Z predchádzajúcej vety plynie, že pre n-ticu $X_1, \dots, X_n$ nezávislých náhodných premenných, ktoré majú konečnú strednú hodnotu a sú všetky diskrétne alebo všetky spojité platí, že diskrétna, resp. spojitá náhodná premenná $\prod_{i=1}^n X_i$ má konečnú strednú hodnotu $E(\prod_{i=1}^n X_i) = \prod_{i=1}^n E(X_i)$. (Skutočne; napríklad pre $n = 3$ máme $E(X_1X_2X_3) = E(X_1X_2)E(X_3) = E(X_1)E(X_2)E(X_3)$, čo plynie kombináciou viet 6.7[10] a 6.10.) Táto vlastnosť nezávislých náhodných premenných má značný teoretický význam, avšak je ju možné využiť aj na priame riešenie úloh.

Príklad 6.12. Komunikačný kanál sa skladá zo série uzlov, pričom vždy i-ty uzol predáva jedno-bitovú informáciu na vstup $i+1$-vému uzlu. Na i-tom uzle dochádza s pravdepodobnosťou p_i k chybe, ktorá sa prejaví tým, že na výstupe tohto uzla bude opačný bit ako na jeho vstupe. Navyše, chyby na jednotlivých uzloch sa vyskytujú navzájom nezávisle. Napíšme vzorec udávajúci pravdepodobnosť, že bit na vstupe prvého uzla bude rovnaký ako bit na výstupe n-tého uzla. (Ide o zovšeobecnenie príkladu 2.23.)

Riešenie:[11] Pre $i = 1, \dots, n$ nech X_i je náhodná premenná, ktorá nadobudne hodnotu -1, ak na uzle i dôjde k chybe a hodnotu 1, ak na uzle i nedôjde k chybe, čiže $P[X_i = -1] = p_i$, $P[X_i = 1] = 1 - p_i$ a $E(X_i) = P[X_i = 1] - P[X_i = -1] = 1 - 2p_i$. Všimnime si, že naša hľadaná pravdepodobnosť je $P[X = 1]$, kde $X = \prod_{i=1}^n X_i$, pričom X je tiež náhodná premenná nadobúdajúca len hodnoty -1 a 1, takže $E(X) = P[X = 1] - P[X = -1]$, z čoho (ak vezmeme do úvahy aj rovnosť $P[X = 1] + P[X = -1] = 1$) dostávame $P[X = 1] = E(X)/2 + 1/2$. S využitím nezávislosti náhodných premenných $X_1, \dots, X_n$ a vety 6.10 teda dostávame

$$P\left[\prod_{i=1}^n X_i = 1\right] = \frac{E\left(\prod_{i=1}^n X_i\right)}{2} + \frac{1}{2} = \frac{\prod_{i=1}^n E\left(X_i\right)}{2} + \frac{1}{2} = \frac{\prod_{i=1}^n (1 - 2p_i)}{2} + \frac{1}{2}.$$

Poznámka 6.13. Vieme, že stredná hodnota súčtu náhodných premenných je rovná súčtu stredných hodnôt týchto náhodných premenných (ak dané stredné hodnoty existujú a sú konečné). Pre rozptyl analogické tvrdenie vo všeobecnosti neplatí, platí však pre nezávislé náhodné premenné, ako spresňuje nasledovná veta.

Veta 6.14 (Rozptyl súčtu nezávislých náhodných premenných)**.** Nech X_1 a X_2 sú nezávislé, obe diskrétne alebo obe spojité náhodné premenné s konečným rozptylom. Potom $X_1 + X_2$ je diskrétna, resp. spojitá náhodná premenná s konečným rozptylom a platí

$$D(X_1 + X_2) = D(X_1) + D(X_2).$$

Dôkaz. Urobme dôkaz pre diskrétne náhodné premenné. Keďže X_1 a X_2 sú nezávislé a majú konečný rozptyl, tak existujú konečné $E(X_1^2)$ a $E(X_2^2)$ (z definície 4.15), ale aj konečná $E(X_1X_2)$ (z vety 6.10), takže aj náhodná premenná $(X_1 + X_2)^2 = X_1^2 + 2X_1X_2 + X_2^2$ má konečnú strednú hodnotu (z poznámky 4.14), a teda má táto náhodná premenná aj konečný rozptyl (opäť z definície 4.15). Z vety 4.20, linearity strednej hodnoty a vety 6.10 dostávame:

$$\begin{aligned} D(X_1 + X_2) &= E\left((X_1 + X_2)^2\right) - E^2(X_1 + X_2) = \\ & E(X_1^2) + 2E(X_1X_2) + E(X_2^2) - E^2(X_1) - 2E(X_1)E(X_2) - E^2(X_2) = \\ & [E(X_1^2) - E^2(X_1)] + [E(X_2^2) - E^2(X_2)] + 2[E(X_1X_2) - E(X_1)E(X_2)] = \\ & D(X_1) + D(X_2). \end{aligned}$$

[10] Všimnite si, že implikácia a)⇒b) vo vete 6.7 zaručuje, že ak sú tri náhodné premenné X_1, X_2, X_3 združene nezávislé, tak sú dve náhodné premenné X_1X_2 a X_3 nezávislé.

[11] Toto elegantné riešenie navrhol študent FMFI UK Slavomír Takáč.

□

Poznámka 6.15. Veta 6.14 znamená, že pre n-ticu $X_1, \ldots, X_n$ nezávislých náhodných premenných, ktoré majú konečný rozptyl a sú všetky diskrétne alebo všetky spojité, má diskrétna, resp. spojitá náhodná premenná $\sum_{i=1}^n X_i$ konečný rozptyl $D(\sum_{i=1}^n X_i) = \sum_{i=1}^n D(X_i)$. (Zdôvodnenie je podobné ako v poznámke 6.11 pre súčin stredných hodnôt všeobecného počtu nezávislých náhodných premenných.) Táto vlastnosť rozptylu je veľmi dôležitá v celej pravdepodobnosti a matematickej štatistike. Jedno z použití vety 6.14 je napríklad v dôkaze „Čebyševovej nerovnosti“, ktorá vedie k fundamentálnemu tvrdeniu nazývanému „slabý zákon veľkých čísel“.

6.2 Markovova, Čebyševova a Černovova nerovnosť

Veta 6.16 (Markovova nerovnosť)**.** Nech Z je diskrétna alebo spojitá náhodná premenná s konečnou strednou hodnotou, ktorá nadobúda nezáporné hodnoty. Potom pre každé $c > 0$ platí:

$$P[Z \geq c] \leq \frac{E(Z)}{c}.$$

Dôkaz. Nech Z je diskrétna náhodná premenná nadobúdajúca hodnoty $(z_i)_{i\in I}$ a nech $c > 0$. Ďalej nech $J = \{i \in I : z_i \geq c\}$. Máme

$$E(Z) = \sum_{i\in I} z_i P[Z = z_i] \geq \sum_{i\in J} z_i P[Z = z_i] \geq c \sum_{i\in J} P[Z = z_i] = cP[Z \geq c].$$

Analogicky, ak Z je spojitá náhodná premenná s hustotou f, máme

$$E(Z) = \int_0^\infty z f(z) dz \geq \int_c^\infty z f(z) dz \geq \int_c^\infty c f(z) dz = cP[Z \geq c].$$

Z nerovnosti $E(Z) \geq cP[Z \geq c]$ dostávame priamo tvrdenie vety. □

Veta 6.17 (Čebyševova nerovnosť)**.** Nech X je diskrétna alebo spojitá náhodná premenná s konečným rozptylom (t.j. aj s konečnou strednou hodnotou). Potom pre každé $a > 0$ platí:

$$P[|X - E(X)| \geq a] \leq \frac{D(X)}{a^2}.$$

Dôkaz. Nech $a > 0$. Udalosti $[|X - E(X)| \geq a]$ a $[(X - E(X))^2 \geq a^2]$ sú totožné, takže $P[|X - E(X)| \geq a] = P[(X - E(X))^2 \geq a^2]$. Dôkaz Čebyševovej nerovnosti je teda dôsledkom Markovovej nerovnosti 6.16, pokiaľ v nej zvolíme $Z = (X - E(X))^2$ a $c = a^2$, pričom si ešte spomenieme, že $E\left((X - E(X))^2\right) = D(X)$. □

Príklad 6.18. Uvažujme istý systém hromadnej obsluhy s náhodným, no stabilným tokom požiadaviek na spracovanie. Z dlhodobých záznamov vieme, že v rozmedzí jednej hodiny je stredná hodnota počtu takýchto požiadaviek rovná 100. a) Použime Markovovu nerovnosť na ohraničenie pravdepodobnosti, že najbližšiu hodinu bude aspoň 250 požiadaviek na spracovanie. b) Predpokladajme navyše, že počet požiadaviek za hodinu má rozptyl 100. Použime Čebyševovu nerovnosť na odhad pravdepodobnosti, že najbližšiu hodinu bude počet požiadaviek aspoň 250. c) Ešte zosilnime predpoklady tak, že počet požiadaviek v priebehu jednej hodiny má Poissonovo rozdelenie so strednou hodnotu 100.[12] Aká je v tomto prípade pravdepodobnosť, že počet požiadaviek za najbližšiu hodinu bude aspoň 250?

Riešenie: Počet požiadaviek počas najbližšej hodiny budeme modelovať ako náhodnú premennú X. a) Keďže X je nezáporná náhodná premenná a predpokladáme $E(X) = 100$, Markovova nerovnosť poskytuje odhad: $P[X \geq 250] \leq 100/250 = 0.4$. b) Ak máme informáciu o rozptyle ($D(X) = 100$), dostávame z Čebyševovej nerovnosti oveľa lepší odhad: $P[X \geq 250] = P[|X - 100| \geq 150] \leq 100/150^2 \approx 0.00444$. c) Ak by sme navyše predpokladali, že X má rozdelenie $Po(100)$, výsledná pravdepodobnosť je ešte oveľa nižšia: $P[X \geq 250] = 1 - \sum_{k=0}^{249} e^{-100}100^k/k!$, čo je prakticky 0.

Markovova a Čebyševova nerovnosť obvykle poskytujú veľmi konzervatívne odhady. Lepší odhad v niektorých situáciách poskytne nasledovná nerovnosť.

Veta 6.19 (Černovova nerovnosť)**.** Nech S je diskrétna alebo spojitá náhodná premenná. Nech $t, a \in \mathbb{R}$ a nech existuje konečná stredná hodnota náhodnej premennej e^{tS}. Potom

$$P[S \geq a] \leq \frac{E(e^{tS})}{e^{ta}}. \tag{6.3}$$

Dôkaz. Nech $t, a \in \mathbb{R}$ sú ľubovoľné čísla. Uvedomme si, že udalosti $P[S \geq a]$ a $[e^{tS} \geq e^{ta}]$ sú totožné, takže $P[S \geq a] = P[e^{tS} \geq e^{ta}]$. Dôkaz dokončíme aplikovaním Markovovej nerovnosti 6.16 na diskrétnu alebo spojitú nezápornú náhodnú premennú $Z = e^{tS}$ a na $c = e^{ta} > 0$. □

Poznámka 6.20. Pravá strana nerovnosti (6.3) je platná pre všeobecné t, pre ktoré je $E(e^{tS}) < \infty$, čo nám dáva priestor na nájdenie „vhodnej“ hodnoty t pre naše konkrétne potreby. Ak by existovala konečná $E(e^{tS})$ pre všetky $t \in (a, b)$, tak formálne najsilnejšiu Černovovu hranicu by sme mohli napísať nasledovne:

$$P[S \geq a] \leq \inf_{t \in (a,b)} \frac{E(e^{tS})}{e^{ta}}.$$

Príklad 6.21. Majme náhodné premenné $X_1, \ldots, X_n$, ktoré sú nezávislé, s rozdelením $P[X_i = 1] = P[X_i = -1] = 1/2$. Pomocou Černovovej nerovnosti nájdime hornú hranicu na pravdepodobnosť udalosti $[\sum_{i=1}^n X_i \geq a]$ pre zadané $a > 0$.

Riešenie: Označme si $S_n = \sum_{i=1}^n X_i$. Podľa Černovovej nerovnosti platí $P[S_n \geq a] \leq E(e^{tS_n})/e^{ta}$ pre akékoľvek t. Lenže pomocou vety 6.10, resp. poznámky 6.11 máme

$$E(e^{tS_n}) = E(e^{\sum_i tX_i}) = \prod_i E(e^{tX_i}) = (e^t + e^{-t})^n/2^n,$$

lebo $E(e^{tX_i}) = (e^t + e^{-t})/2$ pre každé i. Takže

$$P[S_n \geq a] \leq (e^t + e^{-t})^n/(2^n e^{ta})$$

pre akékoľvek t. V literatúre sa v tomto bode ešte typicky využije nerovnosť[13] $(e^t + e^{-t})/2 \leq e^{t^2/2}$ a dosadí sa špeciálne $t = a/n$, čo vedie na ohraničenie

$$P[S_n \geq a] \leq e^{-\frac{a^2}{2n}}.$$

[12] V podobných situáciách požiadavky vznikajú ako realizácie veľkého počtu nezávislých udalostí malej pravdepodobnosti. To má za následok, že Poissonov model na počet požiadaviek môže veľmi dobre vystihovať skutočnosť; porovnaj s vetou 4.48. Tiež si pripomeňme, že náhodná premenná s Poissonovým rozdelením so strednou hodnotou 100 má aj rozptyl 100; pozri vetu 4.53.

[13] Uvedená nerovnosť zjavne platí, ak je funkcia $h(t) = t^2/2 - \ln((e^t + e^{-t})/2)$ nezáporná na celom $\mathbb{R}$. Avšak ľahko sa presvedčíme, že $h'(0) = 0$ a $h''(t) \geq 0$ pre všetky t, čo už vyžadovanú nezápornosť zaručuje.

To demonštruje, že veľké výchylky „symetrickej náhodnej prechádzky“[14] sú málo pravdepodobné. Existuje veľa foriem Černovovej nerovnosti pre rôzne situácie, s ktorými sa možno stretnete v budúcnosti.

6.3 Zákon veľkých čísel a centrálna limitná veta

„Zákon veľkých čísel“ hovorí, že ak budeme sledovať postupnosť realizácií nezávislých náhodných premenných s rovnakou strednou hodnotou μ a spoločne ohraničenými rozptylmi, tak aritmetický priemer zaznamenaných realizácií sa bude približovať k μ. Trochu presnejšie, takzvaný „slabý“[15] zákon veľkých čísel hovorí: k nule konverguje pravdepodobnosť, že sa aritmetický priemer vypočítaný po n realizáciách odchýli od μ o viac ako akékoľvek pevné $\varepsilon > 0$. To má zásadný význam, pretože aritmetický priemer nezávislých náhodných pozorovaní sa veľmi často vyskytuje pri štatistickej analýze dát.

Na druhej strane „centrálna limitná veta“[16] tvrdí, voľne vyjadrené, že ak je n veľké číslo a $X_1, \ldots, X_n$ sú nezávislé, rovnako rozdelené náhodné premenné s konečným no nenulovým rozptylom, tak $S_n = X_1 + \cdots + X_n$ má približne normálne rozdelenie. Presnejšie vyjadrené, ak budeme S_n štandardizovať lineárnou transformáciou na nulovú strednú hodnotu a jednotkový rozptyl, tak distribučná funkcia S_n sa bude blížiť k distribučnej funkcii rozdelenia $N(0, 1)$.

Veta 6.22 (Slabý zákon veľkých čísel)**.** Nech $X_1, X_2, \ldots$ sú nezávislé (diskrétne alebo spojité) náhodné premenné, pričom $D(X_n) \leq \sigma^2$ pre nejaké $\sigma^2 < \infty$ a každé $n \in \mathbb{N}$. Nech $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Potom pre akékoľvek $\varepsilon > 0$ platí:

$$\lim_{n\to\infty} P[|\bar{X}_n - E(\bar{X}_n)| \geq \varepsilon] = 0.$$

Dôkaz. Z predpokladov tvrdenia, vety 4.20 (resp. ekvivalentu tejto vety pre spojité premenné) a poznámky 6.15 k vete 6.14 dostávame

$$D(\bar{X}_n) = D\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n^2} D\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2}\sum_{i=1}^n D(X_i) \leq \frac{\sigma^2}{n}.$$

Dôkaz môžeme ukončiť použitím Čebyševovej nerovnosti (veta 6.17). □

Príklad 6.23. Formalizujme a kriticky zhodnoťme nasledovné voľné tvrdenia: a) Ak budeme donekonečna hádzať mincou, tak podiel výsledkov typu „znak“ sa bude blížiť k 50 percentám; b) Ak máme váhu, ktorej merania majú určitú chybovosť, tak pomocou priemeru z dostatočného počtu vážení vieme s ľubovoľnou presnosťou odhadnúť hmotnosť μ váženého objektu.

