Pravdepodobnosť a štatistika
Poznámky k prednáškam

Radoslav Harman, Lenka Filová,
KAMŠ, FMFI UK

6. októbra 2017

Obsah

1 Úvod **5**

2 Axiomatická definícia pravdepodobnosti **6**
2.1 Priestor udalostí . . . 6
2.2 Pravdepodobnostná miera . . . 8
2.3 Základné vlastnosti pravdepodobnosti . . . 9
2.4 Cvičenia . . . 11

3 Podmieňovanie a nezávislosť udalostí **13**
3.1 Podmienená pravdepodobnosť . . . 13
3.2 Nezávislosť udalostí . . . 16
3.3 Cvičenia . . . 18

4 Všeobecné náhodné premenné **20**
4.1 Základné vlastnosti náhodných premenných . . . 20
4.2 Distribučná funkcia náhodnej premennej . . . 22
4.3 Cvičenia . . . 23

5 Diskrétne náhodné premenné **24**
5.1 Základné vlastnosti diskrétnych náhodných premenných . . . 24
5.2 Číselné charakteristiky diskrétnych náhodných premenných . . . 25
5.3 Základné typy diskrétnych náhodných premenných . . . 27
5.3.1 Alternatívne rozdelenie . . . 27
5.3.2 Binomické rozdelenie . . . 27
5.3.3 Poissonovo rozdelenie . . . 29
5.3.4 Geometrické rozdelenie . . . 31
5.3.5 Hypergeometrické rozdelenie . . . 33
5.4 Využitie indikátorov udalostí a linearity strednej hodnoty . . . 33
5.5 Cvičenia . . . 35

6 Markovovské reťazce **37**
6.1 Klasifikácia stavov markovovského reťazca . . . 38
6.2 Stacionárne rozdelenia . . . 39
6.3 Cvičenia . . . 41

7 Spojité náhodné premenné **43**
7.1 Základné vlastnosti spojitých náhodných premenných 43
7.2 Číselné charakteristiky spojitých náhodných premenných 44
7.3 Základné typy spojitých náhodných premenných 45
7.3.1 Rovnomerné rozdelenie 45
7.3.2 Exponenciálne rozdelenie 46
7.3.3 Paretovo rozdelenie 47
7.3.4 Normálne rozdelenie 48
7.4 Cvičenia 50

8 Náhodné vektory **53**
8.1 Všeobecné náhodné vektory 53
8.1.1 Diskrétne náhodné vektory 54
8.1.2 Spojité náhodné vektory 55
8.2 Nezávislosť náhodných premenných 56
8.3 Základné typy rozdelení náhodných vektorov 61
8.3.1 Multinomické rozdelenie 61
8.3.2 Mnohorozmerné normálne rozdelenie 62
8.4 Cvičenia 64

9 Zákony veľkých čísiel a centrálna limitná veta **66**
9.1 Cvičenia 68

10 Generovanie náhodných premenných a vektorov **70**
10.1 Generovanie realizácií náhodných premenných 70
10.2 Generovanie realizácií náhodných vektorov 73

11 Základy teórie informácie **76**
11.1 Informácia 76
11.2 Entropia 76
11.3 Združená a podmienená entropia 77
11.4 Relatívna entropia a vzájomná informácia 78
11.4.1 Princíp maximálnej entropie 80
11.5 Entropy rate 80
11.6 Cvičenia 81

12 Lineárny regresný model **82**
12.1 Rozdelenia pravdepodobnosti odvodené od normálneho rozdelenia 82
12.2 Náhodný výber a výberové charakteristiky 83
12.3 Lineárny regresný model 84
12.3.1 Lineárny regresný model priamkou 84
12.3.2 Základná veta o lineárnom regresnom modeli 85

13 Bodové a intervalové odhady parametrov štatistických modelov **87**
13.1 Bodové odhady 87
13.1.1 Metóda maximálnej vierohodnosti 87
13.2 Intervalové odhady 88
13.2.1 Odhady parametrov náhodného výberu z $N(\mu, \sigma^2)$ 88

13.3 Odhady parametrov lineárneho regresného modelu priamkou 89
13.4 Cvičenia 90

14 Testovanie štatistických hypotéz 91
14.1 Všeobecný úvod k testovaniu štatistických hypotéz 91
14.2 Testovanie hypotéz o strednej hodnote a disperzii 92
14.3 Testovanie hypotézy o rozdiele stredných hodnôt 93
14.4 Testovanie hypotézy o sklone regresnej priamky 93
14.5 Cvičenia 94

Tento text vzniká pre potreby prednášky ”Pravdepodobnosť a štatistika” (1-INF-435) na Fakulte matematiky, fyziky a informatiky UK. Práca neprešla odbornou recenziou, ani jazykovou kontrolou a môže obsahovať chyby. Vopred ďakujem všetkým študentom, ktorí ma upozornia na nepresnosti, prípadne navrhnú, ako text vylepšiť po obsahovej, alebo formálnej stránke.

Upozornenie: Najdôležitejšie informácie k predmetu nájdete na adrese http://www.iam.fmph.uniba.sk/ospm/Harman/teaching.htm

doc. Mgr. Radoslav Harman, PhD.; harman@fmph.uniba.sk
Katedra aplikovanej matematiky a štatistiky, FMFI UK

Kapitola 1

Úvod

Teória pravdepodobnosti poskytuje základné modely a matematický aparát pre mnohé oblasti teoretickej aj aplikovanej informatiky. Napríklad axiomatizácia samotnej teórie informácie a kódovania je postavená na pravdepodobnosti a na náhodných premenných (pozri [2]). Dôležitou súčasťou teórie zložitosti algoritmov je štúdium správania sa algoritmov pre náhodne generované vstupy. Taktiež, veľa efektívnych algoritmov na riešenie deterministických úloh využíva principiálnym spôsobom náhodnosť, povedzme Millerov-Rabinov test prvočíselnosti, algoritmus na hľadanie najmenšieho rezu v grafe a iné "znáhodnené" algoritmy ([4]). Zaujímavou metódou použitia teórie pravdepodobnosti v kombinatorike je takzvaná "pravdepodobnostná metóda", pomocou ktorej vieme konštruovať prehľadné dôkazy niektorých existenčných tvrdení ([1]). V neposlednom rade, počítačová simulácia a optimalizácia skoro všetkých reálnych systémov a všetky pokročilejšie počítačové hry si vyžadujú použitie vhodných generátorov náhodnosti ([5] a mnoho iných učebníc).

Kapitola 2

Axiomatická definícia pravdepodobnosti

2.1 Priestor udalostí

σ-algebra je jedným zo základných nástrojov teórie pravdepodobnosti. Neskôr na nej budeme definovať pravepodobnostnú mieru.

Definícia 2.1 (σ-algebra)**.** Nech Ω je neprázdna množina a nech $\mathcal{S} \subseteq 2^{\Omega}$. Usporiadanú dvojicu $(\Omega, \mathcal{S})$ nazývame σ-algebra, alebo σ-algebra *udalostí*, ak platí

1. $\Omega \in \mathcal{S}$
2. $A \in \mathcal{S} \Rightarrow \Omega \setminus A \in \mathcal{S}$
3. Ak $(A_i)_{i \in I}$ je postupnosť množín patriacich do systému $\mathcal{S}$, tak $\cup_{i \in I} A_i \in \mathcal{S}$

Poznámka 2.1. Ak použijeme v texte pojem "postupnosť", myslíme tým konečnú, alebo aj nekonečnú spočítateľnú postupnosť. To znamená, že v prechádzajúcej definícii je I konečná, alebo nekonečná spočítateľná *usporiadaná* množina.

Príklad 2.1. Nech Ω je ľubovoľná neprázdna množina a nech $\emptyset \neq A \subsetneq \Omega$. Potom $(\Omega, \{\emptyset, \Omega\})$, $(\Omega, \{\emptyset, A, \Omega/A, \Omega\})$ aj $(\Omega, 2^{\Omega})$ sú σ-algebry. (Symbolom 2^{Ω} značíme množinu všetkých podmnožín množiny Ω.) Ak $\Omega = \{1, 2, 3\}$, potom $(\Omega, \{\emptyset, \{1\}, \{2\}, \{3\}, \Omega\})$ *nie je* σ-algebra.

Veta 2.1 (*Uzavretosť σ-algebry vzhľadom k spočítateľným prienikom*)**.** Ak $(\Omega, \mathcal{S})$ je σ-algebra a $(A_i)_{i \in I}$ je postupnosť prvkov systému $\mathcal{S}$, tak $\cap_{i \in I} A_i \in \mathcal{S}$.

Dôkaz. Ak $A_i \in \mathcal{S}$ pre všetky $i \in I$, tak podľa vlastnosti 2 z definície 2.1 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 Morganovych pravidiel. $\square$

Axiómy definície 2.1 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 komplementov. Ukážme napríklad, že systém $\mathcal{S}$ je uzavretý vzhľadom na množinový rozdiel. Ak $A, B \in \mathcal{S}$, tak podľa vlastnosti 2 definície 2.1 platí $\Omega \setminus A \in \mathcal{S}$. Teda podľa predchádzajúcej vety obsahuje systém $\mathcal{S}$ aj množinu $B \cap (\Omega \setminus A) = B \setminus A$.

V teórii pravdepodobnosti predstavujú prvky množiny Ω najjednoduchšie, ďalej nerozložiteľné výsledky náhodného experimentu (takzvané elementárne výsledky). Tie množiny elementárnych výsledkov, ktoré patria do systému $\mathcal{S}$, reprezentujú "udalosti", ktorým je možné pripísať pravdepodobnosť realizácie. Ak je množina Ω spočítateľná, tak v našom modeli obvykle volíme $\mathcal{S} = 2^\Omega$. Ukazuje sa však, že ak je množina Ω nespočítateľná a zvolili by sme $\mathcal{S} = 2^\Omega$, potom by mohlo byť obtiažne priradiť prvkom takto bohatého systému udalostí pravdepodobnostnú mieru, ktorá by sa nesprávala paradoxne. Napríklad ak je $\Omega = \mathbb{R}^m$, býva vhodným (dostatočne, ale nie *príliš* bohatým) systémom udalostí systém $\mathcal{B}_m$ borelovských množín, definovaný nižšie (pozri definíciu 2.4).

Definícia 2.2 (*Elementárne výsledky a udalosti*)**.** Nech $(\Omega, \mathcal{S})$ je σ-algebra. Prvky množiny Ω nazveme elementárne výsledky a podmnožiny Ω patriace do systému $\mathcal{S}$ nazveme udalosti.

Príklad 2.2. Náš experiment pozostáva z jedného hodu kockou. 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), potom je zmysluplnou množinou elementárnych výsledkov $\Omega = \{1, 2, ..., 6\}$. Systém udalostí môžeme zvoliť $\mathcal{S} = 2^\Omega$. Udalosť $\{1, 2, 3\}$ zodpovedá výroku "padne číslo menšie než 4", udalosť $\{2, 4, 6\}$ zodpovedá výroku "padne párne číslo" a podobne.

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 komplement 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`).

Definícia 2.3 (*σ-algebra generovaná systémom množín*)**.** 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, že $\mathcal{F} \subseteq \mathcal{S}$ a súčasne $(\Omega, \mathcal{S})$ je σ-algebra. Potom σ-algebru $(\Omega, \sigma(\mathcal{F}))$ nazývame σ-algebra podmnožín množiny Ω generovaná systémom množín $\mathcal{F}$, alebo tiež minimálna σ-algebra podmnožín množiny Ω obsahujúca systém $\mathcal{F}$.

Príklad 2.3. V prípade, že je systém $\mathcal{F}$ konečný, je jednoduché nájsť $\sigma(\mathcal{F})$ postupným pridávaním prienikov, zjednotení a komplementov.

Nech napríklad $\Omega = \{1, 2, 3, 4, 5\}$ a $\mathcal{F} = \{\{1, 2\}, \{3, 4\}\}$. Potom

$$\sigma(\mathcal{F}) = \{\emptyset, \Omega, \{1, 2\}, \{3, 4\}, \{1, 2, 3, 4\}, \{3, 4, 5\}, \{1, 2, 5\}, \{5\}\}.$$

Zaujímavejšia je však situácia, keď je systém $\mathcal{F}$ nekonečný, ako v nasledujúcej definícii.

Príklad 2.4. 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 priestor elementárnych výsledkov 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 úplne elementárne. Ukazuje sa, že v tomto prípade je vhodnou σ-algebrou udalostí $\sigma(\mathcal{F})$, 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* bohatý. Napríklad o všetkých nasledovných udalostiach je možné ľahko ukázať, že patria do $\sigma(\mathcal{F})$: $A_1 = \{(i_1, i_2, \ldots) \in \Omega : \forall k \in \mathbb{N}\ (i_k = 0)\}$ ("vo všetkých hodoch padne hlava"), $A_2 = \{(i_1, i_2, \ldots) \in \Omega : \forall k \in \mathbb{N}\ (i_k \neq i_{k+1})\}$ ("budú sa striedať hlavy a znaky"), $A_3 = \{(i_1, i_2, \ldots) \in \Omega : i_1 = \cdots = i_{m-1} = 0 \wedge i_m = 1\}$ ("znak padne prvýkrát v m-tom hode").

Definícia 2.4 (*σ-algebra borelovských podmnožín $\mathbb{R}^m$*)**.** Nech $\mathcal{O}$ je systém všetkých otvorených podmnožín množiny $\mathbb{R}^m$. Potom σ-algebru $(\Omega, \mathcal{B}_m)$ podmnožín $\mathbb{R}^m$ generovanú systémom $\mathcal{O}$ nazývame σ-algebra borelovských podmnožín $\mathbb{R}^m$.

Veta 2.2 (*Základné borelovské podmnožiny množiny $\mathbb{R}$*)**.** Nech $a, b \in \mathbb{R}$, pričom $a < b$. Potom všetky nasledovné množiny patria do $\mathcal{B}_1$: $(-\infty, a)$, (a, ∞), $(-\infty, a]$, $[a, \infty)$ $(a, b]$, $[a, b)$, $[a, b]$, $\{a\}$. Do systému $\mathcal{B}_1$ tiež patrí akákoľvek spočítateľná podmnožina $\mathbb{R}$.

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. □

Príklad 2.5. 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 vhodným priestorom elementárnych výsledkov $\Omega = (0, \infty)$ a príslušným systémom udalostí je $\mathcal{S} = \{B \cap (0, \infty) : B \in \mathcal{B}_1\}$, t.j. tie borelovské množiny, ktoré sú podmnožinou $(0, \infty)$. 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. Samozrejme, rôzne udalosti môžu mať rôznu *pravdepodobnosť* nastatia, ktorú môžeme stanoviť na základe dlhodobých skúseností, fyzikálnych princípov a podobne. (Čo chápeme pod pojmom "pravdepodobnosť" je vysvetlené v nasledujúcej časti.)

2.2 Pravdepodobnostná miera

Definícia 2.5 (*Pravdepodobnostná miera*)**.** Pravdepodobnostná miera na σ-algebre udalostí $(\Omega, \mathcal{S})$ je zobrazenie $P : \mathcal{S} \to \mathbb{R}$ spĺňajúce nasledovné podmienky:

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 z definície 2.5 nazývame aditivita pravdepodobnosti ak je I konečná množina, alebo σ-aditivita ak je I nekonečná spočítateľná množina.

Definícia 2.6 (*Pravdepodobnostný priestor*)**.** Nech $(\Omega, \mathcal{S})$ je σ-algebra udalostí a nech P je pravdepodobnostná miera na $(\Omega, \mathcal{S})$. Potom trojicu $(\Omega, \mathcal{S}, P)$ nazveme pravdepodobnostný priestor.

Príklad 2.6. Nech $\Omega = \{1, \dots, 6\}$ a $\mathcal{S} = 2^\Omega$. Nech $P : \mathcal{S} \to [0, 1]$ 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 hracou kockou, ak nás zaujíma výlučne číselný výsledok hodu. (Pozri príklad 2.2.)

Existencia rovnomernej miery na jednotkovom intervale. Nech $\mathcal{S}$ je množina všetkých borelovských podmnožín kocky $[0, 1]^m$. Podľa úlohy 2.5 je $([0, 1]^m, \mathcal{S})$ σ-algebra. Možno ukázať, že na $([0, 1]^m, \mathcal{S})$ existuje pravdepodobnosť P, ktorá spĺňa nasledovnú podmienku: Pre každý kváder $A = [a_1, b_1] \times \cdots \times [a_m, b_m]$, kde $0 \leq a_i < b_i \leq 1$, platí

$$P(A) = \prod_{i=1}^{m} (b_i - a_i).$$

Pre $m = 1$ môžeme pravdepodobnostný priestor $(\Omega, \mathcal{S}, P)$ považovať za model "rovnomerného náhodného výberu reálneho čísla na úsečke $[0, 1]$". Podobne, pre $m = 2$ je $(\Omega, \mathcal{S}, P)$ model "rovnomerného náhodného výberu bodu na štvorci $[0, 1] \times [0, 1]$". Pravdepodobnosť P možno považovať za definíciu úhrnnej dĺžky (plochy, objemu) borelovských podmnožín jednotkového intervalu (štvorca, kocky).

2.3 Základné vlastnosti pravdepodobnosti

Z definície pravdepodobnostnej miery 2.5 sa dajú odvodiť mnohé užitočné vlastnosti pravdepodobnosti.

Veta 2.3 (*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 2.4 (*Pravdepodobnosť komplementu*)**.** 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 $B = \Omega$. □

Veta 2.5 (*Pravdepodobnosť zjednotenia dvoch udalostí*)**.** Nech A, B sú udalosti pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$. Potom platí

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

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

Veta 2.6 (*Princíp inklúzie-exklúzie (zapojenia-vypojenia)*)**.** Nech $A_1, A_2, \ldots, A_n$ sú udalosti na pravdepodobnostnom priestore $(\Omega, \mathcal{S}, P)$. Potom platí

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

Dôkaz. Pre $n = 2$ je tvrdenie tejto vety také isté ako vo vete 2.5. Pre všeobecné n je možné použiť matematickú indukciu. Na pochopenie princípu dôkazu stačí, ak si ako cvičenie dokážete túto vetu pre $n = 3$. □

Príklad 2.7. Za okrúhlym stolom je rozostavených $n \ge 3$ stoličiek. Náhodne za tento stôl rozsadíme troch ľudí. Aká je pravdepodobnosť p_n, že niektorí dvaja ľudia budú sedieť vedľa seba?

Riešenie: Zrejme $p_3 = 1$; v ďalšom budeme predpokladať, že $n \ge 4$. Označme ľudí ako 1,2,3 a definujme udalosť A (B, C) ako udalosť, že budú veľa 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 $n \ge 4$ 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 $B_1, B_2, ..., B_m$, že budú obsadené susedné stoličky $1, 2$; $2, 3$; ...;resp. $m, 1$. Hľadaná pravdepodobnosť je $P(\cup_i B_i)$, čo sa dá opäť vypočítať pomocou princípu zapojenia-vypojenia.

Poznámka 2.2. 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 2.8. Postupnosť čísiel $(1, 2, \ldots, n)$ dokonale náhodne premiešame. (T.j. každá spomedzi $n!$ permutácií má rovnakú pravdepodobnosť.) Nájdite pravdepodobnosť p_n, že aspoň jedno z čísiel $1, 2, \ldots, n$ bude po premiešaní na svojom pôvodnom mieste. Určte $\lim_{n\to\infty} p_n$.

Riešenie. Nech n je pevné. Nech A je udalosť, že aspoň jedno z čísiel $1, 2, \ldots, n$ 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 $1 \le i_1 < i_2 < \ldots < i_k \le n$ 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 2.6 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 2.3. Vlastnosť pravdepodobnosti z predchádzajúcej úlohy nazývame subaditivita ak je I konečná množina, alebo σ-subaditivita ak je I nekonečná spočítateľná množina.

Príklad 2.9. 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 $f : S \to \{Z, M\}$, 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 2.2. Dokážte, že do tejto kružnice je možné vpísať *rovnostranný* trojuholník tak, aby všetky jeho vrcholy ležali na modrej farbe.

Riešenie. Rovnomerne náhodne zvolíme bod X na jednotkovej kružnici S. Symbolom A označíme udalosť, že X padne do zelenej časti kružnice, symbolom B označme udalosť, že X po otočení o uhol $2\pi/3$ v smere hodinových ručičiek padne do zelenej farby a symbolom C označme udalosť, že 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 taká trojica bodov, čiže rovnostranný trojuholník, ktorého všetky tri vrcholy ležia na modrej časti kružnice.

Veta 2.7 (*Spojitosť pravdepodobnosti zdola*)**.** Nech $(A_i)_{i=1}^{\infty}$ je nekonečná postupnosť udalostí pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$, ktorá je neklesajúca v zmysle $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$. 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 $\cup_{i=1}^{\infty} B_i = \cup_{i=1}^{\infty} A_i$, ďalej že udalosti B_i sú disjunktné a naviac $P(B_i) = P(A_i) - P(A_{i-1})$. Taktiež ľahko overíme, že $\sum_{i=1}^{n} (P(A_i) - P(A_{i-1})) = P(A_n)$. Postupne dostávame $P\left(\cup_{i=1}^{\infty} A_i\right) = P\left(\cup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} (P(A_i) - P(A_{i-1})) = \lim_{n\to\infty} P(A_n)$. □

2.4 Cvičenia

Úloha 2.1 (*Prienik σ-algebier je σ-algebra*)**.** Nech J je neprázdna indexová množina (nemusí byť spočítateľná) a nech $(\Omega, \mathcal{S}_j)$ je σ-algebra pre každé $j \in J$. Potom $(\Omega, \cap_{j\in J} \mathcal{S}_j)$ je tiež σ-algebra.

Úloha 2.2. Uvažujme množinovú operáciu $\uparrow$ definovanú nasledovne: $\uparrow_{i\in I} A_i = \Omega \setminus (\cap_{i\in I} A_i)$, kde I je spočítateľná množina a $A_i \subseteq \Omega \neq \emptyset$. (Logický ekvivalent operácie $\uparrow$ sa zvykne nazývať `nand`.) Ukážte, že $(\Omega, \mathcal{S})$ je σ-algebra vtedy a len vtedy, keď sú splnené nasledovné podmienky: 1) $\Omega \in \mathcal{S}$ a 2) Ak je I spočítateľná množina a $A_i \in \mathcal{S}$ pre všetky $i \in I$, tak $\uparrow_{i\in I} A_i \in \mathcal{S}$.

Úloha 2.3. Nech $a, b, c, d \in \mathbb{R}$, pričom $a < b$ a $c < d$. Zdôvodnite, prečo všetky nasledovné množiny patria do $\mathcal{B}_2$: $(a, b) \times (c, d)$, $[a, b) \times (c, d)$, $[a, b] \times [c, d]$, bod (a, b), akákoľvek spočítateľná podmnožina $\mathbb{R}^2$. Rozmyslite si, prečo $\mathcal{B}_2$ obsahuje akýkoľvek n-uholník, alebo kruh (nezávisle na tom, či hranicu považujeme za súčasť týchto množín, alebo nie).

Úloha 2.4. Sami si definujte nejakú podmnožinu $\mathbb{R}^2$ a dokážte o nej, že je borelovská. (Prakticky akákoľvek množina, ktorú sme schopní popísať priamou konštrukciou, je borelovská. Existujú však aj množiny, ktoré nie sú borelovské; dokonca ich je z hľadiska množinovej mohutnosti "viac".)

Úloha 2.5. Nech $(\Omega, \mathcal{S})$ je σ-algebra a nech $\emptyset \neq A \in \mathcal{S}$. Definujme nasledovný systém množín: $\mathcal{R} = \{A \cap B : B \in \mathcal{S}\}$. Potom $(A, \mathcal{R})$ je tiež σ-algebra. Dokážte!

Úloha 2.6. Nech $\Omega = \{1, 2, 3\}$. Nájdite všetky také systémy $\mathcal{S} \subseteq 2^\Omega$, že $(\Omega, \mathcal{S})$ je σ-algebra.

Úloha 2.7 (*Booleova nerovnosť*)**.** Ukážte, že ak $(A_i)_{i \in I}$ je postupnosť udalostí pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$, tak

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

Úloha 2.8 (*Spojitosť pravdepodobnosti zhora*)**.** Ukážte nasledovné tvrdenie: Nech $(A_i)_{i=1}^{\infty}$ je nekonečná postupnosť udalostí pravdepodobnostného priestoru $(\Omega, \mathcal{S}, P)$, ktorá je nerastúca v zmysle $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$. Potom, rovnako ako vo vete 2.7, $P\left(\cap_{i=1}^{\infty} A_i\right) = \lim_{i \to \infty} P(A_i)$.

Kapitola 3

Podmieňovanie a nezávislosť udalostí

3.1 Podmienená pravdepodobnosť

V tejto kapitole budeme riešiť otázku, aký vplyv má na pravdepodobnosť udalosti A nastanie nejakej inej udalosti B.

Definícia 3.1 (*Podmienená pravdepodobnosť*)**.** 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.

Príklad 3.1. 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 tohoto 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?

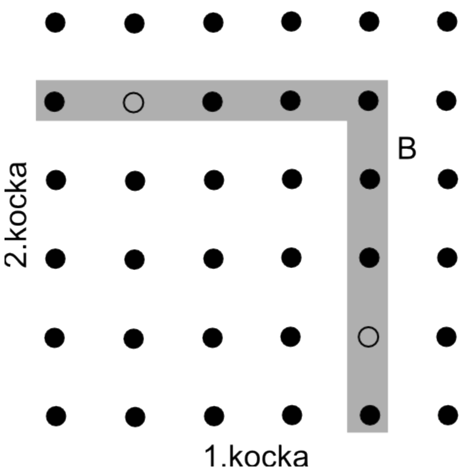
Takéto zadania chápeme nasledovne: Z hľadiska pred výberom vrecka a guľôčky, aká je pravdepodobnosť udalosti $A =$ "vyberieme vrecko s dvomi bielymi guľôčkami" za podmienky udalosti $B =$ "vybratá guľôčka bude biela"? Alebo iná formulácia: Predpokladajme, že priestor $(\Omega, \mathcal{S}, P)$ zodpovedá 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 úlohy 3.7. Aká je pravdepodobnosť udalosti A v tomto novom priestore?

Takže riešenie (nezávisle na tom, 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 $P(A|B) = 2/3$.

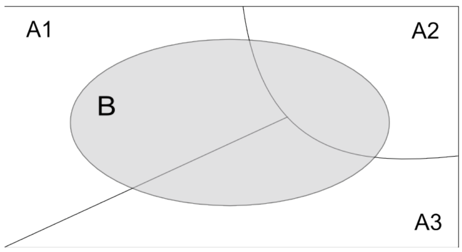
Definícia 3.2 (*Rozklad množiny elementárnych výsledkov*)**.** Budeme hovoriť, že udalosti $A_1, \dots, A_n$ tvoria rozklad množiny Ω elementárnych výsledkov, ak sú tieto udalosti disjunktné, každá z nich má nenulovú pravdepodobnosť a súčasne platí $\cup_{i=1}^n A_i = \Omega$.

Veta 3.1 (*Veta o úplnej pravdepodobnosti*)**.** Nech $A_1, \dots, A_n$ tvoria rozklad množiny elementárnych výsledkov. Nech B je akákoľvek udalosť. Potom platí:

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



Obr. 3.1: Hádžeme naraz dvomi kockami. Ak označíme $B = \{\max(U, V) = 5\}$ udalosť, že väčšie z padnutých čísel je 5 a $A = \{\min(U, V) = 2\}$ udalosť, že menšie z padnutých čísel je 2, potom podmienená pravdepodobnosť $P(A|B) = 2/9$.



Obr. 3.2

Dôkaz. $P(B) = P(B \cap (A_1 \cup \ldots \cup A_n)) = P((B \cap A_1) \cup \ldots \cup (B \cap A_n)) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$. □

Príklad 3.2. Udalosti A_1, A_2, a A_3 na nasledujúcom obrázku tvoria rozklad množiny Ω všetkých elementárnych udalostí, a teda udalosť B môžeme rozložiť na disjunktné podmnožiny typu $A_i \cap B$, $i - 1, 2, 3$. To ale znamená, že $B = (A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B)$ a z aditivity a definície podmienenej pravdepodobnosti je $P(B) = \sum_{i=1}^{3} P(A_i \cap B) = \sum_{i=1}^{3} P(A_i)P(B|A_i)$.