[14] Pozri časť 6.4.

[15] V teórii pravdepodobnosti existuje aj príbuzné tvrdenie, ktoré sa nazýva „silný“ zákon veľkých čísel. Slabý zákon veľkých čísel tvrdí, že postupnosť náhodných premenných $(\bar{X}_n)_{n=1}^\infty$ zo znenia vety 6.22 „konverguje podľa pravdepodobnosti“. Naproti tomu silný zákon veľkých čísel tvrdí, že táto postupnosť konverguje aj v matematicky silnejšom zmysle, takzvane „skoro isto“. Poznamenajme ešte, že konvergencia postupnosti náhodných premenných Y_n z vety 6.24 sa nazýva konvergencia podľa distribučnej funkcie. Rôznymi typmi konvergencie postupností náhodných premenných sa v týchto skriptách nebudeme detailnejšie zaoberať.

[16] Existuje viacero príbuzných tvrdení, ktoré sa nazývajú „centrálna limitná veta“; my sa v týchto skriptách budeme venovať len jednému z nich.

Riešenie: Výrok a) je možné formalizovať nasledovne: Nech $X_1, X_2, \dots$ sú nezávislé náhodné premenné s rozdelením $Alt(0.5)$, pričom $X_i = 0$ znamená, že v i-tom hode padla hlava a $X_i = 1$ znamená, že v i-tom hode padol znak. Potom pre akúkoľvek odchýlku $\varepsilon > 0$ platí $\lim_{n\to\infty} P[|\bar{X}_n - 0.5| \geq \varepsilon] = 0$. Predpokladali sme, že výsledky hodov sa navzájom neovplyvňujú (formálne: $X_1, X_2, \dots$ sú nezávislé) a že minca, ktorou hádžeme, nie je vychýlená v prospech niektorého z dvojice možných výsledkov (formálne: $E(X_i) = 0.5$ pre každé i). Pri reálnom hádzaní mincou by tieto predpoklady mohli byť splnené s vysokou presnosťou. Pozri tiež obrázok 6.1.

Výrok b) sa dá formalizovať nasledovne: Nech $X_1, X_2, \dots$ sú nezávislé náhodné premenné so strednou hodnotou μ, pričom X_i je výsledok i-teho váženia. Rozptyl náhodných premenných X_i je ohraničený nejakou konečnou hodnotou σ^2. Potom pre akúkoľvek odchýlku $\varepsilon > 0$ platí $\lim_{n\to\infty} P[|\bar{X}_n - \mu| \geq \varepsilon] = 0$. Opäť si všimnime, že tento výrok predpokladá, že výsledok jedného merania neovplyvní výsledok druhého merania, čo pri niektorých váhach nemusí byť splnené[17] a že stredná hodnota každého merania je presne μ, čo je splnené len vtedy, ak váha nevykazuje „systematickú chybu“. Tiež môže byť v reálnych situáciách porušený predpoklad nemennej hmotnosti objektu, ktorý vážime. Pri reálnom mnohonásobnom vážení by priemer výsledkov nemusel konvergovať ku skutočnej hmotnosti, ak by použitá váha mala spomínané nedostatky.

Veta 6.24 (Centrálna limitná veta)**.** Nech $X_1, X_2, \dots$ sú nezávislé náhodné premenné s rovnakým rozdelením (t.j. s rovnakou distribučnou funkciou), konečnou strednou hodnotou a nenulovým, konečným rozptylom. Nech $S_n = \sum_{i=1}^n X_i$ a nech

$$Y_n = \frac{S_n - E(S_n)}{\sqrt{D(S_n)}}$$

pre každé $n \in \mathbb{N}$. Nech F_{Y_n} je distribučná funkcia náhodnej premennej Y_n a Φ nech je distribučná funkcia rozdelenia $N(0,1)$. Potom

$$\lim_{n\to\infty} F_{Y_n}(x) = \Phi(x) \text{ pre všetky } x \in \mathbb{R}.$$

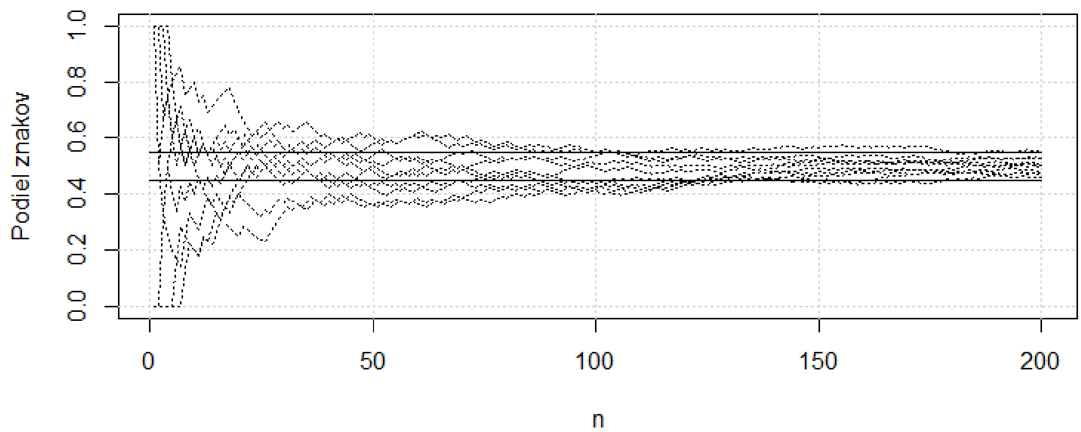
Dôkaz. Dôkaz centrálnej limitnej vety presahuje rámec týchto skrípt; typicky sa centrálna limitná veta dokazuje pomocou takzvaných „charakteristických“ funkcií, ktoré si vyžadujú znalosti z komplexnej analýzy. □

Poznámka 6.25. Všimnime si, že ak si vo vete 6.24 označíme symbolom μ spoločnú strednú hodnotu náhodných premenných $X_1, X_2, \dots$ a symbolom σ^2 spoločný (konečný a nenulový) rozptyl týchto náhodných premenných, tak normalizovaný súčet S_n sa dá vyjadriť nasledovne:

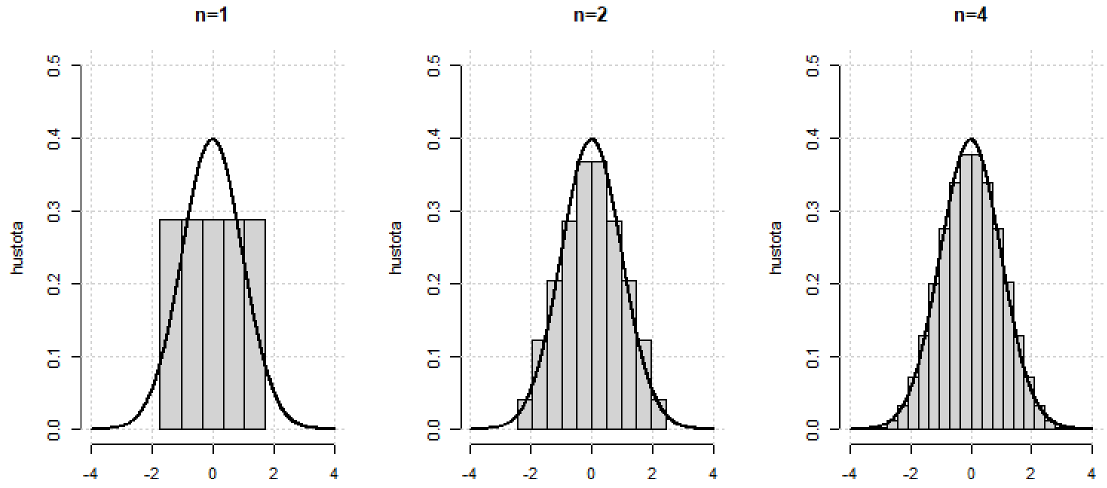
$$Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \sqrt{n}\,\frac{\bar{X}_n - \mu}{\sigma}.$$

Voľne vyjadrené, ak sú splnené predpoklady vety 6.24, tak pre veľké n má $S_n = \sum_{i=1}^n X_i$ približne rozdelenie $N(n\mu, n\sigma^2)$ a priemer $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ má približne rozdelenie $N(\mu, \sigma^2/n)$.

[17] Existujú však aj verzie zákona veľkých čísel, v ktorých náhodné premenné nie sú nezávislé a ich priemer napriek tomu konverguje k spoločnej strednej hodnote, takže určitá forma závislosti medzi meraniami $X_1, X_2, \dots$ nemusí zabrániť tomu, že $\bar{X}_n$ konverguje k μ.



Obr. 6.1: Podiel výsledkov typu „znak“ pri n nezávislých hodoch mincou. Každá z 10 nezávislých simulácií je zobrazená bodkovanou čiarou. Plné vodorovné čiary naznačujú interval ± 0.05 okolo hodnoty 0.5. Simulácie naznačujú, že pravdepodobnosť odchýlky podielu znakov od hodnoty 0.5 o viac ako $\varepsilon = 0.05$ sa blíži k nule. To je v súlade so zákonom veľkých čísel 6.22; pozri tiež časť a) príkladu 6.23. Konvergencia k hodnote 0.5 však nie je veľmi rýchla; na desaťnásobne bližšie priblíženie k tejto hodnote potrebujeme približne stonásobne väčší počet hodov, pretože smerodajná odchýlka priemeru $\bar{X}_n$ klesá s odmocninou z n.



Obr. 6.2: Histogram štandardizovaného súčtu n nezávislých náhodných premenných s rozdelením $R(0,1)$ pre $n = 1, 2, 4$. Čiara označuje hustotu rozdelenia $N(0,1)$.

Poznámka 6.26. Uvedené skutočnosti je možné použiť na počítačové simulovanie realizácií rozdelenia $N(\mu, \sigma^2)$ so zadanými parametrami μ a σ^2. Ak $U_1, \dots, U_n$ sú nezávislé náhodné premenné s rozdelením $R(0,1)$ a $S_n = \sum_{i=1}^n U_i$, potom náhodná premenná

$$Y_n = \sigma \left(\frac{S_n - n/2}{\sqrt{n/12}} + \mu \right)$$

má približne požadované rozdelenie $N(\mu, \sigma^2)$. Na ilustráciu si zobrazme histogram veľkého počtu realizácií náhodných premenných Y_n pre $\mu = 0$, $\sigma^2 = 1$ a $n = 1, 2, 4$; pozri obrázok 6.2. Vidíme, že už súčet štyroch nezávislých náhodných premenných s rovnomerným rozdelením má rozdelenie veľmi podobné normálnemu. Na simulácie z normálneho rozdelenia v praxi obvykle plne postačuje $n = 12$. Realizácie z normálneho rozdelenia a iných rozdelení je možné efektívne generovať mnohými spôsobmi, ktorými sa zaoberá oblasť „stochastických simulačných metód“.

Centrálnu limitnú vetu vieme použiť na približný výpočet pravdepodobnosti v množstve situácií. Vyriešme si pomocou tejto vety dve komplexnejšie úlohy.

Príklad 6.27. Majme pozorovania určitého javu, ktoré je možné reprezentovať nezávislými spojitými náhodnými premennými $Y_1, \dots, Y_n$. Pozorované čísla, zaokrúhlené na m desatinných miest, ukladáme do premenných $Z_1, \dots, Z_n$. Odhadnime rozdelenie odchýlky skutočného priemeru hodnôt $Y_1, \dots, Y_n$ od priemeru zaokrúhlených hodnôt $Z_1, \dots, Z_n$.

Riešenie: Hľadaná diferencia $\frac{1}{n}\sum_{i=1}^n Y_i - \frac{1}{n}\sum_{i=1}^n Z_i$ sa dá vyjadriť ako $\sum_{i=1}^n X_i$, kde $X_i = (Y_i - Z_i)/n$, $i = 1, \dots, n$. Ak predpokladáme, že hodnoty Y_k fluktuujú v rozsahu o mnoho rádov širšom než 10^{-m}, tak odchýlka $Y_i - Z_i$ bude mať približne rozdelenie $R(-10^{-m}/2, 10^{-m}/2)$. To znamená, že X_i bude mať približne rozdelenie $R(-10^{-m}/(2n), 10^{-m}/(2n))$, ktoré má strednú hodnotu 0 a rozptyl $10^{-2m}/(12n^2)$; pozri vetu 5.29. Podľa centrálnej limitnej vety je teda rozdelenie náhodnej premennej $\sum_{i=1}^n X_i$ približne $N(0, \sigma_n^2)$, kde $\sigma_n^2 = n \cdot 10^{-2m}/(12n^2) =$

$10^{-2m}/(12n)$. Smerodajná odchýlka tohto rozdelenia je $\sigma_n = 10^{-m}/\sqrt{12n}$, takže vidíme, že nepresnosť klesá, a to s odmocninou z n.

Príklad 6.28. Z prieskumov vieme, že hmotnosť náhodne zvoleného pasažiera prepravovaného istou leteckou spoločnosťou má rozdelenie so strednou hodnotou $\mu_0 = 78$ kilogramov a rozptylom $\sigma_0^2 = 225$ kilogramov na druhú. Letenku si kúpilo $n = 200$ ľudí, ale z dlhodobých záznamov vieme, že majiteľ letenky sa s pravdepodobnosťou $p = 0.05$ na let nedostaví. Odhadnime pravdepodobnosť, že hmotnosť všetkých pasažierov spolu, ktorí sa na let dostavia, bude viac než $M = 15000$ kilogramov.

Riešenie: Majiteľ letenky k, ktorého telesná hmotnosť je V_k, prispeje k celkovej hmotnosti prepravovaných pasažierov hodnotou X_k. Pritom $X_k = 0$ s pravdepodobnosťou p a $X_k = V_k$ s pravdepodobnosťou $1-p$, teda $X_k = I_k V_k$, kde $I_k \sim Alt(1-p)$. Hľadané rozdelenie súčtu $\sum_{i=1}^n X_i$ je podľa centrálnej limitnej vety približne normálne, no na to, aby sme určili jeho parametre, potrebujeme vypočítať strednú hodnotu a rozptyl samotných náhodných premenných $X_1, \dots, X_n$.

Nech k je akýkoľvek index z $\{1, \dots, n\}$. Z vety 6.10 máme $\mu = E(X_k) = E(I_k)E(V_k) = (1-p)\mu_0$. Z definície rozptylu dostávame $\sigma^2 = D(X_k) = E(X_k^2) - \mu^2 = E(I_k^2 V_k^2) - \mu^2$. Využitím rovnosti $I_k^2 = I_k$ a opäť vety 6.10 máme $E(I_k^2 V_k^2) = E(I_k)E(V_k^2) = E(I_k)(D(V_k) + E^2(V_k)) = (1-p)(\sigma_0^2 + \mu_0^2)$, takže $\sigma^2 = (1-p)(\sigma_0^2 + \mu_0^2) - (1-p)^2\mu_0^2$. Celková hmotnosť prepravovaných pasažierov bude teda mať približne normálne rozdelenie so strednou hodnotou $n\mu$ a rozptylom $n\sigma^2$. Pravdepodobnosť, že toto číslo presiahne M, je

$$P\left[\sum_{i=1}^n X_i > M\right] = P\left[\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} > \frac{M - n\mu}{\sigma\sqrt{n}}\right] \approx 1 - \Phi\left(\frac{M - n\mu}{\sigma\sqrt{n}}\right).$$

Dosadením hodnôt zo zadania dostávame odhad tejto pravdepodobnosti číselne 0.285.

Zamyslime sa ale ešte kriticky nad spoľahlivosťou tohto výsledku. Po prvé, využili sme vetu, ktorá má len limitný charakter a nevieme s istotou tvrdiť, že n je už „dostatočne veľké“. Urobili sme ale aj iné zjednodušenia: Predpokladali sme nezávislosť všetkých náhodných premenných $V_1, \dots, V_n, I_1 \dots, I_n$, kde V_k zodpovedajú hmotnostiam pasažierov a I_k sú indikátormi udalostí, že sa majitelia leteniek dostavia na odlet. To ale v realite určite nie je celkom splnené. (Pokúste sa zdôvodniť prečo.)

6.4 Náhodná prechádzka

Množstvo náhodných dejov nie je možné modelovať ako pozorovania nezávislých náhodných premenných. V tejto časti si uvedieme jednoduchý príklad takého deja.