Veta 3.2 (*Bayesov vzorec*)**.** Nech udalosti $A_1, A_2, \ldots, A_n$ tvoria rozklad množiny elementárnych výsledkov. Nech B je akákoľvek udalosť nenulovej pravdepodobnosti a nech $k \in \{1, \ldots, n\}$. Potom platí:

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

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 dostávame dokazovanú rovnosť. □

Príklad 3.3. Majme systém, ktorý sa môže nachádzať v stavoch $S, A1, A2, A3, B1, B2, B3$, pričom sú možné nasledovné prechody medzi stavmi: $S \to A1$, $S \to A2$, $S \to A3$, $A1 \to B1$, $A1 \to B2$, $A2 \to B1$, $A2 \to B2$, $A2 \to B3$, $A3 \to B2$, $A3 \to B3$. 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 $B2$. Určte 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). 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 $P(A_2|B_2) = 1/4$.

Príklad 3.4. Na vstupe do laboratória sú 2 nezávisle pracujúce biometrické autorizačné zariadenia (typ A: snímanie prstu a typ B: snímanie očnej dúhovky.) Typ A vykazuje $0,6\%$ mylných odmietnutí a $0,1\%$ mylných prijatí. Typ B 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). Určte pravdepodobnosť, že systém ohlási pokus o neautorizovaný vstup pre neautorizovanú osobu. 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 sa naozaj jedná o pokus o neautorizovaný vstup za podmienky, že systém ohlásil pokus o neautorizovaný vstup?

Riešenie: Označme ako N udalosť, že dôjde k neautorizovanému pokusu o vstup a ako H udalosť, že systém *ohlási* neautorizovaný vstup. Označme si tiež ako H_A a H_B udalosti, že zariadenie A, resp. zariadenie B ohlási pokus o neautorizovaný vstup. Tiež si uvedomíme, že pravdepodobnosti $0,6\%$, $0,1\%$, 2%, $0,01\%$, $99,9\%$ a $0,1\%$ v zadaní sú postupne $P(H_A|N^c)$, $P(H_A^c|N)$, $P(H_B|N^c)$, $P(H_B^c|N)$, $P(N^c)$ a $P(N)$.

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

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 $P(H|N^c) = P(H_A \cup H_B|N^c) = P(H_A|N^c) + P(H_B|N^c) - P(H_A|N^c)P(H_B|N^c) = 0,006 + 0,02 - 0,006 \cdot 0,02 = 0,02588$.

Celkovo teda máme

$$P(N|H) = \frac{(1 - 10^{-7})0,001}{(1 - 10^{-7})0,001 + 0,02588 \cdot 0,999} \approx 0,037.$$

3.2 Nezávislosť udalostí

Definícia 3.3 (*Nezávislosť dvoch udalostí*)**.** 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.

Všimnime si, že pre nezávislé udalosti A a B ($P(B) > 0$) platí $P(A|B) = P(A)$, čiže, voľne povedané, znalosť toho, že nastala udalosť B, neovplyvní našu mieru očakávania, že nastane aj udalosť A.

Naopak, ak platí $P(A|B) = P(A)$ (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 3.4 (*Združená nezávislosť n-tice udalostí*)**.** $A_1, \ldots, A_n$ nech 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 sú tieto udalosti združene závislé.

Príklad 3.5. 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 čísiel $1, 2, \ldots, 8$, alebo modelu hádzania tromi mincami, ak každú z ôsmich rôznych kombinácií výsledkov na jednotlivých minciach označíme jedným z čísiel $1, 2, \ldots, 8$.)

Definujme nasledovné udalosti: $A = \{1, 2, 3, 4\}$, $B = \{1, 2, 5, 6\}$, $C = \{1, 3, 5, 7\}$, $D = \{1, 2, 7, 8\}$ a $E = \{1, 3, 4, 8\}$. 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 $P(A \cap B \cap E) = P(A)P(B)P(E)$.

Veta 3.3 (*Nezávislosť komplementov*)**.** 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 $A_1, \ldots, A_n$ sú združene nezávislé udalosti. Pre každé $i = 1, \ldots, n$ zvoľme za A'_i buď A_i, alebo $\Omega \setminus A_i$. Potom aj udalosti $A'_1, \ldots, A'_n$ 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 vety. □

Nezávislosť po skupinách: Predpokladajme, že máme združene nezávislé udalosti $A_1, \ldots, A_n$. Nech $n_0 = 0 < n_1 < n_2 < \ldots < n_{k-1} < n_k = n$. Pre každé $j = 1, \ldots, k$ vytvorme B_j z $A_{n_{j-1}+1}, \ldots, A_{n_j}$ akokoľvek, pomocou operácií komplementu, zjednotenia, alebo prieniku. Potom sú aj udalosti $B_1, \ldots, B_k$ združene nezávislé.

Veta 3.4 (*Binomická formula*)**.** Nech $A_1, A_2, \ldots, A_n$ sú 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 platí
$$P(A) = \binom{n}{k} p^k (1-p)^{n-k}$$

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.
$$B_{i_1,\ldots,i_k} = \bigcap_{i \in \{i_1,\ldots,i_k\}} A_i \bigcap_{j \notin \{i_1,\ldots,i_k\}} (\Omega \setminus A_j)$$
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 3.6. Komunikačný kanál sa skladá zo série uzlov, pričom vždy i-ty uzol predáva jednobitovú informáciu na vstup $i+1$-vému uzlu. Na každom uzle však s pravdepodobnosťou p dochádza k chybe, ktorá sa prejaví tým, že na výstupe tohoto uzla bude opačný bit ako na jeho vstupe. Naviac, chyby na jednotlivých uzloch sa vyskytujú navzájom nezávisle. Napíšte 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 $(k = 0, ..., n)$, tak podľa binomickej formule je hľadaná pravdepodobnosť
$$P(A^{(0)} \cup A^{(2)} \cup ... \cup A^{(n)}) = \sum_{0 \le j \le n/2} \binom{n}{2j} p^{2j} (1-p)^{n-2j} = \frac{(1-2p)^n}{2} + \frac{1}{2}.$$

Príklad 3.7. 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 priestoru elementárnych výsledkov a pravdepodobnosť každej z nich je 1/2, dostávame podľa vety o úplnej pravdepodobnosti a binomickej formule nasledovné rovnosti:
$$P(B) = \sum_{i=1}^{2} P(B \cap C_i) P(C_i) = \frac{1}{2}\left(2\binom{6}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2 + \binom{6}{4}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^2\right),$$

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

Po mechanických úpravách zisťujeme, že

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3}{5} = 0,6.$$

Veta 3.5 (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 množiny elementárnych výsledkov, 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 platí

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

Príklad 3.8. V urne máme 10 guličiek, z ktorých je 5 bielych, 3 sú modré a 2 sú červené. Z urny 7-krát vyberieme guličku, pričom každú vybranú guličku vrátime naspäť do urny ešte pred výberom ďalšej guličky. Vypočítajte pravdepodobnosť, že takto vyberieme spolu 3 biele guličky, 2 modré guličky a 2 červené guličky (nezávisle na tom, 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 guličku j-tej farby (farba 1 je biela, farba 2 modrá a farba 3 červená), tak máme situáciu z predchádzajúcej vety. Dostávame

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

3.3 Cvičenia

Úloha 3.1. Hádžeme 100-krát mincou. Aká je pravdepodobnosť, že padne rovnaký počet hláv a znakov?

Úloha 3.2. Vykonáme dva nezávislé hody kockou. Aká je pravdepodobnosť, že súčet bodiek na oboch kockách je párne číslo, ak vieme, že v prvom hode padlo číslo 2?

Úloha 3.3. Uvažujme systém zložený z troch nezávislých komponentov, ktorý je funkčný, ak aspoň dva z týchto troch komponentov pracujú bez poruchy. Aká je pravdepodobnosť, že systém je funkčný, ak každý z komponentov funguje s pravdepodobnosťou 0,98?

Úloha 3.4. Máme dve vrecká: v prvom je a_1 čiernych a b_1 bielych guličiek a v druhom je a_2 čiernych a b_2 bielych guličiek. Z každéhé vrecka vyberieme po jednej guličke a z týchto dvoch potom jednu guličku. Aká je pravdepodobnosť, že vybratá gulička bude biela?

Úloha 3.5. V urne je 12 loptičiek, po troch loptičkách z každej zo štyroch farieb. Postupne vyberieme štyri loptičky (s vracaním, t.j. vybratú loptičku vrátime naspäť do urny ešte pred výberom nasledujúcej loptičky). Aká je pravdepodobnosť, že všetky štyri vybraté loptičky budú mať rôznu farbu?

Úloha 3.6. Športovec 3 krát nezávisle vystrelil na cieľ. Pravdepodobnosti zásahov sú postupne 0.5, 0.6 a 0.7. Nájdite pravdepodobnosť toho, že
a) v cieli bude aspoň jeden zásah,
b) v cieli bude práve jeden zásah,
c) v cieli budú práve dva zásahy.

Úloha 3.7. Nech $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor a nech $B \in \mathcal{S}$, $P(B) > 0$. Definujme funkciu $P_B : \mathcal{S} \to \mathbb{R}$ nasledovne: $P_B(A) = P(A|B)$. Potom P_B je pravdepodobnostná miera na $(\Omega, \mathcal{S})$.

Úloha 3.8. Uvažujme pravdepodobnostný priestor $(\Omega, \mathcal{S}, P)$ modelujúci hod dvomi kockami, čiže $\Omega = \{1, ..., 6\}^2$, $\mathcal{S} = 2^\Omega$ a $P(M) = |M|/36$, kde $|M|$ je počet prvkov množiny M. Formálne zapíšte udalosť A zodpovedajúcu výroku, že na prvej kocke padne párne číslo a udalosť B zodpovedajúcu výroku, že na druhej kocke padne nepárne číslo. Presvedčte sa, že udalosti A, B sú nezávislé. Uvažujme udalosť C, že na prvej kockej padne číslo menšie ako 5 a súčasne na druhej kocke padne číslo menšie ako 4. Uvažujme tiež udalosť D, že na prvej kocke padne jedno z čísiel $3, 4, 5$ a súčasne na druhej kocke padne jedno z čísiel $2, 3, 4, 5$. Formálne zapíšte udalosti C, D ako podmnožiny množiny Ω. Sú udalosti C, D nezávislé?

Úloha 3.9 (Voľne prevzaté z [3]). Predpokladajme, že sme napísali (klasický, "deterministický") algoritmus na overovanie nasledovnej rovnosti dvoch polynómov:

$$\prod_{i=1}^{d}(a_i x - b_i) = c_d x^d + ...c_1 x + c_0. \tag{3.1}$$

(Vstupom algoritmu sú koeficienty $a_1, ..., a_d, b_1, ..., b_d, c_0, ..., c_d$ a výstupom je logická hodnota `true`, alebo `false`, podľa toho, či (3.1) platí alebo nie.) Všimnite si, že tento algoritmus potrebuje $\Theta(d^2)$ násobení.

Uvažujme tiež nasledovný "znáhodnený" algoritmus: rovnomerne náhodne vygenerujeme prirodzené čísla $X_1, ..., X_k$ z množiny $\{1, 2, ..., 100d\}$. Nech Q je polynóm na ľavej strane rovnosti (3.1) a R nech je polynóm na pravej strane rovnosti (3.1). Pokiaľ bude platiť $Q(X_i) = R(X_i)$ pre každú hodnotu $X_1, ..., X_k$, znáhodnený algoritmus vráti hodnotu `true`, v opačnom prípade (t.j. ak $Q(X_i) \neq R(X_i)$ pre čo i len jednu hodnotu $X_1, ..., X_k$), znáhodnený algoritmus vráti hodnotu `false`. Aká je pravdepodobnosť, že sa znáhodnený algoritmus "pomýli" v prípade, že rovnosť (3.1) platí a aká je táto pravdepodobnosť v prípade, že (3.1) neplatí? Aký počet násobení vyžaduje náš znáhodnený algoritmus?

Kapitola 4

Všeobecné náhodné premenné

V mnohých situáciách je výsledkom experimentu nejaká numerická hodnota. Napríklad ak náhodne vyberáme športovca z družstva, môže nás zaujímať jeho výkon. V takýchto prípadoch je užitočné daným hodnotám priradiť príslušné pravdepodobnosti. Toto priradenie budeme robiť pomocou reálnych funkcií definovaných na priestore Ω, ktoré nazveme náhodné premenné.

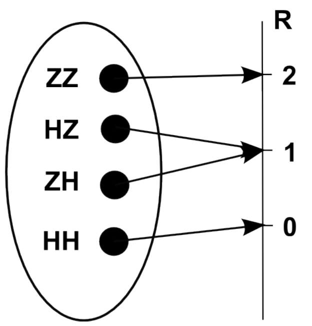
4.1 Základné vlastnosti náhodných premenných

Definícia 4.1 (*Náhodná premenná*)**.** Nech $(\Omega, \mathcal{S}, P)$ je pravdepodobnostný priestor. Budeme hovoriť, že funkcia $X : \Omega \to \mathbb{R}$ je náhodná premenná, ak pre každé $x \in \mathbb{R}$ platí

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

Príklad 4.1. Príklady náhodných premenných:

- Hádžeme naraz piatimi kockami. Súčet bodiek na všetkých kockách, najväčší počet bodiek na kocke, aj počet kociek, na ktorých padla šestka, sú náhodné premenné.
- Hráme nasledovnú hru: hádžeme mincou a vyhráme 1 euro, ak padne znak a prehráme 2 eurá, ak padne hlava. Výška výhry v tejto hre je náhodná premenná.
- Hádžeme dvakrát mincou. Počet znakov v týchto dvoch hodoch je náhodná premenná.



Pre jednoduchosť budeme množinu $\mathcal{B}_1$ borelovských podmnožín množiny $\mathbb{R}$ označovať symbolom $\mathcal{B}$.

Veta 4.1. 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 platí $X^{-1}(H) \in \mathcal{S}$. Ľahko overíme, že $(\mathbb{R}, \mathcal{H})$ je σ-algebra, 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 $(\mathbb{R}, \mathcal{B})$ je najmenšia σ-algebra 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 $\mathcal{B}$. □

Poznámka 4.1. Pre $B \in \mathcal{B}$ budeme udalosť $\{\omega \in \Omega : X(\omega) \in \mathcal{B}\}$ 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.

Definícia 4.2. Funkciu $g : \mathbb{R}^n \to \mathbb{R}$ nazveme borelovskou, 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 vety majú predovšetkým teoretický význam a ich dôkazy sú technicky pomerne zdĺhavé, preto ich dokazovať nebudeme.

Veta 4.2. Nech $B_1, B_2, ... \in \mathcal{B}_n$ sú navzájom disjunktné a také, že $\cup_{i=1}^{\infty} B_i = \mathbb{R}^n$. Nech $g_1, g_2, ...$ 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 (funkcia g je "po častiach spojitá").

Veta 4.3. Nech $X_1, ..., X_n$ sú náhodné premenné na priestore udalostí $(\Omega, \mathcal{S})$ a nech g je borelovská funkcia. Potom aj zobrazenie $g(X_1, ..., X_n) : \Omega \to \mathbb{R}$ definované

$$g(X_1, ..., X_n)(\omega) = g(X_1(\omega), ..., X_n(\omega));\ \omega \in \Omega$$

je náhodnou premennou na $(\Omega, \mathcal{S})$.

Dôležité je uvedomiť si to, že prakticky každá "slušná" funkcia z $\mathbb{R}^n$ do $\mathbb{R}$ sa dá napísať v tvare po častiach spojitej 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 funkcia $g(x_1, x_2) = x_1 + x_2$ je spojitá na $\mathbb{R}^2$), ale aj napríklad $X_1^{\lceil |X_2| \rceil}$ 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 vety ako funkcia po častiach spojitá na spočítateľnom systéme dvojrozmerných borelovských množín).

Definícia 4.3 (*Nezávislosť náhodných premenných*)**.** Nezávislými nazývame náhodné premenné $X_1, ..., X_n$ vtedy, keď pre akékoľvek $B_1, ..., B_n \in \mathcal{B}$ platí

$$P[X_1 \in B_1, ..., X_n \in B_n] = \prod_{i=1}^{n} P[X_i \in B_i]. \tag{4.1}$$

Veta 4.4. Nech $X_1, \dots, X_n$ sú náhodné premenné a nech systém množín $\mathcal{F} \subseteq \mathcal{B}$ generuje na $\mathbb{R}$ σ-algebru $(\mathrm{R}, \mathcal{B})$. Potom $X_1, \dots, X_n$ sú nezávislé vtedy a len vtedy, keď rovnosť (4.1) platí pre akékoľvek $B_1, ..., B_n \in \mathcal{F}$.

Veta 4.5 (*Nezávislosť náhodných premenných po skupinách*)**.** Predpokladajme, že $X_1, X_2, \dots, X_n$ sú nezávislé náhodné premenné. Nech $n_0 = 0 < n_1 < n_2 < \dots < n_{k-1} < n_k = n$. Pre každé $j = 1, \dots, k$ nech $g_j : \mathbb{R}^{n_j - n_{j-1}} \to \mathbb{R}$ je borelovská funkcia a

$$Y_j = g_j(X_{n_{j-1}+1}, X_{n_{j-1}+2}, \dots, X_{n_j}).$$

Potom sú náhodné premenné $Y_1, Y_2, ..., Y_k$ nezávislé.

Teda napríklad, ak X_1, X_2, X_3 sú nezávislé náhodné premenné, potom aj náhodné premenné $Y_1 = X_1$ a $Y_2 = X_2^2 - X_3$ sú nezávislé.

4.2 Distribučná funkcia náhodnej premennej

Definícia 4.4 (*Distribučná funkcia*)**.** Distribučnou funkciou náhodnej premennej X nazývame funkciu $F : \mathbb{R} \to \mathbb{R}$, ktorá je v bode $x \in \mathbb{R}$ definovaná

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

Príklad 4.2. Hádžeme dvakrát mincou. Nech náhodná premenná X počet znakov v týchto dvoch hodoch. Potom distribučná funkcia náhodnej premennej X je

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

Veta 4.6 (*Základné vlastnosti distribučnej funkcie*)**.** Nech F je distribučná funkcia akejkoľvek náhodnej premennej X. Potom platí:

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

Dokážeme spojitosť zľava. Nech $a \in \mathbb{R}$. Pre každé prirodzené číslo n platí $[X < a - 1/n] \subseteq [X < a - 1/(n+1)]$, takže z neklesajúcosti distribučnej funkcie a vety 2.7 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. $\square$

Veta 4.7 (*Pravdepodobnosti intervalov vyjadrené pomocou F*)**.** Nech F je distribučná funkcia náhodnej premennej X. Nech $a, b \in \mathbb{R}$, pričom $a < b$. Potom platí

1. $P[a \le X < b] = F(b) - F(a)$, $P[a \le X] = 1 - F(a)$
2. $P[a \le X \le b] = \lim_{x \to b^+} F(x) - F(a)$, $P[X = a] = \lim_{x \to a^+} F(x) - F(a)$
3. $P[a < X < b] = F(b) - \lim_{x \to a^+} F(x)$, $P[a < X] = 1 - \lim_{x \to a^+} F(x)$
4. $P[a < X \le b] = \lim_{x \to b^+} F(x) - \lim_{x \to a^+} F(x)$, $P[X \le b] = \lim_{x \to b^+} F(x)$

Dôkaz. Priamo z definície máme $P[a \le X < b] = P[X < b] - P[X < a] = F(b) - F(a)$. 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. $P[X = a] = P(\cap_{n=1}^{\infty} [a - 1/n \le X < a + 1/n]) = \lim_{n \to \infty} P[a - 1/n \le X < a + 1/n] = \lim_{n \to \infty} (F(a + 1/n) - F(a - 1/n)) = \lim_{x \to a^+} F(x) - F(a)$. Druhá rovnosť plynie zo spojitosti pravdepodobnosti, posledná rovnosť plynie z toho, že F je spojitá zľava. □

4.3 Cvičenia

Úloha 4.1. Nech $(\Omega, \mathcal{S})$ je priestor udalostí (σ-algebra) a nech $A \subseteq \Omega$. Definujme zobrazenie $I_A : \Omega \to \mathbb{R}$ nasledovne: $I_A(\omega) = 1$ pre všetky $\omega \in A$ a $I_A(\omega) = 0$ pre všetky $\omega \in \Omega / A$. (Zobrazenie I_A nazývame identifikátorom množiny A.) Potom I_A je náhodná premenná na tomto priestore vtedy a len vtedy, keď $A \in \mathcal{S}$. Dokážte!

Úloha 4.2. Presvedčte sa, že na priestore $(\Omega, 2^{\Omega})$ je každé zobrazenie $X : \Omega \to \mathbb{R}$ náhodnou premennou a na priestore $(\Omega, \{\Omega, \emptyset\})$ sú náhodnými premennými len konštantné zobrazenia. Popíšte množinu všetkých náhodných premenných na σ-algebre $(\Omega, \{\Omega, \{0, 1\}, \{2\}, \emptyset\})$, kde $\Omega = \{0, 1, 2\}$.

Úloha 4.3. Uvažujme pravdepodobnostný priestor $(\Omega, 2^{\Omega}, P)$, kde $\Omega = \{1, ..., 6\}$ a $P(A) = |A|/6$ pre $A \subseteq \Omega$. Nájdite distribučnú funkciu náhodných premenných $X : \Omega \to \mathbb{R}$ definovaných
a) $X(\omega) = \omega$;
b) $X(\omega) = \omega \mod 4$ pre všetky $\omega \in \Omega$.

Úloha 4.4. Uvažujme pravdepodobnostný priestor $(\Omega, \mathcal{S}, P)$, kde $\Omega = \{a, b, c\}$, $\mathcal{S} = 2^{\Omega}$ a pre každé $A \in \mathcal{S}$ platí $P(A) = |A|/3$. ($|A|$ je počet prvkov množiny A.) Načrtnite distribučnú funkciu náhodnej veličiny X, nájdite $E(X)$ a $D(X)$, ak
a) $X(a) = X(b) = X(c) = 0$;
b) $X(a) = X(b) = 0, X(c) = 1$;
c) $X(a) = 0, X(b) = 1, X(c) = 2$.

Kapitola 5

Diskrétne náhodné premenné

5.1 Základné vlastnosti diskrétnych náhodných premenných

Definícia 5.1 (*Diskrétna náhodná premenná*)**.** Náhodnú premennú X na pravdepodobnostnom priestore $(\Omega, \mathcal{S}, P)$ nazývame diskrétna, ak jej obor hodnôt $X(\Omega) \subset \mathbb{R}$ je spočítateľná množina.

Niekedy hovoríme, že diskrétna náhodná premenná X "nadobúda" spočítateľne veľa hodnôt. Hovoríme tiež, že diskrétna náhodná premenná X nadobúda číslo x "s pravdepodobnosťou" $P[X = x]$.

Poznámka 5.1. Všimnite si, že ak je $X_1, ..., X_n$ sú diskrétne náhodné premenné a $g : \mathbb{R}^n \to \mathbb{R}$ je akákoľvek borelovská funkcia, tak $g(X_1, ..., X_n)$ je tiež diskrétna náhodná premenná, pretože jej obor hodnôt musí byť spočítateľný.

Veta 5.1 (*Nezávislosť diskrétnych náhodných premenných*)**.** Diskrétne náhodné premenné $X_1, X_2, ..., X_n$ sú nezávislé vtedy a len vtedy, keď platí

$$P[X_1 = x_1, X_2 = x_2, ..., X_n = x_n] = \prod_{i=1}^{n} P[X_i = x_i] \tag{5.1}$$

pre všetky $x_1, x_2, ..., x_n \in \mathbb{R}$.

Dôkaz. Dôkaz implikácie $\Rightarrow$ plynie priamo z definície nezávislosti a z toho, že jednoprvkové množiny $B_1 = \{x_1\}$,...,$B_n = \{x_n\}$ sú borelovské.

Ukážme opačnú implikáciu. Nech $B_1, ..., B_n$ sú akékoľvek borelovské množiny. Pre každé $i = 1, ..., n$ označíme $C_i = X_i(\Omega) \cap B_i$ množinu tých čísiel z oboru hodnôt náhodnej premennej X_i, ktoré patria do B_i. Keďže náhodné premenné X_i sú diskrétne, tak množiny C_i sú spočítateľné. Z aditivity (prípadne σ-aditivity) pravdepodobnosti a rovnosti (5.1) dostávame:

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

□

5.2 Číselné charakteristiky diskrétnych náhodných premenných

Definícia 5.2 (*Stredná hodnota diskrétnej náhodnej premennej*)**.** Nech X je diskrétna náhodná premenná a nech rad $\sum_{x \in X(\Omega)} xP[X = x]$ absolútne konverguje. Potom hovoríme, že náhodná premenná X má konečnú strednú hodnotu $E(X)$ a kladieme:

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

Ak rad na pravej strane absolútne nekonverguje, hovoríme, že náhodná premenná X nemá konečnú strednú hodnotu.[1]

Poznámka 5.2. 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 $x_1, x_2, ...$ je (akékoľvek) očíslovanie prvkov množiny $\mathbb{X}$. Absolútna konvergencia zaručuje, že hodnota $\sum_{x \in \mathbb{X}} r(x)$ je konečná a rovnaká nezávislé na tom, v akom poradí sčitujeme členy tohoto radu.

Veta 5.2 (*Linearita strednej hodnoty*)**.** Nech X a Y sú diskrétne náhodné premenné, ktoré majú konečnú strednú hodnotu. Nech a, b sú reálne čísla. Potom aj diskrétna náhodná premenná $aX + bY$ 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. Obor hodnôt náhodnej premennej $aX + bY$ je určite podmnožinou množiny $aX(\Omega) + bY(\Omega)$, kde $X(\Omega)$ a $Y(\Omega)$ sú obory hodnôt náhodných premenných X a Y. Z definície strednej

[1] Ak je X nezáporná náhodná premenná a $\sum_{x \in X(\Omega)} xP[X = x] = \infty$, tak sa niekedy hovorí, že náhodná premenná X má nekonečnú strednú hodnotu.

hodnoty a aditivity (prípade σ-aditivity) pravdepodobnosti dostávame:

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

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

Veta 5.3 (*Stredná hodnota funkcie diskrétnej náhodnej premennej*)**.** Nech X je diskrétna náhodná premenná a nech $g:\mathbb{R}\to\mathbb{R}$ je funkcia. Potom $g(X)$ je diskrétna náhodná premenná a jej stredná hodnota je konečná, ak rad $\sum_{x\in X(\Omega)} g(x)P[X=x]$ absolútne konverguje. V takom prípade platí:

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

Definícia 5.3 (*Disperzia diskrétnej náhodnej premennej*)**.** Nech X je diskrétna náhodná premenná a nech náhodná premenná X^2 má konečnú strednú hodnotu. Potom hovoríme, že náhodná premenná X má konečnú disperziu $D(X)$ a kladieme:

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

Ak náhodná premenná X^2 nemá konečnú strednú hodnotu, tak hovoríme, že X má nekonečnú disperziu.

Veta 5.4 (*Základné vlastnosti disperzie diskrétnej náhodnej premennej*)**.** Nech X je náhodná premenná, ktorá má konečnú disperziu a nech a,b sú reálne čísla. Potom aj diskrétna náhodná premenná $aX+b$ má konečnú disperziu a platí:

$$\begin{aligned} D(X) &= E(X^2)-E(X)^2, \\ D(aX+b) &= a^2D(X). \end{aligned}$$

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

Ak diskrétna náhodná premenná X nadobúda hodnoty $(x_i)_{i\in I}$ s nenulovou pravdepodobnosťou a ak má konečnú strednú hodnotu, tak z vety 5.3 a predchádzajúcej vety plynie, že disperziu X môžeme vypočítať podľa ktoréhokoľvek z nasledujúcich dvoch vzorcov

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

5.3 Základné typy diskrétnych náhodných premenných

5.3.1 Alternatívne rozdelenie

Definícia 5.4 (*Alternatívne rozdelenie*)**.** 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 $X \sim Alt(p)$.

Príklad 5.1. Hádžeme mincou, na ktorej padne znak s pravdepodobnosťou p. 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 5.5. Nech $X \sim Alt(p)$. Potom $E(X) = p$ a $D(X) = p(1-p)$.

Dôkaz. Dôkaz je elementárne cvičenie. □

5.3.2 Binomické rozdelenie

Definícia 5.5 (*Binomické rozdelenie*)**.** 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 $X \sim Bin(n,p)$.

Príklad 5.2. Hádžeme mincou z predchádzajúceho príkladu 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 5.6. Nech $X \sim Bin(n,p)$. Potom $E(X) = np$ a $D(X) = np(1-p)$.

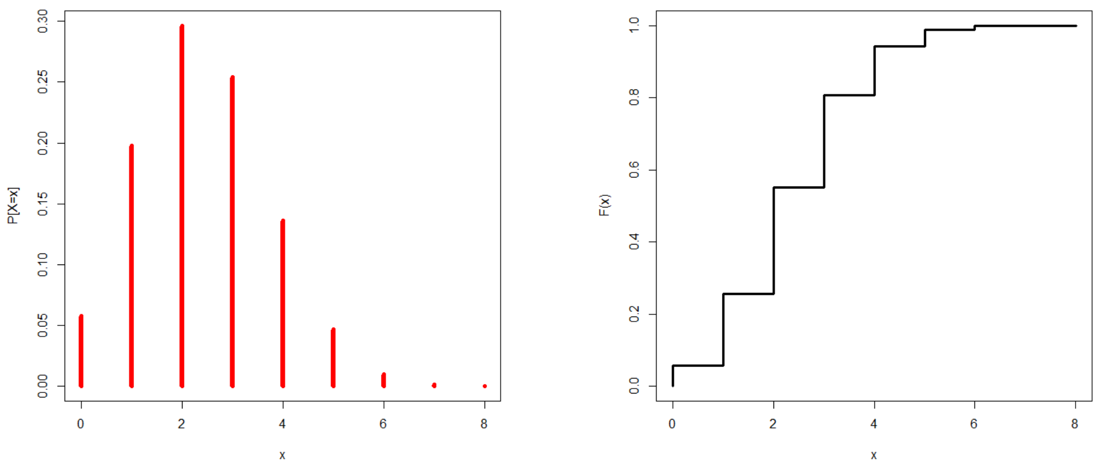
Dôkaz. Nech $X \sim Bin(n,p)$, $q = 1-p$. Platí:

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

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

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

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



Obr. 5.1: Rozdelenie pravdepodobnosti a distribučná funkcia náhodnej premennej $X \sim Bin(8, 0.3)$

Poznámka 5.3. Náhodná premenná X s rozdelením $Bin(n,p)$ zodpovedá počtu úspechov, ak robíme n nezávislých experimentov a pravdepodobnosť úspechu v každom experimente je p. Všimnite si tiež, že $X \sim Alt(p)$ vtedy a len vtedy, keď $X \sim Bin(1,p)$.

Príklad 5.3. 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. Nájdite strednú hodnotu počtu jednotiek, ktoré bude obsahovať chromozóm y.

Riešenie: Ak za "úspech" budeme považovať to, že dôjde k preklopeniu bitu, tak 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 pri mutácii 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 5.6 dostávame:

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

Príklad 5.4. Hráme hru, v ktorej sa ťahajú 4 čísla zo 40. Vyhráme, ak uhádneme všetky štyri čísla. Určte pravdepodobnosť, že ak sa hry zúčastníme 500 krát, vyhráme raz alebo dvakrát. Ak pravdepodobnosť výhry v jednej hre označíme p, máme

$$p = \frac{\binom{4}{4}\binom{36}{0}}{\binom{40}{4}} \approx 1.1 \times 10^{-5}$$

Potom náhodná premenná X, ktorá označuje počet výhier v 500 hrách, má binomické rozdelenie s parametrami 500 a p a

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

V takomto prípade, keď je n 'veľké' a p 'veľmi malé', je výhodnejšie aproximovať binomické rozdelenie nasledovným spôsobom.

Veta 5.7. Majme postupnosť náhodných premenných $X_{n_1}, X_{n_1+1}, \ldots$ pričom $X_n \sim Bin(n, \lambda/n)$, kde $0 < \lambda \le n_1$. Potom pre každé $k \in \{0, 1, 2, \ldots\}$ platí

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

Dôkaz.

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

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

Avšak zrejme platí

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

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

z predchádzajúcich rovností dostávame požadované tvrdenie. $\square$

5.3.3 Poissonovo rozdelenie

Definícia 5.6 (*Poissonovo rozdelenie*)**.** 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 $X \sim Po(\lambda)$.

Príklad 5.5. V predchádzajúcom príklade môžeme binomické rozdelenie aproximovať Poissonovým rozdelením s parametrom $\lambda = 500p \approx 0.00547$ a dostávame Potom náhodná premenná X, ktorá označuje počet výhier v 500 hrách, má binomické rozdelenie s parametrami 500 a p a

$$P[X = 1 \vee X = 2] = e^{-\lambda}\left(\frac{\lambda}{1} + \frac{\lambda^2}{2}\right) \approx 0.00547.$$

Veta 5.8. Nech $X \sim Po(\lambda)$. Potom $E(X) = \lambda$ a $DX = \lambda$.

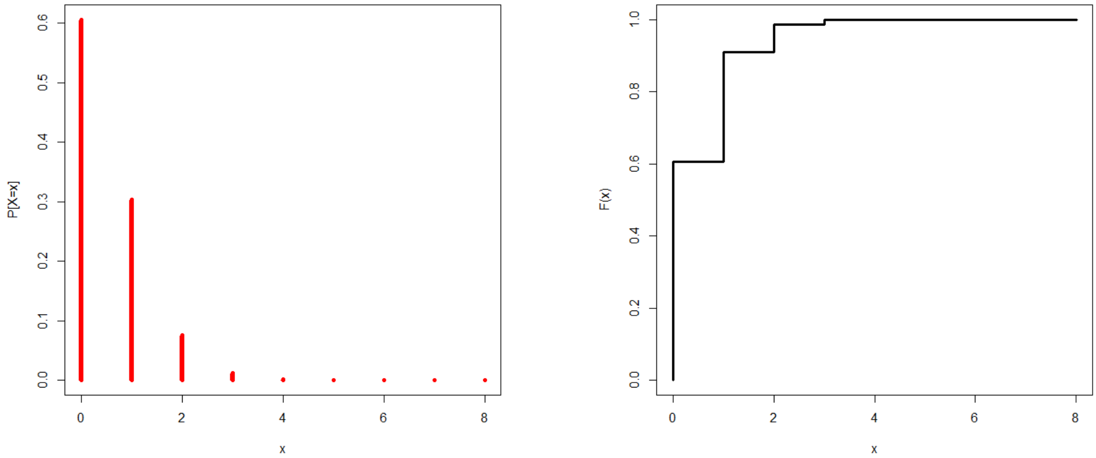
Dôkaz.

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

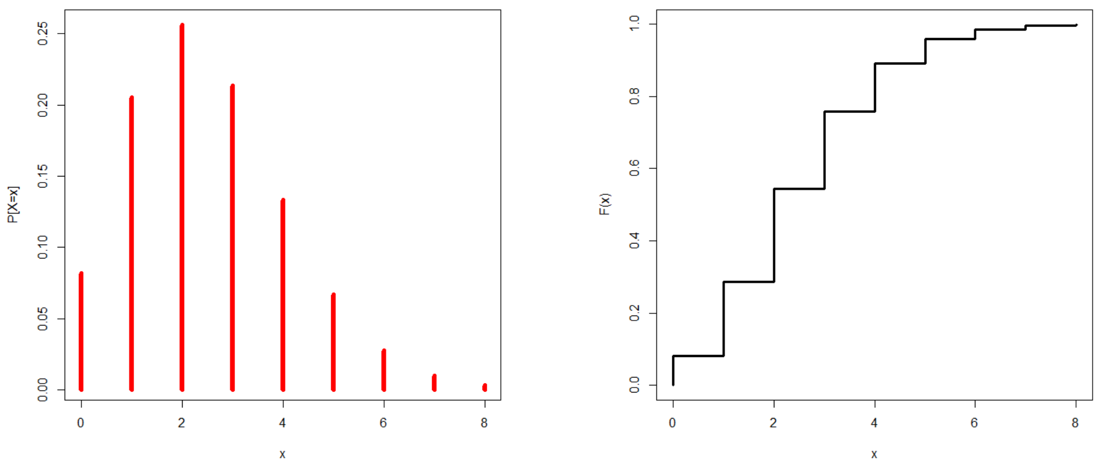
Podobne máme

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

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



Obr. 5.2: Rozdelenie pravdepodobnosti a distribučná funkcia náhodnej premennej $X \sim Po(0.5)$



Obr. 5.3: Rozdelenie pravdepodobnosti a distribučná funkcia náhodnej premennej $X \sim Po(2.5)$

Poissonovo rozdelenie sa používa napríklad na modelovanie počtu rozpadov atómov rádioaktívnej látky za určitý čas, volaní na telefónnu ústredňu, impulzov prichádzajúcich na neurónovú bunku a podobne. Poissonovo rozdelenie je tiež limitným rozdelením niektorých postupností náhodných premenných (pozri vetu 5.7 a príklad 5.6).

Príklad 5.6. Nech $n \in \mathbb{N}$ a nech X_n znamená počet prvkov, ktoré zostanú na svojom pôvodnom mieste po dokonalej náhodnej permutácii postupnosti $(1, ..., n)$. Dokážte, 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 2.8. Položme $q_n = P[X_n = 0] = 1 - p_n$, t.j. q_n je pravdepodobnosť, že po náhodnej permutácii n čísiel nezostane ani jedno z nich na svojom pôvodnom mieste. Dodefinujme tiež $q_0 = 1$. Pre všeobecné $k \geq 1$ máme

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

kde $B_{i_1,...,i_k}$ znamená udalosť, že čísla $i_1, ..., i_k$ 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 disjkunktné a pravdepodobnosť každej takejto udalosti je

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

Spojením týchto výsledkov dostávame

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

Použitím riešenia príkladu 2.8 máme

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

5.3.4 Geometrické rozdelenie

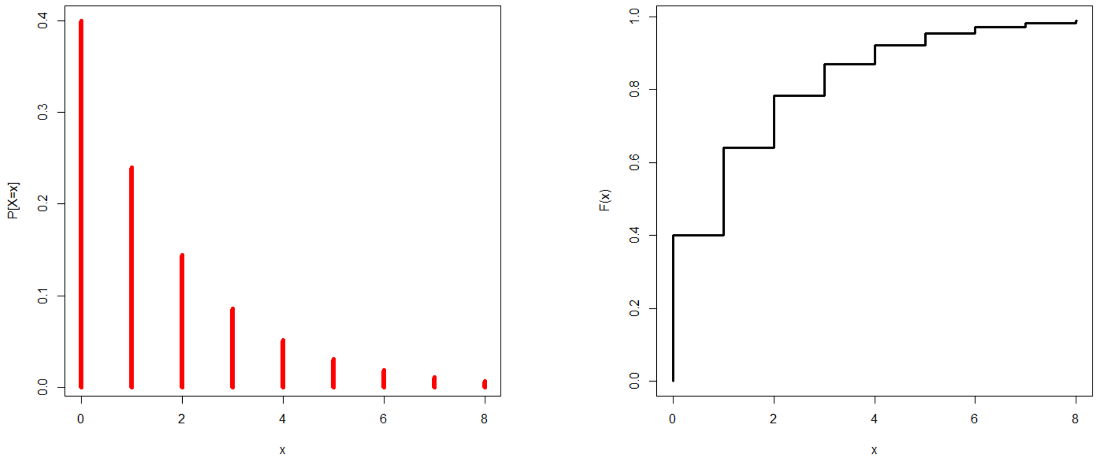
Definícia 5.7 (*Geometrické rozdelenie*)**.** Hovoríme, že diskrétna náhodná premenná X má geometrické rozdelenie s parametrom $p \in [0, 1]$, ak pre každé $k = 0, 1, 2, \ldots$ platí

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

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

Príklad 5.7. Hádžeme mincou, na ktorej padne znak s pravdepodobnosťou p. Potom náhodná premenná, ktorá reprezentuje počet hodov, kým nepadne prvý krát znak, má geometrické rozdelenie s parametrom p.

Veta 5.9. Nech $X \sim Geo(p)$. Potom platí $E(X) = (1 - p)/p$ a $D(X) = (1 - p)/p^2$.



Obr. 5.4: Rozdelenie pravdepodobnosti a distribučná funkcia náhodnej premennej $X \sim Geo(0.4)$

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

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

Dvojnásobným derivovaním funkcie f dostaneme vzorec pre súčet nekonečného radu, pomocou ktorého odvodíme disperziu veľmi podobne, ako sme vypočítali strednú hodnotu. $\square$

Náhodná premenná X s rozdelením $Geo(p)$ zodpovedá počtu neúspešných experimentov predchádzajúcich prvý úspešný experiment, ak pravdepodobnosť úspechu v jednotlivom experimente je p.

Príklad 5.8. 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 formule

$$p = \binom{4}{3}(0,1)^3(0,9)^1 + \binom{4}{4}(0,1)^4(0,9)^0 = 0,0037.$$

Počet "neúspešných experimentov" 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 5.9:

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

Príklad 5.9. 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. Určte strednú hodnotu času výpočtu T.

5.3.5 Hypergeometrické rozdelenie

Definícia 5.8 (*Hypergeometrické rozdelenie*)**.** Hovoríme, že diskrétna náhodná premenná X má hypergeometrické rozdelenie s parametrami $M, N, n \in \mathbb{N}$, spĺňajúcimi $N < M$ a $n < M$, ak

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

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

Poznámka 5.4. Predpokladajme, že máme urnu s M guličkami, z ktorých je N čiernych. Z urny (súčasne alebo, ekvivalentne, bez vrátenia) vyberieme n guličiek. Ak X znamená počet čiernych guličiek v tomto výbere, tak X má rozdelenie $Hyp(M, N, n)$. Hypergeometrické rozdelenie sa používa napríklad pri pravdepodobnostnej analýze lotérií a v štatistickej kontrole kvality.

Veta 5.10. Nech $X \sim Hyp(M, N, n)$. Potom $E(X) = n\frac{N}{M}$.

Dôkaz. Pozri príklad 5.10. □

5.4 Využitie indikátorov udalostí a linearity strednej hodnoty

Príklad 5.10. Nech $X \sim Hyp(M, N, n)$. Určte EX.

Riešenie: Ak sa chceme vyhnúť výpočtu komplikovaných súm, môžeme urobiť dôkaz s využitím linearity strednej hodnoty a to nasledovne: Uvažujme urnovú schému z poznámky 5.4. Označme pre $i = 1, ..., n$ ako U_i náhodnú premennú, ktorá nadobúda hodnotu 1 v prípade, že i-ta vybratá guľôčka je čierna a hodnotu 0 ak je biela. 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 disperziu; výpočet je však zdĺhavejší.

Príklad 5.11. 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 $(1, ..., n)$, rovnako ako v príklade 5.6. Určte $E(X_n)$ a $D(X_n)$!

Riešenie: Nech $U_1, ..., U_n$ 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 5.5. Keďže $X = U_1 + ... + U_n$, dostávame z tvrdenia 5.2 $E(X) = \sum_{i=1}^n E(U_i) = 1$.

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

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

Z odvodených rovností a vety 5.4 dostávame

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

Z príkladu 5.6 (a vety 5.8) vieme, že limitné rozdelenie náhodných premenných X_n má strednú hodnotu aj disperziu 1. Práve sme sa však presvedčili, že $E(X_n) = D(X_n) = 1$ pre *každé* n.

Príklad 5.12. Nech $\sigma = (\sigma_1, ..., \sigma_n)$ je permutácia postupnosti $(1, ..., n)$. Index $i \in \{1, ..., n\}$ nazveme rekordom v permutácii σ, ak buď $i = 1$, alebo ak $i > 1$ a $\sigma_i > \sigma_j$ pre všetky $j \in \{1, ..., i-1\}$. Nech X_n označuje počet rekordov v dokonalej náhodnej permutácii σ postupnosti $(1, ..., n)$. (Pozn.: Kvôli jednoduchosti značíme v tomto príklade aj pevnú aj náhodnú permutáciu symbolom σ.) Nájdite $E(X_n)$ a $D(X_n)$.

Riešenie: Pre všetky $i = 1, ..., n$ označme ako U_i náhodnú premennú, ktorá nadobúda hodnotu 1 práve vtedy, keď v náhodnej permutácii čísiel $(1, ..., n)$ bude i rekordom a hodnotu 0 inak. Zjavne $X_n = U_1 + ... + U_n$. Tiež je zrejmé, že pre každé $i \in \{1, ..., n\}$ je udalosť [i je rekord] ekvivalentná udalosti $[\sigma_i = \max(\sigma_1, ..., \sigma_i)]$, ktorej pravdepodobnosť je $1/i$, pretože z dôvodov symetrie má každá z i hodnôt $\sigma_1, ..., \sigma_i$ rovnakú pravdepodobnosť, že bude najväčšia v množine $\{\sigma_1, ..., \sigma_i\}$. 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. Aby sme vypočítali varianciu, overte si najprv[2], že pre všetky $i, j \in \{1, ..., n\}; i < j$ sú udalosti [i je rekord] a [j je rekord] nezávislé a teda $E(U_i U_j) = P[U_i U_j = 1] = P[U_i = 1, U_j = 1] = P[U_i = 1]P[U_j = 1] = \frac{1}{ij}$. Máme preto

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

[2] ako cvičenie

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

Riešenie: Najprv si definujeme (nenáhodné) konštanty R_k a V_k pre $k = 0, ..., 16$ a to nasledovne: R_k označuje počet žien, ktoré sedia na stoličkách s číslami $k, k+1, ..., k+6$ modulo 17. V_k je 1 ak na stoličke s číslom k sedí žena a v opačnom prípade V_k definujeme ako 0.

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

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

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

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

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 $3, 4, ...$ Z toho ale plynie, že nutne musí existovať aspoň jedna sedmica susedných stoličiek, na ktorých sedia aspoň 3 ženy.

Poznámka 5.5. Predchádzajúci príklad ukazuje, že pravdepodobnostné metódy môžeme použiť aj na riešenie niektorých deterministických úloh (napríklad kombinatorických, alebo grafových). Súhrnne sa takéto techniky nazývajú "probabilistic method".

5.5 Cvičenia

Úloha 5.1. Ukážte: Nech B je borelovská množina reálnych čísiel a nech X je diskrétna náhodná premenná so (spočítateľným) oborom hodnôt H. Potom

$$P[X \in B] = \sum_{x \in B \cap H} P[X = x].$$

Úloha 5.2. Nech X je diskrétna náhodná premenná s oborom hodnôt $\{x_1, ..., x_n\}$, pričom $x_1 < ... < x_n$. Načrtnite distribučnú funkciu náhodnej premennej X.

Úloha 5.3. Hodíme n-krát mincou. Sériou nazveme postupnosť za sebou idúcich rovnakých výsledkov, pred a za ktorými je výsledok opačný, alebo žiadny (t.j. začiatok, alebo koniec). Napríklad pri $n = 8$ obsahuje výsledok "HZZZHHHH" tri série, výsledok "HZHZZHZH" 7 sérií. Nech X znamená výsledný počet sérií. Nájdite rozdelenie náhodnej premennej X, t.j. hodnoty $P[X = k]$ pre $k = 1, \ldots, 8$. Nájdite $E(X)$ a $D(X)$.

Úloha 5.4. Na začiatku hry SCRABBLE hráč náhodne volí 7 rôznych písmeniek spomedzi 100. Z týchto 100 písmeniek je 44 samohlások. Nech X znamená počet samohlások, ktoré si vyberieme. Nájdite rozdelenie (t.j. pravdepodobnosti $P[X = k]$ pre $k = 0, ..., 7$) a strednú hodnotu náhodnej premennej X.
Riešenie: $P[X = k] = \binom{44}{k}\binom{56}{7-k}/\binom{100}{7}$; $E(X) = \frac{7.44}{100}$.

Úloha 5.5. Vo vrecku máme všetkých 2^n podmnožín množiny $\{1, 2, ..., n\}$ (vrátane prázdnej množiny). Z vrecka najprv náhodne vyberieme množinu E, vrátime ju späť a po chvíľke vyberieme ďalšiu množinu F. Nech náhodná premenná X znamená počet prvkov zjednotenia E a F; Y nech znamená počet prvkov prieniku E a F. Nájdite rozdelenie a strednú hodnotu
a) náhodnej veličiny X;
b) náhodnej veličiny Y.

Úloha 5.6. Z interpretácie binomického rozdelenia (pozri poznámku 5.3) je zrejmé, že ak $X \sim Bin(n, p)$, tak X má rovnaké rozdelenie ako náhodná premenná $X' = U_1 + ... + U_n$, kde $U_i \sim Alt(p)$ nadobudne hodnotou 0 v prípade neúspechu v i-tom experimente a 1 v prípade úspechu v i-tom experimente. Pomocou tohoto vyjadrenia odvoďte $E(X)$ a $D(X)$.

Kapitola 6

Markovovské reťazce

Rozšírením pojmu náhodnej premennej je náhodný proces, ktorý je tvorený postupnosťou náhodných premenných $X_1, X_2, X_3, \ldots$. Majme postupnosť náhodných premenných $\{X_t\}_{t=1}^{\infty}$. Hodnoty náhodných premenných $X_t, t \geq 0$ budeme nazývať stavmi systému v čase t a množinu možných stavov systému označíme S. Ak množina všetkých stavov je konečná alebo spočítateľná, postupnosť $\{X_t\}_{t=1}^{\infty}$ sa nazýva reťazec. Ďalej budeme uvažovať len reťazce s konečnou množinou stavov $S = \{1, 2, \ldots, m\}$.

Definícia 6.1 (*Markovovský reťazec*)**.** Hovoríme, že reťazec $\{X_t\}_{t=1}^{\infty}$ s množinou stavov $S = \{1, 2, \ldots, m\}$ má Markovovu vlastnosť, ak pre všetky $n \in \mathbb{N}$ a pre ľubovoľné stavy $i_1, \ldots, i_n \in S$ platí

$$P(X_{n+1} = i_{n+1}|, X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1) = P(X_{n+1} = i_{n+1}|, X_n = i_n)$$

V tomto prípade sa $\{X_t\}_{t=1}^{\infty}$ nazýva markovovský reťazec.

Podmienené pravdepodobnosti

$$P(X_{n+1} = i|, X_n = j) =: p_{ij}(n),\ n \in \mathbb{N}$$

nazývame *pravdepodobnosti prechodu* zo stavu j v čase n do stavu i v čase $n + 1$.

Ak pravdepodobnosti prechodu nezávisia od času, t.j. ak pre ľubovoľné n je $P(X_{n+1} = i|, X_n = j) = P(X_n = i|, X_{n-1} = j) =: p_{ij}$, hovoríme o *stacionárnom* reťazci.

Pravdepodobnosti prechodu stacionárneho markovovského reťazca môžeme zapísať do tzv. *matice pravdepodobností prechodu*

$$\mathrm{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}$$

Z definície pravdepodobností prechodu vyplýva, že $\sum_j p_{ij} = 1$ pre všetky i.
Pravdepodobnosť prechodu zo stavu i do stavu j po t krokoch definujeme ako

$$p_{ij}(t) = P(X_{m+t} = j|X_t = i)$$

Dá sa ľahko ukázať, že

$$p_{ij}^{(}t) = \sum_u p_{iu} p_{uj}^{t-1}$$

Odtiaľ vyplýva, že ak označíme $\mathrm{P}(t)$ maticu pravdepodobností prechodu po t krokoch, potom

$$\mathrm{P}(t) = \mathrm{P}^t.$$

Tento výsledok zovšobecňuje nasledovná veta:

Veta 6.1. (*Chapmanova-Kolmogorovova rovnosť*): Označme $p_{ij}(t)$ prvky matice prechodu $P(t)$. Potom pre všetky $t, s \geq 0$ a pre každé dva stavy $i, j \in S$ platí: $p_{ij}(t+s) = \sum_{k \in S} p_{ik}(s) p_{kj}(t)$, a teda $P(t+s) = P(t)P(s)$.

Rozdelenie pravdepodobnosti $p(0) = (p_1(0), p_2(0), \ldots, p_m(0))$ také, že $p_i(0) = P[X_0 = i]$ pre $i = 1, \ldots, m$ sa nazýva počiatočné rozdelenie reťazca. Dá sa ukázať, že matice prechodu $P(t), t = 1, 2, 3, \ldots$ a počiatočné rozdelenie $p(0)$ jednoznačne určujú rozdelenia reťazca $\{X_t\}_{t=1}^{\infty}$.

Markovovský reťazec môžeme znázorniť pomocou orientovaného váženého grafu (V, E, w), kde množina vrcholov V reprezentuje stavy reťazca. Medzi vrcholmi i, j existuje hrana $(i, j) \in E$ práve vtedy, keď $p_{ij} > 0$, a potom $w(i, j) = p_{ij}$.

Príklad 6.1. Mucha sa pohybuje po priamke. V každej sekunde sa pohne s pravdepodobnosťou 0.3 o centimeter doľava, s pravdepodobnosťou 0.3 o centimeter doprava, a s pravdepodobnosťou 0.4 ostane na mieste. Na pozíciách 1 a m na ňu striehne pavúk: ak mucha pristane na jednom z týchto dvoch miest, pavúk ju zožerie.

Zostrojme markovovský reťazec pre túto situáciu za predpokladu, že mucha je na začiatku v niektorej z pozícií $2, \ldots, m$. Nech stavy $1, \ldots, m$ reprezentujú možné pozície muchy. Pravdepodobnosti prechodu potom sú

$$p_{11} = 1, \; p_{mm} = 1,$$

$$p_{ij} = \begin{cases} 0.4 & \text{ak } i = j \\ 0.3 & \text{ak } j = \pm i, \; i = 2, \ldots, m-1. \end{cases}$$

Matica pravdepodobností prechodu je

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0 & 0.3 & 0.4 & 0.3 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

6.1 Klasifikácia stavov markovovského reťazca

Definícia 6.2 (*Ireducibilný markovovský reťazec*)**.** Hovoríme, že stav j je *dosiahnuteľný* zo stavu i, ak pre systém v stave i je nenulová pravdepodobnosť, že niekedy dosiahne stav j. Markovovský reťazec, pre ktorý platí že ľubovoľný stav je dosiahnuteľný z ľubovoľného iného stavu v konečnom čase, sa nazýva *ireducibilný*. Formálne

$$\forall i, j \in S \; \exists m \in \mathbb{N}, m < \infty : P(X_{n+m} = i | X_n = j) > 0$$

Definícia 6.3. Množina stavov C sa nazýva uzavretá, ak pre ľubovoľné $i \in C$ a $j \notin C$ je $p_{ij} = 0$, teda ak zo stavov v C nie je dosiahnuteľný žiadny stav mimo C.

Definícia 6.4 (*Absorbujúci stav*)**.** Stav i sa nazýva absorbujúci, ak $p_{ii} = 1$.

Poznámka 6.1. Vidíme teda, že absorbujúci stav je jednoprvková uzavretá množina stavov. V ireducibilnom markovovskom reťazci je jedinou uzavretou množinou množina všetkých stavov.

Príklad 6.2.

Označme $f_{ij}^{(t)}$ pravdepodobnosť toho, že prvý prechod do stavu j, ak sme začali v stave i, nastáva po k krokoch v čase t, t.j.

$$f_{ij}^{(t)} = P(X_t = j \text{ a pre všetky } 1 \leq s \leq t-1 : X_s \neq j | X_0 = i)$$

Definícia 6.5 (*Trvalé a prechodné stavy*)**.** Stav i sa nazýva trvalý, ak $\sum_{t\geq 1} f_{ii}^{(t)} = 1$. Ak $\sum_{t\geq 1} f_{ii}^{(t)} < 1$, stav sa nazýva prechodný.

Ďalej označme h_{ij} očakávaný čas, za ktorý sa reťazec dostane zo stavu i do stavu j:

$$h_{ij} = \sum_{t\geq 1} t f_{ij}^{(t)}$$

Definícia 6.6. Trvalý stav sa nazýva kladný, ak $h_{ii} < \infty$.

Definícia 6.7 (*Periodické stavy*)**.** Hovoríme, že stav i má periódu $d > 1$, ak $p_{ii}^{(m)} = 0$ pre každé m, ktoré nie je násobkom čísla d a d je najväčšie číslo s touto vlastnosťou.

Definícia 6.8 (*Ergodické stavy*)**.** Neperiodický, kladný, trvalý stav sa nazýva ergodický. Markovovský reťazec sa nazýva ergodický, ak sú všetky jeho stavy ergodické.

Príklad 6.3.

6.2 Stacionárne rozdelenia

Označme $p_i(t)$ pravdepodobnosť, že reťazec je v čase t v stave i. Potom

$$p_i(t) = \sum_{j\geq 0} p_j(t-1)p_{ji},$$

a teda $\bar{p}(t) = \bar{p}(t-1)\mathrm{P}$, kde $\bar{p}(t) = (p_0(t), p_1(t), p_2(t), \ldots)$.