Príklad 6.29 (Náhodná prechádzka)**.** Uvažujme nasledovnú situáciu. Nech $X_1, X_2, \dots$ je postupnosť nezávislých diskrétnych náhodných premenných s rozdelením určeným rovnosťami $P[X_n = 1] = p$ a $P[X_n = -1] = q$ pre všetky $n \in \mathbb{N}$, kde $p \in (0, \frac{1}{2})$[18] je známe a $q = 1-p$. Definujme postupnosť náhodných premenných $S_0, S_1, S_2, \dots$ nasledovne: S_0 nadobúda hodnotu 0 s pravdepodobnosťou 1 a pre ostatné náhodné premenné platí

$$S_n = S_{n-1} + X_n, \ \ n \geq 1.$$

[18] V tomto príklade zámerne uvažujeme prípad $p \neq 1/2$, čiže „asymetrickú“ náhodnú prechádzku. V pravdepodobnostnom modelovaní sa častejšie vyskytuje „symetrická“ náhodná prechádzka s $p = 1/2$, pozri príklad 6.21.

Modelovaným dejom môže byť napríklad pozícia častice pohybujúca sa vertikálnym smerom, ktorú sledujeme v „diskrétnych časoch“ $0, 1, 2, \ldots$. V čase 0 je častica v bode 0; v každom ďalšom čase „preskočí“ s pravdepodobnosťou p o jednotku nahor a s pravdepodobnosťou q o jednotku nadol. Náhodné premenné $S_0, S_1, S_2, \ldots$ nie sú nezávislé.[19] Ohľadom postupnosti $S_1, S_2, \ldots$ sa môžeme snažiť zodpovedať veľa otázok. Ak je k zadané celé nezáporné číslo, môžeme sa napríklad opýtať: a) Aká je pravdepodobnosť, že v čase $n \geq 1$ bude častica vo výške k? b) Aká je pravdepodobnosť, že maximum všetkých náhodných premenných $S_0, S_1, S_2, \ldots$ (čiže maximálna výška častice, ktorú častica dosiahne za celú dobu svojho pohybu) bude k?

Riešenie: a) Najprv si uvedomme, že n a k musia mať rovnakú paritu a $|k| \leq n$, lebo v opačnom prípade $P[S_n = k] = 0$. Predpokladajme teda, že k a $n \geq |k|$ sú obe párne; prípad nepárnych k, n sa rieši analogicky. Na to, aby nastala udalosť $[S_n = k]$ sa musí spomedzi náhodných premenných $X_1, \ldots, X_n$ realizovať o k viac vo forme $+1$ než vo forme -1, čiže $n/2 + k/2$ z týchto náhodných premenných sa musí realizovať ako $+1$ a $n/2 - k/2$ z nich sa musí realizovať ako -1. Lenže samotné X_i sú nezávislé, takže môžeme použiť binomickú formulu (2.2) a dostávame

$$P[S_n = k] = \binom{n}{\frac{n}{2} + \frac{k}{2}} p^{\frac{n}{2} + \frac{k}{2}} q^{n - \frac{n}{2} - \frac{k}{2}}.$$

Rozdelenie náhodnej premennej S_n teda nie je, striktne vzaté, binomické, ale je binomickému veľmi podobné (pozri definíciu 4.39). Presnejšie, $S_n = 2Z_n - n$, kde $Z_n \sim Bin(n, p)$. To nám umožňuje využiť teóriu binomického rozdelenia a znalosti o lineárnych transformáciách náhodných premenných na odvodenie mnohých zaujímavých charakteristík náhodnej premennej S_n. V rámci cvičenia si môžete napríklad premyslieť, aká je stredná hodnota S_n a aký je rozptyl S_n.

b) Riešenie tejto časti urobíme mierne neporiadne, ale dostatočne presvedčivo. Častica má tendenciu pohybovať sa viac nadol ako nahor, takže sa dá ľahko uveriť, že s pravdepodobnosťou 1 v niektorom konečnom čase nadobudne svoju najvyššiu pozíciu, ktorú už nikdy neprekoná. Takže $M = \max_{n \geq 0} S_n$ je dobre definovaná náhodná premenná.[20] Podľa vety o úplnej pravdepodobnosti

$$P[M \geq k] = P[M \geq k | S_1 = +1] P[S_1 = +1] + P[M \geq k | S_1 = -1] P[S_1 = -1]. \tag{6.4}$$

Uvedomme si, že keby sme mali zaručené, že $S_1 = +1$, tak by rozdelenie náhodnej premennej M bolo o jednotku posunuté nahor, čiže pre všetky $k \geq 0$ platí $P[M \geq k | S_1 = +1] = P[M \geq k - 1]$. Ak by sme naopak mali zaručené, že $S_1 = -1$, tak po krátkej úvahe sa presvedčíme, že pre všetky $k \geq 1$ máme $P[M \geq k | S_1 = -1] = P[M \geq k + 1]$. Takže rovnosť (6.4) môžeme pre všetky $k \geq 1$ prepísať do tvaru

$$P[M \geq k] = P[M \geq k - 1] P[S_1 = +1] + P[M \geq k + 1] P[S_1 = -1].$$

[19] Nie sú nezávislé z pohľadu pravdepodobnostného modelu, ktorý máme pred začiatkom pozorovania celého deja. Dá sa ukázať, že pokiaľ by sme boli informovaní o hodnote náhodnej premennej S_n, tak by táto informácia spôsobila zmenu nášho pravdepodobnostného modelu na taký, v ktorom by boli nezávislé všetky dvojice náhodných premenných S_{n_-} a S_{n_+}, kde $n_- \leq n \leq n_+$.

[20] Na množine pravdepodobnosti 0, ktorá zodpovedá tomu, že častica utečie do $+\infty$, definujeme M ľubovoľne.

Ak si označíme $p_k = P[M \geq k]$, tak dostávame rekurentné lineárne rovnice

$$\begin{aligned} p_k &= pp_{k-1} + qp_{k+1}, k \geq 1, \text{ resp.} \\ p_n &= \frac{1}{q}p_{n-1} - \frac{p}{q}p_{n-2}, n \geq 2. \end{aligned} \tag{6.5}$$

Podľa teórie lineárnych rekurentných rovníc je všeobecné riešenie systému (6.5) tvaru $p_n = A\lambda_1^n + B\lambda_2^n$, pre všetky $n \geq 0$, kde λ_1, λ_2 sú korene „charakteristického polynómu" systému (6.5) $\lambda^2 - \frac{1}{q}\lambda + \frac{p}{q}$, čiže $\lambda_1 = 1$ a $\lambda_2 = \frac{p}{q}$. Všeobecné riešenie teda dostávame $p_n = A + B(p/q)^n$. Aké je A a B? V prvom rade $p_0 = P[M \geq 0] = 1$, takže $A = 1 - B$, čiže $p_n = 1 - B + B(p/q)^n$. Zároveň sme však uverili, že M je dobre definovaná náhodná premenná, takže musí platiť $\lim_{n\to\infty} P[M \geq n] = 0$. Preto musíme nutne mať $B = 1$. Dostali sme $p_n = (p/q)^n$, teda

$$P[M = n] = P[M \geq n] - P[M \geq n+1] = p_n - p_{n+1} = \left(1 - \frac{p}{q}\right)\left(\frac{p}{q}\right)^n.$$

Takže (pozri definíciu 4.56) M má presne geometrické rozdelenie s parametrom $1 - p/q$. Tým pádom vieme o maximálnej výške M, ktorú celkovo naša skákajúca častica dosiahne, aj veľa iných poznatkov, napríklad strednú hodnotu a rozptyl.

Samozrejme ohľadom náhodnej prechádzky je možné položiť veľa otázok, ktoré sú omnoho zložitejšie ako uvedené dve.

Poznámka 6.30. Náhodná prechádzka je elementárnym príkladom „markovovského reťazca". Vo všeobecnom markovovskom reťazci hypotetická častica preskakuje zo stavu i do iného stavu j s určitou (často známou) pravdepodobnosťou p_{ij}, pričom $p_{ij} \geq 0$ sú akékoľvek, musia len spĺňať $\sum_j p_{ij} = 1$ pre každý stav i. Markovovské reťazce tvoria dôležitú oblasť teórie pravdepodobnosti, s aplikáciami napríklad v „teórii kódovania", v „systémoch hromadnej obsluhy" a v „umelej inteligencii". Všeobecnejším modelom sú takzvané „časové rady", v ktorých je charakter závislosti pozorovaných náhodných premenných komplikovanejší; časové rady sa využívajú v ekonomickom modelovaní, v medicínskych aplikáciách, v meteorológii a podobne. Ešte zložitejšie modely musíme použiť v prípadoch, v ktorých pozorujeme číselné hodnoty indexované nie prirodzenými číslami, ale všetkými reálnymi číslami (v určitom intervale). Takýmito súbormi náhodných premenných sa zaoberá všeobecná „teória náhodných procesov" s aplikáciami napríklad vo fyzike a vo finančnej matematike. Spomenuté dôležité súčasti pokročilej teórie pravdepodobnosti však presahujú rámec týchto skrípt; študenti sa s nimi môžu oboznámiť na špecializovaných prednáškach.

Kapitola 7

Náhodné vektory

7.1 Základy teórie náhodných vektorov

Koncept náhodného vektora nám pomáha analyzovať vzťahy medzi náhodnými premennými. Umožňuje mám tiež narábať s viacerými náhodnými premennými súčasne. Konečný počet závislých náhodných premenných je typickou situáciou pri modelovaní a spracovávaní mnohorozmerných dát. Napríklad skupina „vysvetľujúcich“ premenných meraných na klasifikovaných objektoch, v istých aplikáciách nazývaných „prediktory“, je často konečná množina komplikovane závislých náhodných premenných.

7.1.1 Všeobecné náhodné vektory

Definícia 7.1. Nech $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor. Nech $X_1, \ldots, X_m : \Omega \to \mathbb{R}$ sú náhodné premenné. Potom[1] $\mathbf{X} = (X_1, \ldots, X_m)^T$ nazývame m-rozmerný **náhodný vektor**.[2]

Poznámka 7.2. Náhodný vektor $\mathbf{X} = (X_1, \ldots, X_m)^T$ sa dá chápať aj ako zobrazenie z Ω do $\mathbb{R}^m$, ktoré elementárnemu výsledku ω priraďuje vektor $\mathbf{X}(\omega) = (X_1(\omega), \ldots, X_m(\omega))^T$. Z tohto pohľadu je $\mathbf{X}(\Omega) = \{\mathbf{X}(\omega) : \omega \in \Omega\}$ oborom hodnôt náhodného vektora $\mathbf{X}$, čiže množinou m-rozmerných vektorov, ktoré $\mathbf{X}$ „nadobúda“. Dá sa ukázať, že v tomto zmysle je zobrazenie $\mathbf{X} : \Omega \to \mathbb{R}^m$ náhodným vektorom práve vtedy, ak sú všetky množiny $\{\omega \in \Omega : \mathbf{X}(\omega) \in B\}$, kde $B \in \mathcal{B}_m$, udalosťami z $\mathcal{S}$; porovnaj s vetou 3.3.

Poznámka 7.3. Pre značenie udalostí určených náhodným vektorom $\mathbf{X} = (X_1, \ldots, X_m)^T$ používame podobnú konvenciu ako v prípade náhodných premenných. Napríklad, ak $x_1, \ldots, x_m \in \mathbb{R}$ sú fixné čísla a $\mathbf{x} = (x_1, \ldots, x_m)^T$, tak všetky nasledovné výrazy označujú tú istú udalosť: $\{\omega \in \Omega : \mathbf{X}(\omega) = \mathbf{x}\}$, $[\mathbf{X} = \mathbf{x}]$, $\cap_{i=1}^m [X_i = x_i]$, $[X_1 = x_1, \ldots, X_m = x_m]$.

Príklad 7.4. Hádžeme dvomi kockami. Nech X_1 je náhodná premenná reprezentujúca minimum padnutých čísel a X_2 je náhodná premenná reprezentujúca maximum padnutých čísel. Potom $(X_1, X_2)^T$ je náhodný vektor, ktorý nadobúda vektory z množiny $\mathbf{X}(\Omega) = \{(i_1, i_2)^T : 1 \le i_1 \le i_2 \le 6\}$. Pravdepodobnosti udalostí $[X_1 = x_1, X_2 = x_2]$, kde $x_1, x_2 \in \{1, \ldots, 6\}$,

[1] Symbol T znamená transpozíciu riadkového vektora $(X_1, \ldots, X_m)$. V tomto texte, tak ako vo väčšine učebníc a odborných článkov v oblasti teórie pravdepodobnosti a matematickej štatistiky, chápeme vektory ako stĺpce.

[2] Formálne je teda klasická náhodná premenná jednorozmerný náhodný vektor.

zobrazuje tabuľka 7.1. Samotná náhodná premenná X_1 nadobúda hodnoty $1, \dots, 6$ s pravdepodobnosťami 11/36, 9/36, 7/36, 5/36, 3/36 a 1/36, ktoré môžeme získať ako stĺpcové súčty pravdepodobností v tabuľke 7.1. Podobne, náhodná premenná X_2 nadobúda hodnoty $1, \dots, 6$ s pravdepodobnosťami 1/36, 3/36, 5/36, 7/36, 9/36 a 11/36, ktoré môžeme získať ako riadkové súčty pravdepodobností v tabuľke 7.1. Všimnime si, že náhodné premenné X_1 a X_2 nie sú nezávislé, čo plynie napríklad z vety 6.9.

$x_2 \downarrow$ $x_1 \rightarrow$	1	2	3	4	5	6
6	2/36	2/36	2/36	2/36	2/36	1/36
5	2/36	2/36	2/36	2/36	1/36	0
4	2/36	2/36	2/36	1/36	0	0
3	2/36	2/36	1/36	0	0	0
2	2/36	1/36	0	0	0	0
1	1/36	0	0	0	0	0

Tabuľka 7.1: Pravdepodobnosti $P[X_1 = x_1, X_2 = x_2]$ z príkladu 7.4.

Poznámka 7.5. Náhodné vektory môžeme rôznymi spôsobmi transformovať, pričom výsledok je opäť náhodný vektor. Presnejšie, ak by sme mali m-rozmerný náhodný vektor $\mathbf{X}$ a borelovské funkcie $g_1, \dots, g_k : \mathbb{R}^m \to \mathbb{R}$ (pozri definíciu 3.9), tak $\mathbf{Y} = (g_1(\mathbf{X}), \dots, g_k(\mathbf{X}))^T$ je k-rozmerný náhodný vektor. Keďže všetky „slušné“ funkcie sú borelovské (pozri napríklad lemu 3.10), tak v zásade akákoľvek „slušná“ transformácia náhodného vektora je opäť náhodný vektor.

Definícia 7.6. Distribučná funkcia náhodného vektora $\mathbf{X} = (X_1, \dots, X_m)^T$ je funkcia $F : \mathbb{R}^m \to \mathbb{R}$, ktorá je definovaná

$$F(\mathbf{x}) = P[X_1 < x_1, \dots, X_m < x_m] \text{ pre všetky } \mathbf{x} = (x_1, \dots, x_m)^T \in \mathbb{R}^m.$$

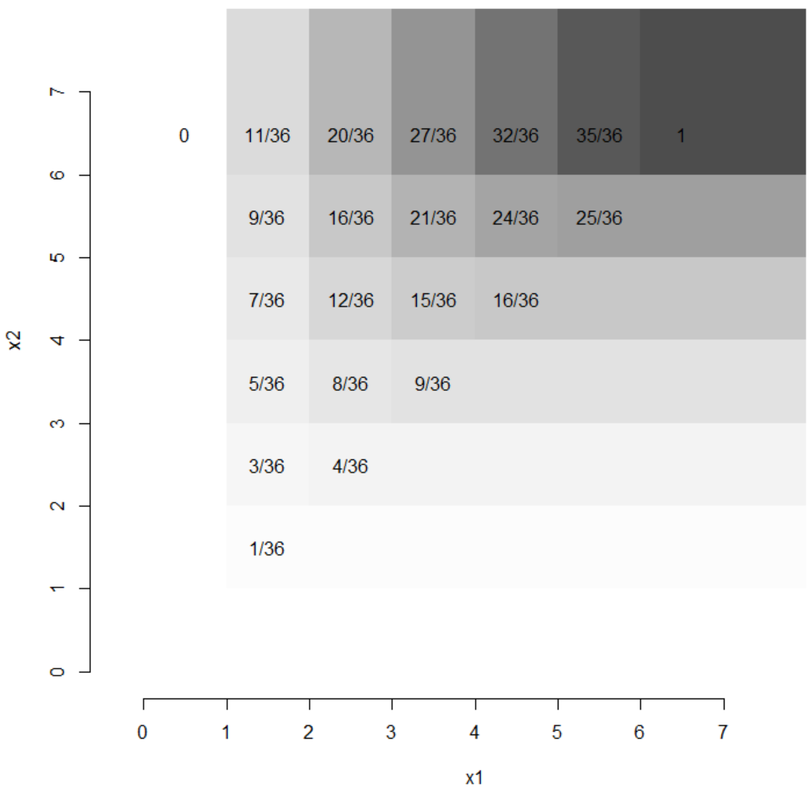
Poznámka 7.7. Je možné ukázať, že distribučná funkcia m-rozmerného náhodného vektora $\mathbf{X}$ jednoznačne určuje pravdepodobnosť $P[\mathbf{X} \in B]$ pre akúkoľvek m-rozmernú borelovskú množinu B. V nasledujúcej vete uvedieme dôležitý špeciálny prípad, v ktorom je B mnohorozmerný kváder.