Definícia 6.9 (*Stacionárne rozdelenie*)**.** Stacionárne rozdelenie markovovského reťazca je také pravdepodobnostné rozdelenie π, pre ktoré platí $\pi = \pi\mathrm{P}$.

Ak sa reťazec dostane do stacionárneho rozdelenia, už v ňom zotrvá v každom ďalšom čase. Nasledovná veta popisuje reťazce, ktoré konvergujú ku stacionárnemu rozdeleniu.

Veta 6.2. Ľubovoľný konečný, ireducibilný a ergodický markovovský reťazec má nasledovné vlastnosti:
1. Existuje práve jedno stacionárne rozdelenie $\pi = (\pi_0, \ldots, \pi_n)$
2. Pre ľubovoľné i, j existuje limita $\lim_{t\to\infty} p_{ji}^{(t)}$ a táto limita nezávisí na j
3. $\pi_i = \lim_{t\to\infty} p_{ji}^{(t)} = h_{ii}^{-1}$

Poznámka 6.2. Požiadavka na neperiodickosť reťazca nie je nutnou k existencii stacionárneho rozdelenia. Každý konečný markovovský reťazec má stacionárne rozdelenie, ale v prípade periodického stavu i stacionárna pravdepodobnosť π_i nie je limitnou pravdepodobnosťou.

Stacionárne rozdelenie reťazca je riešením sústavy rovníc $\pi \mathrm{P} = \pi$.

Príklad 6.4. Uvažujme markovovský reťazec s dvoma stavmi 0 a 1 a s maticou prechodu

$$\mathrm{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

Tento reťazec môže predstavovať model prenosu bitov, pri ktorom stav 0 v čase t znamená, že bit bol prenesený bezchybne a stav 1 znamená, že bit bol preklopený na opačný. Stacionárne rozdelenie nájdeme riešením nasledovnej sústavy rovníc:

$$\pi_0(1-p) + \pi_1 q = \pi_0$$

$$\pi_0 p + \pi_1(1-q) = \pi_1$$

$$\pi_0 + \pi_1 = 0,$$

teda $\pi_0 = \frac{q}{p+q}$ a $\pi_1 = \frac{p}{p+q}$.

V niektorých prípadoch vieme stacionárne rozdelenie reťazca vypočítať použitím nasledovnej vety.

Veta 6.3. Uvažujme konečný, ireducibilný a ergodický markovovský reťazec s maticou pravdepodobností prechodu P. Ak existujú nezáporné čísla $\pi = (\pi_0, \ldots, \pi_n)$ také, že $\sum_{i=0}^{n} \pi_i = 1$ a ak pre ľubovoľné stavy i,j platí $\pi_i p_{ij} = \pi_j p_{ji}$, potom π je stacionárne rozdelenie reťazca s maticou P.

Dôkaz. j-ty prvok vektora $\pi \mathrm{P}$ sa dá napísať ako

$$\sum_{i=0}^{n} \pi_i p_{ij} = \sum_{i=0}^{n} \pi_j p_{ji} = \pi_j,$$

a teda $\pi = \pi \mathrm{P}$. Keďže $\sum_{i=0}^{n} \pi_i = 1$, z Vety 6.2 vyplýva, že π je jednoznačne určené stacionárne rozdelenie reťazca. □

Príklad 6.5. (Algoritmus Google Page Rank) Reprezentujme systém webových stránok ako graf s množinou vrcholov V, kde vrcholy predstavujú jednotlivé webové stránky. Ak zo stránky i vedie odkaz na stránku j, v grafe bude orientovaná hrana vedúca z i do j. Označme $N(i)$ množinu stránok, na ktoré odkazuje stránka i a definujme pravdepodobnosti prechodu pre reťazec s množinou stavov V nasledovne:

$$p_{ij} = \begin{cases} |N(i)|^{-1}, & \text{ak } j \in N(i), \\ |V|^{-1}, & \text{ak } N(i) = \emptyset, \\ 0 & \text{inak.} \end{cases}$$

Ak $\{X_n\}$ je markovovský reťazec s maticou prechodu $P = (p_{ij})$, nemusí byť nerozložiteľný a neperiodický. Preto uvažujeme reťazec s pravdepodobnosťami prechodu

$$r_{ij} = (1-\alpha)p_{ij} + \frac{\alpha}{|V|},$$

kde α je malé nenulové číslo, teda používateľ si väčšinou vyberá rovnomerne jeden z odkazov na stránke, ale niekedy vyberie rovnomerne náhodne zo všetkých možných webových stránok. Takýto reťazec má jednoznačne určené stacionárne rozdelenie π a vyhľadávač priradí stránke i vyššie poradie ako stránke j, ak $\pi_i > \pi_j$.

Príklad 6.6 (*Čakanie na obsluhu*)**.** Uvažujme model zákazníkov čakajúcich v rade, pričom v každom kroku nastane jedna z troch možností:
1. ak je v rade menej ako n zákazníkov, nový zákazník príde do rady s pravdepodobnosťou λ
2. ak rad nie je prázdny, bude s pravdepodobnosťou μ obslúžený prvý zákazník
3. s pravdepodobnosťou $1 - \lambda - \mu$ ostane rad nezmenený.

Ak označíme X_t počet zákazníkov v rade v čase t, potom X_t je markovovský reťazec s nasledovnými pravdepodobnosťami prechodu:

$$p_{i,i+1} = \lambda \text{ ak } i < n$$

$$p_{i,i-1} = \mu \text{ ak } i > 0$$

$$p_{ii} = \begin{cases} 1 - \lambda & \text{ak } i = 0 \\ 1 - \lambda - \mu & \text{ak } 1 \leq i \leq n - 1 \\ 1 - \mu & \text{ak } i = n \end{cases}$$

Tento markovovský reťazec je konečný, ireducibilný a neperiodický, teda existuje jeho stacionárne rozdelenie π, ktoré nájdeme riešením sústavy

$$\pi_0 = (1 - \lambda)\pi_0 + \mu\pi_1$$

$$\pi_i = \lambda\pi_{i-1} + (1 - \lambda - \mu)\pi_i + \mu\pi_{i+1}, \text{ pre } 1 \leq i \leq n - 1$$

$$pi_n = \lambda\pi_{n-1} + (1 - \mu)\pi_n$$

$$\sum_{i=0}^{n} \pi_i = 1$$

Ľahko overíme, že

$$\pi_i = \frac{(\lambda/\mu)^i}{\sum j = 0^n (\lambda/\mu)^j}.$$

6.3 Cvičenia

Úloha 6.1. Žaba skáče po priamke spôsobom, že v každej sekunde skočí s pravdepodobnosťou b napravo a s pravdepodobnosťou $1 - b$ naľavo. Začne v jednej z pozícií $1, \ldots, m$, ale z 0 môže skočiť len do 1 a z $m + 1$ len do m. Modelujte tento proces ako markovovský reťazec a nájdite jeho stacionárne rozdelenie.

Úloha 6.2. Roztržitý profesor má dva dáždniky, ktoré používa na cestu do práce a z práce. Ak prší a tam, kde sa nachádza, má k dispozícii dáždnik, vezme si ho. Ak neprší, dáždnik si nikdy nevezme. Predpokladajme, že v každom čase prší s pravdepodobnosťou p. Aká je stacionárna pravdepodobnosť, že profesor zmokne?

Úloha 6.3. Z dlhodobých pozorovaní počasia vieme, že ak v nejaký deň prší, na ďalší deň bude pršať s pravdepodobnosťou 0.4. Naopak, slnečný deň je nasledovaný ďalším slnečným dňom s pravdepodobnosťou 0.7. Urobte 'predpoveď počasia' na utorok a stredu, ak viete, že v pondelok pršalo.

Úloha 6.4. Na daný počítač pristupujú vzdialene navzájom nezávisle dvaja používatelia. Počas konkrétneho časového úseku sa momentálne prihlásený používateľ odhlási s pravdepodobnosťou 0.5 a neprihlásený používateľ sa prihlási s pravdepodobnosťou 0.2. Vypočítajte pravdepodobnosti prechodu reťazca X_t, ktorý reprezentuje počet prihlásených používateľov v čase t a nájdite jeho stacionárne rozdelenie.

Kapitola 7

Spojité náhodné premenné

7.1 Základné vlastnosti spojitých náhodných premenných

Definícia 7.1 (*Spojitá náhodná premenná*)**.** Nech X je náhodná premenná a nech existuje integrovateľná funkcia $f : \mathbb{R} \to [0, \infty)$ taká, že pre každé $x \in \mathbb{R}$ platí

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

Potom hovoríme, že X je spojitá náhodná premenná s hustotou f.

Veta 7.1. Nech X je spojitá náhodná premenná s hustotou f. Potom

$$\int_{-\infty}^{\infty} f(t)dt = 1$$

. Naviac, pre akékoľvek $a, b \in \mathbb{R}$, $a < b$, platí $P[X = a] = 0$, $P[a \leq X < b] = P[a < X \leq b] = P[a < X < b] = P[a \leq X \leq b] = \int_a^b f(t)dt$.

Dôkaz. Zrejme $\int_{-\infty}^{\infty} f(t)dt = \lim_{x\to\infty} \int_{-\infty}^{x} f(t)dt = \lim_{x\to\infty} F(x) = 1$ podľa vety 4.6. Rovnosť $P[X = a] = 0$ plynie zo základných vlastností integrálu. Ak je F distribučná funkcia náhodnej premennej X, tak platí $P[a \leq X < b] = P[X < b] - P[X < a] = F(b) - F(a) = \int_{-\infty}^{b} f(t)dt - \int_{-\infty}^{a} f(t)dt = \int_a^b f(t)dt$. Keďže $P[X = a] = P[X = b] = 0$, tak $P[a \leq X < b] = P[a < X < b] = P[a < X \leq b] = P[a \leq X \leq b]$. □

Vzťah medzi distribučnou funkciou a hustotou: Nech X je spojitá náhodná premenná s hustotou f, pričom funkcia f je spojitá s výnimkou maximálne konečného počtu bodov. Potom distribučná funkcia náhodnej premennej X je spojitá na celom $\mathbb{R}$ a spojite diferencovateľná maximálne s výnimkou konečného počtu bodov (tých, v ktorých je funkcia f nespojitá).

Naopak, 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 $f(x) = dF(x)/dx$ pre všetky $x \notin H$ je hustotou X.

Veta 7.2 (*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 $a < 0$. □

Príklad 7.1. Nech X je spojitá náhodná premenná so spojitou hustotou f_X. Ukážme, že náhodná premenná $Y = X^2$ je tiež spojitá a nájdime jej hustotu f_Y.

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

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

Vieme, že funkcia diskrétnej náhodnej premennej je opäť diskrétna náhodná premenná (Poznámka 5.1). Avšak ak je X spojitá náhodná premenná, 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!

7.2 Číselné charakteristiky spojitých náhodných premenných

Definícia 7.2. Nech X je spojitá náhodná premenná s hustotou f a nech existuje konečný integrál $\int_{-\infty}^{\infty} |x| f(x) dx$. Potom hovoríme, že náhodná premenná X má konečnú strednú hodnotu

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

Veta 7.3 (*Linearita strednej hodnoty*)**.** Nech X a Y sú spojité náhodné premenné, ktoré majú konečnú strednú hodnotu. Nech a, b sú akékoľvek reálne čísla také, že náhodná premenná $aX + bY$ je spojitá, alebo diskrétna. Potom $aX + bY$ má konečnú strednú hodnotu a platí

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

Veta 7.4 (*Veta o transformácii*)**.** Nech X je spojitá náhodná premenná s hustotou f. Nech g je po častiach spojitá funkcia taká, že $g(X)$ je spojitá, alebo diskrétna náhodná premenná.

Nech naviac existuje konečný integrál $\int_{-\infty}^{\infty} |g(x)| f(x)dx$. Potom náhodná premenná $g(X)$ má konečnú strednú hodnotu a platí:

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

Pre spojitú náhodnú premennú X definujeme disperziu rovnako ako pre diskrétnu náhodnú premennú (definícia 5.3). Platia pre ňu rovnaké vzťahy ako vo vete 5.4.

Ak spojitá náh. premenná X má hustotu f a konečnú strednú hodnotu, tak náhodná premenná $(X - E(X))^2$ je spojitá a platí

- $D(X) = \int_{-\infty}^{\infty}(x - E(X))^2 f(x)dx$,
- $D(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - (E(X))^2$,

ak sú príslušné integrály konečné.

7.3 Základné typy spojitých náhodných premenných

7.3.1 Rovnomerné rozdelenie

Definícia 7.3 (*Rovnomerné rozdelenie*)**.** Hovoríme, že náhodná premenná X má 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 $X \sim R(a, b)$. Ak $X \sim R(0, 1)$, tak hovoríme, že X má štandardizované rovnomerné rozdelenie.

Veta 7.5. Nech $X \sim R(a, b)$. Potom pre distribučnú funkciu F náhodnej premennej X platí

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

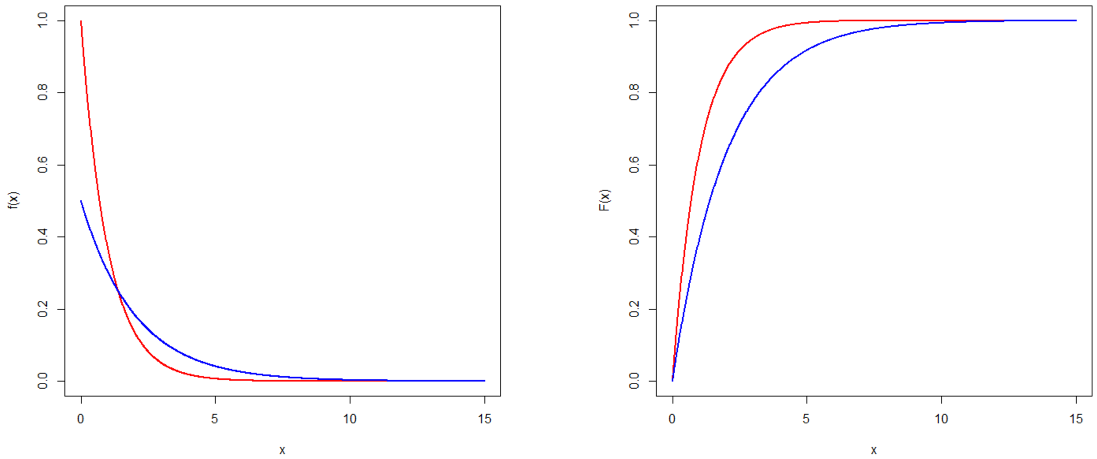
Naviac, ak $c, d \in \mathbb{R}$, $c > 0$, tak $cX + d \sim R(ca + d, cb + d)$. Špeciálne, $\frac{X-a}{b-a} \sim R(0, 1)$.

Dôkaz. Dôkaz je jednoduché cvičenie. □

Veta 7.6. Nech $X \sim R(a, b)$. Potom $E(X) = (a + b)/2$ a $D(X) = (b - a)^2/12$.

Dôkaz. Dôkaz je jednoduché cvičenie. □

Rovnomerné rozdelenie sa používa napríklad na generovanie realizácií náhodných premenných z iných typov rozdelení pomocou vhodných transformácií.



Obr. 7.1: Hustota pravdepodobnosti a distribučná funkcia náhodných premenných $X_1 \sim Exp(1)$ (červená) a $X_2 \sim Exp(2)$ (modrá).

7.3.2 Exponenciálne rozdelenie

Definícia 7.4 (*Exponenciálne rozdelenie*)**.** Hovoríme, že náhodná premenná X má exponenciálne rozdelenie s parametrom $\lambda > 0$, ak je X spojitá náhodná premenná s hustotou

$$f(x) = \frac{1}{\lambda} e^{-x/\lambda} \text{ pre } x \geq 0$$

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

Veta 7.7. Nech $X \sim Exp(\lambda)$. Potom pre distrib. funkciu F náhodnej premennej X platí

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

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

Dôkaz. Dôkaz je jednoduché cvičenie. □

Veta 7.8. Nech $X \sim Exp(\lambda)$. Potom $E(X) = \lambda$ a $D(X) = \lambda^2$.

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

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

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

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 $X_\lambda/\lambda \sim Exp(1)$, podľa vety 7.7. Z vlastností disperzie a strednej hodnoty, s použitím vzťahov $E(X_1) = 1$ a $D(X_1) = 1$ dostávame

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

□

Veta 7.9. Nech kladná náhodná premenná U má rozdelenie $R(0,1)$ a nech $\lambda > 0$. Potom $-\lambda \ln(U) \sim Exp(\lambda)$.

Dôkaz. Ukážeme, že ak $X = -\ln(U)$, tak $X \sim Exp(1)$. Nech F_X je distribučná funkcia X a nech F_U je distribučná funkcia U. Keďže $F_U(u) = u$ pre $u \in [0,1]$, máme pre každé $x > 0$: $F_X(x) = P[X < x] = P[-\ln(U) < x] = P[U > e^{-x}] = 1 - F_U(e^{-x}) = 1 - e^{-x}$. Použijúc vetu 7.7, môžeme dôkaz uzavrieť. □

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, v neurofyziológii na modelovanie doby medzi príchodmi impulzov na neurónovú bunku a podobne.

7.3.3 Paretovo rozdelenie

Definícia 7.5 (*Paretovo rozdelenie*)**.** 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 7.10. Nech $X \sim Par(\alpha, k)$. Potom pre distribučnú funkciu F náhodnej premennej X platí

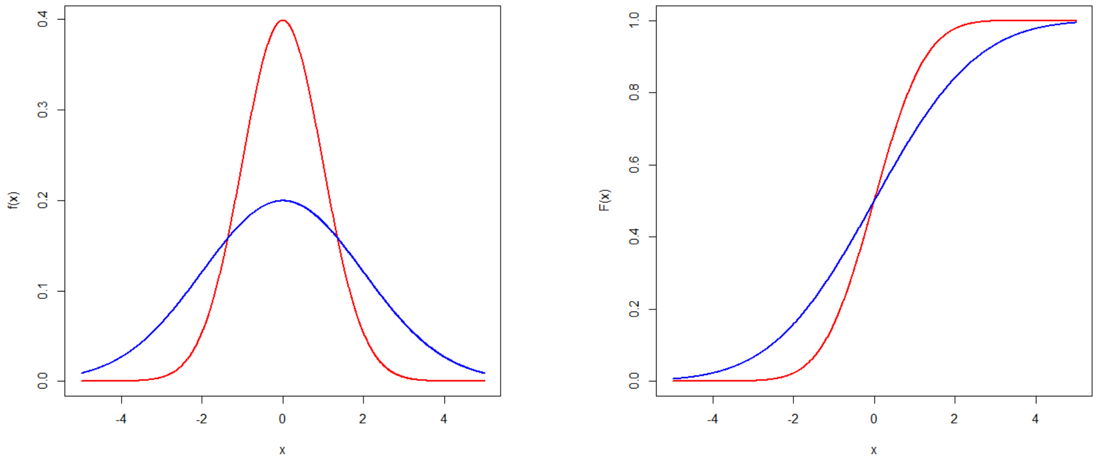
$$F(x) = \begin{cases} 0 \text{ pre } x \leq k \\ 1 - (k/x)^\alpha \text{ pre } x > k. \end{cases}$$

Dôkaz. Dôkaz je jednoduché cvičenie. □

Veta 7.11. Nech $X \sim Par(\alpha, k)$, pričom $\alpha > 1$. Potom $E(X) = \frac{\alpha \cdot k}{\alpha - 1}$. (Pre $\alpha \in (0,1]$ má náhodná premenná s rozdelením $Par(\alpha, k)$ nekonečnú strednú hodnotu.)

Dôkaz. Dôkaz je jednoduché cvičenie. □

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



Obr. 7.2: Hustota pravdepodobnosti a distribučná funkcia náhodných premenných $X_1 \sim N(0,1)$ (červená) a $X_2 \sim N(0,4)$ (modrá).

7.3.4 Normálne rozdelenie

Definícia 7.6 (*Normálne rozdelenie*)**.** Nech $\mu \in \mathbb{R}$ a $\sigma > 0$. Hovoríme, že náhodná premenná X má normálne rozdelenie 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 (alebo tiež normalizované) normálne rozdelenie. Distribučnú funkciu náhodnej premennej X s rozdelením $N(0,1)$ budeme označovať symbolom Φ.

Poznámka 7.1. Pre hustotu f normálneho rozdelenia je možné ukázať platnosť $\int_{-\infty}^{\infty} f(x)dx = 1$ pomocou takzvaného Poissonovho integrálu

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

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

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

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

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

Pomocou predchádzajúcich vzťahov určíme strednú hodnotu a disperziu normálneho rozdelenia (pozri vetu 7.13).

Normálne rozdelenie má ústredné postavenie v teoretickej aj aplikovanej teórii pravdepodobnosti a matematickej štatistike (budeme sa s ním často stretávať až do konca prednášky).

Distribučnú funkciu normálneho rozdelenia s akýmikoľvek parametrami vieme jednoducho vyjadriť pomocou Φ - distribučnej funkcie rozdelenia $N(0,1)$, 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 získať z tabuliek, alebo vypočítať numericky, napríklad ako čiastočné súčty vhodných nekonečných radov. Viac podrobností možno nájsť na stránkach www.mathworld.com pod heslom "Normal Distribution Function".

Veta 7.12 (*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

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

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 7.2 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

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

Tým sme ukázali aj poslednú časť tvrdenia. □

Príklad 7.2. Z dlhodobých štatistík o konkrétnom užívateľovi vieme, že pri vzdialenom prihlásení sa na server má počet sektorov na disku, na ktoré zapisuje alebo ktoré prečíta, približne normálne rozdelenie so strednou hodnotou 500 a disperziou 15^2. Monitorovací program nám oznámil, že pri poslednom prihlásení došlo k zápisu alebo čítaniu na 535 sektoroch. Aká je pravdepodobnosť, že došlo k zneužitiu tohto konta?

Hľadajme pravdepodobnosť $P[X \geq 535]$, kde náhodná premenná X označuje počet sektorov na disku.

$$P[X \geq 535] = P\left[U \geq \frac{535 - 500}{15}\right] = 1 - \Phi(35/15) \approx 0.01,$$

kde U je náhodná premenná so štandardizovaným normálnym rozdelením.

Veta 7.13. 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 7.1 máme

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

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

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 7.12. 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 $D(X)$. □

Príklad 7.3. (*Detekcia signálu.*) Prenášame binárny signál S, ktorý môže byť buď -1 alebo 1. Komunikačný kanál tento signál zašumí aditívnym normálnym šumom X s nulovou strednou hodnotou a disperziou σ^2. Prijímač usúdi, že bol prenesený signál $+1$ (alebo -1), ak sa doňho dostane hodnota väčšia ako 0 (alebo menšia ako 0). Aká je pravdepodobnosť chyby?

Riešenie: K chybe dochádza, akje $S = -1$ a šum je aspoň 1, vtedy $S + X = X - 1 \geq 0$ alebo ak $S = 1$ a šum je menší ako -1, vtedy $S + X < 0$. Pravdepodobnosť chyby v prvom prípade je

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P\left(\frac{X-\mu}{\sigma} < \frac{1-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{1-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{1}{\sigma}\right).$$

V druhom prípade je pravdepodobnosť vďaka symetrii rovnaká.

Praktický význam normálneho rozdelenia ukazujú takzvané centrálne limitné vety, ktoré hovoria, že za istých podmienok postupnosti súčtov náhodných premenných konvergujú k normálnemu rozdeleniu.

Veta 7.14 (*Integrálna Moivrova-Laplaceova veta*)**.** Nech $p \in (0,1)$ a nech $X_1, X_2, \ldots$ sú náhodné premenné, $X_n \sim Bin(n,p)$ pre každé $n \in \mathbb{N}$. Nech $x \in \mathbb{R}$. Potom

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

Predchádzajúca veta sa používa na aproximáciu pravdepodobností týkajúcich sa binomického rozdelenia v tom zmysle, že rozdelenie $Bin(n,p)$ je možné aproximovať rozdelením $N(np, np(1-p))$ v prípade, že n je "veľké" a p nie je "blízko" nuly ani jednotky.

7.4 Cvičenia

Úloha 7.1. Náhodná premenná X má distribučnú funkciu tvaru

$$F(x) = \begin{cases} 0 \text{ ak } x \leq 0 \\ cx^2 \text{ ak } 0 < x \leq 1 \\ 1 \text{ ak } x > 1 \end{cases}$$

Určte:
a) konštantu c tak, aby $F(x)$ bola distribučnou funkciou,
b) hustotu pravdepodobnosti náhodnej premennej X,
c) pravdepodobnosť $P[X \in (1/3, 1/2)]$.

Úloha 7.2. Pre distribučnú funkciu F náhodnej premennej X platí: $F(x) = x^2/2 + x + 1/2$ ak $x \in (-1, 0)$ a $F(x) = -x^2/2 + x + 1/2$ ak $x \in (0, 1)$. Nájdite hustotu, strednú hodnotu a disperziu náhodnej premennej X.

Úloha 7.3. Náhodná premenná má hustotu

$$f(x) = \begin{cases} cx \text{ ak } 0 \leq x < 1 \\ 2 - x \text{ ak } 1 \leq x \leq 2 \\ 0 \text{ ak } x \notin [0, 2] \end{cases}$$

Určte:
a) konštantu c tak, aby $f(x)$ bola hustotou pravdepodobnosti,
b) distribučnú funkciu $F(x)$,
c) pravdepodobnosť $P[1 < X < 2]$.

Úloha 7.4. Nech $a, b \in \mathbb{R}$, pričom $a > 1, b > 0$. Náhodná premenná X má distribučnú funkciu $F(x) = 0$ pre $x \leq b$ a $F(x) = 1 - (b/x)^a$ pre $x > b$. Určte hodnoty a, b. Nájdite hustotu a strednú hodnotu náhodnej premennej X.

Úloha 7.5. Nech $\mu, \Delta \in \mathbb{R}$, kde $\Delta > 0$ a nech spojitá náhodná premenná X má hustotu $f(x) = (\Delta - |x - \mu|)/\Delta^2$ pre $x \in [\mu - \Delta, \mu + \Delta]$ a $f(x) = 0$ inak. Nájdite distribučnú funkciu, strednú hodnotu a disperziu X.

Úloha 7.6. Bod B volíme rovnomerne náhodne na kružnici so stredom v $(0, 0)$ a polomerom 1. (T.j. B volíme na hranici kruhu so stredom v $(0, 0)$ a polomerom 1 a to tak, že uhol určený bodmi $(1, 0), (0, 0), B$ má rovnomerné rozdelenie na intervale $[0, 2\pi]$.) Nech S znamená obsah a trojuholníka určeného bodmi $(-1, 0), (1, 0), B$. Nájdite distribučnú funkciu, hustotu a strednú hodnotu náhodnej premennej S.

Úloha 7.7. Bod B volíme rovnomerne náhodne na jednotkovej kružnici v rovine. Nech L znamená vzdialenosť bodu B od bodu $(1, 0)$. Nájdite distribučnú funkciu, hustotu a strednú hodnotu náhodnej premennej L.

Úloha 7.8. Nech $X \sim Exp(\lambda)$, kde $\lambda > 0$. **a)** Ukážte, že náhodná premenná $\lfloor X \rfloor$ (dolná celá časť X) má geometrické rozdelenie. **b)**[dú] Nájdite distribučnú funkciu, hustotu a strednú hodnotu náhodnej premennej $Y = e^{-X}$.

Úloha 7.9. Nech náhodná premenná X má normalizované normálne rozdelenie. Nájdite hustotu náhodných premenných
a) e^X;
b) X^2;
c) $|X|$;
d) $\sqrt{|X|}$.

Úloha 7.10. Vo štvorci náhodne rovnomerne nezávisle vygenerujeme 4000 bodov. Pomocou aproximácie binomického rozdelenia normálnym odhadnite pravdepodobnosť, že menej ako 3000 z týchto bodov padne do kruhu, ktorý je vpísaný danému štvorcu. Výsledok zapíšte pomocou Φ - distribučnej funkcie rozdelenia $N(0, 1)$.

Úloha 7.11. Ukážte, že exponenciálne rozdelenie "nemá pamäť", t.j. dokážte, že ak $X \sim Exp(\lambda)$ a $0 < a < b$, potom $P([X > b]|[X > a]) = P[X > b - a]$.

Úloha 7.12. Rozmer vyrábanej súčiastky má približne normálne rozdelenie so strednou hodnotou $\mu = 1000$ mm. Výrobok považujeme za dobrý, ak sa jeho rozmer nelíši od 1000 mm o viac ako 1 mm. Aká musí byť disperzia rozmeru súčiastok, aby pomer nepodarkov neprekračoval 1 percento? Výsledok vyjadrite pomocou kvantilovej funkcie Φ^{-1} rozdelenia $N(0, 1)$.

Kapitola 8

Náhodné vektory

V pravdepodobnostných modeloch nás často zaujíma viac náhodných premenných súčasne. Tieto náhodné premenné pochádzajú z rovnakého pravdepodobnostného priestoru a to, aké hodnoty nadobúda jedna z nich, je často ovplyvnené hodnotami ostatných. Aby sme mohli s takýmito skupinami náhodných premenných pohodlne pracovať, zavedieme pojem náhodného vektora.

8.1 Všeobecné náhodné vektory

Definícia 8.1 (*Náhodný vektor*). Nech $X_1, \ldots, X_m$ sú náhodné premenné na spoločnom pravdepodobnostnom priestore. Potom $\mathbf{X} = (X_1, \ldots, X_m)^T$ nazývame m-rozmerným náhodným vektorom.

Príklad 8.1. Hádžeme dvomi kockami. Definujme nasledujúce náhodné premenné:
U: maximum padnutých čísel
V: súčet padnutých čísel
Rozdelenie náhodného vektora $(U, V)'$ potom môžeme popísať nasledovnou tabuľkou:

U/V	2	3	4	5	6	7	8	9	10	11	12
1	1/36										
2		2/36	1/36								
3			2/36	2/36	1/36						
4				2/36	2/36	2/36	1/36				
5					2/36	2/36	2/36	2/36	1/36		
6						2/36	2/36	2/36	2/36	2/36	1/36

Riadkové, resp. stĺpcové súčty v tejto tabuľke zodpovedajú marginálnym pravdepodobnostiam pre náhodnú premennú U, resp. V.

Definícia 8.2 (*Distribučná funkcia náhodného vektora*). Distribučnou funkciou náhodného vektora $\mathbf{X} = (X_1, \ldots, X_m)^T$ nazývame funkciu $F : \mathbb{R}^m \to \mathbb{R}$, ktorá je definovaná

$$F(\mathbf{x}) = P[X_1 < x_1, \ldots, X_m < x_m] \text{ pre všetky } \mathbf{x} = (x_1, \ldots, x_m)^T \in \mathbb{R}^m$$

Je možné ukázať, že distribučná funkcia F m-rozmerného náhodného vektora $\mathbf{X}$ jednoznačne určuje pravdepodobnosť $P[\mathbf{X} \in B]$ pre akúkoľvek m-rozmernú borelovskú množinu B.

Veta 8.1 (*Základné vlastnosti distribučnej funkcie náhodného vektora*)**.** Nech F je distribučná funkcia m-rozmerného náhodného vektora $\mathbf{X}$. Potom platí:

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. $\lim_{x_1,\ldots,x_m \to \infty} F(x_1, \ldots, x_m) = 1$
4. $\lim_{x_i \to -\infty} F(x_1, \ldots, x_m) = 0$ pre každé $i = 1, ..., m$ a $x_1, ..., x_{i-1}, x_{i+1}, ..., x_m \in \mathbb{R}$

Dôkaz. Tieto vlastnosti sa dokážu analogicky ako základné vlastnosti distribučnej funkcie jednorozmernej náhodnej premennej. □

Veta 8.2 (Existencia náh. vektora s danou distribučnou funkciou)**.** Nech funkcia $F : \mathbb{R}^m \to \mathbb{R}$ spĺňa všetky štyri vlastnosti z vety 8.1. Potom existuje pravdepodobnostný priestor a náhodné premenné $X_1, \ldots X_m$ na tomto pravdepodobnostnom priestore také, že F je distribučnou funkciou náhodného vektora $\mathbf{X} = (X_1, \ldots, X_m)^T$.

Distribučná funkcia náhodného vektora $(X_1, \ldots, X_m)^T$ jednoznačne určuje distribučné funkcie náhodných premenných $X_1, \ldots, X_m$ (marginálne distribučné funkcie), ako ukazuje nasledujúca veta. Distribučné funkcie náhodných premenných $X_1, \ldots, X_m$ však nemusia jednoznačne určovať distribučnú funkciu náhodného vektora $(X_1, \ldots, X_m)^T$.

Veta 8.3 (*Marginálna distribučná funkcia*)**.** Nech $F_\mathbf{X}$ je distribučná funkcia náhodného vektora $\mathbf{X} = (X_1, \ldots, X_m)^T$ a nech $1 \leq i_1 < ... < i_k \leq m$. Potom pre distribučnú funkciu $F_\mathbf{Y}$ náhodného vektora $\mathbf{Y} = (X_{i_1}, \ldots, X_{i_k})^T$ platí

$$F_\mathbf{Y}(x_{i_1}, ..., x_{i_k}) = \lim F_\mathbf{X}(x_1, ..., x_m)$$

kde limitu berieme pre $x_i \to \infty$ pre všetky $i \notin \{i_1, ..., i_k\}$.

Dôkaz. Dôkaz je jednoduché cvičenie. □

8.1.1 Diskrétne náhodné vektory

Definícia 8.3 (*Diskrétny náhodný vektor*)**.** Náhodný vektor $\mathbf{X} = (X_1, \ldots, X_m)^T$ nazývame diskrétny, ak je spočítateľný jeho obor hodnôt

$$\mathbf{X}(\Omega) = \{(x_1, ..., x_m) \in \mathbb{R}^m : x_i \in X_i(\Omega), \forall i = 1, ..., m\}.$$

Definícia 8.4 (*Stredná hodnota diskrétneho náhodného vektora*)**.** Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je diskrétny náhodný vektor a nech existuje stredná hodnota každej (diskrétnej) náhodnej premennej X_i pre $i = 1, ..., m$. Potom povieme, že náhodný vektor $\mathbf{X}$ má strednú hodnotu

$$E(\mathbf{X}) = (E(X_1), \ldots, E(X_m))^T$$

Veta 8.4 (*Linearita strednej hodnoty náhodného vektora*)**.** Nech $\mathbf{X}_1, \mathbf{X}_2$ sú m-rozmerné diskrétne náhodné vektory a nech existuje konečná stredná hodnota (každej komponenty) náhodných vektorov $\mathbf{X}_1, \mathbf{X}_2$. Nech A, B sú matice typu $k \times m$. Potom $\mathbf{Y} = A\mathbf{X}_1 + B\mathbf{X}_2$ je k-rozmerný diskrétny náhodný vektor, ktorý má konečnú strednú hodnotu danú vzťahom

$$E(\mathbf{Y}) = AE(\mathbf{X_1}) + BE(\mathbf{X_2})$$

Dôkaz. Nech $\mathbf{X}_i = (X_{i,1}, ..., X_{i,m})^T$ pre $i = 1, 2$. Potom j-tu komponentu Y_j vektora $\mathbf{Y}$ môžeme pre každé $j \in \{1, ..., k\}$ zapísať nasledovne:

$$Y_j = A_{j,1}X_{1,1} + ... + A_{j,m}X_{1,m} + B_{j,1}X_{2,1} + ... + B_{j,m}X_{2,m}$$

Z linearity strednej hodnoty diskrétnych náhodných premenných (Veta 5.2) dostávame

$$E(Y_j) = A_{j,1}E(X_{1,1}) + ... + A_{j,m}E(X_{1,m}) + B_{j,1}E(X_{2,1}) + ... + B_{j,m}E(X_{2,m})$$

Maticový zápis platnosti predchádzajúcej rovnosti pre každé $j = 1, ..., k$ je práve $E(\mathbf{Y}) = AE(\mathbf{X_1}) + BE(\mathbf{X_2})$. □

Veta 8.5 (*Stredná hodnota funkcie diskrétneho náhodného vektora*)**.** Nech m-rozmerný 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 spojitá funkcia. Potom stredná hodnota diskrétnej náhodnej premennej $g(\mathbf{X})$ existuje a je konečná vtedy a len vtedy, keď existuje a je konečná suma $\sum_{i\in I} g(\mathbf{x}_i)P[\mathbf{X} = \mathbf{x}_i]$. V takomto prípade platí:

$$E(g(\mathbf{X})) = \sum_{i\in I} g(\mathbf{x}_i)P[\mathbf{X} = \mathbf{x}_i]$$

8.1.2 Spojité náhodné vektory

Definícia 8.5. Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je náhodný vektor s distribučnou funkciou $F_\mathbf{X}$. Nech existuje integrovateľná funkcia $f : \mathbb{R}^m \to [0, \infty)$ taká, že

$$F_\mathbf{X}(x_1, ..., x_m) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} f(t_1, ..., t_m)dt_1...dt_m \text{ pre všetky } x_1, ..., x_m \in \mathbb{R}$$

Potom hovoríme, že $\mathbf{X}$ je spojitý náhodný vektor a f je hustota $\mathbf{X}$.

Veta 8.6. Nech $\mathbf{X}$ je m-rozmerný spojitý náhodný vektor s hustotou f. Nech $B \subseteq \mathbb{R}^m$ je borelovská množina. Potom

$$P[\mathbf{X} \in B] = \int \cdots \int_B f(t_1, ..., t_m)dt_1...dt_m$$

Príklad 8.2. Nech náhodný vektor $\mathbf{X} = (X_1, X_2)$ je daný hustotou

$$f(x_1, x_2) = \begin{cases} 1 & \text{ak } (x_1, x_2) \in [0, 1] \times [0, 1] \\ 0 & \text{inak.} \end{cases}$$

Nájdite pravdepodobnosť, že $\mathbf{X}$ nadobúda hodnoty zo štvorca $[0, 1/2]^2$.

$$P[\mathbf{X} \in [0, 1/2] \times [0, 1/2]] = \int_0^{1/2} \int_0^{1/2} f(x_1, x_2)dx_1dx_2 = \int_0^{1/2} \int_0^{1/2} 1dx_1dx_2 = \frac{1}{2}\frac{1}{2} = \frac{1}{4}.$$

Veta 8.7. Nech $\mathbf{X} = (X_1, ..., X_m)^T$ je m-rozmerný spojitý náhodný vektor s hustotou f. Potom je pre každé $i = 1, ..., m$ náhodná premenná X_i spojitá s hustotou danou v bode $x_i \in \mathbb{R}$ predpisom

$$f_i(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, ..., x_m)dx_1 \ldots dx_{i-1}dx_{i+1} \ldots dx_m$$

Definícia 8.6 (*Stredná hodnota spojitého náhodného vektora*)**.** Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je spojitý náhodný vektor a nech existuje stredná hodnota každej (spojitej) náhodnej premennej X_i pre $i = 1, ..., m$. Potom povieme, že náhodný vektor $\mathbf{X}$ má strednú hodnotu

$$E(\mathbf{X}) = (E(X_1), \ldots, E(X_m))^T$$

Veta 8.8 (*Linearita strednej hodnoty spojitého náhodného vektora*)**.** Nech $\mathbf{X}_1, \mathbf{X}_2$ sú m-rozmerné spojité náhodné vektory a nech existuje konečná stredná hodnota (každej komponenty) náhodných vektorov $\mathbf{X}_1, \mathbf{X}_2$. Nech A, B sú matice typu $k \times m$, pričom $\mathbf{Y} = A\mathbf{X}_1 + B\mathbf{X}_2$ je spojitý alebo diskrétny náhodný vektor. Potom $\mathbf{Y}$ má konečnú strednú hodnotu danú vzťahom

$$E(\mathbf{Y}) = AE(\mathbf{X_1}) + BE(\mathbf{X_2})$$

Veta 8.9 (*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 spojitá funkcia taká, že $g(\mathbf{X})$ je spojitá náh. premenná. Potom stredná hodnota náhodnej premennej $g(\mathbf{X})$ existuje a je konečná vtedy a len vtedy, keď existuje a je konečný integrál $\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, ..., x_m) f(x_1, ..., x_m) dx_1 ... dx_m$. V takomto prípade platí:

$$E(g(\mathbf{X})) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, ..., x_m) f(x_1, ..., x_m) dx_1 ... dx_m$$

8.2 Nezávislosť náhodných premenných

Definícia 8.7. Budeme hovoriť, že náhodné premenné $X_1, \ldots, X_n$ sú združene nezávislé, ak pre všetky $x_1, \ldots, x_n \in \mathbb{R}$ platí

$$P[X_1 < x_1, \ldots, X_n < x_n] = P[X_1 < x_1] \ldots P[X_n < x_n]$$

alebo, ekvivalentne, ak

$$F_{\mathbf{X}}(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n)$$

kde $F_{\mathbf{X}}$ je distribučná funkcia náhodného vektora $\mathbf{X} = (X_1, \ldots, X_n)^T$ a F_{X_i} je distribučná funkcia náhodnej premennej X_i pre $i = 1, \ldots, n$.

Veta 8.10. Náhodné premenné $X_1, \ldots, X_n$ sú združene nezávislé vtedy a len vtedy, keď pre akýkoľvek výber množín $B_1, \ldots, B_n \in \mathcal{B}$ platí

$$P[X_1 \in B_1, \ldots, X_n \in B_n] = P[X_1 \in B_1] \ldots P[X_n \in B_n]$$

Veta 8.11 (*Nezávislosť diskrétnych náhodných premenných*)**.** Nech $X_1, \ldots, X_n$ sú diskrétne náhodné premenné. Potom $X_1, \ldots, X_n$ sú nezávislé vtedy a len vtedy, keď pre akékoľvek reálne čísla $x_1, \ldots, x_n$ platí

$$P[X_1 = x_1, \ldots, X_n = x_n] = P[X_1 = x_1] \ldots P[X_n = x_n]$$

Dôkaz. Dôkaz je jednoduché cvičenie (pozri riešenie úlohy 8.6, časť b). □

Veta 8.12 (*Nezávislosť spojitých náh. premenných*). Nech náhodné premenné $X_1, \ldots, X_n$ sú spojité s hustotami $f_1, ..., f_n$. Ak sú $X_1, \ldots, X_n$ nezávislé, tak funkcia $f : \mathbb{R}^n \to [0, \infty)$ definovaná

$$f(x_1, ..., x_n) = f_1(x_1)...f_n(x_n) \tag{8.1}$$

je hustotou náhodného vektora $\mathbf{X} = (X_1, ..., X_n)^T$. Naopak, ak pre nejakú hustotu f náhodného vektora $\mathbf{X}$ a hustoty $f_1, ..., f_n$ náhodných premenných $X_1, \ldots, X_n$ platí (8.1), potom sú náhodné premenné $X_1, \ldots, X_n$ nezávislé.

Príklad 8.3. Komunikačný kanál sa skladá zo série uzlov, pričom vždy i-ty uzol predáva jednobitovú 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 tohoto uzla bude opačný bit ako na jeho vstupe. Naviac, chyby na jednotlivých uzloch sa vyskytujú navzájom nezávisle. Napíšte vzorec udávajúci pravdepodobnosť, že bit na vstupe prvého uzla bude rovnaký ako bit na výstupe n-tého uzla. (Porovnajte s príkladom 3.6.)

Riešenie.[1] Pre $i = 1, ..., n$ 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 tiež, 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 dostávame $P[X = 1] = E(X)/2 + 1/2$. S využitím nezávislosti náhodných premenných $X_1, ..., X_n$ a vety 8.13 teda dostávame

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

Veta 8.13 (*Stredná hodnota súčinu nezávislých náh. premenných*). Nech X_1 a X_2 sú nezávislé, obe diskrétne alebo obe spojité náhodné premenné s konečnou strednou hodnotou. Potom X_1X_2 je diskrétna, resp. spojitá náhodná premenná s konečnou strednou hodnotou a platí

$$E(X_1X_2) = E(X_1)E(X_2)$$

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 $Z = X_1X_2$ je zrejme diskrétna, lebo jej obor hodnôt $Z(\Omega)$ je spočítateľná množina súčinov $z = x_1x_2$, kde $x_1 \in X_1(\Omega)$ a $x_2 \in X_2(\Omega)$. Uvedomíme 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 $x_1 \in X_1(\Omega)$ a $x_2 \in X_2(\Omega)$, pre ktoré platí $x_1x_2 = z$. Naviac, 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 úlohu 8.6, časť b. Máme teda:

$$E(X_1X_2) = E(Z) = \sum_{z \in Z(\Omega)} zP[Z = z] = \sum_{x_1 \in A_1} \sum_{x_2 \in A_2} x_1x_2P[X_1 = x_1, X_2 = x_2] =$$

$$\sum_{x_1 \in A_1} \sum_{x_2 \in A_2} x_1x_2P[X_1 = x_1]P[X_2 = x_2] = \sum_{x_1 \in A_1} x_1P[X_1 = x_1] \sum_{x_2 \in A_2} x_2P[X_2 = x_2] =$$

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

□

[1] Toto elegantné riešenie navrhol študent FMFI UK Stanislav Takáč.

Z predchádzajúcej vety okamžite plynie, že pre n-ticu $X_1, ..., X_n$ nezávislých náhodných premenných, ktoré majú konečnú strednú hodnotu, platí $E(\prod_{i=1}^n X_i) = \prod_{i=1}^n E(X_i)$. Toto tvrdenie má značný teoretický význam, avšak je ho možné využiť aj na riešenie príkladov, ako ukážeme v ďalšom.

Veta 8.14 (*Disperzia súčtu nezávislých náhodných premenných*)**.** Nech X_1 a X_2 sú nezávislé, obe diskrétne alebo obe spojité náhodné premenné s konečnou disperziou. Potom $X_1 + X_2$ je diskrétna, resp. spojitá náhodná premenná s konečnou disperziou a platí

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

Dôkaz. Z definície disperzie, linearity strednej hodnoty a vety 8.13 máme: $D(X_1 + X_2) = E((X_1+X_2)^2) - (E(X_1+X_2))^2 = E(X_1^2) + 2E(X_1X_2) + E(X_2^2) - (E(X_1))^2 - 2(E(X_1)E(X_2)) - (E(X_2))^2 = (E(X_1^2) - (E(X_1))^2) + (E(X_2^2) - (E(X_2))^2) + 2(E(X_1X_2) - E(X_1)E(X_2)) = D(X_1) + D(X_2)$. □

Príklad 8.4. Čas výpočtu T znáhodneného algoritmu má distribučnú funkciu $F(t) = t^\delta$, kde $t \in (0, 1)$ a δ je kladná konštanta. Nech Y je (náhodná) doba výpočtu tohoto algoritmu na dvoch nezávislých procesoroch, t.j. $Y = \min\{T_1, T_2\}$, kde T_1, T_2 sú nezávislé náhodné premenné s distribučnou funkciou F. Pre náhodnú premennú Y určte distribučnú funkciu, hustotu a strednú hodnotu.

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

$$F_Y(y) = P[\min\{T_1, T_2\} < y] = P([T_1 < y] \cup [T_2 < y]) = 1 - P[T_1 \geq y, T_2 \geq y] =$$

$$1 - P[T_1 \geq y]P[T_2 \geq y] = 1 - (1 - F(y))^2 = 1 - (1 - t^\delta)^2.$$

Hustotu dostaneme derivovaním: $f_Y(y) = 2\delta(1 - t^\delta)t^{\delta-1}$, opäť pre $y \in (0, 1)$; inde je samozrejme $f_Y(y) = 0$. Nakoniec dostávame

$$E(Y) = \int_0^1 t f_Y(t)dt = 2\delta \int_0^1 t^\delta - t^{2\delta}dt = 2\delta((\delta + 1)^{-1} - (2\delta + 1)^{-1}).$$

Jeden z možných zápisov je

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

Definícia 8.8 (*Kovariancia náhodných premenných*)**.** Nech X_1, X_2 sú náhodné premenné s konečnou disperziou, obe diskrétne alebo obe spojité. Kovarianciou náhodných premenných X_1, X_2 nazývame hodnotu

$$cov(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2)$$

Ak X_1, X_2 sú premenné s konečnou disperziou, tak je možné dokázať, že aj náhodná premenná X_1X_2 má konečnú strednú hodnotu.

Výpočet kovariancie v tvare sumy: Nech 2-rozmerný náhodný vektor $\mathbf{X} = (X_1, X_2)^T$ nadobúda hodnoty $(\mathbf{x}_i)_{i \in I}$ s nenulovou pravdepodobnosťou. Potom podľa vety o strednej hodnote funkcie náhodného vektora môžeme stredné hodnoty a kovarianciu X_1, X_2 vypočítať podľa vzťahov

- $E(X_1) = \sum_{i \in I} (\mathbf{x}_i)_1 P[\mathbf{X} = \mathbf{x}_i]$
- $E(X_2) = \sum_{i \in I} (\mathbf{x}_i)_2 P[\mathbf{X} = \mathbf{x}_i]$
- $E(X_1 X_2) = \sum_{i \in I} (\mathbf{x}_i)_1 (\mathbf{x}_i)_2 P[\mathbf{X} = \mathbf{x}_i]$

kde $(\mathbf{x}_i)_1$ a $(\mathbf{x}_i)_2$ sú komponenty vektora $\mathbf{x}_i$.

Výpočet kovariancie v tvare integrálu: Majme spojitý 2-rozmerný náhodný vektor $\mathbf{X} = (X_1, X_2)^T$ s hustotou f. Potom podľa vety o strednej hodnote funkcie spojitého náhodného vektora môžeme stredné hodnoty a kovarianciu X_1, X_2 vypočítať podľa vzťahov

- $E(X_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_i f(x_1, x_2) dx_1 dx_2$ pre $i = 1, 2$
- $E(X_1 X_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2) dx_1 dx_2$

Veta 8.15 (*Základné vlastnosti kovariancie*)**.** Nech X_1, X_2 sú náhodné premenné s konečnou disperziou, obe diskrétne alebo obe spojité. Potom platí

1. $cov(X_1, X_2) = cov(X_2, X_1)$
2. $cov(X_i, X_i) = D(X_i)$ pre $i = 1, 2$
3. $cov(X_1, X_2) = E((X_1 - E(X_1))(X_2 - E(X_2)))$
4. $cov(aX_1 + b, cX_2 + d) = a.c.cov(X_1, X_2)$ pre každé $a, b, c, d \in \mathbb{R}$
5. $(cov(X_1, X_2))^2 \leq D(X_1) D(X_2)$
6. $cov(X_1, X_2) = 0$ ak sú X_1, X_2 nezávislé

Dôkaz. Prvé štyri rovnosti plynú priamo z definícií a z linearity strednej hodnoty. Dokážeme tvrdenie 5). Pripomeňme, že podľa Cauchyho nerovnosti platí pre akékoľvek reálne čísla a_i, b_i, $i = 1, ..., n$: $\left(\sum_{i=1}^n a_i b_i\right)^2 \leq \sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2$.

Nech náhodný vektor $\mathbf{X} = (X_1, X_2)^T$ nadobúda vektory $(\mathbf{x}_i)_{i \in I}$ s nenulovou pravdepodobnosťou. Použijúc vetu o strednej hodnote transformácie náh. vektora a Cauchyho nerovnosť máme

$$(cov(X_1, X_2))^2 = \left(\sum_{i \in I} ((\mathbf{x}_i)_1 - E(X_1))((\mathbf{x}_i)_2 - E(X_2)) P[\mathbf{X} = \mathbf{x}_i] \right)^2 \leq$$
$$\sum_{i \in I} ((\mathbf{x}_i)_1 - E(X_1))^2 P[\mathbf{X} = \mathbf{x}_i] \sum_{i \in I} ((\mathbf{x}_i)_2 - E(X_2))^2 P[\mathbf{X} = \mathbf{x}_i] = D(X_1) D(X_2)$$

Posledné tvrdenie plynie priamo z definície kovariancie a vety 8.13. □

Definícia 8.9 (*Kovariančná matica náhodného vektora*)**.** Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je náhodný vektor a nech existuje konečná disperzia každej náhodnej premennej X_i pre $i = 1, ..., m$. Potom povieme, že náhodný vektor $\mathbf{X}$ má konečnú kovariančnú maticu $Cov(\mathbf{X})$, pričom i, j-tu komponentu tejto matice definujeme

$$(Cov(\mathbf{X}))_{i,j} = cov(X_i, X_j)$$

Veta 8.16 (*Kovariančná matica lineárnej transformácie náhodného vektora*). Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je diskrétny 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 diskrétny náhodný vektor s konečnou kovariančnou maticou a platí

$$Cov(\mathbf{Y}) = ACov(\mathbf{X})A^T$$

Dôkaz. To, že $\mathbf{Y} = A\mathbf{X} + b$ je diskrétny náhodný vektor, plynie z úlohy 8.3. Naviac, j-tu komponentu Y_j vektora $\mathbf{Y}$ môžeme pre každé $j \in \{1, ..., k\}$ zapísať nasledovne:

$$Y_j = A_{j,1}X_1 + ... + A_{j,m}X_m + b_j$$

kde $A_{j,r}$ je prvok matice A v j-tom riadku a r-tom stĺpci a b_j je j-ta komponenta vektora b. Z definície kovariancie, kovariančnej matice a z linearity strednej hodnoty ľahko overíme, že pre každé $i, j \in \{1, ..., k\}$ platí:

$(Cov(\mathbf{Y}))_{i,j} = cov(Y_i, Y_j) = cov(A_{i,1}X_1 + ... + A_{i,m}X_m + b_i, A_{j,1}X_1 + ... + A_{j,m}X_m + b_j) = \sum_{r,s=1}^{m} A_{i,r}A_{j,s}cov(X_r, X_s) = (ACov(\mathbf{X})A^T)_{i,j}$

Tým je dôkaz dokončený. □

Veta 8.17 (*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 disperziami 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 kovar. matice a prvých dvoch vzťahov vo vete 8.15. Pozitívnu semidefinitnosť matice $Cov(\mathbf{X})$ dokážeme takto: Nech $a \in \mathbb{R}^m$ je ľubovoľný vektor. Z nezápornosti disperzie náhodnej premennej $a^T\mathbf{X}$ a z vety o kovariančnej matici lineárnej transformácie náhodného vektora máme $0 \leq D(a^T\mathbf{X}) = Cov(a^T\mathbf{X}) = a^TCov(\mathbf{X})a$. Tým je pozitívna semidefinitnosť $Cov(\mathbf{X})$ dokázaná. □

Definícia 8.10 (*Korelačný koeficient*). Nech X_1 a X_2 sú obe diskrétne, alebo obe spojité náhodné premenné s konečnou disperziou. Korelačným koeficientom (alebo stručne koreláciou) náhodných premenných X_1, X_2 nazývame hodnotu

$$\rho(X_1, X_2) = \frac{cov(X_1, X_2)}{\sqrt{D(X_1)}\sqrt{D(X_2)}}$$

Ak $\rho(X_1, X_2) = 0$, potom hovoríme, že X_1, X_2 sú nekorelované.

Veta 8.18 (*Základné vlastnosti korelačného koeficientu*). Nech X_1 a X_2 sú obe diskrétne, alebo obe spojité náhodné premenné s konečnou a nenulovou disperziou. Potom platí:

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. $\rho(X_1, X_2) \in [-1, 1]$
4. $\rho(X_1, X_2) = 0$ ak sú X_1, X_2 nezávislé

Dôkaz. Dôkaz všetkých častí tvrdenia plynie priamo z definície korelácie a základných vlastností kovariancie. □

8.3 Základné typy rozdelení náhodných vektorov

8.3.1 Multinomické rozdelenie

Príklad 8.5. Hádžeme n guličiek do m krabíc, pričom pravdepodobnosť, že nejaká gulička padne do j-tej krabice je p_j, $j = 1, \ldots, m$ a platí $p_j > 0$, $\sum_{j=1}^m p_j = 1$. Zostrojme náhodný vektor $\mathbf{X} = (X_1, \ldots, X_m)$ tak, že náhodná premenná X_j bude reprezentovať výsledný počet guličiek v j-tej krabici. Aké rozdelenie pravdepodobnosti má náhodný vektor X?

Definícia 8.11 (*Multinomické rozdelenie*)**.** 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 nezáporné reálne čísla také, že $\sum_{j=1}^m p_j = 1$. Nech platí

$$P[X_1 = k_1, \ldots, X_m = k_m] = \frac{n!}{k_1! \ldots k_m!} p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m}$$

pre všetky tie $k_1, \ldots, k_m \in \{0, 1, \ldots, n\}$, ktoré spĺňajú $\sum_{j=1}^m k_j = n$. Potom hovoríme, že náhodný vektor $\mathbf{X}$ má multinomické rozdelenie s parametrami $n, p_1, ..., p_m$. Túto skutočnosť značíme $\mathbf{X} \sim Mult(n, p_1, \ldots, p_m)$.