Veta 7.8 (Vyjadrenie pravdepodobnosti kvádrov pomocou distribučnej funkcie náhodného vektora)**.** Nech F je distribučná funkcia náhodného vektora $\mathbf{X} = (X_1, \dots, X_m)^T$ a nech $a_i^{(0)} \in \mathbb{R}$, $a_i^{(1)} \in \mathbb{R} \cup \{\infty\}$, pričom $a_i^{(0)} < a_i^{(1)}$ pre každé $i = 1, \dots, m$. Potom

$$P[a_1^{(0)} \le X_1 < a_1^{(1)}, a_2^{(0)} \le X_2 < a_2^{(1)}, \dots, a_m^{(0)} \le X_m < a_m^{(1)}] = \sum_{k_1, \dots, k_m \in \{0,1\}} (-1)^{m-k_1-\dots-k_m} F(a_1^{(k_1)}, \dots, a_m^{(k_m)}). \tag{7.1}$$

Dôkaz. Pre jednoduchosť označme $A_i = [a_i^{(0)} \le X_i < a_i^{(1)}]$ a $B_i^{(k_i)} = [X_i < a_i^{(k_i)}]$ pre všetky $i = 1, \dots, m$. Všimnime si, že pre akúkoľvek udalosť C (špeciálne pre $C = \emptyset$) a akékoľvek i platí

$$P(A_i \cap C) = P(B_i^{(1)} \cap C) - P(B_i^{(0)} \cap C); \tag{7.2}$$



Obr. 7.1: Distribučná funkcia $F(x_1, x_2)$ náhodného vektora z príkladu 7.4. Ako každá distribučná funkcia dvojrozmerného náhodného vektora, F je neklesajúca a spojitá zľava v oboch premenných (spojitosť zľava ilustrácia nezachycuje). Platí $F(x_1, x_2) \to 1$, ak $x_1 \to \infty$ a súčasne $x_2 \to \infty$. Taktiež $F(x_1, x_2) \to 0$, ak $x_1 \to -\infty$ alebo $x_2 \to -\infty$. Pre fixné x_1 je $\lim_{x_2 \to \infty} F(x_1, x_2)$ rovná hodnote distribučnej funkcie náhodnej premennej X_1 v bode x_1. Podobne, pre fixné x_2 je $\lim_{x_1 \to \infty} F(x_1, x_2)$ rovná hodnote distribučnej funkcie náhodnej premennej X_2 v bode x_2. Pre diskrétny náhodný vektor (pozri časť 7.1.2), akým je náhodný vektor z príkladu 7.4, je distribučná funkcia „schodovitá“. Pre spojité náhodné vektory (pozri časť 7.1.3) je však distribučná funkcia spojitá.

pozri vetu 1.38. Pomocou uvedeného značenia je možné rovnosť (7.1) prepísať nasledovne:

$$P\left(\cap_{i=1}^m A_i\right) = \sum_{k_1,\dots,k_m \in \{0,1\}} (-1)^{m-k_1-\dots-k_m} P\left(\cap_{i=1}^m B_i^{(k_i)}\right).$$

Indukciou dokážeme silnejšie tvrdenie, a to že pre každé $r \in \{1,\dots,m\}$ a každú udalosť C_r:

$$P(\cap_{i=1}^r A_i \cap C_r) = \sum_{k_1,\dots,k_r \in \{0,1\}} (-1)^{r-k_1-\dots-k_r} P\left(\cap_{i=1}^r B_i^{(k_i)} \cap C_r\right). \tag{7.3}$$

Pre $r = 1$ táto rovnosť platí, pretože $P\left(A_1 \cap C_1\right) = P(B_1^{(1)} \cap C_1) - P(B_1^{(0)} \cap C_1)$ je špeciálny prípad rovnosti (7.2). Predpokladajme, že $m \geq 2$. Ukážme nasledovný výrok: ak (7.3) platí pre akékoľvek $r \in \{1,\dots,m-1\}$ a pre akúkoľvek udalosť C_r, tak (7.3) platí aj pre $r+1$ a akúkoľvek udalosť C_{r+1}.

$$\begin{aligned}
P(\cap_{i=1}^{r+1} A_i \cap C_{r+1}) &= P\left(\cap_{i=1}^r A_i \cap A_{r+1} \cap C_{r+1}\right) = \\
& \sum_{k_1,\dots,k_r \in \{0,1\}} (-1)^{r-k_1-\dots-k_r} P\left(\cap_{i=1}^r B_i^{(k_i)} \cap A_{r+1} \cap C_{r+1}\right) = \\
& \sum_{k_1,\dots,k_r \in \{0,1\}} (-1)^{l-k_1-\dots-k_r} P\left(B_{r+1}^{(1)} \cap_{i=1}^r B_i^{(k_i)} \cap C_{r+1}\right) - \\
& \sum_{k_1,\dots,k_r \in \{0,1\}} (-1)^{l-k_1-\dots-k_r} P\left(B_{r+1}^{(0)} \cap_{i=1}^r B_i^{(k_i)} \cap C_{r+1}\right) = \\
& \sum_{k_1,\dots,k_{r+1} \in \{0,1\}} (-1)^{(r+1)-k_1-\dots-k_{r+1}} \left(\cap_{i=1}^{r+1} B_i^{(k_i)} \cap C_{r+1}\right).
\end{aligned}$$

Prvá rovnosť je triviálna, v druhej sme využili indukčný predpoklad s $C_r = A_{r+1} \cap C_{r+1}$, v tretej sme opäť využili (7.2) s $i = r+1$ a $C = \cap_{i=1}^r B_i^{(k_i)} \cap C_{r+1}$ a posledná rovnosť vznikla len prepisom do kompaktnejšieho tvaru. □

Príklad 7.9. Vyjadrime pomocou distribučnej funkcie F náhodného vektora $(X_1, X_2)^T$ pravdepodobnosť toho, že $(X_1, X_2)^T$ padne do obdĺžnika $[a_1, b_1) \times [a_2, b_2)$, kde $a_1 < b_1$, $a_2 < b_2$.

Riešenie: Ide o špeciálny prípad vety 7.8:

$$P[a_1 \leq X_1 < b_1, a_2 \leq X_2 < b_2] = F(b_1, b_2) - F(a_2, b_1) - F(a_1, b_2) + F(a_1, a_2).$$

Ak ste preskočili všeobecný dôkaz vety 7.8, pokúste sa dokázať aspoň tento špeciálny prípad.

Veta 7.10 (Základné vlastnosti distribučnej funkcie náhodného vektora)**.** Nech F je distribučná funkcia m-rozmerného náhodného vektora $\mathbf{X}$. Potom 1) $0 \leq F(\mathbf{x}) \leq 1$ pre všetky $\mathbf{x} \in \mathbb{R}^m$, 2) F je neklesajúca a spojitá zľava v každej premennej, 3) $\lim_{x_1,\dots,x_m \to \infty} F(x_1,\dots,x_m) = 1$, 4) $\lim_{x_i \to -\infty} F(x_1,\dots,x_m) = 0$ pre každé $i = 1,\dots,m$ a $x_1,\dots,x_{i-1},x_{i+1},\dots,x_m \in \mathbb{R}$.

Dôkaz. Tieto vlastnosti sa dajú dokázať analogicky ako základné vlastnosti distribučnej funkcie jednorozmernej náhodnej premennej (veta 3.16). □

Poznámka 7.11. Distribučná funkcia náhodného vektora $(X_1,\dots,X_m)^T$ jednoznačne určuje distribučné funkcie náhodných premenných $X_1,\dots,X_m$, niekedy tiež nazývané „marginálne" distribučné funkcie. Distribučné funkcie náhodných premenných $X_1,\dots,X_m$ však nemusia jednoznačne určovať distribučnú funkciu náhodného vektora $(X_1,\dots,X_m)^T$.

Veta 7.12 (Marginálne distribučné funkcie)**.** Nech $F_{\mathbf{X}}$ je distribučná funkcia náhodného vektora $\mathbf{X} = (X_1, \dots, X_m)^T$ a nech $1 \leq i_1 < \dots < i_k \leq m$. Potom pre distribučnú funkciu $F_{\mathbf{Y}}$ náhodného vektora $\mathbf{Y} = (X_{i_1}, \dots, X_{i_k})^T$ platí

$$F_{\mathbf{Y}}(x_{i_1}, \dots, x_{i_k}) = \lim F_{\mathbf{X}}(x_1, \dots, x_m),$$

kde limitu berieme pre $x_i \to \infty$ pre všetky $i \notin \{i_1, \dots, i_k\}$.

Dôkaz. Dôkaz je jednoduché cvičenie. □

Veta 7.13 (Distribučná funkcia náhodného vektora s nezávislými zložkami)**.** Nech X_i je náhodná premenná s distribučnou funkciou F_i pre $i = 1, \dots, m$. Nech F je distribučná funkcia náhodného vektora $(X_1, \dots, X_m)^T$. Potom platí: Náhodné premenné $X_1, \dots, X_m$ sú združene nezávislé vtedy a len vtedy, keď

$$F(x_1, \dots, x_m) = F_1(x_1) \cdots F_m(x_m)$$

pre všetky $x_1, \dots, x_m \in \mathbb{R}$.

Dôkaz. Veta vyplýva z ekvivalencie a)⇔d) vo vete 6.7 a z definície 7.6. □

Poznámka 7.14. Na modelovanie niektorých situácií, najmä v mnohorozmernej štatistike, je vhodné používať celú postupnosť náhodných vektorov. Podobne ako náhodné premenné, aj náhodné vektory môžu byť nezávislé alebo závislé. Nezávislosť náhodných vektorov definujeme analogicky ako bola definovaná nezávislosť náhodných premenných v definícii 6.1:

Definícia 7.15. Náhodné vektory $\mathbf{X}_1, \dots, \mathbf{X}_n$ rozmerov $m_1, \dots, m_n$ sa nazývajú **združene nezávislé**, ak pre akýkoľvek výber množín $B_1 \in \mathcal{B}_{m_1}, \dots, B_n \in \mathcal{B}_{m_n}$ platí, že udalosti $[\mathbf{X}_1 \in B_1], [\mathbf{X}_2 \in B_2], \dots, [\mathbf{X}_n \in B_n]$ sú združene nezávislé. Náhodné vektory $\mathbf{X}_1, \mathbf{X}_2, \dots$ nekonečnej postupnosti sa nazývajú združene nezávislé, ak sú združene nezávislé náhodné vektory $\mathbf{X}_1, \dots, \mathbf{X}_n$ pre každé n. Ak náhodné vektory $\mathbf{X}_1, \dots, \mathbf{X}_n$, resp. $\mathbf{X}_1, \mathbf{X}_2, \dots$ nie sú združene nezávislé, hovoríme, že sú **združene závislé**.[3]

Poznámka 7.16. Nezávislosť náhodných vektorov sa dá vyjadriť viacerými ekvivalentnými spôsobmi. Nechávame na čitateľa, aby sa pokúsil pre náhodné vektory formulovať tvrdenie analogické tvrdeniu 6.7.

7.1.2 Diskrétne náhodné vektory

Definícia 7.17. Náhodný vektor $\mathbf{X} = (X_1, \dots, X_m)^T$ nazývame m-rozmerný **diskrétny náhodný vektor**, ak sú všetky náhodné premenné $X_1, \dots, X_m$ diskrétne.

Poznámka 7.18. V súlade s poznámkou 7.2 môžeme náhodný vektor $\mathbf{X} = (X_1, \dots, X_m)^T$ na priestore $(\Omega, \mathcal{S}, P)$ chápať ako zobrazenie z Ω do $\mathbb{R}^m$. Ľahko si premyslíme, že takto chápaný náhodný vektor je diskrétny vtedy a len vtedy, ak je spočítateľný jeho obor hodnôt $\mathbf{X}(\Omega)$.

[3] Upozornime, že náhodné vektory $\mathbf{X}_1 = (X_{11}, \dots, X_{1m_1})^T$, $\mathbf{X}_2 = (X_{21}, \dots, X_{2m_2})^T$ môžu byť nezávislé, aj ak sú náhodné premenné $X_{11}, \dots, X_{1m_1}$ združene závislé a súčasne sú náhodné premenné $X_{21}, \dots, X_{2m_2}$ združene závislé. Analogická poznámka pochopiteľne platí nielen pre dvojicu, ale aj pre n-ticu $\mathbf{X}_1, \dots, \mathbf{X}_n$ nezávislých náhodných vektorov. Často napríklad $\mathbf{X}_1, \dots, \mathbf{X}_n$ tvoria pozorovania m závislých číselných charakteristík (napríklad tlaku a teploty, $m_i \equiv m = 2$) na n nezávislých objektoch (napríklad pacientoch).

Poznámka 7.19. Analogicky ako pri (jednorozmerných) náhodných premenných, hovoríme, že diskrétny náhodný vektor $\mathbf{X}$ „nadobúda“ spočítateľne veľa hodnôt a že diskrétny náhodný vektor $\mathbf{X}$ nadobúda vektor $\mathbf{x}$ „s pravdepodobnosťou“ $P[\mathbf{X} = \mathbf{x}]$. Funkciu, ktorá priraďuje pravdepodobnosť $P[\mathbf{X} = \mathbf{x}]$ vektorom $\mathbf{x} \in \mathbf{X}(\Omega)$, budeme nazývať **pravdepodobnostná funkcia** diskrétneho náhodného vektora $\mathbf{X}$.

Definícia 7.20. Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je diskrétny náhodný vektor a nech existuje konečná stredná hodnota diskrétnej náhodnej premennej X_i pre každé $i = 1, \ldots, m$. Potom hovoríme, že náhodný vektor $\mathbf{X}$ má konečnú **strednú hodnotu**

$$E(\mathbf{X}) = (E(X_1), \ldots, E(X_m))^T.$$

Príklad 7.21. Náhodný vektor $\mathbf{X} = (X_1, X_2)^T$ z príkladu 7.4 je diskrétny; nájdime jeho strednú hodnotu. V príklade 7.4 sme uviedli s akými pravdepodobnosťami nadobúdajú náhodné premenné X_1 a X_2 hodnoty $1, \ldots, 6$. Tieto pravdepodobnosti a definícia 4.6 umožňujú priamočiaro vypočítať, že $E(X_1) = 91/36 \approx 2.528$ a $E(X_2) = 161/36 \approx 4.472$. Takže stredná hodnota celého náhodného vektora $(X_1, X_2)^T$ je $E(\mathbf{X}) = (91, 161)^T/36 \approx (2.528, 4.472)^T$.

Definícia 7.22 (Lineárna transformácia náhodného vektora)**.** Predpokladajme, že $\mathbf{X} = (X_1, \ldots, X_m)^T$ je m-rozmerný náhodný vektor a nech A je matica rozmeru $k \times m$ s prvkami a_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, m$. Potom náhodný vektor $\mathbf{Y} = A\mathbf{X}$ nazývame **lineárna transformácia** náhodného vektora $\mathbf{X}$, čím rozumieme k-rozmerný náhodný vektor $\mathbf{Y} = (Y_1, \ldots, Y_k)^T$ so zložkami

$$Y_i = a_{i1}X_1 + a_{i2}X_2 + \ldots + a_{im}X_m, \; i = 1, \ldots, k. \tag{7.4}$$

Veta 7.23 (Linearita strednej hodnoty diskrétneho náhodného vektora)**.** Nech $\mathbf{X}_1, \ldots, \mathbf{X}_n$ sú m-rozmerné diskrétne náhodné vektory, z ktorých každý má konečnú strednú hodnotu a nech $A_1, \ldots, A_n$ sú matice typu $k \times m$. Potom k-rozmerný diskrétny[4] náhodný vektor $\sum_{l=1}^n A_l \mathbf{X}_l$ má konečnú strednú hodnotu a táto stredná hodnota sa dá vyjadriť v tvare

$$E\left(\sum_{l=1}^n A_l \mathbf{X}_l\right) = \sum_{l=1}^n A_l E(\mathbf{X}_l).$$

Dôkaz. Pomocou linearity strednej hodnoty diskrétnych náhodných premenných 4.12 a základných definícií maticovej algebry je možné priamočiaro ukázať, že i-ta zložka vektora $E(\sum_{l=1}^n A_l \mathbf{X}_l)$ je presne i-ta zložka vektora $\sum_{l=1}^n A_l E(\mathbf{X}_l)$, pre všetky $i = 1, \ldots, k$. □

Poznámka 7.24. Najdôležitejší špeciálny prípad vety 7.23 je: ak m-rozmerný diskrétny náhodný vektor $\mathbf{X}$ má konečnú strednú hodnotu a ak A je matica typu $k \times m$, tak $A\mathbf{X}$ je k-rozmerný diskrétny náhodný vektor, ktorý má konečnú strednú hodnotu $AE(\mathbf{X})$.