Veta 8.19 (*Vzťah multinomického a binomického rozdelenia*)**.** Nech $\mathbf{X} = (X_1, \ldots, X_m)^T \sim Mult(n, p_1, \ldots, p_m)$. Potom $X_j \sim Bin(n, p_j)$ pre každé $j = 1, \ldots, m$.

Dôkaz. Nech $\mathbf{X} = (X_1, \ldots, X_m)^T \sim Mult(n, p_1, \ldots, p_m)$. Dokážeme, že $X_1 \sim Bin(n, p_1)$. Pre každé $k_1 \in \{0, 1, ..., n\}$ platí

$$P[X_1 = k_1] = \sum P[X_1 = k_1, \ldots, X_m = k_m] = \sum \frac{n!}{k_1! \ldots k_m!} p_1^{k_1} p_2^{k_2} \ldots p_m^{k_m}$$

pričom všade v tomto dôkaze sumujeme cez tie $k_2, \ldots, k_m \in \{0, 1, \ldots, n\}$, ktoré spĺňajú $\sum_{j=2}^m k_j = n - k_1$. Z multinomického rozvoja

$$(1 - p_1)^{n-k_1} = (p_2 + ... + p_m)^{n-k_1} = \sum \frac{(n-k_1)!}{k_2! \ldots k_m!} p_2^{k_2} \ldots p_m^{k_m}$$

dostávame

$$P[X_1 = k_1] = \frac{n! p_1^{k_1}}{k_1!(n-k_1)!} \sum \frac{(n-k_1)!}{k_2! \ldots k_m!} p_2^{k_2} \ldots p_m^{k_m} = \binom{n}{k_1} p_1^{k_1} (1-p_1)^{n-k_1}$$

Tvrdenie $X_j \sim Bin(n, p_j)$ môžeme dokázať analogicky pre každé $j = 1, \ldots, m$. □

Predpokladajme, že robíme n nezávislých experimentov, pričom výsledkom každého z týchto experimentov je práve jedna z m rôznych možností a to zakaždým s pravdepodobnosťami $p_1, ..., p_m$. Ak pre $j = 1, ..., m$ bude X_j znamenať počet tých experimentov, ktoré skončili j-tym výsledkom, potom $\mathbf{X} = (X_1, \ldots, X_m)^T \sim Mult(n, p_1, \ldots, p_m)$.

Všimnite si, že ak označíme ako V_{ij} náhodnú premennú, ktorá nadobúda 1, ak skončil i-ty experiment výsledkom j a 0 inak, potom pre každý výber indexov $j_1, j_2, \ldots, j_n \in \{1, \ldots, m\}$ sú náhodné premenné $V_{1j_1}, \ldots, V_{nj_n}$ nezávislé. Naviac, pre každé $j = 1, ..., m$ platí $X_j = V_{1j} + ... + V_{nj}$.

Veta 8.20 (*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

$$(Cov(\mathbf{X}))_{j,j} = np_j(1-p_j) \text{ pre } j = 1, ..., m$$

$$(Cov(\mathbf{X}))_{j,k} = -np_jp_k \text{ pre } j, k \in \{1, ..., m\}, j \neq k.$$

Dôkaz. Vlastnosť $E(X_j) = np_j$ a $(Cov(\mathbf{X}))_{j,j} = D(X_j) = -np_j(1-p_j)$ pre $j = 1, ..., m$ plynie z viet 8.19 a 5.6.

Dokážeme, že $(Cov(\mathbf{X}))_{j,k} = -np_jp_k$ pre pevné $j \neq k$. Nech $X_j = V_{1j} + ... + V_{nj}$ pre každé $j = 1, ..., m$ ako v predchádzajúcej poznámke. Potom

$$(Cov(\mathbf{X}))_{j,k} = cov(X_j, X_k) = cov(V_{1j} + ... + V_{nj}, V_{1k} + ... + V_{nk}) =$$

$$\sum_{i,l=1}^{n} cov(V_{ij}, V_{lk}) = \sum_{i \neq l} cov(V_{ij}, V_{lk}) + \sum_{i=1}^{n} cov(V_{ij}, V_{ik}) = \sum_{i=1}^{n} cov(V_{ij}, V_{ik}),$$

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

$$cov(V_{ij}, V_{ik}) = E(V_{ij}V_{ik}) - E(V_{ij})E(V_{ik}) = -p_jp_k.$$

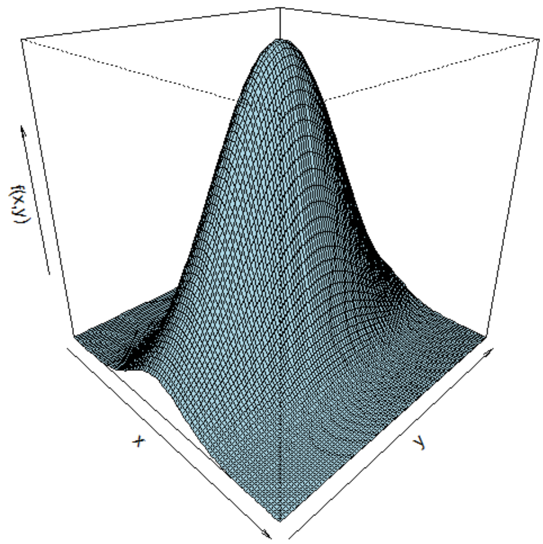
Tým je dôkaz ukončený. □

8.3.2 Mnohorozmerné normálne rozdelenie

Definícia 8.12 (*Regulárne mnohorozmerné normálne rozdelenie*)**.** Nech $\mu \in \mathbb{R}^m$ a nech Σ je pozitívne definitná (t.j. regulárna pozitívne semidefinitná) matica typu $m \times m$. Nech $\mathbf{X} = (X_1, \ldots, X_m)^T$ je spojitý náhodný vektor s hustotou

$$f(\mathbf{x}) = (2\pi)^{-m/2}(\det \Sigma)^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)\right)$$

Potom hovoríme, že $\mathbf{X}$ má regulárne m-rozmerné normálne rozdelenie s parametrami μ a Σ. Túto skutočnosť značíme $\mathbf{X} \sim N_m(\mu, \Sigma)$.



Veta 8.21. Nech $\mathbf{X} \sim N_m(\mu, \Sigma)$, kde $\mu \in \mathbb{R}^m$ a Σ je pozitívne definitná matica typu $m \times m$. Nech $b \in \mathbb{R}^k$, 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 volíme $k = 1$, $A = (0, ..., 0, 1, 0, ..., 0)$ (jednotka je na i-tej pozícii) a $b = 0$ tak dostávame: $X_i \sim N(\mu_i, (\Sigma)_{ii})$.

Veta 8.22. Nech $\mathbf{X} = (X_1, ..., X_m)^T \sim N_m(\mu, \Sigma)$, 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 8.21 vidíme, že pre hustotu f vektora $\mathbf{X}$ a hustoty f_i komponent X_i náhodného vektora $\mathbf{X}$ platí:

$$f(x_1, ..., x_m) = \frac{1}{(2\pi)^{m/2}} \exp(-(x_1^2 + ... + x_m^2)/2) = \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi}} \exp(-x_i^2/2) = \prod_{i=1}^{m} f_i(x_i)$$

Na základe vety 8.12 usudzujeme, že $X_1, ..., X_m$ sú nezávislé a preto $cov(X_i, X_j) = 0$ ak $i \neq j$. 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 $Cov(\mathbf{X}) = I$. 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 8.21 má náhodný vektor $\Sigma^{-1/2}(\mathbf{X} - \mu)$ rozdelenie $N(0, I)$ a podľa prvej časti dôkazu, podľa vety o linearite strednej hodnoty a vety o kovariančnej matici lineárnej transformácie máme rovnosti:

$$0 = E(\Sigma^{-1/2}(\mathbf{X} - \mu)) = \Sigma^{-1/2}(E(\mathbf{X}) - \mu),$$

$$I = Cov(\Sigma^{-1/2}(\mathbf{X} - \mu)) = \Sigma^{-1/2} Cov(\mathbf{X}) \Sigma^{-1/2},$$

z ktorých dostávame tvrdenie vety. $\square$

Veta 8.23. Platí ekvivalencia nasledovných dvoch výrokov:
A) $X_1, ..., X_m$ sú nezávislé náhodné premenné, pričom $X_i \sim N(\mu_i, \sigma_i^2)$;
B) Náhodný vektor $\mathbf{X} = (X_1, ..., X_m)^T$ má rozdelenie $N_m(\mu, \Sigma)$, kde $\mu = (\mu_1, ..., \mu_m)^T$ a Σ je diagonálna matica s prvkami $\sigma_1^2, ..., \sigma_m^2$ na diagonále.

Dôkaz. Veta sa dá jednoducho ukázať pomocou viet 8.21 a 8.12. □

8.4 Cvičenia

Úloha 8.1. Náhodný vektor $(X, Y)'$ má združenú hustotu tvaru

$$f(x, y) = \begin{cases} k(x + y) \text{ ak } (x, y) \in S \\ 0 \text{ ak } (x, y) \notin S, \end{cases}$$

kde S je štvorec ohraničený priamkami $x = 0$, $x = 3$, $y = 0$ a $y = 3$. Určte koeficient k tak, aby $f(x, y)$ bola hustota. Vypočítajte marginálne hustoty náhodných premenných X a Y a pravdepodobnosť, že náhodný vektor $(X, Y)'$ padne do štvorca S_1 ohraničeného priamkami $x = 1$, $x = 2$, $y = 1$ a $y = 2$,

Úloha 8.2. Združená hustota náhodného vektora $(X, Y)'$ má tvar

$$f(x, y) = \begin{cases} 8xy \text{ ak } 0 < y \leq x < 1 \\ 0 \text{ ak } inak, \end{cases}$$

Vypočítajte marginálne hustoty náhodných premenných X a Y.

Úloha 8.3. Presvedčte sa o platnosti nasledovných tvrdení:
a) $\mathbf{X} = (X_1, \dots, X_m)^T$ je diskrétny náhodný vektor vtedy a len vtedy, keď samotné $X_1, \dots, X_m$ sú diskrétne náhodné premenné;
b) Ak je $\mathbf{X} = (X_1, \dots, X_m)^T$ diskrétny náhodný vektor a $g : \mathbb{R}^m \to \mathbb{R}^k$ je akákoľvek funkcia, tak $g(\mathbf{X})$ je tiež diskrétny náhodný vektor. Špeciálne $X_i + X_j$ ako aj $X_i X_j$ pre $i, j \in \{1, ..., m\}$ sú diskrétne náhodné premenné.

Úloha 8.4. Nech $m = 2$, $\mathbf{X} = (X, Y)^T$ a $B = [a, b) \times [c, d)$. Ukážte, že platí

$$P[\mathbf{X} \in B] = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

Úloha 8.5. Nech $X_1, \dots, X_n$ sú združene nezávislé náhodné premenné (v zmysle horeuvedenej definície). Ukážte, že potom sú združene nezávislé aj náhodné premenné $X_{i_1}, X_{i_2}, \dots, X_{i_k}$ pre akékoľvek indexy $1 \leq i_1 < i_2 < ... < i_k \leq n$.

Úloha 8.6. Dokážte implikáciu "⇒" vo vete 8.10 pre tieto špeciálne prípady:
a) $n = 2$, $B_1 = [a, b)$, $B_2 = [c, d)$ a X_1, X_2 sú akékoľvek nezávislé náhodné premenné. (Môžete použiť tvrdenie z úlohy 8.4.)
b) $n = 2$, $B_1 = \{x_1\}$, $B_2 = \{x_2\}$ a X_1, X_2 sú diskrétne nezávislé náhodné premenné.

Úloha 8.7. Z interpretácie binomického rozdelenia plynie, že náhodná premenná X_n s rozdelením $Bin(n, p)$ je súčet n nezávislých náhodných premenných s rozdelením $Alt(p)$. Pomocou tohto faktu ukážte, že Veta 7.14 je dôsledkom vety 9.4.

Úloha 8.8. Budeme počítať súčet nezávislých náhodných čísiel $X_1, ..., X_{1000}$, ktoré pred spočítaním zaokrúhlime na celé čísla $Y_1, ..., Y_{1000}$. Diferencie $D_i = X_i - Y_i$, pre všetky $i = 1, ..., 1000$, sú nezávislé náhodné premenné s rovnakým rozdelením so strednou hodnotou 0 a disperziou 1/12. Preto podľa centrálnej limitnej vety možno celkovú chybu súčtu $Err = \sum_{i=1}^{1000} D_i$, ktorej sa dopustíme zaokrúhľovaním, aproximovať normálnym rozdelením. Pomocou distribučnej funkcie Φ rozdelenia $N(0,1)$ zapíšte pravdepodobnosť udalosti $[|Err| < 10]$, t.j., že absolútna hodnota rozdielu súčtu zaokrúhlených čísiel sa od súčtu pôvodných čísiel nebude líšiť o viac ako 10.

Úloha 8.9. Nech $\mathbf{X} \sim N_2(\mu, \Sigma)$, kde

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \text{ a } \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

pričom Σ je regulárna. Presvedčte sa, že potom hustota X je daná vzťahom:

$$f_{\mathbf{X}}(x_1, x_2) = c^{-1} \exp\left(-\frac{1}{1-\rho^2} \left[\frac{(x_1-\mu_1)^2}{2\sigma_1^2} - \frac{\rho(x_1-\mu_1)(x_1-\mu_1)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{2\sigma_2^2} \right] \right),$$

kde $c = 2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}$.

Kapitola 9

Zákony veľkých čísiel a centrálna limitná veta

Veta 9.1 (*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 $(z_i)_{i\in I}$. Ďalej nech $J = \{i \in I : z_i \geq c\}$. Máme

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

, z čoho dostávame priamo tvrdenie vety.
Ak Z je spojitá náhodná premenná s hustotou f, máme

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

, z čoho dostávame priamo tvrdenie vety. □

Príklad 9.1. Pošta denne spracuje 10000 listov. Aká je pravdepodobnosť, že zajtra budú musieť spracovať aspoň 15000 listov?

Nech náhodná premenná X označuje počet listov, ktoré sa na pošte spracujú. Potom zo zadania príkladu vieme, že $E(X) = 10000$. Použitím Markovovej vety dostávame

$$P[X \geq 15000] \leq \frac{10000}{15000} = \frac{2}{3}.$$

Veta 9.2 (*Čebyševova nerovnosť*)**.** Nech X je diskrétna alebo spojitá náhodná premenná s konečnou disperziou (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. Veta je dôsledkom Markovovej nerovnosti 9.1 ak v nej zvolíme $Z = (X - E(X))^2$ a $c = a^2$. □

Príklad 9.2. Pomocou Čebyševovej nerovnosti vypočítajme horné ohraničenie pre počet listov, ktoré bude musieť pošta spracovať, ak navyše vieme, že $D(X) = 10000$:

$$P[X \geq 15000] = P[X - 10000 \geq 5000] \leq P[(X - 10000 \geq 5000) \cup (X - 10000 \leq -5000)] =$$

$$= P[|X - 10000| \geq 5000] \leq \frac{10000}{5000^2} = \frac{1}{2500}.$$

Príklad 9.3. Nech p je percento voličov podporujúcich nejakého kandidáta vo voľbách. Spýtame sa n náhodne (rovnomerne nezávisle z celej populácie) vybraných respondentov, či by daného kandidáta volili a zaznamenáme K_n kladných odpovedí. Nakoľko presne vieme odhadnúť p? Odpoveď každého respondenta môžeme chápať ako náhodnú premennú $X_i \sim Alt(p)$. Z Čebyševovej nerovnosti máme

$$P[|K_n - p| \geq \varepsilon] \leq \frac{p(1-p)}{n\varepsilon} \leq \frac{1}{4n\varepsilon^2}.$$

Všeobecnejšou formou Čebyševovej nerovnosti je tzv. slabý zákon veľkých čísel.

Veta 9.3 (*Slabý zákon veľkých čísiel*)**.** 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 $\epsilon > 0$ platí:

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

Dôkaz. Z predpokladov tvrdenia, vety 5.4 (resp. ekvivalentu tejto vety pre spojité premenné) a vety 8.14 dostávame

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

. Dôkaz môžeme ukončiť použitím Čebyševovej nerovnosti (Veta 9.2). □

Veta 9.4 (*Centrálna limitná veta*)**.** Nech $X_1, X_2, \ldots$ sú nezávislé náhodné premenné s rovnakým rozdelením (t.j. s rovnakou distribučnou funkciou), konečnou strednou hodnotou a nenulovou a konečnou disperziou. 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 platí

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

Dôkaz. Dôkaz tejto vety presahuje rámec prednášky. □

Všimnite si, že predchádzajúcu vetu môžeme formulovať aj nasledovne: Nech $X_1, X_2, X_3, ...$ sú nezávislé náhodné premenné s rovnakým rozdelením, konečnou strednou hodnotou μ a konečnou disperziou σ^2, kde $\sigma > 0$. Potom

$$\lim_{n\to\infty} P\left[\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} < x\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \text{ pre všetky } x \in \mathbb{R}$$

Príklad 9.4. Na pamäťovej karte máme voľných 330 MB. Aká je pravdepodobnosť, že sa na ňu zmestí 300 fotografií s priemernou veľkosťou 1 MB a disperziou 0.5^2?

$$P[S \leq 330] = P\left[U \leq \frac{330 - 300 \cdot 1}{0.5\sqrt{300}}\right] = \Phi(3.46) \approx 0.9997.$$

Príklad 9.5. Vieme, že počet chýb na 1000 riadkov kódu má Poissonovo rozdelenie so strednou hodnotou 5. Aká je pravdepodobnosť, že v tridsiatich 1000-riadkových programoch bude viac ako 170 chýb?

Označme náhodnou premennou X_i počet chýb v i-tom programe, teda $X_i \sim Po(5)$. Potom, ak S je celkový počet chýb v 30 programoch, S má približne normálne rozdelenie so strednou hodnotou a varianciou $30 \cdot 5$. Máme teda

$$P[S \geq 170] = P\left[U \geq \frac{170 - 150}{\sqrt{150}}\right] = 1 - \Phi(2.45) \approx 0.05.$$

Príklad 9.6. Do lietadla naložíme 300 kusov batožiny, pričom hmotnosť jedného kusu batožiny je náhodná premenná s rovnomerným rozdelením na intervale od 10 do 100 kg. Aká je pravdepodobnosť, že celková hmotnosť batožiny v lietadle presiahne 17000 kg? Približnú odpoveď na túto otázku dostaneme pomocou centrálnej limitnej vety: chceme vypočítať $P[S_{300} > 17000]$, kde S_{300} je súčet hmotností 300 kusov batožiny. Stredná hodnota a disperzia hmotnosti jedného kusu batožiny je

$$\mu = \frac{10 + 100}{2} = 55, \; \sigma^2 = (100 - 10)^2/12 = 675$$

Potom máme

$$P[S_{300} \leq 17000] = \Phi\left(\frac{17000 - 300 \cdot 55}{\sqrt{300 \cdot 675}}\right) \approx 0.8667,$$

a teda

$$P[S_{300} > 17000] \approx 0.1333.$$

9.1 Cvičenia

Úloha 9.1. Letecká spoločnosť predala 410 leteniek na let, v ktorom je 400 miest. Vypočítajte pravdepodobnosť, že každý cestujúci bude mať kde sedieť, ak vieme, že cestujúci so zakúpenou letenkou príde na check-in s pravdepodobnosťou 0.96.

Úloha 9.2. Dve agentúry na prieskum verejnej mienky vykonávajú prieskum volebných preferencií kandidáta, každá na vzorke 1000 voličov. Určte hornú hranicu pravdepodobnosti, že odhad preferencií u oboch agentúr sa bude odlišovať' o viac ako 5%.

Úloha 9.3. Vo voľbách sú dvaja kandidáti, pričom prieskum verejnej mienky na vzorke 1500 voličov ukázal, že 52% voličov podporuje prvého kandidáta a 48% voličov podporuje druhého kandidáta. Aká je pravdepodobnost', že prieskum správne odhadol víťaza volieb?

Úloha 9.4. Uvažujme prieskum preferencií medzi dvomi kandidátmi, pričom podľa prieskumu verejnej mienky vykonaného na n respondentoch prvého kandidáta podporuje 100p% voličov a druhého kandidáta podporuje 100(1-p)% voličov. Aké veľké musí byť n, aby sme so spoľahlivosťou aspoň 95% vedeli určiť skutočný výsledok volieb s chybou menšou ako 3%?

Úloha 9.5. Poisťovňa poistila 1000 osôb. Pravdepodobnosť úmrtia v priebehu jedného roka je pre každého z nich 0,005. Zistite, s akou pravdepodobnosťou bude poiťovňa zisková, ak poistné je 22 eur a poistná suma je 7000 eur.

Úloha 9.6. Pravdepodobnosť narodenia chlapca je 0,515. Aká je pravdepodobnosť, že medzi 10000 novorodencami bude viac dievčat ako chlapcov?

Kapitola 10

Generovanie náhodných premenných a vektorov

10.1 Generovanie realizácií náhodných premenných

V tejto kapitole predpokladáme, že vieme sekvenčne generovať nezávislé realizácie z rovnomerného rozdelenia na intervale $(0, 1)$. V praxi obvykle postačuje postupnosť ”pseudonáhodných” čísiel získaných napríklad pomocou špeciálnych (deterministických) kongruenčných generátorov. Rozsiahlou teóriou kongruenčných generátorov sa však nebudeme hlbšie zaoberať (niektoré základné pojmy boli spomenuté na prednáške).

Definícia 10.1 (*Kvantilová funkcia*)**.** Nech $F : \mathbb{R} \to [0, 1]$ je distribučná funkcia. Kvantilovou funkciou distribučnej funkcie F nazývame funkciu $G : (0, 1) \to \mathbb{R}$ definovanú

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

Ak je distribučná funkcia F spojitá a rastúca na celom $\mathbb{R}$, potom príslušná kvantilová funkcia G je inverzná funkcia k F.

Veta 10.1 (*Metóda inverznej transformácie*)**.** Nech F je distribučná funkcia a nech G je kvantilová funkcia funkcie F. Nech náhodná premenná U má rozdelenie $R(0, 1)$. Potom náhodná premenná $G(U)$ má distribučnú funkciu F.

Dôkaz. Dôkaz pre všeobecnú distribučnú funkciu je technicky zdĺhavý. Uvedieme si len dôkaz pre prípad, že funkcia F je spojitá a rastúca na celom $\mathbb{R}$. V takomto prípade $G = F^{-1}$ (inverzná funkcia). Nech $U \sim R(0, 1)$. Potom pre distribučnú funkciu F_Y náhodnej premennej $Y = G(U)$ a pre každé $y \in \mathbb{R}$ platí:

$$F_Y(y) = P[Y < y] = P[G(U) < y] = P[U < F(y)] = F_U(F(y)) = F(y)$$

pretože distribučná funkcia F_U premennej U spĺňa $F_U(u) = u$ pre každé $u \in [0, 1]$. □

Príklad 10.1 (*Generovanie rozdelenia $Exp(\lambda)$*)**.** Distribučná funkcia rozdelenia $Exp(\lambda)$ je $F(x) = 0$ pre $x \leq 0$ a $F(x) = 1 - e^{-x/\lambda}$ pre $x > 0$. Kvantilová funkcia je $G(y) = -\lambda ln(1 - y)$ pre $y \in (0, 1)$, čiže podľa predchádzajúcej vety platí: Ak $U \sim R(0, 1)$, tak $-\lambda ln(1 - U) \sim Exp(\lambda)$. Keďže aj $1 - U \sim R(0, 1)$, tak dostávame tvrdenie vety 7.9: $-\lambda ln(U) \sim Exp(\lambda)$

Príklad 10.2. Nech $m \in \mathbb{N}$. Uvažujme distribučnú funkciu $F(x) = 0$ pre $x \leq 0$, $F(x) = x^m$ pre $x \in (0,1)$ a $F(x) = 1$ pre $x \geq 1$. (Táto distribučná funkcia charakterizuje rozdelenie euklidovej normy m-rozmerného náhodného vektora s rovnomerným rozdelením vo vnútri m-rozmernej gule so stredom v počiatku súradnicovej sústavy a polomerom 1.) Keďže kvantilová funkcia je $G(y) = y^{1/m}$ pre $y \in (0,1)$, tak podľa vety 10.1 dostávame: Ak $U \sim R(0,1)$, potom $U^{1/m}$ má distribučnú funkciu F.

Príklad 10.3 (*Generovanie diskrétneho rozdelenia s konečným nosičom*)**.** Nech $p_1, p_2, ..., p_n$ sú nezáporné čísla také, že $\sum_{i=1}^n p_i = 1$ a nech $x_1 < x_2 < ... < x_n$. Uvažujme distribučnú funkciu $F(x) = 0$ pre $x \leq x_1$ a $F(x) = \sum_{i:x_i<x} p_i$ pre $x > x_1$. (F je distribučná funkcia diskrétnej náhodnej premennej, ktorá nadobúda hodnoty $x_1, ..., x_n$ s pravdepodobnosťami $p_1, ..., p_n$. Takéto je napríklad binomické, alebo hypergeometrické rozdelenie.) 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 $k = 2, ..., n$. Realizácie príslušnej diskrétnej náhodnej premennej preto môžeme generovať ako $G(U)$, kde $U \sim R(0,1)$.

Konkrétnejšie, interval $(0,1)$ rozbijeme na intervaly $I_1 = (0, p_1)$, $I_2 = [p_1, p_1 + p_2)$, $I_3 = [p_1+p_2, p_1+p_2+p_3)$,..., $I_n = [1-p_n, 1)$. 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 diskrétnej náhodnej premennej s distribučnou funkciou F.

Veta 10.2 (*Zamietacia metóda*)**.** Nech $f, h : \mathbb{R} \to [0, \infty)$ sú dve funkcie hustoty, ktoré sú kladné na intervale $I \subseteq \mathbb{R}$ a inde nulové, pričom

$$c = \sup \{f(x)/h(x) : x \in I\} < \infty$$

Nech $Y_1, U_1, Y_2, U_2, ...$ sú nezávislé náhodné premenné, pričom $Y_1, Y_2, ...$ majú hustotu h a $U_1, U_2, ...$ majú rozdelenie $R(0,1)$. Označme

$$N = \min \{n \in \mathbb{N} : c.h(Y_n)U_n < f(Y_n)\}$$

Potom náhodná premenná Y_N má hustotu f.

Dôkaz. Pre jednoduchosť urobíme dôkaz len pre tento prípad: $I = (0,1)$ a h je hustota rozdelenia $R(0,1)$, t.j. $h(x) = 1$ pre $x \in (0,1)$, $h(x) = 0$ pre $x \notin (0,1)$.

Zvoľme ľubovoľné $z \in (0,1)$. Zrejme platí

$$P[Y_N < z] = \cup_{k=1}^{\infty} P[Y_N < z, N = k] = \cup_{k=1}^{\infty} P[Y_k < z, N = k].$$

Keďže pre akékoľvek k má náhodný vektor $(Y_k, U_k)^T$ dvojrozmerné rovnomerné rozdelenie na množine $[0,1] \times [0,1]$, je $P[Y_k < z, cU_k < f(Y_k)]$ rovná ploche množiny $B_z = \{(y,u)' \in [0,1] \times [0,1] : y < z \text{ a } c.u < f(y)\}$, teda

$$P[Y_k < z, cU_k < f(Y_k)] = \frac{1}{c} \int_0^z f(y)dy = r_z.$$

Preto $P[Y_1 < z, N = 1] = P[Y_1 < z, c.U_1 < f(Y_1)] = r_z$ a pre $n \geq 2$ je $P[Y_k < z, N = k] = P[Y_k < z, c.U_1 \geq f(Y_1), ..., c.U_{k-1} \geq f(Y_{k-1}), c.U_k < f(Y_k)] = P[Y_k < z, c.U_k < f(Y_k)]P[c.U_1 \geq f(Y_1)]...P[c.U_{k-1} \geq f(Y_{k-1})] = r_z(1-r_1)^{k-1}$. Takže

$$P[Y_N < z] = r_z \sum_{k=1}^{\infty} (1 - 1/c)^{k-1} = r_z c = \int_0^z f(y)dy.$$

Z toho plynie, že f je hustota Y_N. $\square$

Najčastejšie používaný špeciálny prípad predchádzajúcej vety je nasledovný: f je ohraničená hustota kladná len na intervale $I = (a, b)$, kde a, b sú konečné čísla a h je hustota rozdelenia $R(a, b)$. V tomto prípade platí $N = \min\{n \in \mathbb{N} : d.U_n < f(Y_n)\}$, kde $d = \sup\{f(x) : x \in I\}$.

Príklad 10.4. Uvažujme hustotu $f(x) = (8/\pi)\sqrt{x}\sqrt{1-x}$ pre $x \in (0, 1)$ a $f(x) = 0$ pre $x \notin (0, 1)$. (Jedná sa o tzv. beta-rozdelenie s parametrami 3/2 a 3/2.) Ak zvolíme h hustotu rovnomerného rozdelenia na intervale $(0, 1)$, tak realizáciu z rozdelenia s hustotou f dostaneme nasledovne:

Budeme generovať nezávislé realizácie $Y_1, U_1, Y_2, U_2, \ldots$ z rozdelenia $R(0, 1)$ až pokým nenastane prípad $(4/\pi)U_n < f(Y_n)$. V tomto okamihu bude Y_n reprezentovať realizáciu z rozdelenia s hustotou f.

Príklad 10.5 (*Generovanie z normálneho rozdelenia zamietacou metódou*)**.** Pomocou zamietacej metódy je možné generovať aj realizácie z rozdelenia $N(0, 1)$ (a následne, vhodnou lineárnou transformáciou, aj z rozdelenia $N(\mu, \sigma^2)$).

Uvažujme najprv hustotu $f(x) = \sqrt{2/\pi} \times e^{-x^2/2}$ pre $x > 0$ a $f(x) = 0$ pre $x \leq 0$, čo zodpovedá hustote náhodnej premennej $|X|$, ak $X \sim N(0, 1)$. Zvolíme za h hustotu rozdelenia $Exp(1)$, t.j. $h(x) = e^{-x}$ pre $x > 0$ a $h(x) = 0$ pre $x \leq 0$. Pomocou štandardných metód matematickej analýzy vypočítame, že $c = \sup\{f(x)/h(x) : x > 0\} = \sqrt{2e/\pi}$. Podmienka určujúca index n akceptovanej realizácie je

$$\sqrt{2e/\pi} \times e^{-Y_n} U_n < \sqrt{2/\pi} \times e^{-Y_n^2/2},$$

čo je možné ekvivalentne zapísať v tvare

$$U_n < \exp\left(-(Y_n - 1)^2/2\right)$$

Vygenerovať realizáciu premennej z rozdelenia s hustotou f môžeme teda nasledovne: generujeme nezávislé realizácie $Y_1, U_1, Y_2, U_2, \ldots$, pričom $Y_1, Y_2, \ldots$ majú rozdelenie $Exp(1)$ a $U_1, U_2, \ldots$ majú rozdelenie $R(0, 1)$. Akonáhle bude splnené $U_n < \exp\left(-(Y_n - 1)^2/2\right)$, bude $Y = Y_n$ konečným výsledkom.

Pomocou realizácie Y s hustotou f môžeme vygenerovať realizáciu z rozdelenia $N(0, 1)$ jednoducho tak, že Y vynásobíme s pravdepodobnosťou 1/2 hodnotou -1 a s pravdepodobnosťou 1/2 hodnotou 1 (t.j. číslu Y dáme náhodné znamienko).

Veta 10.3 (*Boxov-Müllerov generátor normálneho rozdelenia*)**.** Nech U, V sú nezávislé náhodné premenné s rozdelením $R(0, 1)$. Nech

$$X = \sqrt{-2ln(U)}\cos(2\pi V) \text{ a } Y = \sqrt{-2ln(U)}\sin(2\pi V)$$

Potom X a Y sú nezávislé náhodné premenné, obe s rozdelením $N(0, 1)$.

Veta 10.4 (*Generovanie Poissonovho rozdelenia*)**.** Nech $X_1, X_2, \ldots$ sú nezávislé premenné s rozdelením $Exp(1)$ a $N = \min\{n \in \mathbb{N} : \sum_{i=1}^n X_i > \lambda\}$. Potom $N - 1 \sim Po(\lambda)$. Ekvivalentne: Nech $U_1, U_2, \ldots$ sú nezávislé náhodné premenné s rozdelením $R(0, 1)$. Nech $N = \min\left\{n \in \mathbb{N} : \prod_{i=1}^n U_i < e^{-\lambda}\right\}$. Potom $N - 1 \sim Po(\lambda)$.

Veta 10.5 (*Generovanie geometrického rozdelenia*)**.** Nech $p \in (0, 1)$ a nech $U \sim R(0, 1)$. Potom

$$\lfloor \ln(U)/\ln(1-p) \rfloor \sim Geo(p)$$

kde $\lfloor . \rfloor$ označuje dolnú celú časť.

Žiadna z uvedených metód generovania zo špecifických typov rozdelení (v tomto, ani v predchádzajúcom odstavci) nie je vo všeobecnosti najrýchlejšia. Najrýchlejšie známe metódy sú obvykle algoritmicky veľmi komplikované (napríklad využívajú rozsiahle tabuľky špeciálnych konštánt), ich programová implementácia je zdĺhavá a oplatí sa iba v prípade, že je potrebné generovať obrovské počty realizácií.

10.2 Generovanie realizácií náhodných vektorov

Veta 10.6. Nech $\mathbf{X} \sim Mult(n, p_1, \ldots, p_m)$, kde $p_1, ..., p_m > 0$. Potom $X_1 \sim Bin(n, p_1)$. Nech ďalej $k_1, ..., k_m \in \{0, 1, ..., n\}$ sú také, že $\sum_{j=1}^m k_j = n$. Nech $i \in \{2, 3, ..., m\}$. Označme $n_i = \sum_{j=i}^m k_j$ a $r_i = p_i / \sum_{j=i}^m p_j$. Potom

$$P[X_i = k_i | X_1 = k_1, \ldots, X_{i-1} = k_{i-1}] = \binom{n_i}{k_i} r_i^{k_i} (1 - r_i)^{n_i - k_i}$$

Predchádzajúca veta znamená, že "podmienené rozdelenie" náhodnej premennej X_i za podmienky $[X_1 = k_1, \ldots, X_{i-1} = k_{i-1}]$ je rozdelenie $Bin(\sum_{j=i}^m k_j, p_i / \sum_{j=i}^m p_j)$. Realizáciu $(k_1, ..., k_m)^T$ náhodného vektora $\mathbf{X} \sim Mult(n, p_1, \ldots, p_m)$ teda vygenerujeme tak, že najprv vygenerujeme realizáciu k_1 z rozdelenia $Bin(n, p_1)$, potom vygenerujeme realizáciu k_2 z rozdelenia $Bin(n - k_1, p_2 / \sum_{j=2}^m p_j)$, potom realizáciu k_3 z rozdelenia $Bin(n - k_1 - k_2, p_3 / \sum_{j=3}^m p_j)$ a tak ďalej.

Veta 10.7. Nech $X_1, ..., X_m$ sú nezávislé náhodné premenné s rozdelením $N(0, 1)$. Nech $\mu \in \mathbb{R}^m$ a nech Σ je pozitívne definitná matica typu $m \times m$. Predpokladajme, že $\Sigma = \mathbf{C}\mathbf{C}^T$, kde $\mathbf{C}$ je matica typu $m \times m$. Potom náhodný vektor $\mathbf{Y} = \mathbf{C}(X_1, ..., X_m)^T + \mu$ má rozdelenie $N_m(\mu, \Sigma)$.

Dôkaz. Veta je jednoduchým dôsledkom viet 8.21 a 8.23. □

Maticu $\mathbf{C}$ z predošlej vety je možné určiť viacerými metódami; obvykle sa však používa dolná trojuholníková matica, ktorú je možné z matice Σ vypočítať relatívne rýchlym algoritmom a naviac mierne zjednodušuje aj výpočet súčinu $\mathbf{C}(X_1, ..., X_m)^T$. Takúto dolnú trojuholníkovú maticu vypočítame napríklad pomocou nasledovného algoritmu (prvky matice Σ označíme s_{ij} a prvky matice C označíme c_{ij})

1 Prvkom matice C nad diagonálou priraď hodnotu 0.

2 $c_{11} \leftarrow \sqrt{s_{11}}$ a pre $i = 2$ až m vykonaj: $c_{i1} \leftarrow s_{i1}/c_{11}$

4 $c_{22} \leftarrow \sqrt{s_{22} - c_{21}^2}$

5 Pre $i = 3$ až m vykonaj:

5.1 Pre $j = 2$ až $i - 1$ vykonaj: $c_{ij} \leftarrow (s_{ij} - \sum_{k=1}^{j-1} c_{ik} c_{jk})/c_{jj}$

5.2 $c_{ii} \leftarrow \sqrt{s_{ii} - \sum_{k=1}^{i-1} c_{ik}^2}$

6 Výsledok: C

Veta 10.8. Nech $G_m(r)$ je m-rozmerná guľa so stredom v počiatku súradnicovej sústavy a polomerom $r > 0$:

$$G_m(r) = \left\{ (x_1, ..., x_m) \in \mathbb{R}^m : \sqrt{\sum_{i=1}^{m} x_i^2} \leq r \right\}.$$

Nech $X_1, ..., X_m, U$ sú nezávislé náhodné premenné, pričom $X_i \sim N(0,1)$ a $U \sim R(0,1)$. Potom náhodný vektor

$$\mathbf{Y} = r(\sum_{i=1}^{m} X_i^2)^{-1/2}(X_1, ..., X_m)^T$$

má rovnomerné rozdelenie na hranici (povrchu) gule $G_m(r)$ a náhodný vektor $\mathbf{Z} = U^{1/m}\mathbf{Y}$ má rovnomerné rozdelenie na vo vnútri (t.j. na) $G_m(r)$.

Dôkaz. Veta plynie z rotačnej symetrie hustoty rozdelenia $N_m(0, I)$ a z faktu, že ak náhodný vektor má rovnomerné rozdelenie vo vnútri $G_m(1)$, tak jeho Euklidova norma má rovnaké rozdelenie ako $U^{1/m}$ pre $U \sim R(0,1)$. □

Existuje viacero alternatívnych, často efektívnejších spôsobov ako generovať náhodné vektory rovnomerne na povrchu $G_m(r)$, alebo na $G_m(r)$, najmä pre dimenzie $m = 2, 3$ (ak vynecháme triviálny prípad $m = 1$).

Zo širokého spektra rôznych metód spomenieme napríklad nasledovné "zamietacie" metódy pre $m = 2, 3$. Generujme $X_1, Y_1, X_2, Y_2, ...$ ako nezávislé realizácie z rozdelenia $R(-1, 1)$, až kým prvýkrát nenastane prípad $X_n^2 + Y_n^2 < 1$. Označme $X = X_n, Y = Y_n$. Potom platí:

- $(X, Y)^T$ má rovnomerné rozdelenie na $G_2(1)$
- $(X^2 + Y^2)^{-1/2}(X, Y)^T$ má rovnomerné rozdelenie na hranici $G_2(1)$
- $(X^2 + Y^2)^{-1}(X^2 - Y^2, 2XY)^T$ má rovnomerné rozdelenie na hranici $G_2(1)$
- $(2X\sqrt{1 - X^2 - Y^2}, 2Y\sqrt{1 - X^2 - Y^2}, 1 - 2(X^2 + Y^2))^T$ má rovnomerné rozdelenie na hranici $G_3(1)$.

Veta 10.9. Nech $S_m(h)$, $h > 0$, je m-rozmerný simplex typu:

$$S_m(h) = \left\{ (x_1, ..., x_m)^T \in \mathbb{R}^m : \sum_{i=1}^{m} x_i \leq h \text{ a } x_i \geq 0 \text{ pre všetky } i = 1, ..., m \right\}$$

Predpokladajme, že $U_1, ..., U_m$ sú nezávislé náhodné premenné s rozdelením $R(0,1)$ a nech $Y_1, ..., Y_m$ je usporiadanie týchto náhodných premenných od najmenšej po najväčšiu. Definujme $Z_1 = Y_1$ a $Z_i = Y_i - Y_{i-1}$ pre $i = 2, ..., m$. Potom náhodný vektor $\mathbf{Z} = h.(Z_1, ..., Z_m)^T$ má rovnomerné rozdelenie na simplexe $S_m(h)$.

Dôkaz. Technicky presný dôkaz pre všeobecné m je zdĺhavý, avšak základná myšlienka je jednoduchá; popíšeme ju pre $m = 2$ a $h = 1$. V tomto prípade generujeme U_1, U_2 nezávisle rovnomerne na intervale $(0, 1)$, teda bod $\mathbf{U} = (U_1, U_2)^T$ padá rovnomerne náhodne do štvorca $S = (0, 1) \times (0, 1)$. Ak $\mathbf{U}$ padne do trojuholníka $T = (0, 0), (0, 1), (1, 1)$, tak položíme $\mathbf{Y} = (Y_1, Y_2)^T = \mathbf{U}$, ak padne $\mathbf{U}$ do trojuholníka $S \setminus T$, tak $\mathbf{Y}$ dostaneme preklopením bodu $\mathbf{U}$ okolo priamky $y = x$. Je teda zrejmé, že $\mathbf{Y}$ má rovnomerné rozdelenie na trojuholníku T. Z $\mathbf{Y}$ dostaneme $\mathbf{Z} = (Z_1, Z_2)^T$ už len jednoduchým lineárnym zobrazením transformujúcim T na simplex $S_2(1)$. □

Vetu 10.9 je možné použiť na generovanie realizácií s rovnomerným rozdelením na akomkoľvek simplexe (napr. na ľubovoľnom nedegenerovanom trojuholníku v $\mathbb{R}^2$) pomocou vhodnej lineárnej transformácie. Takisto nie je ťažké použiť predchádzajúcu vetu na generovanie realizácií náhodných vektorov s rovnomerným rozdelením na ohraničenej polyedrickej množine v $\mathbb{R}^m$, ak poznáme rozklad tejto množiny na všeobecné simplexy.

Kapitola 11

Základy teórie informácie

11.1 Informácia

V tejto kapitole sa pokúsime kvantifikovať pojem informácie.

Definícia 11.1. Uvažujme pravdepodobnostný priestor $(\Omega, \mathcal{S}, P)$ taký, že $|\Omega| = m$. Pod informáciou obsiahnutou v udalosti E budeme rozumieť veličinu

$$I(E) = -\log_2(P(E)).$$

Informáciu budeme merať v bitoch.

Všimnime si, že I je klesajúcou funkciou $P(E)$, t.j. ak pre $E, F \in \mathcal{S}$: $P(E) \leq P(F)$, potom $I(E) \geq I(F)$. Tiež platí, že ak E, F sú nezávislé, potom $I(E \cap F) = I(E) + I(F)$ a pre všetky $E \in \mathcal{S}$ je $I(E) \geq 0$.

Poznámka 11.1. V tejto kapitole budeme pod označením $\log(x)$ rozumieť logaritmus so základom 2.

Príklad 11.1. Uvažujme náhodnú premennú $X \sim Alt(1/2)$. Máme

$$I(X = 0) = I(X = 1) = -\log_2(1/2) = 1.$$

Teda výberom jednej z dvoch rovnako pravdepodobných možností získame jeden bit informácie.

11.2 Entropia

Ak máme diskrétnu náhodnú premennú, ktorá nadobúda hodnoty $\{x_1, \dots, x_n\}$ s pravdepodobnosťami $p_1, \dots, p_n$, nevieme s určitosťou povedať, aké veľké sú informácie $I(p_i) := I(X = x_i)$, $i = 1, \dots, n$. Môžeme teda aj informáciu $I(X)$ chápať ako náhodnú premennú. Stredná hodnota tejto náhodnej premennej sa nazýva entropia.

Definícia 11.2. [*Entropia*]. Nech X je diskrétna náhodná premenná, ktorá nadobúda hodnoty $\{x_1, \dots, x_n\}$ s pravdepodobnosťami $p_1, \dots, p_n$. Potom entropiou H premennej X nazývame 'mieru neurčitosti'

$$H(X) = E(I(X)) = -\sum_{i=1}^{n} p_i \log p_i$$

V prípade, že $p_i = 0$, definujeme $p_i \log p_i = 0$. Všimnime si, že hodnota entropie nezávisí na tom, aké hodnoty nadobúda náhodná premenná X, ale len na ich pravdepodobnostiach.

Príklad 11.2. Majme náhodnú premennú

$$X = \begin{cases} u_1 & \text{s pravdepodobnosťou } 1/4 \\ u_2 & \text{s pravdepodobnosťou } 1/2 \\ u_3 & \text{s pravdepodobnosťou } 1/4 \end{cases}$$

Potom

$$H(X) = -\frac{1}{4}\log\frac{1}{4} - \frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} \approx 1.04$$

Entropia nadobúda hodnoty z intervalu $[0, \log n]$. Minimum nastáva, ak pre niektoré $j \in \{1, \dots, n\}$ je $p_j = 1$, teda $P[X = x_j] = 1$. Naopak, maximálnu hodnotu entropie dostaneme pre $p_i = 1/n$, $i = 1, \dots, n$, teda ak X nadobúda všetky hodnoty s rovnakou pravdepodobnosťou.

Veta 11.1. Nech X je diskrétna náhodná premenná, ktorá nadobúda hodnoty $\{x_1, \dots, x_n\}$ s pravdepodobnosťami $p_1, \dots, p_n$. Potom:

1. $H(X) \geq 0$ a rovnosť nastáva práve vtedy, keď existuje také x, že $P[X = x] = 1$.

2. $H(X) \leq \log(n)$ a rovnosť nastáva práve vtedy, keď X má rovnomerné rozdelenie.

Dôkaz. 1. Nezápornosť je zrejmá z definície entropie. Aby $H(X) = 0$, musí byť $p_i \log p_i = 0$ pre všetky i. To ale znamená, že existuje nejaké k, pre ktoré je $p_k = 1$.

2. Predpokladajme najskôr, že $p_i > 0$ pre všetky i. Z definície máme

$$H(X) - \log(n) = -\frac{1}{\ln(2)}\left(\sum_{i=1}^{n} p_i \ln p_i + \ln(n)\right) = -\frac{1}{\ln(2)}\left(\sum_{i=1}^{n} p_i \ln(p_i n)\right) =$$

$$= \frac{1}{\ln(2)}\left(\sum_{i=1}^{n} p_i \ln\left(\frac{1}{p_i n}\right)\right) \leq \frac{1}{\ln(2)}\left(\sum_{i=1}^{n} p_i \left(\frac{1}{p_i n} - 1\right)\right) \leq$$

$$\leq \frac{1}{\ln(2)}\left(\sum_{i=1}^{n}\left(\frac{1}{n} - p_i\right)\right) = 0.$$

Použili sme vzťah $\ln(x) \leq x - 1$. Rovnosť v tomto vzťahu platí práve vtedy, keď $x = 1$, a teda v našom prípade $1/(p_i n) - 1 = 0$, t.j. $p_i = 1/n$, $i = 1, \dots, n$. □

11.3 Združená a podmienená entropia

Definícia 11.3. Nech X a Y sú dve diskrétne náhodné premenné definované na rovnakom pravdepodobnostnom priestore, X nadobúda hodnoty $\{x_1, \dots, x_n\}$ s pravdepodobnosťami $p_1, \dots, p_n$, Y nadobúda hodnoty $\{y_1, \dots, y_m\}$ s pravdepodobnosťami $q_1, \dots, q_m$. Označme združené pravdepodobnosti $p_{ij} = P[X = x_i, Y = y_j]$, $i = 1, \dots, n$, $j = 1, \dots, m$. Potom definujeme združenú entropiu

$$H(X, Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} \log p_{ij}.$$

Všimnite si, že $H(X,Y) = H(Y,X)$.

Aby sme vedeli lepšie popísať vzťah medzi závislosťou náhodných premenných a entropiou, zavedieme pojem podmienenej entropie.

Definícia 11.4. V kontexte predchádzajúcej definície označme podmienené pravdepodobnosti $p_{j|i} = P[Y = y_j | X = x_i]$. Potom podmienenou entropiou Y za podmienky $X = x_i$ nazývame

$$H_i(Y) = -\sum_{j=1}^{m} p_{j|i} \log(p_{j|i}).$$

$H_i(Y)$ meria našu mieru neurčitosti o Y, ak vieme, že nastala udalosť $X = x_i$. Uvažujme teraz náhodnú premennú $H_.(Y)$, ktorá nadobúda hodnoty $H_1(Y), \ldots, H_n(Y)$ s pravdepodobnosťami $p_1, \ldots, p_n$; $H_.(Y)$ je teda funkciou X.

Definícia 11.5. Pod podmienenou entropiou náhodnej premennej Y pri danej X budeme rozumieť

$$H_X(Y) = E(H_.(Y)) = \sum_{i=1}^{n} p_i H_i(Y).$$

Veta 11.2.

$$H_X(Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \log(p_{j|i}).$$

Veta 11.3. Ak X a Y sú nezávislé, potom $H_X(Y) = H(Y)$.

Príklad 11.3.

Veta 11.4. $H(X,Y) = H(X) + H_X(Y)$.

Dôkaz.

$$H(X,Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \log(p_{j|i} p_i) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \log(p_{j|i}) - -\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \log(p_i).$$

□

Dôsledok 11.1. Ak X a Y sú nezávislé náhodné premenné, potom $H(X,Y) = H(X) + H(Y)$.

11.4 Relatívna entropia a vzájomná informácia

Videli sme, že $H_X(Y)$ je miera informácie Y, ktorá nie je obsiahnutá v X. Teda informácia Y, ktorá je obsiahnutá aj v X, potom je $H(Y) - H_X(Y)$.

Definícia 11.6 (*Vzájomná informácia*)**.** Nech X a Y sú dve diskrétne náhodné premenné so združeným rozdelením $p_{ij} = P[X = x_i, Y = y_j]$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Vzájomnou informáciou X a Y nazývame

$$I(X,Y) = H(Y) - H_X(Y).$$

Veta 11.5 (*Vlastnosti vzájomnej informácie*). 1.

$$I(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}\log\left(\frac{p_{ij}}{p_i q_j}\right)$$

2. $I(X,Y) = I(Y,X)$

3. Ak X a Y sú nezávislé, potom $I(X,Y) = 0$.

Dôkaz. Dôkazy tvrdení 2 a 3 sú triviálne.
1. $H(Y) = -\sum_{j=1}^{m} q_j \log(q_j) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}\log(q_j)$ a odtiaľ

$$I(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}\log(q_j) + -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}\log(p_{j|i}).$$

Dosadením $p_{j|i} = p_{ij}/p_i$ dostávame požadované tvrdenie. □

Príklad 11.4.

Definícia 11.7 (*Kullback-Leiblerova vzdialenosť*). Relatívna entropia alebo Kullback-Leiblerova vzdialenosť medzi dvoma rozdeleniami pravdepodobnosti p a q je definovaná ako

$$D(p \parallel q) = \sum_x p(x)\frac{p(x)}{q(x)} = \mathrm{E}_p \log\frac{p(X)}{q(X)}.$$

Poznámka 11.2. V predchádzajúcej definícii $0\log\frac{0}{q} = 0$ a $p\log\frac{p}{0} = \infty$.

Relatívna entropia je miera vzdialenosti medzi dvoma rozdeleniami pravdepodobnosti alebo miera 'straty' pri predpoklade správnosti rozdelenia q, ak je v skutočnosti správne rozdelenie p.

Veta 11.6. Nech $p(x)$, $q(x)$ sú dve rozdelenia pravdepodobnosti. Potom $D(p \parallel q) \geq 0$ a rovnosť nastáva práve vtedy, keď $p(x) = q(x)$ pre všetky x.

Dôkaz. Označme $A = \{x : p(x) > 0\}$. Potom

$$-D(p \parallel q) = -\sum_{x\in A} p(x)\frac{p(x)}{q(x)} = \sum_{x\in A} p(x)\log\frac{q(x)}{p(x)} \leq \log\sum_{x\in A} p(x)\frac{q(x)}{p(x)} \leq \log\sum_x q(x) = 0.$$

Keďže $t \to \log T$ je striktne konkávna funkcia, rovnosť nastáva len ak $\frac{q(x)}{p(x)} = 1$. □

Príklad 11.5.

11.4.1 Princíp maximálnej entropie

Nech X je diskrétna náhodná premenná, ktorá nadobúda hodnoty $\{x_1, \ldots, x_n\}$ s pravdepodobnosťami $p_1, \ldots, p_n$. Predpokladajme, že o X vieme, že $E(X) = K$, kde K je nejaká konštanta. Ak $K \neq \frac{1}{n}\sum_{i=1}^n x_i$, vieme, že X nemá rovnomerné rozdelenie.

Hľadajme maximum entropie $H(X)$ vzhľadom na ohraničenia $\sum_{i=1}^n p_i = 1$, $\sum_{i=1}^n p_i x_i = K$.

Maximalizujeme teda funkciu $(n+2)$ premenných

$$L(p_1, \ldots, p_n; \lambda, \mu) = -\sum_{i=1}^n p_i \log(p_i) + \lambda\left(\sum_{i=1}^n p_i - 1\right) + \mu\left(\sum_{i=1}^n p_i x_i - K\right),$$

kde λ a μ sú Lagrangeove multiplikátory.

Ak túto funkciu zderivujeme podľa premenných a deriváciu položíme rovnú 0, dostávame vyjadrenie pre p_i:

$$p_i = \exp \ln(2)\lambda - 1 + \ln(2)\mu x_i, \ i = 1, \ldots, n.$$

Pretože $\sum_{i=1}^n p_i = 1$, musí platiť

$$\lambda = \frac{1}{\ln(2)}\left(1 - \ln\left(\sum_{i=1}^n \exp ln(2)\mu x_i\right)\right) \tag{11.1}$$

a odtiaľ dostávame tzv. Gibbsovo rozdelenie

$$p_i = \frac{\exp ln(2)\mu x_i}{\sum_{i=1}^n \exp ln(2)\mu x_i}, \ i = 1, \ldots, n. \tag{11.2}$$

Gibbsovo rozdelenie je teda prirodzenou alternatívou k rovnomernému rozdeleniu, ak poznáme len strednú hodnotu náhodnej premennej.

11.5 Entropy rate

V prípade postupnosti n náhodných premenných je prirodzené pýtať sa, ako rastie entropia s rastúcim n.

Definícia 11.8. Miera entropie náhodného procesu $\{X_i\}$ je definovaná ako

$$H(S) = \lim_{n\to\infty} \frac{1}{n} H(X_1, \ldots, X_n)$$

Príklad 11.6. Nech $X_1, X_2, \ldots$ sú nezávislé a rovnako rozdelené náhodné premenné. Potom

$$H(S) = \lim_{n\to\infty} \frac{H(X_1, \ldots, X_n)}{n} = \lim_{n\to\infty} \frac{nH(X_1)}{n} = H(X_1)$$

Veta 11.7. Ak je $\{X_i\}$ stacionárny markovovský reťazec, potom

$$H(S) = \lim_{n\to\infty} H(X_n|X_{n-1}, \ldots, X_1) = \lim_{n\to\infty} H(X_n|X_{n-1}) = H(X_2|X_1),$$

kde podmienenú entropiu počítame pri danom stacionárnom rozdelení π, teda

$$H(S) = -\sum_{i,j} \pi_i p_{ij} \log p_{ij}.$$

Príklad 11.7. Entropia markovovského reťazca z príkladu 6.4 je

$$H(S) = H(X_2|X_1) = \frac{p}{p+q} H(q) + \frac{q}{p+q} H(p).$$

11.6 Cvičenia

Úloha 11.1. Nech $S = \sum_{k=1}^{1} 01/k^2$. Uvažujme náhodnú premennú X takú, že $P[X = k] = 1/S \cdot k^2$ pre $k = 1, \ldots, 10$. Nájdite $H(X)$.

Úloha 11.2. Hádžeme vyváženou mincou až pokiaľ nepadne hlava. Nájdite entropiu náhodnej premennej X, ktorá reprezentuje počet hodov v tomto experimente.

Kapitola 12

Lineárny regresný model

12.1 Rozdelenia pravdepodobnosti odvodené od normálneho rozdelenia

Definícia 12.1. Nech $X_1, ..., X_k$ sú nezávislé náhodné premenné s rozdelením $N(0,1)$. Potom rozdelenie náhodnej premennej $Z = \sum_{i=1}^k X_i^2$ nazývame chíkvadrát rozdelenie s k-stupňami voľnosti. Značíme $Z \sim \chi_k^2$.

Dá sa ukázať, že $Z \sim \chi_k^2$ vtedy a len vtedy, keď je Z spojitá náhodná premenná s hustotou danou predpisom

$$f(z) = \frac{z^{k/2-1}e^{-z/2}}{2^{k/2}\Gamma(k/2)}$$

pre $z > 0$ a $f(z) = 0$ pre $z \leq 0$. Vo vyjadrení tejto hustoty je Γ takzvaná gama funkcia, pre ktorú platí $\Gamma(1/2) = \sqrt{\pi}$ a pre každé $k \in \mathbb{N}$: $\Gamma(k) = (k-1)!$ a $\Gamma(k+1/2) = \frac{\sqrt{\pi}}{2^k}(2k-1)!!$. Pre naše potreby však nie je tvar hustoty rozdelenia χ_k^2 dôležitý.