Príklad 7.25. Hodíme dvomi hracími kockami. Nájdime strednú hodnotu náhodného vektora $\mathbf{Y} = (Y_1, Y_2)^T$, kde Y_1 je absolútna hodnota rozdielu výsledkov hodov a Y_2 je súčet výsledkov hodov. Tento príklad by sme mohli riešiť detailným spôsobom, analogicky, ako sme riešili príklady 7.4 a 7.21. Výsledok príkladu 7.21 a veta 7.23 nám však poskytujú nasledovnú skratku. Nech $\mathbf{X} = (X_1, X_2)^T$ je náhodný vektor z príkladov 7.4 a 7.21. Základným pozorovaním je, že $\mathbf{Y} = A\mathbf{X}$, kde A je 2×2 matica s číslami $-1, 1$ na diagonále a číslami $1, 1$ mimo diagonály. Takže na základe vety 7.23 a výsledku príkladu 7.21 platí $E(\mathbf{Y}) = AE(\mathbf{X}) = (70/36, 252/36)^T \approx (1.944, 7)^T$.

[4]Pozri poznámku 4.10.

Veta 7.26 (Stredná hodnota funkcie diskrétneho náhodného vektora)**.** Nech m-rozmerný diskrétny náhodný vektor $\mathbf{X}$ nadobúda vektory $(\mathbf{x}_i)_{i\in I}$ s nenulovou pravdepodobnosťou a nech $g:\mathbb{R}^m\to\mathbb{R}$ je funkcia. Potom stredná hodnota diskrétnej[5] náhodnej premennej $g(\mathbf{X})$ existuje a je konečná, ak je rad $\sum_{i\in I} g(\mathbf{x}_i)P[\mathbf{X}=\mathbf{x}_i]$ absolútne konvergentný. V takomto prípade platí:

$$E(g(\mathbf{X}))=\sum_{i\in I} g(\mathbf{x}_i)P[\mathbf{X}=\mathbf{x}_i].$$

Dôkaz. Analogicky ako dôkaz vety 4.11. □

Príklad 7.27. Vypočítajme $E(X_1X_2)$ pre náhodný vektor z príkladu 7.4. Podľa vety 7.26 platí $E(X_1X_2)=\sum x_1x_2P[X_1=x_1,X_2=x_2]$, kde sumujeme cez všetky také dvojice x_1,x_2, pre ktoré $P[X_1=x_1,X_2=x_2]>0$. Túto sumu vieme vypočítať priamo z tabuľky 7.1 ako $1\cdot1\cdot(1/36)+1\cdot2\cdot(2/36)+2\cdot2\cdot(2/36)+\ldots+6\cdot6\cdot(1/36)=441/36=12.25$.[6] Stredná hodnota súčinu dvoch náhodných premenných sa využíva napríklad na výpočet kovariancie a korelácie, s ktorými sa zoznámime neskôr.

Poznámka 7.28. Nezávislosť diskrétnych náhodných vektorov vieme overiť podobným spôsobom, aký nám poskytuje veta 6.9.

7.1.3 Spojité náhodné vektory

Definícia 7.29. Nech $\mathbf{X}=(X_1,\ldots,X_m)^T$ je náhodný vektor s distribučnou funkciou F. Nech existuje integrovateľná funkcia $f:\mathbb{R}^m\to[0,\infty)$ taká, že

$$F(x_1,\ldots,x_m)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_m} f(t_1,\ldots,t_m)\mathrm{d}t_1\ldots\mathrm{d}t_m \text{ pre všetky } x_1,\ldots,x_m\in\mathbb{R}.$$

Potom hovoríme, že $\mathbf{X}$ je **spojitý náhodný vektor** a f je **hustota** náhodného vektora $\mathbf{X}$.

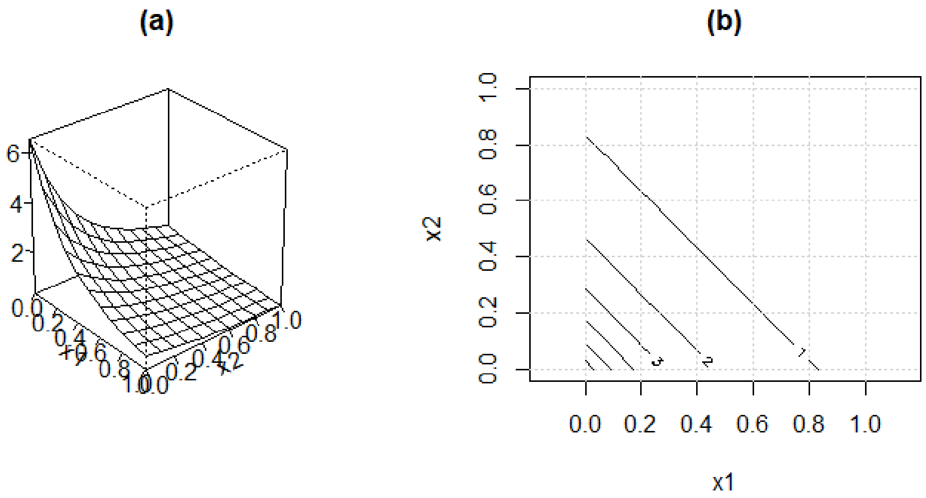
Veta 7.30 (Vyjadrenie pravdepodobnosti borelovských množín pomocou hustoty spojitého náhodného vektora)**.** Nech $\mathbf{X}$ je m-rozmerný spojitý náhodný vektor s hustotou f. Nech $B\subseteq\mathbb{R}^m$ je borelovská množina. Potom

$$P[\mathbf{X}\in B]=\int\!\cdots\!\int_B f(t_1,\ldots,t_m)\mathrm{d}t_1\cdots\mathrm{d}t_m.$$

Poznámka 7.31. Podobne ako v kapitole 5, viaceré tvrdenia tejto časti nebudeme dokazovať, pretože dôkazy závisia od detailov toho, ako mali študenti vybudovaný pojem integrálu. Spojité náhodné vektory sú žiaľ podobne zaťažené zložitými výnimkami ako jednorozmerné spojité náhodné premenné (pozri napríklad poznámky 5.4 a 5.11), čo tiež nebudeme v týchto skriptách podrobnejšie rozoberať.

[5] Pozri poznámku 4.10.

[6] V tomto konkrétnom príklade vieme na výpočet $E(X_1X_2)$ použiť aj nasledovné pozorovanie: $X_1X_2=Z_1Z_2$, kde Z_1,Z_2 sú priamo výsledky hodu na prvej, resp. druhej kocke. Náhodné premenné Z_1,Z_2 sú však nezávislé a ich stredné hodnoty sú 3.5, takže použitím vety 6.10 dostávame $E(X_1X_2)=E(Z_1Z_2)=E(Z_1)E(Z_2)=3.5^2=12.25$.



Obr. 7.2: Náhodný vektor z príkladu 7.32 s parametrom $a = 2.1$. a) trojrozmerný pohľad na hustotu, b) úrovňové množiny („vrstevnice“) hustoty.

Príklad 7.32. Model využívajúci spojitý náhodný vektor zodpovedá súčasnému pozorovaniu či generovaniu viacerých spojitých náhodných premenných, niekedy zložito previazaných a závislých.[7]

Uvažujme napríklad nasledovnú situáciu. Od výrobcu dostávame každý týždeň zásielku veľkého množstva kusov istého výrobku. Z každej zásielky náhodne vyberáme dva výrobky, pričom na oboch uskutočňujeme test doby životnosti. Výsledky sú dvojice kladných reálnych čísel, ktoré môžeme modelovať ako realizácie $(X_1^{(1)}, X_2^{(1)})^T, (X_1^{(2)}, X_2^{(2)})^T, \ldots$ nezávislých spojitých náhodných vektorov. Z dlhodobých pozorovaní sme zistili,[8] že hustota každého z týchto náhodných vektorov je

$$f(x_1, x_2) = a(a+1)(x_1 + x_2 + 1)^{-(a+2)},$$

pre všetky kladné x_1, x_2 a $f(x_1, x_2) = 0$ pre ostatné x_1, x_2. Číslo $a > 2$ je známa konštanta. Takýto model nám umožňuje vypočítať pravdepodobnosť rôznych udalostí týkajúcich sa dvojice dôb životnosti výrobkov, ktoré náhodne vyberieme na testovanie z nasledujúcej zásielky. Napríklad pravdepodobnosť, že pre oba testované výrobky nameriame životnosť väčšiu ako t, je

$$P[X_1 > t, X_2 > t] = \int_t^\infty \int_t^\infty a(a+1)(x_1 + x_2 + 1)^{-(a+2)} \mathrm{d}x_2 \mathrm{d}x_1 = (2t+1)^{-a}.$$

[7] Platí, že ak je $\mathbf{X} = (X_1, \ldots, X_m)^T$ spojitý náhodný vektor, tak každá z náhodných premenných $X_1, \ldots, X_m$ je spojitá, avšak opačne toto tvrdenie neplatí: ak je každá z náhodných premenných $X_1, \ldots, X_m$ spojitá, tak celý náhodný vektor $\mathbf{X} = (X_1, \ldots, X_m)^T$ ešte nemusí byť spojitý. Napríklad ak $X_1 \sim N(0,1)$, tak pre náhodný vektor $\mathbf{X} = (X_1, X_1)^T$ neexistuje hustota v zmysle definície 7.29, takže $\mathbf{X}$ nie je spojitý. Pochopiteľne, takýto náhodný vektor $\mathbf{X}$ nie je ani diskrétny.

[8] Špecifikovanie vhodného modelu a vyjadrenie našej neistoty o rôznych aspektoch modelu je už záležitosť matematickej štatistiky.

Poznámka 7.33. Ak je $\mathbf{X}$ spojitý m-rozmerný náhodný vektor s distribučnou funkciou F, tak za určitých podmienok vieme z F jednoducho vyčítať, že $\mathbf{X}$ je spojitý náhodný vektor, pričom príslušnú hustotu dostaneme pomocou parciálnych derivácií funkcie F. Napríklad ak je distribučná funkcia F spojite diferencovateľná na celom $\mathbb{R}^m$, tak $\mathbf{X}$ je spojitý náhodný vektor a pre jeho hustotu platí

$$f(x_1, \ldots, x_m) = \frac{\partial^m F(x_1, \ldots, x_m)}{\partial x_1 \cdots \partial x_m}$$

pre všetky $x_1, \ldots, x_m \in \mathbb{R}$.

Veta 7.34 (Marginálne hustoty)**.** Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je m-rozmerný spojitý náhodný vektor s hustotou f. Potom je pre každé $i = 1, \ldots, m$ náhodná premenná X_i spojitá s hustotou danou v bode $x_i \in \mathbb{R}$ predpisom

$$f_i(x_i) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_m) \mathrm{d}x_1 \ldots \mathrm{d}x_{i-1} \mathrm{d}x_{i+1} \ldots \mathrm{d}x_m.$$

Príklad 7.35. Hustota jednotlivých náhodných premenných z príkladu 7.32 je

$$f_1(x_1) = \int_0^{\infty} a(a+1)(x_1 + x_2 + 1)^{-(a+2)} \mathrm{d}x_2 = a(x_1 + 1)^{-(a+1)}$$

pre kladné x_1 a $f_1(x_1) = 0$ pre $x_1 \leq 0$. Analogicky $f_2(x_2) = a(x_2 + 1)^{-(a+1)}$ pre $x_2 > 0$ a $f_2(x_2) = 0$ pre $x_2 \leq 0$. Všimnime si, že rozdelenie týchto náhodných premenných je posunuté Paretovo rozdelenie 5.38 s parametrami $k = 1$, $\alpha = a$.

Veta 7.36 (Nezávislosť spojitých náhodných premenných)**.** Uvažujme spojité náhodné premenné $X_1, \ldots, X_n$ s hustotami $f_1, \ldots, f_n$. Ak sú $X_1, \ldots, X_n$ nezávislé, tak funkcia $f : \mathbb{R}^n \to [0, \infty)$ definovaná

$$f(x_1, \ldots, x_n) = f_1(x_1) \cdots f_n(x_n) \tag{7.5}$$

je hustotou náhodného vektora $\mathbf{X} = (X_1, \ldots, X_n)^T$. Naopak, ak pre nejakú hustotu f náhodného vektora $\mathbf{X}$ a hustoty $f_1, \ldots, f_n$ náhodných premenných $X_1, \ldots, X_n$ platí (7.5), potom sú náhodné premenné $X_1, \ldots, X_n$ nezávislé.[9]

Príklad 7.37. Náhodné premenné z príkladu 7.32 nie sú nezávislé, pretože súčin hustôt z príkladu 7.35 nie je hustotou celého náhodného vektora.

Definícia 7.38. Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je spojitý náhodný vektor a nech existuje konečná stredná hodnota každej (spojitej) náhodnej premennej X_i pre všetky $i = 1, \ldots, m$. Potom hovoríme, že náhodný vektor $\mathbf{X}$ má konečnú **strednú hodnotu**

$$E(\mathbf{X}) = (E(X_1), \ldots, E(X_m))^T.$$

Príklad 7.39. Odvoďme strednú hodnotu náhodného vektora $\mathbf{X} = (X_1, X_2)^T$ z príkladu 7.32. Platí $E(\mathbf{X}) = (E(X_1), E(X_2))^T$, kde pre $i = 1, 2$:

$$E(X_i) = \int_0^{\infty} a x_i (x_i + 1)^{-(a+1)} \mathrm{d}x_i = \frac{1}{a(1-a)} [(a x_i + 1)(x_i + 1)^{-a}]_0^{\infty} = \frac{1}{a-1}.$$

[9] Všimnite si, že toto tvrdenie pre spojité náhodné premenné je analogické tvrdeniu 6.9 platnému pre diskrétne náhodné premenné.

Veta 7.40 (Linearita strednej hodnoty spojitého náhodného vektora)**.** Nech $\mathbf{X}_1, \ldots, \mathbf{X}_n$ sú m-rozmerné spojité náhodné vektory, z ktorých každý má konečnú strednú hodnotu a nech $A_1, \ldots, A_n$ sú matice rozmeru $k \times m$. Nech $\mathbf{Y} = \sum_{l=1}^{n} A_l \mathbf{X}_l$. Potom k-rozmerný náhodný vektor $\mathbf{Y}$[10] má konečnú strednú hodnotu danú vzťahom[11]

$$E\left(\sum_{l=1}^{n} A_l \mathbf{X}_l\right) = \sum_{l=1}^{n} A_l E(\mathbf{X}_l).$$

Poznámka 7.41. Špeciálny prípad vety 7.23 je: ak m-rozmerný spojitý náhodný vektor $\mathbf{X}$ má konečnú strednú hodnotu a ak A je matica typu $k \times m$, tak $A\mathbf{X}$ je k-rozmerný náhodný vektor, ktorý má konečnú strednú hodnotu $AE(\mathbf{X})$.

Veta 7.42 (Stredná hodnota funkcie spojitého náhodného vektora)**.** Nech m-rozmerný spojitý náhodný vektor $\mathbf{X}$ má hustotu f. Nech $g : \mathbb{R}^m \to \mathbb{R}$ je borelovská funkcia a nech existuje a je konečný integrál

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |g(x_1, \ldots, x_m)| f(x_1, \ldots, x_m) \mathrm{d}x_1 \ldots \mathrm{d}x_m.$$

Potom náhodná premenná $g(\mathbf{X})$ má konečnú strednú hodnotu a platí[12]

$$E(g(\mathbf{X})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_m) f(x_1, \ldots, x_m) \mathrm{d}x_1 \ldots \mathrm{d}x_m.$$

Príklad 7.43. Vypočítajme $E(X_1 X_2)$ pre náhodný vektor $(X_1, X_2)^T$ z príkladu 7.32. Podľa vety 7.42 platí

$$E(X_1 X_2) = \int_0^{\infty} \int_0^{\infty} x_1 x_2 a(a+1)(x_1 + x_2 + 1)^{-(a+2)} \mathrm{d}x_1 \mathrm{d}x_2 = \frac{1}{(a-1)(a-2)}.$$

7.2 Rozdelenie funkcie náhodných premenných

V teoretických odvodeniach aj v aplikáciách sa často vyskytuje úloha vypočítať funkciu pravdepodobnosti alebo hustotu borelovskej funkcie $h(X_1, X_2)$ dvoch náhodných premenných X_1 a X_2. Typicky sú X_1 a X_2 nezávislé náhodné premenné, obe diskrétne alebo obe spojité, a funkcia $h : \mathbb{R}^2 \to \mathbb{R}$ je súčet, rozdiel, súčin, prípadne podiel.

Uvažujme najprv uvedenú úlohu v prípade, že náhodné premenné X_1, X_2 sú diskrétne. Potom je $H = h(X_1, X_2)$ diskrétna náhodná premenná, ktorá nadobúda hodnoty $s \in I$, kde I je konečná alebo nekonečná spočítateľná množina. Očividne pre každé $s \in I$ platí $P[H = s] = \sum P[X_1 = x_1, X_2 = x_2]$, kde suma prebieha cez všetky také $x_1 \in X_1(\Omega)$ a $x_2 \in X_2(\Omega)$, pre ktoré $h(x_1, x_2) = s$. Tento súčet určite existuje a patrí do intervalu $[0, 1]$,[13] avšak vyjadriť ho ako jednoduchú funkciu premennej s môže byť ťažké.

[10] Tento náhodný vektor nemusí byť ani spojitý, ani diskrétny.
[11] Toto tvrdenie platí, nech by bol náhodný vektor $\mathbf{Y}$ akéhokoľvek typu.
[12] Uvedená rovnosť platí nezávisle od toho, akého typu je náhodná premenná $g(\mathbf{X})$.
[13] Tento súčet navyše nezávisí od poradia sčítavania, ani ak má príslušná suma nekonečne veľa sčítancov.

Príklad 7.44. Budeme hádzať klasickou vyváženou mincou. Nájdime rozdelenie náhodnej premennej H, ktorá znamená, koľkokrát hodíme „znak“, pokým zaznamenáme dvakrát výsledok „hlava“. Tento príklad je možné vyriešiť aj kombinatoricky, použime však postup vysvetlený vyššie. Nech X_1 znamená počet padnutí znaku, kým prvýkrát zaznamenáme výsledok hlava a X_2 znamená počet padnutí znaku medzi zaznamenaním prvého a druhého výsledku hlava. Zrejme $H = X_1 + X_2$. Náhodné premenné X_1 a X_2 sú nezávislé, obe s rozdelením $Geo(1/2)$. Náhodná premenná H nadobúda hodnoty $s = 0, 1, 2, \dots$ a platí

$$P[H = s] = \sum_{x_1, x_2 \geq 0, x_1 + x_2 = s} P[X_1 = x_1, X_2 = x_2] = \sum_{x_1=0}^{s} P[X_1 = x_1, X_2 = s - x_1] =$$
$$\sum_{x_1=0}^{s} P[X_1 = x_1] P[X_2 = s - x_1] = \sum_{x_1=0}^{s} \frac{1}{2^{x_1+1}} \frac{1}{2^{s-x_1+1}} = \frac{s+1}{2^{s+2}}.$$

Našli sme analytický predpis pre rozdelenie náhodnej premennej H.[14]

Ukážme si teraz základnú stratégiu konštrukcie rozdelenia náhodnej premennej $H = h(X_1, X_2)$ v prípade, že $(X_1, X_2)^T$ je spojitý náhodný vektor so známou hustotou f. Najprv vypočítame distribučnú funkciu F_H náhodnej premennej H, a to nasledovne:

$$F_H(t) = P[h(X_1, X_2) < t] = P[(X_1, X_2)^T \in B_t] = \iint_{B_t} f(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2, \tag{7.6}$$

kde $B_t = \{(x_1, x_2)^T : h(x_1, x_2) < t\}$ je množina jednoznačne určená funkciou h. Množiny B_t pre $t = 1$ a pre najčastejšie sa vyskytujúce funkcie h sú znázornené na obrázku 7.3. Pochopiteľne, vypočítať integrál na pravej strane rovnosti (7.6) pre všeobecnú množinu B_t môže byť ťažké, dokonca aj v prípade nezávislosti X_1 a X_2, čiže ak má hustota f špeciálne jednoduchý tvar $f(x_1, x_2) = f_1(x_1) f_2(x_2)$,[15] kde f_1 a f_2 sú známe hustoty náhodných premenných X_1, resp. X_2. Ak sa nám podarí odvodiť vzťah pre $F_H(t)$ v každom t, tak môžeme výpočet hustoty náhodnej premennej $h(X_1, X_2)$ dokončiť derivovaním; presnejšie pozri vetu 5.9.

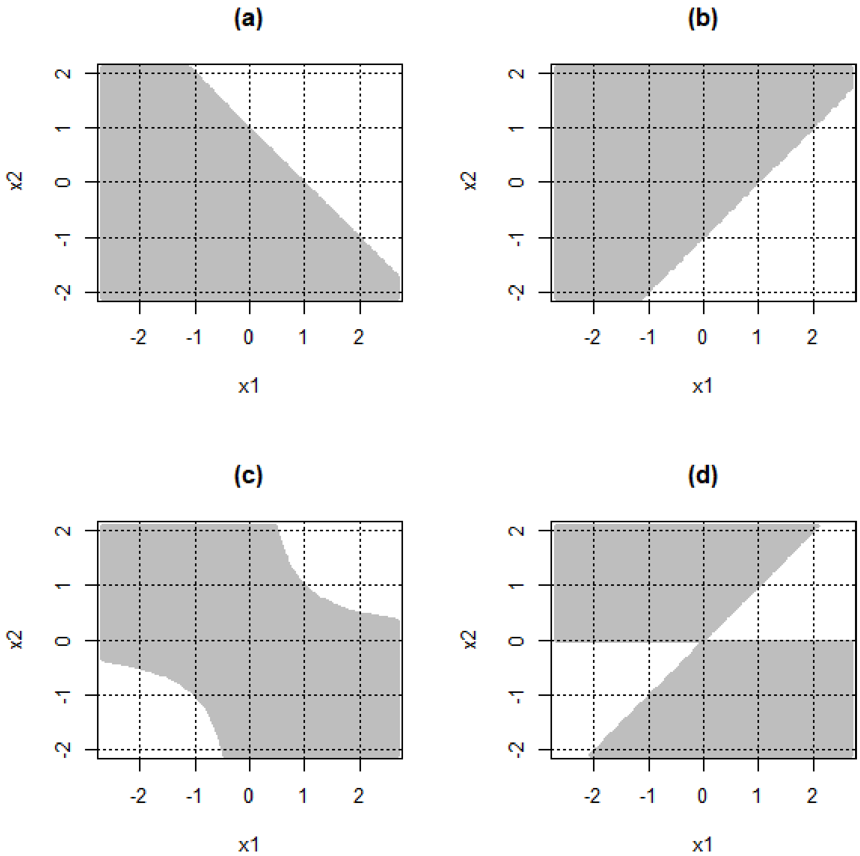
Príklad 7.45. Spustíme dva výpočty znáhodneného algoritmu; prvý bude trvať X_1 minút a druhý X_2 minút. Z teoretickej analýzy vieme, že X_1 a X_2 môžeme považovať za nezávislé náhodné premenné, obe s rozdelením $Exp(\lambda)$, kde $\lambda > 0$ je známy parameter. Vypočítajme hustotu náhodnej premennej $X_1 + X_2$. V prípade, že výpočty spúšťame za sebou, tak $X_1 + X_2$ reprezentuje dobu od spustenia prvého výpočtu po ukončenie druhého výpočtu.

Riešenie: Keďže X_1 a X_2 sú nezávislé, obe s rozdelením $Exp(\lambda)$, tak hustota náhodného vektora $(X_1, X_2)^T$ je súčinom hustôt f_1 a f_2, kde $f_1(x_1) = \lambda e^{-\lambda x_1}$ pre $x_1 > 0$ a $f_1(x_1) = 0$ pre $x_1 \leq 0$, resp. $f_2(x_2) = \lambda e^{-\lambda x_2}$ pre $x_2 > 0$ a $f_2(x_2) = 0$ pre $x_2 \leq 0$. Keďže nás zaujíma súčet $X_1 + X_2$, tak oblasť B_t je polrovina bodov $(x_1, x_2)^T$, ktoré spĺňajú $x_1 + x_2 < t$, ilustrovaná pre $t = 1$ na obrázku 7.3, panel (a). Distribučná funkcia náhodného vektora $(X_1, X_2)^T$ v bode $t > 0$ je preto

$$\begin{aligned} F_H(t) &= \iint_{B_t} \left(\lambda e^{-\lambda x_1}\right) \left(\lambda e^{-\lambda x_2}\right) \mathrm{d}x_2 \mathrm{d}x_1 \\ &= \int_0^t \lambda e^{-\lambda x_1} \left(\int_0^{x_1 - t} \lambda e^{-\lambda x_2} \mathrm{d}x_2 \right) \mathrm{d}x_1 = 1 - e^{-\lambda t} - \lambda t e^{-\lambda t}. \end{aligned}$$

[14] Toto rozdelenie má aj meno: ide o „negatívne binomické“ rozdelenie s parametrami 2 a 1/2.

[15] Pozrite vetu 7.36.



Obr. 7.3: Sivá oblasť znázorňuje množiny B_t pre $t = 1$. (a) $h(x_1, x_2) = x_1 + x_2$, (b) $h(x_1, x_2) = x_1 - x_2$, (c) $h(x_1, x_2) = x_1 x_2$, (d) $h(x_1, x_2) = x_1/x_2$ pre $x_2 \neq 0$ a $h(x_1, x_2) = 0$ inak.

Derivovaním F_H dostávame hustotu náhodnej premennej X_1+X_2 v tvare $f_H(t) = \lambda^2 t e^{-\lambda t}$ pre $t > 0$ a $f_H(t) = 0$ pre $t \leq 0$.[16]

Podobne by sme vedeli vypočítať napríklad hustotu náhodnej premennej $|X_1 - X_2|$; ak by sme oba výpočty spustili súčasne, tak táto náhodná premenná reprezentuje, koľko budeme musieť čakať od ukončenia kratšieho výpočtu po ukončenie dlhšieho výpočtu. Tiež je možné vypočítať napríklad hustotu náhodnej premennej $\sqrt{X_1X_2}$, čiže geometrického priemeru časov výpočtov alebo aj hustotu náhodnej premennej X_1/X_2, ktorá reprezentuje, koľkokrát dlhšie bude trvať prvý výpočet v porovnaní s druhým výpočtom.

7.3 Kovariancia a korelácia náhodných premenných

„Kovariancia“ a „korelácia“ sú jednoduché číselné miery závislosti dvoch náhodných premenných. V pravdepodobnosti a štatistike sa veľmi často využívajú, a to napriek tomu, že vo všeobecnosti nevedia plne charakterizovať závislosť náhodných premenných.

Takzvaná „kovariančná matica“ náhodného vektora a príbuzná „korelačná matica“ náhodného vektora sú kľúčové pre mnohorozmernú štatistiku a dátovú vedu. V tomto úvodnom texte si uvedieme len definície a základné vlastnosti; študenti sa však s kovariančnými a korelačnými maticami stretnú na pokročilejších prednáškach.

Definícia 7.46. Nech X_1, X_2 sú náhodné premenné s konečným rozptylom, obe diskrétne alebo obe spojité. Potom hovoríme, že X_1, X_2 majú konečnú kovarianciu a číslo

$$cov(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2).$$

nazývame **kovariancia** náhodných premenných X_1, X_2.

Poznámka 7.47. Ak X_1, X_2 sú premenné s konečným rozptylom, tak je možné dokázať, že náhodná premenná X_1X_2 má konečnú strednú hodnotu.

Veta 7.48 (Základné vlastnosti kovariancie)**.** Nech X_1, X_2 sú náhodné premenné s konečným rozptylom.[17] Potom 1) $cov(X_1, X_2) = cov(X_2, X_1)$; 2) $cov(aX_1+b, cX_2+d) = ac \cdot cov(X_1, X_2)$ pre každé $a, b, c, d \in \mathbb{R}$; 3) $cov(X_1 + X_3, X_2) = cov(X_1, X_2) + cov(X_3, X_2)$;[18] 4) $cov(X_i, X_i) = D(X_i)$ pre $i = 1, 2$; 5) $cov(X_1, X_2) = E\left((X_1 - E(X_1))(X_2 - E(X_2))\right)$; 6) $cov(X_1, X_2) = 0$ ak sú X_1, X_2 nezávislé; 7) $(cov(X_1, X_2))^2 \leq D(X_1)D(X_2)$.

Dôkaz. Prvé štyri rovnosti plynú priamo z definícií a z linearity strednej hodnoty. Tvrdenie 6) plynie priamo z definície kovariancie a vety 6.10. Dokázať tvrdenie 7) je zložitejšie; urobme dôkaz len pre prípad, že X_1 a X_2 sú obe diskrétne náhodné premenné. Pripomeňme si, že podľa Cauchyho-Schwarzovej nerovnosti platí pre akékoľvek $a_i, b_i \in \mathbb{R}$, $i = 1, \ldots, n$:[19]

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2.$$

[16] Hustota, ktorú sme získali, zodpovedá „Erlangovmu“ rozdeleniu s parametrami 2 a λ, ktoré sa využíva napríklad pri modelovaní „systémov hromadnej obsluhy“.

[17] Veta platí nezávisle na tom, akého typu sú náhodné premenné X_1, X_2.

[18] V tejto časti tvrdenia predpokladáme, že aj X_3 má konečný rozptyl. Dá sa dokázať, že ak X_1 a X_3 majú konečný rozptyl, tak aj $X_1 + X_3$ má konečný rozptyl.

[19] Cauchyho-Schwarzovu nerovnosť nebudeme dokazovať; predpokladáme, že študenti sa s ňou zoznámili v základnejších kurzoch matematiky.

Nech diskrétny náhodný vektor $\mathbf{X} = (X_1, X_2)^T$ má obor hodnôt$(\mathbf{x}_i)_{i \in I}$, pričom zložky vektora $\mathbf{x}_i$ označíme $(\mathbf{x}_i)_1$ a $(\mathbf{x}_i)_2$. Použijúc časť 3) znenia tejto vety, vetu 7.26 a Cauchyho-Schwarzovu nerovnosť s

$$\begin{aligned} a_i &= [(\mathbf{x}_i)_1 - E(X_1)] \sqrt{P[\mathbf{X} = \mathbf{x}_i]}, \\ b_i &= [(\mathbf{x}_i)_2 - E(X_2)] \sqrt{P[\mathbf{X} = \mathbf{x}_i]}, \end{aligned}$$

dostávame

$$(cov(X_1, X_2))^2 = \left(\sum_{i \in I} [(\mathbf{x}_i)_1 - E(X_1)] [(\mathbf{x}_i)_2 - E(X_2)] P[\mathbf{X} = \mathbf{x}_i] \right)^2 \leq$$
$$\sum_{i \in I} ((\mathbf{x}_i)_1 - E(X_1))^2 P[\mathbf{X} = \mathbf{x}_i] \sum_{i \in I} ((\mathbf{x}_i)_2 - E(X_2))^2 P[\mathbf{X} = \mathbf{x}_i] = D(X_1) D(X_2).$$

□

Poznámka 7.49. Rovnosti 1) až 3) vo vete 7.48 znamenajú, že kovariancia je symetrická a lineárna. Vlastnosti 2) a 3) implikujú, že ak sú $X_1, \ldots, X_m$ a $Y_1, \ldots, Y_k$ akékoľvek náhodné premenné s konečným rozptylom a $a_0, a_1, \ldots, a_m \in \mathbb{R}$, $b_0, b_1, \ldots, b_k \in \mathbb{R}$, tak

$$cov \left(\sum_{j=1}^{m} a_j X_j + a_0, \sum_{i=1}^{k} b_i Y_i + b_0 \right) = \sum_{j=1}^{m} \sum_{i=1}^{k} a_j b_i cov(X_j, Y_i). \tag{7.7}$$

Upozornime ešte špeciálne na to, že opačná implikácia v časti 6) neplatí; existujú náhodné premenné, ktoré sú závislé, no majú nulovú kovarianciu.

Definícia 7.50. Nech X_1 a X_2 sú náhodné premenné s konečným a nenulovým rozptylom. **Korelačný koeficient**[20] (alebo stručne korelácia) náhodných premenných X_1, X_2 je hodnota

$$\rho(X_1, X_2) = \frac{cov(X_1, X_2)}{\sqrt{D(X_1)} \sqrt{D(X_2)}}.$$

Ak $\rho(X_1, X_2) = 0$, potom hovoríme, že X_1, X_2 sú **nekorelované**. Ak $\rho(X_1, X_2) > 0$ hovoríme, že náhodné premenné X_1, X_2 sú **kladne korelované**. Ak $\rho(X_1, X_2) < 0$ hovoríme, že náhodné premenné X_1, X_2 sú **záporne korelované**.

Veta 7.51 (Základné vlastnosti korelačného koeficientu)**.** Nech X_1 a X_2 sú náhodné premenné s konečným a nenulovým rozptylom.[21] Potom 1) $\rho(X_1, X_2) = \rho(X_2, X_1)$; 2) $\rho(aX_1 + b, cX_2 + d) = \rho(X_1, X_2)$ pre každé $a > 0, c > 0, b, d \in \mathbb{R}$; 3) $\rho(X_1, X_2) = 0$, ak sú X_1, X_2 nezávislé; 4) $\rho(X_1, X_2) \in [-1, 1]$.

Dôkaz. Dôkaz všetkých častí tvrdenia plynie priamo z definície korelácie a základných vlastností kovariancie uvedených vo vete 7.48. □

Poznámka 7.52. Nech diskrétny náhodný vektor $\mathbf{X} = (X_1, X_2)^T$ má obor hodnôt $(\mathbf{x}_i)_{i \in I}$. Veta 7.26 nám umožňuje vypočítať $cov(X_1, X_2)$ a $\rho(X_1, X_2)$ pomocou vzťahov

[20]Uvedený korelačný koeficient sa niekedy presnejšie nazýva „Pearsonov“; existujú aj iné typy korelačných koeficientov.

[21]Veta platí nezávisle na tom, akého typu sú náhodné premenné X_1, X_2.

1. $E(X_1) = \sum_{i\in I}(\mathbf{x}_i)_1 P[\mathbf{X} = \mathbf{x}_i]$, $E(X_1^2) = \sum_{i\in I}(\mathbf{x}_i)_1^2 P[\mathbf{X} = \mathbf{x}_i]$,
2. $E(X_2) = \sum_{i\in I}(\mathbf{x}_i)_2 P[\mathbf{X} = \mathbf{x}_i]$, $E(X_2^2) = \sum_{i\in I}(\mathbf{x}_i)_2^2 P[\mathbf{X} = \mathbf{x}_i]$,
3. $E(X_1X_2) = \sum_{i\in I}(\mathbf{x}_i)_1(\mathbf{x}_i)_2 P[\mathbf{X} = \mathbf{x}_i]$,

kde $(\mathbf{x}_i)_1$ a $(\mathbf{x}_i)_2$ sú zložky vektora $\mathbf{x}_i$.

Príklad 7.53. Vypočítajme kovarianciu a koreláciu náhodných premenných X_1, X_2 z príkladu 7.4. V príkladoch 7.21 a 7.27 sme vypočítali, že $E(X_1) = \frac{91}{36}$, $E(X_2) = \frac{161}{36}$ a $E(X_1X_2) = \frac{441}{36}$. Kovariancia X_1 a X_2 je potom $cov(X_1, X_2) = \frac{441}{36} - \frac{91}{36}\frac{161}{36} = \frac{35^2}{36^2}$. Ďalej dopočítame rozptyly X_1, X_2: $D(X_i) = (\sum_{j=1}^{6} P[X_i = j]j^2) - E^2(X_i) = \frac{2555}{36^2}$ pre $i = 1, 2$, a teda $\rho(X_1, X_2) = \frac{35^2}{2555} \approx 0.48$.

Poznámka 7.54. Nech spojitý náhodný vektor $\mathbf{X} = (X_1, X_2)^T$ má hustotu f. Potom podľa vety 7.42 môžeme $cov(X_1, X_2)$ a $\rho(X_1, X_2)$ vypočítať pomocou vzťahov

1. $E(X_1) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2$, $E(X_1^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1^2 f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2$,
2. $E(X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2 f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2$, $E(X_2^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2^2 f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2$,
3. $E(X_1X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2$.

Príklad 7.55. Vypočítajme kovarianciu a koreláciu náhodných premenných X_1, X_2 z príkladu 7.32. Vieme už, že $E(X_i) = \frac{1}{a-1}$ (z príkladu 7.39) a $E(X_1X_2) = \frac{1}{(a-1)(a-2)}$ (z príkladu 7.43). Potom teda $cov(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2) = \frac{1}{(a-1)^2(a-2)}$. Na to, aby sme vypočítali korelačný koeficient, potrebujeme ešte rozptyly marginálnych rozdelení, ktoré odvodíme napríklad integrovaním uvedeným v poznámke 7.54: $D(X_i) = E(X_i^2) - (E(X_i))^2 = \frac{a}{(a-1)^2(a-2)}$. Dosadením do definície korelačného koeficientu dostávame $\rho(X_1, X_2) = \frac{1}{a}$.

Definícia 7.56. Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je náhodný vektor a nech každá náhodná premenná X_i, $i = 1, \ldots, m$, má konečný rozptyl. Potom povieme, že náhodný vektor $\mathbf{X}$ má konečnú **kovariančnú maticu** $Cov(\mathbf{X})$, pričom i, j-ty prvok tejto matice definujeme

$$(Cov(\mathbf{X}))_{i,j} = cov(X_i, X_j).$$

Veta 7.57 (Kovariančná matica lineárnej transformácie náhodného vektora)**.** Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je náhodný vektor s konečnou kovariančnou maticou $Cov(\mathbf{X})$, nech A je matica typu $k \times m$ a nech $b \in \mathbb{R}^k$. Potom $\mathbf{Y} = A\mathbf{X} + b$ je k-rozmerný náhodný vektor s konečnou kovariančnou maticou a platí

$$Cov(\mathbf{Y}) = ACov(\mathbf{X})A^T.$$

Dôkaz. Ak $\mathbf{Y} = (Y_1, \ldots, Y_k)^T = A\mathbf{X} + b$, potom pre každé $j = 1, \ldots, k$ platí

$$Y_j = a_{j1}X_1 + \cdots + a_{jm}X_m + b_j,$$

kde a_{jr} je prvok matice A v j-tom riadku a r-tom stĺpci a b_j je j-ta zložka vektora b. Z definície kovariancie, rovnosti (7.7) a definície kovariančnej matice dostávame pre každé $i, j \in \{1, \ldots, k\}$:

$$\begin{aligned}(Cov(\mathbf{Y}))_{i,j} &= cov(a_{i1}X_1 + \cdots + a_{im}X_m + b_i,\ a_{j1}X_1 + \cdots + a_{jm}X_m + b_j) \\ &= \sum_{r,s=1}^{m} a_{ir}a_{js}cov(X_r, X_s) = (ACov(\mathbf{X})A^T)_{i,j}.\end{aligned}$$

□

Veta 7.58 (Základné vlastnosti kovariančnej matice)**.** $\mathbf{X} = (X_1, \ldots, X_m)^T$ nech je náhodný vektor s konečnou kovariančnou maticou $Cov(\mathbf{X})$. Potom matica $Cov(\mathbf{X})$ je symetrická, pozitívne semidefinitná, s rozptylmi náhodných premenných $X_1, \ldots, X_m$ na diagonále.

Dôkaz. Symetrickosť $Cov(\mathbf{X})$ a vzťah $(Cov(\mathbf{X}))_{i,i} = D(X_i)$ plynú z definície kovariančnej matice a prvých dvoch vzťahov vo vete 7.48. Pozitívnu semidefinitnosť matice $Cov(\mathbf{X})$ dokážeme takto: Nech $a \in \mathbb{R}^m$ je ľubovoľný vektor. Z nezápornosti rozptylu náhodnej premennej $a^T\mathbf{X}$ a z vety o kovariančnej matici lineárnej transformácie náhodného vektora máme[22] $0 \leq D(a^T\mathbf{X}) = cov(a^T\mathbf{X}, a^T\mathbf{X}) = Cov(a^T\mathbf{X}) = a^T Cov(\mathbf{X})a$. Použili sme pri tom nezápornosť rozptylu (akejkoľvek náhodnej premennej), vlastnosť 4 z vety 7.48 a vetu 7.57. Tým sme pre maticu $Cov(\mathbf{X})$ dokázali definičnú vlastnosť pozitívnej semidefinitnosti. $\square$

Definícia 7.59. (Korelačná matica) Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je náhodný vektor a nech každá náhodná premenná $X_1, \ldots, X_m$ má konečný a nenulový rozptyl. Potom povieme, že náhodný vektor $\mathbf{X}$ má konečnú **korelačnú maticu** $R(\mathbf{X})$, pričom i, j-ty prvok tejto matice definujeme

$$(R(\mathbf{X}))_{i,j} = \rho(X_i, X_j).$$

Poznámka 7.60. Ako sa ľahko presvedčíme, korelačná matica náhodného vektora $\mathbf{X} = (X_1, \ldots, X_m)^T$ je kovariančnou maticou náhodného vektora

$$\mathbf{Y} = \left(\frac{X_1}{\sqrt{D(X_1)}}, \ldots, \frac{X_m}{\sqrt{D(X_m)}} \right)^T,$$

teda je to tiež symetrická a pozitívne semidefinitná matica. Uvedomme si, že korelácia akejkoľvek náhodnej premennej (ktorá má konečný a nenulový rozptyl) so sebou samou je rovná 1, takže korelačná matica má na diagonále jednotky.

Príklad 7.61. Náhodný vektor z príkladu 7.53 má kovariančnú a korelačnú maticu

$$\frac{1}{36^2}\begin{pmatrix} 2555 & 1225 \\ 1225 & 2555 \end{pmatrix}, \text{ resp. } \begin{pmatrix} 1 & \frac{1225}{2555} \\ \frac{1225}{2555} & 1 \end{pmatrix}.$$

Náhodný vektor z príkladu 7.55 má kovariančnú a korelačnú maticu

$$\frac{1}{(a-1)^2(a-2)}\begin{pmatrix} a & 1 \\ 1 & a \end{pmatrix}, \text{ resp. } \begin{pmatrix} 1 & 1/a \\ 1/a & 1 \end{pmatrix}.$$

7.4 Typy rozdelení náhodných vektorov

7.4.1 Multinomické rozdelenie

Definícia 7.62. Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je náhodný vektor. Nech $n \in \mathbb{N}$ a nech $p_1, \ldots, p_m$ sú kladné reálne čísla také, že $\sum_{j=1}^m p_j = 1$. Nech platí

$$P[X_1 = k_1, \ldots, X_m = k_m] = \frac{n!}{k_1! \ldots k_m!} p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m}$$

[22] V tomto špeciálnom prípade je $a^T\mathbf{X}$ náhodná premenná a súčasne jednorozmerný náhodný vektor, takže $cov(a^T\mathbf{X}, a^T\mathbf{X})$ je jediný prvok 1×1 matice $Cov(a^T\mathbf{X})$.

pre všetky tie $k_1, \dots, k_m \in \{0, 1, \dots, n\}$, ktoré spĺňajú $\sum_{j=1}^m k_j = n$. Potom hovoríme, že náhodný vektor $\mathbf{X}$ má **multinomické rozdelenie** s parametrami $n, p_1, \dots, p_m$. Túto skutočnosť značíme $\mathbf{X} \sim Mult(n, p_1, \dots, p_m)$.

Poznámka 7.63. Predpokladajme, že budeme realizovať n nezávislých pokusov, pričom výsledkom každého z nich bude práve jeden z m typov výsledku. V každom pokuse sa realizuje j-ty typ výsledku s pravdepodobnosťou p_j. Nech X_j znamená počet tých pokusov, ktoré skončia j-tym typom výsledku. Potom $\mathbf{X} = (X_1, \dots, X_m)^T \sim Mult(n, p_1, \dots, p_m)$. Náhodný vektor s multinomickým rozdelením je zjavne diskrétny. Multinomické rozdelenie sme už nepriamo používali v súvislosti s multinomickou formulou 2.26.

Príklad 7.64. Nech $Y_1, \dots, Y_n$ sú nezávislé náhodné premenné s distribučnou funkciou F. Nech $m > 2$ a $a_1 < \dots < a_{m-1}$ nech sú reálne čísla. Nech $X_1, \dots, X_m$ znamenajú to, koľko z hodnôt $Y_1, \dots, Y_n$ padne do intervalov $(-\infty, a_1]$, $(a_1, a_2]$, $\dots$, (a_{m-1}, ∞). Potom $\mathbf{X} = (X_1, \dots, X_m)^T \sim Mult(n, p_1, \dots, p_m)$, kde $p_1 = F(a_1)$, $p_2 = F(a_2) - F(a_1)$, $\dots$, $p_m = 1 - F(a_{m-1})$. Uvedomme si, že náhodné premenné $X_1, \dots, X_m$ reprezentujú „výšku dielikov“ v „histograme“, ktorý skonštruujeme z dát $Y_1, \dots, Y_n$.

Veta 7.65 (Vzťah multinomického a binomického rozdelenia)**.** Nech $\mathbf{X} = (X_1, \dots, X_m)^T \sim Mult(n, p_1, \dots, p_m)$. Potom $X_j \sim Bin(n, p_j)$ pre každé $j = 1, \dots, m$.

Dôkaz. Nech $\mathbf{X} = (X_1, \dots, X_m)^T \sim Mult(n, p_1, \dots, p_m)$. Dokážeme, že $X_1 \sim Bin(n, p_1)$. Pre každé $k_1 \in \{0, 1, \dots, n\}$ platí

$$P[X_1 = k_1] = \sum P[X_1 = k_1, \dots, X_m = k_m] = \sum \frac{n!}{k_1! \dots k_m!} p_1^{k_1} p_2^{k_2} \dots p_m^{k_m},$$

pričom všade v tomto dôkaze sumujeme cez tie $k_2, \dots, k_m \in \{0, 1, \dots, n\}$, ktoré spĺňajú $\sum_{j=2}^m k_j = n - k_1$. Z multinomického rozvoja

$$(1 - p_1)^{n-k_1} = (p_2 + \dots + p_m)^{n-k_1} = \sum \frac{(n-k_1)!}{k_2! \dots k_m!} p_2^{k_2} \dots p_m^{k_m}$$

dostávame

$$P[X_1 = k_1] = \frac{n! p_1^{k_1}}{k_1!(n-k_1)!} \sum \frac{(n-k_1)!}{k_2! \dots k_m!} p_2^{k_2} \dots p_m^{k_m} = \binom{n}{k_1} p_1^{k_1} (1 - p_1)^{n-k_1}.$$

Tvrdenie $X_j \sim Bin(n, p_j)$ môžeme dokázať analogicky pre každé $j = 1, \dots, m$. □

Poznámka 7.66. Predchádzajúca veta znamená, že multinomické rozdelenie môžeme považovať za „viacrozmerné binomické rozdelenie“, hoci treba mať na pamäti, že zložky multinomického rozdelenia sú vždy *závislé* binomické premenné, pozri vetu 7.68.

Poznámka 7.67. Uvažujme interpretáciu z poznámky 7.63. Všimnime si, že ak označíme ako V_{ij} náhodnú premennú, ktorá nadobúda 1, ak skončil i-ty pokus výsledkom j a 0 inak, potom pre každý výber indexov $j_1, j_2, \dots, j_n \in \{1, \dots, m\}$ sú náhodné premenné $V_{1j_1}, \dots, V_{nj_n}$ nezávislé. Navyše, pre každé $j = 1, \dots, m$ platí $X_j = V_{1j} + \dots + V_{nj}$. Toto pozorovanie nám umožní jednoducho ukázať nasledovnú vetu.

Veta 7.68 (Stredná hodnota a kovariančná matica multinomického rozdelenia)**.** Nech $\mathbf{X} \sim Mult(n, p_1, \ldots, p_m)$. Potom pre strednú hodnotu vektora $\mathbf{X}$ platí $E(\mathbf{X}) = (np_1, \ldots, np_m)^T$ a pre kovariančnú maticu $Cov(\mathbf{X})$ platí

$$(Cov(\mathbf{X}))_{j,j} = np_j(1-p_j) \text{ pre } j = 1, \ldots, m,$$

$$(Cov(\mathbf{X}))_{j,k} = -np_jp_k \text{ pre } j, k \in \{1, \ldots, m\}, j \neq k.$$

Dôkaz. Vlastnosť $E(X_j) = np_j$ a $(Cov(\mathbf{X}))_{j,j} = D(X_j) = -np_j(1-p_j)$ pre $j = 1, \ldots, m$ plynie z viet 7.65 a 4.43. Dokážeme, že $(Cov(\mathbf{X}))_{j,k} = -np_jp_k$ pre pevné $j \neq k$. Nech $X_j = V_{1j} + \cdots + V_{nj}$ pre každé $j = 1, \ldots, m$ ako v predchádzajúcej poznámke. Potom

$$(Cov(\mathbf{X}))_{j,k} = cov(X_j, X_k) = cov(V_{1j} + \cdots + V_{nj}, V_{1k} + \cdots + V_{nk}) =$$

$$\sum_{i,l=1}^{n} cov(V_{ij}, V_{lk}) = \sum_{i \neq l} cov(V_{ij}, V_{lk}) + \sum_{i=1}^{n} cov(V_{ij}, V_{ik}) = \sum_{i=1}^{n} cov(V_{ij}, V_{ik}),$$

kde posledná rovnosť plynie z toho, že V_{ij}, V_{lk} sú nezávislé, ak $i \neq l$, teda $\sum_{i \neq l} cov(V_{ij}, V_{lk}) = 0$. Avšak $j \neq k$, preto $V_{ij}V_{ik} = 0$ pre každé i, takže

$$cov(V_{ij}, V_{ik}) = E(V_{ij}V_{ik}) - E(V_{ij})E(V_{ik}) = -p_jp_k.$$

□

Príklad 7.69. Lotérie sa účastní $n = 100$ ľudí. Každý človek získa práve jednu cenu; cenu v prvom poradí získa človek s pravdepodobnosťou 1/10, cenu v druhom poradí s pravdepodobnosťou 2/10 a v treťom poradí s pravdepodobnosťou 7/10. Nech X_i je náhodná premenná, ktorá znamená počet výhier v i-tom poradí, $i = 1, 2, 3$. Nájdeme všetky vzájomné korelácie náhodných premenných X_1, X_2, X_3.

Riešenie: Náhodný vektor $\mathbf{X} = (X_1, X_2, X_3)^T$ má rozdelenie $Mult(100, \frac{1}{10}, \frac{2}{10}, \frac{7}{10})$. Podľa vety 7.68 máme $D(X_1) = 9$, $D(X_2) = 16$, $D(X_3) = 21$ a $Cov(X_1, X_2) = -2$, $Cov(X_1, X_3) = -7$, $Cov(X_2, X_3) = -14$. Odtiaľ dostaneme korelačné koeficienty $\rho(X_1, X_2) = -\frac{1}{6}$, $\rho(X_1, X_3) = -\frac{7}{3\sqrt{21}}$, $\rho(X_2, X_3) = -\frac{7}{2\sqrt{21}}$.[23]

7.4.2 Viacrozmerné normálne rozdelenie

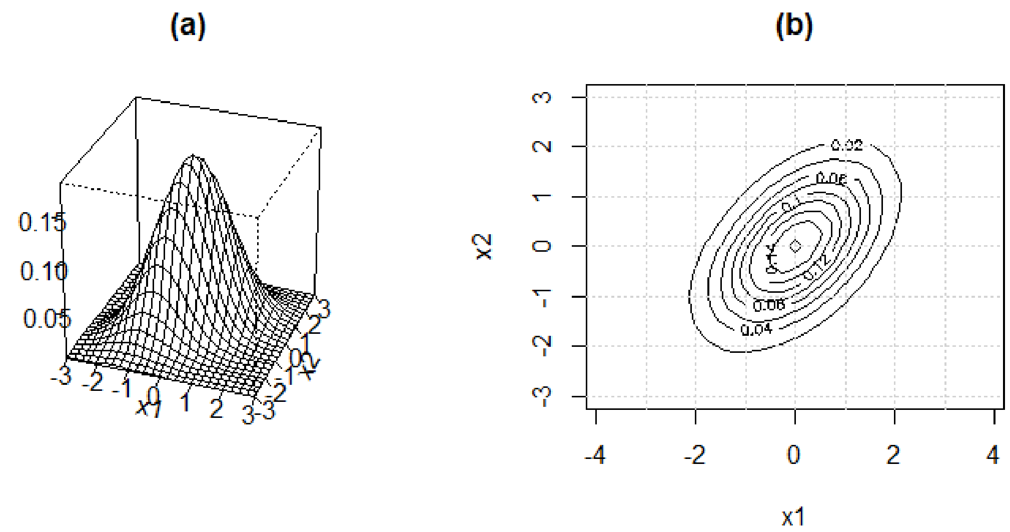
Viacrozmerné normálne rozdelenie[24] má ústredné postavenie v pravdepodobnostnom modelovaní a vo viacrozmerných štatistických metódach, ako napríklad „regresná analýza“, „analýza rozptylu“, „diskriminačná analýza“ a „analýza zhlukov“. Dôvodom sú jeho veľmi výhodné analytické vlastnosti a aj skutočnosť, že vektorové pozorovania sa často správajú približne tak, ako keby pochádzali z viacrozmerného normálneho rozdelenia.

Definícia 7.70. Nech $\mu \in \mathbb{R}^m$ a nech Σ je pozitívne definitná matica typu $m \times m$. Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je spojitý náhodný vektor s hustotou

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{m/2}\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)\right).$$

[23] Uvedomme si, že korelácie zložiek multinomického rozdelenia sú vždy záporné; pokúste sa intuitívne zdôvodniť prečo.

[24] Niekedy sa môžeme stretnúť s pojmom „mnohorozmerné“ normálne/gaussovské rozdelenie alebo „združené“ normálne/gaussovské rozdelenie.



Obr. 7.4: Dvojrozmerné normálne rozdelenie so strednou hodnotou $\mu = (0,0)^T$ a kovariančnou maticou s jednotkami na diagonále, a hodnotami 0.5 mimo diagonály. (a) trojrozmerný pohľad na hustotu; (b) úrovňové množiny („vrstevnice“) hustoty.

Potom hovoríme, že $\mathbf{X}$ má regulárne[25] **m-rozmerné normálne rozdelenie** s parametrami μ a Σ. Túto skutočnosť značíme $\mathbf{X} \sim N_m(\mu, \Sigma)$.

Veta 7.71 (Lineárna transformácia náhodného vektora s regulárnym m-rozmerným normálnym rozdelením)**.** Nech $\mathbf{X} = (X_1, \ldots, X_m)^T \sim N_m(\mu, \Sigma)$, kde $\mu \in \mathbb{R}^m$ a Σ nech je pozitívne definitná matica typu $m \times m$. Nech $b \in \mathbb{R}^k$, kde $k \leq m$. Nech A je matica typu $k \times m$ a hodnosti k. Potom náhodný vektor $A\mathbf{X} + b$ má k-rozmerné regulárne normálne rozdelenie $N_k(A\mu + b, A\Sigma A^T)$. Špeciálne, ak $k = 1$, $b = 0$ a $A = (a_1, \ldots, a_m)$, kde $a_i \neq 0$ pre aspoň jeden index i, tak dostávame: $\sum_{i=1}^m a_i X_i \sim N(\sum_{i=1}^m \mu_i, A\Sigma A^T)$.[26] Ak v predošlom volíme $a_i = 1$ pre nejaký index i a $a_j = 0$ pre všetky $j \neq i$, tak dostávame: $X_i \sim N(\mu_i, D(X_i))$.

Dôkaz. Dôkaz sme sa rozhodli do týchto skrípt nezaradiť, pretože využíva relatívne zložitý technický aparát mnohorozmernej matematickej analýzy, prípadne komplexnú analýzu (charakteristické funkcie). □

Poznámka 7.72. Rozpíšme ešte slovne obsah (veľmi silnej a dôležitej) vety 7.71: Lineárna transformácia regulárneho viacrozmerného normálneho rozdelenia je opäť regulárne viacrozmerné normálne rozdelenie.[27] Nenulová lineárna kombinácia zložiek regulárneho viacrozmerného normálneho rozdelenia je náhodná premenná s (jednorozmerným) normálnym rozdelením. Špeciálne, každá zložka regulárneho viacrozmerného normálneho náhodného vektora má (jednorozmerné) normálne rozdelenie.

[25] Okrem pojmu „regulárne“ sa používa aj pojem „nedegenerované“. Existuje aj definícia širšej triedy viacrozmerných normálnych rozdelení, zahŕňajúca aj takzvané „degenerované“ normálne rozdelenia, ktoré však nie sú spojité, čiže nemajú v našom zmysle hustotu. Degenerovanými normálnymi rozdeleniami sa v týchto skriptách nezaoberáme.

[26] Všimnite si, že v tomto prípade je $A\Sigma A^T$ kladné reálne číslo.

[27] Pokiaľ transformačná matica spĺňa určitú podmienku nedegenerovanosti.

Veta 7.73 (Základné číselné charakteristiky regulárneho m-rozmerného normálneho náhodného vektora)**.** Nech $\mathbf{X} = (X_1, \ldots, X_m)^T \sim N_m(\mu, \Sigma)$, kde $\mu \in \mathbb{R}^m$ a Σ je pozitívne definitná matica typu $m \times m$. Potom $E(\mathbf{X}) = \mu$ a $Cov(\mathbf{X}) = \Sigma$.

Dôkaz. Najprv predpokladajme, že Σ je jednotková matica a μ je nulový vektor. Z definície normálneho náhodného vektora a vety 7.71 vidíme, že pre hustotu f vektora $\mathbf{X}$ a hustoty f_i zložiek X_i náhodného vektora $\mathbf{X}$ platí:

$$f(x_1, \ldots, x_m) = \frac{1}{(2\pi)^{m/2}} \exp\left(-\frac{x_1^2 + \ldots + x_m^2}{2}\right) = \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_i^2}{2}\right) = \prod_{i=1}^{m} f_i(x_i).$$

Na základe vety 7.36 usudzujeme, že $X_1, \ldots, X_m$ sú nezávislé, a preto $cov(X_i, X_j) = 0$ pre všetky $i \neq j$. Keďže $X_i \sim N(0,1)$, tak $E(X_i) = 0$ a $cov(X_i, X_i) = D(X_i) = 1$. Preto $E(\mathbf{X}) = 0$ a $Cov(\mathbf{X}) = I_m$.

Teraz predpokladajme, že $\mu \in \mathbb{R}^m$ a Σ je akákoľvek pozitívne definitná matica typu $m \times m$. Nech $\Sigma^{-1/2}$ je odmocnina z matice Σ^{-1}, t.j. taká pozitívne definitná matica, že $\Sigma^{-1/2}\Sigma^{-1/2} = \Sigma^{-1}$. (Existencia takej matice sa štandardne ukazuje v teórii matíc.) Podľa vety 7.71 má náhodný vektor $\Sigma^{-1/2}(\mathbf{X} - \mu)$ rozdelenie $N_m((0, \ldots, 0)^T, I_m)$ a podľa prvej časti dôkazu, podľa vety o linearite strednej hodnoty a vety o kovariančnej matici lineárnej transformácie platí:

$$0 = E(\Sigma^{-1/2}(\mathbf{X} - \mu)) = \Sigma^{-1/2}(E(\mathbf{X}) - \mu),$$
$$I_m = Cov(\Sigma^{-1/2}(\mathbf{X} - \mu)) = \Sigma^{-1/2} Cov(\mathbf{X}) \Sigma^{-1/2},$$

z ktorých dostávame tvrdenie vety. □

Veta 7.74 (Združene normálny náhodný vektor s nezávislými zložkami)**.** Platí ekvivalencia nasledovných dvoch výrokov: 1) $X_1, \ldots, X_m$ sú nezávislé náhodné premenné, pričom $X_i \sim N(\mu_i, \sigma_i^2)$; 2) Náhodný vektor $\mathbf{X} = (X_1, \ldots, X_m)^T$ má rozdelenie $N_m(\mu, \Sigma)$, kde $\mu = (\mu_1, \ldots, \mu_m)^T$ a Σ je diagonálna matica s prvkami $\sigma_1^2, \ldots, \sigma_m^2$ na diagonále.

Dôkaz. Veta sa dá jednoducho ukázať pomocou viet 7.71 a 7.36. □

Príklad 7.75 (Dvojrozmerné normálne rozdelenie)**.** Vezmime v definícii 7.70 $m = 2$, uvažujeme teda regulárne dvojrozmerné normálne rozdelenie náhodného vektora $\mathbf{X} = (X_1, X_2)^T$ so strednou hodnotou $\mu = (\mu_1, \mu_2)^T$ a kovariančnou maticou

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix},$$

kde $\sigma_1^2 = D(X_1)$, $\sigma_2^2 = D(X_2)$ a $\sigma_{12} = \sigma_{21} = cov(X_1, X_2)$. Označme ďalej $\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}$ korelačný koeficient náhodných premenných X_1 a X_2. Hustota $\mathbf{X} \sim N_2(\mu, \Sigma)$ sa dá potom napísať nasledovne

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}}{2(1-\rho^2)}\right\}.$$

Dá sa jednoducho ukázať, že úrovňové množiny $\{(x_1, x_2)^T : f(x_1, x_2) \leq r\}$ sú elipsy.[28] Tvar hustoty f a jej úrovňové množiny sú ilustrované na obrázku 7.4. Ak $\Sigma = I_2$, potom sú úrovňové

[28] Pre $r \in (0, y)$, kde $y = f(\mu_1, \mu_2)$.

množiny hustoty f kružnice. Táto situácia zodpovedá štandardizovanému dvojrozmernému rozdeleniu, v ktorom majú obe zložky štandardizované jednorozmerné rozdelenie a sú nezávislé (a teda aj nekorelované a $\rho = 0$).

Príklad 7.76. Dve agentúry na prieskum verejnej mienky vykonávajú prieskum volebných preferencií istého kandidáta na prezidenta, každá na náhodnej vzorke n voličov. Predpokladajme, že podiel všetkých voličov, ktorí preferujú tohto kandidáta, je p. Pomocou aproximácie binomického rozdelenia normálnym odhadnite hornú hranicu pravdepodobnosti, že odhad preferencií u oboch agentúr sa bude odlišovať o menej ako 5 %.

Riešenie: Nech X_1 je počet hlasov za kandidáta, ktoré nameria prvá agentúra a nech X_2 je počet hlasov za kandidáta, ktoré nameria druhá agentúra. Zaujíma nás $P[|X_1/n - X_2/n| > 0.05]$, čiže $P[-0.05n < X_1 - X_2 < 0.05n]$. Nájdime preto približné rozdelenie náhodnej premennej $X_1 - X_2$. Predpokladajme, že obe agentúry realizujú výber respondentov dokonale náhodne na množine všetkých potenciálnych voličov, a to navzájom nezávisle. Potom náhodné premenné X_1 a X_2 majú rozdelenie $Hyp(M, n, pM)$, kde M je počet všetkých potenciálnych voličov. Ak je počet voličov M mnohonásobne väčší ako počet respondentov n, tak sú X_1 a X_2 nezávislé náhodné premenné s približne binomickým rozdelením $Bin(n, p)$. Ak je navyše n dostatočne veľké a p nie príliš extremálne (povedzme $n = 1000$ a $0.1 < p < 0.9$), tak môžeme použiť De Moivrovu-Laplaceovu vetu 5.53, ktorá implikuje, že X_1 a X_2 budú náhodné premenné s rozdelením približne $N(np, np(1-p))$. Podľa vety 7.74 platí, že náhodný vektor $\mathbf{X}$ má združene normálne rozdelenie so strednou hodnotou $\mu = (np, np)^T$ a diagonálnou kovariančnou maticou Σ s diagonálnymi prvkami rovnými hodnote $np(1-p)$. Následne, keďže $X_1 - X_2 = (1, -1)\mathbf{X}$, tak $X_1 - X_2$ má podľa viet 7.71 a 7.73 jednorozmerné normálne rozdelenie so strednou hodnotou $(1, -1)\mu = 0$ a rozptylom $(1, -1)\Sigma(1, -1)^T = 2np(1-p)$. Pre hľadanú pravdepodobnosť teda dostávame

$$P[-0.05n < X_1 - X_2 < 0.05n] = P\left[\frac{-0.05n}{\sqrt{2np(1-p)}} < \frac{X_1 - X_2}{\sqrt{2np(1-p)}} < \frac{0.05n}{\sqrt{2np(1-p)}}\right] = \\ = 2\Phi\left(\frac{0.05\sqrt{n}}{\sqrt{2p(1-p)}}\right) - 1.$$

Napríklad pre $n = 1000$ voličov a $p = 0.5$ je táto pravdepodobnosť približne 0.975. Stochastické simulácie, v ktorých nie je potrebné robiť vyššie uvedené zjednodušenia, sú v dobrej zhode s týmto číselným odhadom. Pre p menšie alebo väčšie ako 1/2, prípadne pre $n > 1000$, vyjde odhadovaná pravdepodobnosť ešte vyššia.

Poznámka 7.77. Existuje aj množstvo iných typov rozdelení pravdepodobnosti diskrétnych či spojitých náhodných vektorov. Spomeňme napríklad „negatívne multinomické“ alebo „Dirichletovo rozdelenie“.

Literatúra

[1] N Alon, JH Spencer: The Probabilistic Method, 4th Edition, Wiley 2016

[2] DP Bertsekas, JN Tsitsiklis: Introduction to Probability, 2nd Edition, Athena Scientific 2008

[3] JK Blitzstein, J Hwang: Introduction to Probability, 2nd Edition, Chapman and Hall/CRC 2019

[4] TM Cover, JA Thomas: Elements of Information Theory, Wiley-Interscience 2006

[5] GR Grimmett, DR Stirzaker: Probability and Random Processes, 3rd Edition, Oxford University Press 2001

[6] M Mitzenmacher, E Upfal: Probability and Computing, Cambridge University Press 2005

[7] SM Ross: Simulation, Academic Press 2012

[8] AN Shiryaev: Probability, 2nd Edition, Graduate Texts in Mathematics, Springer 1996