Veta 12.1. Nech $Z \sim \chi_k^2$. Potom $E(Z) = k$ a $D(Z) = 2k$.

Dôkaz. Ak $Z \sim \chi_k^2$, potom $Z = \sum_{i=1}^k X_i^2$, kde $X_1, ..., X_k$ sú nezávislé náhodné premenné s rozdelením $N(0,1)$. Keďže $E(X_i^2) = D(X_i) + (E(X_i))^2 = 1 + 0 = 1$, tak platí $E(Z) = \sum_{i=1}^k E(X_i^2) = k$. Na odvodenie disperzie Z počítajme: $E(X_i^4) = \int_{-\infty}^{\infty} x^4 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 3$ (pozri poznámku 7.1). To znamená, že $D(X_i^2) = E(X_i^4) - (E(X_i^2))^2 = 3 - 1 = 2$. S použitím vety 8.14 dostávame $D(Z) = \sum_{i=1}^k D(X_i^2) = 2k$. □

Definícia 12.2. Nech Y, Z sú nezávislé náhodné premenné, pričom $Y \sim N(0,1)$ a $Z \sim \chi_k^2$. Potom rozdelenie náhodnej premennej $T = Y/\sqrt{Z/k}$ nazývame t-rozdelenie (alebo aj Studentove rozdelenie) s k-stupňami voľnosti. Značíme $T \sim t_k$.

Podobne ako v prípade rozdelenia χ_k^2, nie je pre naše potreby dôležité poznať tvar hustoty rozdelenia t_k. Uvedieme ho len v tejto poznámke:

$$f(t) = \frac{\Gamma((k+1)/2)}{\sqrt{\pi k}\Gamma(k/2)} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}$$

pre všetky $t \in \mathbb{R}$. Rozdelenie t_1 sa nazýva Cauchyho rozdelenie a z teoretického hľadiska je zaujímavé napríklad tým, že nemá strednú hodnotu. Pre $k \geq 2$ rozdelenie t_k má strednú hodnotu 0. Disperzia rozdelenia t_k je konečná pre $k \geq 3$ a rovná $k/(k-2)$, čo však nebudeme dokazovať. Je tiež možné ukázať, že pre $k \to \infty$ konvergujú hustoty rozdelení t_k k hustote rozdelenia $N(0,1)$.

12.2 Náhodný výber a výberové charakteristiky

Máme konečnú množinu čísel $\{y_1, \dots, y_n\}$, ktoré sme získali meraním nejakej veličiny v danej populácii (napr. krvný tlak u pacientov užívajúcich konkrétny liek).

Definícia 12.3. Nech $Y_1, ..., Y_m$ sú nezávislé náhodné premenné, všetky s rovnakým rozdelením daným distribučnou funkciou F. Potom hovoríme, že náhodné premenné $Y_1, ..., Y_m$ tvoria náhodný výber rozsahu m z rozdelenia s distribučnou funkciou F.

Aby sme v skratke popísali tento, často rozsiahly, súbor dát, používame súhrnné charakteristiky. Načastejšie používanou charakteristikou polohy je výberový priemer a charakteristikou variability výberový rozptyl.

Definícia 12.4. Nech $Y_1, ..., Y_m$, $m \geq 2$ tvoria náhodný výber (z nejakého rozdelenia). Nech

$$\bar{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i \quad \text{a} \quad S^2 = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \bar{Y})^2$$

Potom náhodnú premennú $\bar{Y}$ nazveme výberový priemer a náhodnú premennú S^2 výberový rozptyl náhodných premenných $Y_1, ..., Y_m$.

Veta 12.2. Nech $Y_1, ...Y_m$ tvoria náhodný výber z rozdelenia $N(\mu, \sigma^2)$, pričom $m \geq 2$. Nech $\bar{Y}$ je výberový priemer a S^2 je výberový rozptyl náhodného výberu $Y_1, ..., Y_m$. Potom platí:

$$\bar{Y} \sim N(\mu, \sigma^2/m) \qquad E(\bar{Y}) = \mu$$

$$\frac{(m-1)S^2}{\sigma^2} \sim \chi^2_{m-1} \qquad E(S^2) = \sigma^2$$

$$T = \frac{\bar{Y} - \mu}{S}\sqrt{m} \sim t_{m-1}$$

Dôkaz. Ak $Y_1, ..., Y_m$ tvoria náhodný výber z rozdelenia $N(\mu, \sigma^2)$, potom pre náhodný vektor $\mathbf{Y} = (Y_1, ..., Y_m)^T$ platí: $\mathbf{Y} \sim N_m(F\theta, \sigma^2 I)$, kde $\theta = \mu$ a $F = (1, 1, ..., 1)^T$. Máme teda špeciálny prípad vety 12.5 s $k = 1$. Všimneme si, že

$$\hat{\theta} = (F^T F)^{-1} F^T \mathbf{Y} = \bar{Y}$$

a že náhodná premenná S^2 z vety 12.5 je práve výberový rozptyl náhodného výberu $Y_1, ..., Y_m$. Tým dostávame tvrdenie o rozdelení a strednej hodnote $\bar{Y}$ a S^2. Posledné tvrdenie dostaneme tak, že vo vete 12.5 položíme $c = 1$. $\square$

12.3 Lineárny regresný model

12.3.1 Lineárny regresný model priamkou

Príklad 12.1. Chceme určiť závislosť medzi dávkou radiácie a zmenšovaním nádoru v nasledovnej situácii:

x: dávka radiácie	1	2	3	4	5
y: veľkosť nádoru	14	11	10	8.5	6.5

Predpokladajme,že závislosť sa riadi jednoduchým lineárnym modelom

$$y_(x_i) = \theta_0 + \theta_1 x_i + \varepsilon_i,$$

kde ε_i je náhodná chyba i-teho merania.

Našim cieľom je určiť neznámy vektor $\theta = (\theta_0, \theta_1)^T$ tak, aby výsledná priamka čo najviac zodpovedala meraniam. Chceme teda minimalizovať štvorce odchýlok našej priamky od nameraných hodnôt, a teda nájsť také $\hat{\theta}$, aby platilo

$$\hat{\theta} = \arg\min_{\theta\in\mathbb{R}^2} \sum_{i=1}^{5} [y_(x_i) - (\theta_0 + \theta_1 x_i)]^2.$$

Definícia 12.5. Ak $Y_1, ..., Y_m$ je postupnosť nezávislých náhodných premenných, pričom $Y_i \sim N(ax_i + b, \sigma^2)$, kde $x_1, ..., x_m \in \mathbb{R}$, $a, b \in \mathbb{R}$ a $\sigma^2 > 0$, tak hovoríme, že náhodné premenné $Y_1, ..., Y_m$ spĺňajú lineárny regresný model priamkou s normálnymi a nezávislými chybami s konštantným rozptylom σ^2.

Všimnite si, že $Y_1, ..., Y_m$ spĺňajú predpoklady prechádzajúcej definície vtedy a len vtedy, keď $Y_i = ax_i + b + \epsilon_i$, kde "chyby meraní" $\epsilon_1, ..., \epsilon_m$ tvoria náhodný výber z rozdelenia $N(0, \sigma^2)$.

Veta 12.3. Nech $Y_1, ..., Y_m$, $m \geq 3$ je postupnosť nezávislých náhodných premenných, pričom $Y_i \sim N(ax_i + b, \sigma^2)$, kde $x_1, ..., x_m \in \mathbb{R}$ nie sú všetky rovnaké, $a, b \in \mathbb{R}$, $\sigma^2 > 0$. Nech $\bar{x} = (1/m)\sum_{i=1}^m x_i$. Definujme nasledovné náhodné premenné:

$$\hat{a} = \frac{\sum_{i=1}^m Y_i(x_i - \bar{x})}{\sum_{i=1}^m (x_i - \bar{x})^2}, \quad \hat{b} = \bar{Y} - \hat{a}\bar{x}, \quad S^2 = \frac{1}{m-2}\sum_{i=1}^m (Y_i - \hat{a}x_i - \hat{b})^2$$

Potom platí:

$$(\hat{a}, \hat{b})^T \sim N_2((a, b)^T, \Sigma) \quad E((\hat{a}, \hat{b})^T) = (a, b)^T$$

kde

$$\Sigma = \frac{\sigma^2}{m\sum_{i=1}^m x_i^2 - m^2\bar{x}^2} \begin{pmatrix} m & -m\bar{x} \\ -m\bar{x} & \sum_{i=1}^m x_i^2 \end{pmatrix}$$

$$\frac{(m-2)S^2}{\sigma^2} \sim \chi^2_{m-2} \quad E(S^2) = \sigma^2$$

$$\frac{\hat{a} - a}{S}\sqrt{\sum_{i=1}^m x_i^2 - m\bar{x}^2} \sim t_{m-2}$$

Dôkaz. Tvrdenie vety je priamym dôsledkom vety 12.5. Skutočne, ak označíme $\mathbf{Y} = (Y_1, ..., Y_m)^T$ a $\theta = (a, b)^T$, potom platí: $\mathbf{Y} \sim N_m(F\theta, \sigma^2 I)$, kde

$$F = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ 1 & 1 & \dots & 1 \end{pmatrix}^T$$

Všimnime si, že ak $x_1, ..., x_m$ nie sú všetky rovnaké, tak hodnosť matice F je $k = 2$. Mechanickým výpočtom sa ľahko presvedčíme, že platí

$$(\hat{a}, \hat{b})^T = (F^T F)^{-1} F^T \mathbf{Y}$$

ďalej že náhodné premenné S^2 v tejto vete a vo vete 12.5 sú (pre naše špecifické F) rovnaké a že

$$(F^T F)^{-1} = \frac{1}{m \sum_{i=1}^m x_i^2 - m^2 \bar{x}^2} \begin{pmatrix} m & -m\bar{x} \\ -m\bar{x} & \sum_{i=1}^m x_i^2 \end{pmatrix}$$

Takto z vety 12.5 priamo dostávame rozdelenie a stredné hodnoty $(\hat{a}, \hat{b})^T$ a S^2. Posledný výrok dostaneme tak, že vo vete 12.5 položíme $c = (1, 0)^T$. $\square$

K predpisom pre $\hat{a}$ a $\hat{b}$ je možné dospieť pomocou takzvanej metódy najmenších štvorcov, t.j. $\hat{a}$ a $\hat{b}$ minimalizujú súčet $\sum_{i=1}^m (Y_i - ax_i - b)^2$. Náhodné premenné $\hat{a}$, $\hat{b}$ sa preto nazývajú odhady koeficientov a, b pomocou metódy najmenších štvorcov (MNŠ-odhady).

12.3.2 Základná veta o lineárnom regresnom modeli

Vo všeobecnosti, ak v bodoch $\xi_1, \dots, x_m$ nameriame dáta $y_1, \dots, y_m$, lineárnym regresným modelom nazývame vzťah

$$y_i = f^T(x_i)\theta + \varepsilon_i,$$

kde f je daná vektorová funkcia, $\theta = (\theta_1, \dots, \theta_k)^T$ je vektor neznámych parametrov a pre náhodné chyby platí $E(\varepsilon_i) = 0, D(\varepsilon_i) = \sigma^2$ a $cov(\varepsilon_i, \varepsilon_j) = 0$ pre $i \neq j$. Našou úlohou je z dát odhadnúť priebeh regresnej funkcie $x \to f^T(x)\theta$.

Veta 12.4. Nech F je matica typu $m \times k$, pričom $k < m$ a hodnosť matice F je k, $\theta \in \mathbb{R}^k$. Potom minimalizačný problém

$$\hat{\theta} = \arg\min_{\theta \in \mathbb{R}^k} \sum_{i=1}^m [Y_i - (F\theta)_i]^2$$

má práve jedno riešenie

$$\hat{\theta} = (F^T F)^{-1} F^T \mathbf{Y}$$

Veta 12.5. Nech pre náhodný vektor $\mathbf{Y} = (Y_1, ..., Y_m)^T$ platí:

$$\mathbf{Y} \sim N_m(F\theta, \sigma^2 I)$$

kde F je matica typu $m \times k$, pričom $k < m$ a hodnosť matice F je k, ďalej $\theta \in \mathbb{R}^k$, $\sigma^2 > 0$ a I je jednotková matica typu $m \times m$. Nech

$$\hat{\theta} = (F^T F)^{-1} F^T \mathbf{Y}$$

$$S^2 = \frac{1}{m-k}\left\|\mathbf{Y} - F\hat{\theta}\right\|^2 = \frac{1}{m-k}\sum_{i=1}^{m}\left(Y_i - (F\hat{\theta})_i\right)^2$$

Potom platí

$$\hat{\theta} \sim N_k(\theta, \sigma^2(F^TF)^{-1}) \qquad E(\hat{\theta}) = \theta$$

$$\frac{(m-k)S^2}{\sigma^2} \sim \chi^2_{m-k} \qquad E(S^2) = \sigma^2$$

$$T = \frac{c^T(\hat{\theta}-\theta)}{S\sqrt{c^T(F^TF)^{-1}c}} \sim t_{m-k}$$

pre každý nenulový vektor $c \in \mathbb{R}^k$.

Veta 12.6. Nech $Y_{1,1}, ...Y_{1,m_1}, Y_{2,1}, ..., Y_{2,m_2}$ sú združene nezávislé náhodné premenné. Pre $i = 1$ aj pre $i = 2$ nech $Y_{i,1}, ...Y_{i,m_i}$ tvoria náhodný výber z rozdelenia $N(\mu_i, \sigma^2)$, pričom $m_i \geq 2$, $\sigma^2 > 0$ a nech $\bar{Y}_i$ a S_i^2 sú výberový priemer, resp. výberový rozptyl daného náhodného výberu. Označme $m = m_1 + m_2$ a

$$S^2 = \frac{1}{m-2}\left((m_1-1)S_1^2 + (m_2-1)S_2^2\right)$$

Potom platí:

$$\frac{(m-2)S^2}{\sigma^2} \sim \chi^2_{m-2} \qquad E(S^2) = \sigma^2$$

$$T = \frac{\bar{Y}_1 - \bar{Y}_2 - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{m_1} + \frac{1}{m_2}}} \sim t_{m-2}$$

Dôkaz. Veta je špeciálnym prípadom tvrdenia 12.5 pre $\mathbf{Y} = (Y_{1,1}, ...Y_{1,m_1}, Y_{2,1}, ..., Y_{2,m_2})^T$, $m = m_1 + m_2$, $k = 2$, $\theta = (\mu_1, \mu_2)^T$

$$F = \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 1 \end{pmatrix}^T$$

(vľavo hore je m_1 a vpravo dole je m_2 jednotiek) a pre $c = (1, -1)^T$. $\square$

Kapitola 13

Bodové a intervalové odhady parametrov štatistických modelov

13.1 Bodové odhady

Definícia 13.1 (Nevychýlený bodový odhad parametra)**.** Nech $\Theta \subseteq \mathbb{R}^k$ je parametrický priestor a nech $\{F_\theta : \mathbb{R}^m \to [0,1]; \theta \in \Theta\}$ je parametrická trieda distribučných funkcií. Nech $g_i : \mathbb{R}^m \to \mathbb{R}$ je taká funkcia, že pre každé $\theta \in \Theta$ platí: Ak náhodný vektor $(X_1, ..., X_m)^T$ má distribučnú funkciu F_θ, tak náhodná premenná $T_i = g_i(X_1, ..., X_m)$ má strednú hodnotu θ_i. Náhodnú premennú T_i potom nazývame nevychýlený odhad parametra θ_i.

Definícia 13.2. Hovoríme, že odhad T^* je rovnomerne lepší ako odhad T, ak $D_\theta(T^*) \leq D_\theta(T)$ pre všetky $\theta \in \Theta$.

Definícia 13.3 (Konzistentný odhad)**.** Odhad T_n, ktorý dostaneme na základe náhosného výberu rozsahu n, nazývame konzistentný, ak

$$\forall \varepsilon > 0 \ \forall \theta : \lim_{n \to \infty} P_\theta[|T_n - \theta| < \varepsilon] = 1.$$

13.1.1 Metóda maximálnej vierohodnosti

Definícia 13.4. Nech $x_1, \ldots, x_n$ je realizácia náhodného výberu z rozdelenia s $P_\theta[X = x]$ alebo s hustotou $f(x;\theta)$. Odhadom parametra θ metódou maximálnej vierohodnosti rozumieme tú hodnotu $\hat{\theta}$, ktorá maximalizuje vierohodnostnú funkciu

$$L(\theta) = \begin{cases} P_\theta[X_1 = x_1] \cdots P_\theta[X_n = x_n] & \text{pre diskrétne rozdelenie} \\ f(x_1;\theta) \cdots f(x_n;\theta) & \text{pre spojité rozdelenie.} \end{cases}$$

Funkcia vierohodnosti L je teda združená hustota $X_1, \ldots, X_n$ chápaná ako funkcia θ:

$$L(x;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

a

$$\hat{\theta} = \arg\max_{\theta} L(x;\theta)$$

Väčšinou je ľahšie narábať s prirodzeným logaritmom tejto funkcie (tzv. log-likelihood)

$$l(x;\theta) = \ln L(x;\theta)$$

Príklad 13.1. Máme realizáciu $x_1, \ldots, x_n$ z alternatívneho rozdelenia s parametrom $p \in (0,1)$. Odhadnite tento parameter metódou maximálnej vierohodnosti.

$$L(x;p) = P[X_1 = x_1] \cdots P[X_n = x_n] = p^{\sum x_i}(1-p)^{n-\sum x_i}$$

Logaritmus vierohodnostnej funkcie je

$$l(x;p) = \ln L(x;p) = \sum x_i \ln p + (n - \sum x_i) \ln(1-p)$$

Aby sme určili maximum, položíme deriváciu rovnú 0:

$$l'(x;p) = \frac{1}{p}\sum x_i + \frac{1}{1-p}\sum x_i = 0 \Longrightarrow \hat{p} = \frac{1}{n}\sum x_i$$

13.2 Intervalové odhady

Definícia 13.5 (Interval spoľahlivosti)**.** Nech $\Theta \subseteq \mathbb{R}^k$ je parametrický priestor, nech $\{F_\theta : \mathbb{R}^m \to [0,1]; \theta \in \Theta\}$ je parametrická trieda distribučných funkcií a nech $\alpha \in (0,1)$. Nech $g_* : \mathbb{R}^m \to \mathbb{R}$ a $g^* : \mathbb{R}^m \to \mathbb{R}$ sú také funkcie, že $g_* \leq g^*$ na celom $\mathbb{R}^k$ a pre každé $\theta \in \Theta$ platí: Ak náhodný vektor $(X_1, ..., X_m)^T$ má distribučnú funkciu F_θ, tak

$$P[g_*(X_1, ..., X_m) < \theta_i < g^*(X_1, ..., X_m)] = 1 - \alpha$$

Potom povieme, že (náhodný) interval

$$[g_*(X_1, ..., X_m), g^*(X_1, ..., X_m)]$$

je $100(1-\alpha)$-percentným intervalom spoľahlivosti pre parameter θ_i.

13.2.1 Odhady parametrov náhodného výberu z $N(\mu, \sigma^2)$

Veta 13.1. Nech $Y_1, ..., Y_m$ je náhodný výber z rozdelenia $N(\mu, \sigma^2)$, kde $m \geq 2$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$, nech $\bar{Y}$ je výberový priemer a S^2 je výberový rozptyl náhodného výberu $Y_1, ..., Y_m$. Potom $\bar{Y}$ je nevychýlený odhad parametra μ, S^2 je nevychýlený odhad parametra σ^2. Naviac, pre $\alpha \in (0,1)$ je

$$\left[\bar{Y} - \Delta_\mu, \bar{Y} + \Delta_\mu\right], \text{ kde } \Delta_\mu = t_{m-1}(1-\alpha/2)\frac{S}{\sqrt{m}}$$

$100(1-\alpha)$-percentný interval spoľahlivosti pre parameter μ a

$$\left[\frac{(m-1)S^2}{\chi^2_{m-1}(1-\alpha/2)}, \frac{(m-1)S^2}{\chi^2_{m-1}(\alpha/2)}\right]$$

$100(1-\alpha)$-percentný interval spoľahlivosti pre parameter σ^2.

Dôkaz. Veta je dôsledkom vety 12.2 a definícií intervalov spoľahlivosti a kvantilov. □

Poznámka 13.1. Všimnite si, že dĺžka intervalu spoľahlivosti klesá s rastúcim "počtom pozorovaní" m, ale rastie so zmenšujúcou sa hodnotu α, t.j. ak rastú naše požiadavky na spoľahlivosť tohoto intervalu.

Príklad 13.2. V analýze bezpečnosti siete sa testoval čas potrebný na prenos paketov aplikácie v pevnej a bezdrôtovej sieti. Nájdite 95%-ný interval spoľahlivosti pre rozdiel stredných hodnôt, ak máme k dispozícii nasledovné údaje:

	výb.priemer	smer. odchýlka	veľkosť výberu
pevná sieť	2.000	6.299	436
bezdrôtová sieť	11.520	9.939	344

Riešenie:

$$11.520 - 2.000 \pm 1.96\sqrt{\frac{6.299^2}{436} + \frac{9.939^2}{344}} = 9.52 \pm 1.22$$

13.3 Odhady parametrov lineárneho regresného modelu priamkou

Veta 13.2. Nech $Y_1, ..., Y_m$, $m \geq 3$ je postupnosť nezávislých náhodných premenných, pričom $Y_i \sim N(ax_i + b, \sigma^2)$, kde $x_1, ..., x_m \in \mathbb{R}$ nie sú všetky rovnaké, $a, b \in \mathbb{R}$, $\sigma^2 > 0$. Nech $\bar{x} = (1/m)\sum_{i=1}^m x_i$ a nech $x \in \mathbb{R}$. Nech $\hat{a}$, $\hat{b}$ a S^2 sú náhodné premenné definované vo vete 12.3 Potom $\hat{a}$ je nevychýlený odhad parametra a, $\hat{b}$ je nevychýlený odhad parametra b a S^2 je nevychýlený odhad parametra σ^2. Naviac, pre $\alpha \in (0, 1)$ je

$$[\hat{a} - \Delta_a, \hat{a} + \Delta_a], \text{ kde } \Delta_a = \frac{t_{m-2}(1-\alpha/2)S}{\sqrt{\sum_{i=1}^m x_i^2 - m\bar{x}^2}}$$

$100(1-\alpha)$-percentný interval spoľahlivosti pre parameter a,

$$\left[\hat{b} - \Delta_b, \hat{b} + \Delta_b\right], \text{ kde } \Delta_b = t_{m-2}(1-\alpha/2)S\sqrt{\frac{1}{m} + \frac{\bar{x}^2}{\sum_{i=1}^m x_i^2 - m\bar{x}^2}}$$

$100(1-\alpha)$-percentný interval spoľahlivosti pre parameter b,

$$\left[\frac{(m-2)S^2}{\chi^2_{m-2}(1-\alpha/2)}, \frac{(m-2)S^2}{\chi^2_{m-2}(\alpha/2)}\right]$$

$100(1-\alpha)$-percentný interval spoľahlivosti pre parameter σ^2 a pre akékoľvek $x \in \mathbb{R}$ je

$$\left[\hat{a}x + \hat{b} - \Delta_x, \hat{a}x + \hat{b} + \Delta_x\right], \text{ kde}$$

$$\Delta_x = t_{m-2}(1-\alpha/2)S\sqrt{\frac{1}{m} + \frac{(x-\bar{x})^2}{\sum_{i=1}^m x_i^2 - m\bar{x}^2}}$$

$100(1-\alpha)$-percentný interval spoľahlivosti pre hodnotu $ax + b$.

Dôkaz. Veta je dôsledkom vety 12.3 a definícií intervalov spoľahlivosti a kvantilov. □

Poznámka 13.2. Uvažujme značenie z predchádzajúcej vety. Potom množina

$$H = \left\{(x, y)^T \in \mathbb{R}^2 : \hat{a}x + \hat{b} - \Delta_x \leq y \leq \hat{a}x + \hat{b} + \Delta_x\right\}$$

sa nazýva pás spoľahlivosti pre hodnoty $ax + b$. Pás spoľahlivosti je najužší v bode $x = \bar{x}$ a rozširuje sa pre $x \to \infty$ a $x \to -\infty$.

13.4 Cvičenia

Úloha 13.1. 100 náhodne vybraných programátorov sme sa spýtali, či je C++ ich obľúbený programovací jazyk a 15 z nich odpovedalo kladne. Nájdite 95%-ný interval spoľahlivosti pre podiel programátorov, ktorých obľúbeným jazykom je C++.

Úloha 13.2. Do siete pošleme 500 paketov v čase medzi 10 a 11 hodinou a zistíme priemerné oneskorenie 0.8s so smerodajnou odchýlkou 0.1. Pre 200 paketov v čase medzi 22 a 23 hodinou sme zistili priemerné oneskorenie 0.5s so smerodajnou odchýlkou 0.08. Zostrojte 99%-ný interval spoľahlivosti pre rozdiel medzi oneskoreniami.

Úloha 13.3. Merali sme výšku 12 rôznych detí vo vekoch od $x_1 = 18, x_2 = 19, ..., x_{12} = 29$ mesiacov. Predpokladáme, že stredná výška (v danom vekovom rozmedzí) má lineárny trend v závislosti od veku, pričom individuálne odchýlky od strednej výšky majú normálne rozdelenie s rozptylom nezávislým od veku. Formálne: pre dieťa vo veku x_i je nameraná výška y_i realizáciou náhodnej premennej $Y_i = ax_i + b + \epsilon_i$, kde $\epsilon_i \sim N(0, \sigma^2)$. Nájdite regresnú priamku $\hat{a}x + \hat{b}$, kde $\hat{a}$ a $\hat{b}$ sú odhady parametrov a, b metódou najmenších štvorcov. Platí: $\sum_{i=1}^{12} x_i = 282$, $\sum_{i=1}^{12} x_i^2 = 6770$, $\sum_{i=1}^{12} y_i = 961$, $\sum_{i=1}^{12} y_i x_i = 22751,6$.

Úloha 13.4. Meriame rýchlosť pohybujúceho sa objektu v časoch $t_i = i$ sekúnd, pre $i = 1, 2, 3, 4, 5$. Z fyzikálnej podstaty problému a vlastností meracích zariadení vieme, že nameraná rýchlosť objektu v čase t je náhodná premenná s rozdelením $N(at + b, \sigma^2)$, kde a je zrýchlenie objektu, b je jeho počiatočná rýchlosť a σ^2 je konštanta charakterizujúca chybovosť meraní. Naviac vieme, že merania sú navzájom nezávislé. Konkrétne namerané hodnoty rýchlosti boli $v_1 = 99,27$; $v_2 = 108,80$; $v_3 = 119,39$; $v_4 = 128,30$ a $v_5 = 139,88$ m/s. Platí: $\sum_{i=1}^{5} v_i = 595,64$ a $\sum_{i=1}^{5} t_i v_i = 1887,64$. Nájdite odhad $\hat{a}$ zrýchlenia a $\hat{b}$ počiatočnej rýchlosti metódou najmenších štvorcov.

Úloha 13.5. Za účelom testovania špeciálnej váhy sme vykonali 10 nezávislých vážení závažia, ktorého hmotnosť je presne 1 gram. V každom vážení váha ukázala mierne odlišný výsledok. Z výsledkov sme vypočítali výberový priemer 1,004581 gramu a výberový rozptyl 0,000119 g^2. Predpokladáme, že výsledky jednotlivých meraní majú normálne rozdelenie $N(\mu, \sigma^2)$. Vykonajte test hypotézy $H_0 : \mu = 1$ voči $H_1 : \mu \neq 1$ na hladine významnosti $\alpha = 0,1$. Vieme, že 95-percentný kvantil Studentovho rozdelenia s 9-timi stupňami voľnosti je $t_9(0,95) \approx 1,833$.

Kapitola 14

Testovanie štatistických hypotéz

14.1 Všeobecný úvod k testovaniu štatistických hypotéz

Majme daný pravdepodobnostný model na rozdelenie dát $Y_1, ..., Y_m$ vyjadrený (napríklad) systémom distribučných funkcií $\{F(\cdot, \theta); \theta \in \Theta\}$, kde $\Theta \subseteq \mathbb{R}^k$ je parametrický priestor. To znamená, že vieme, že náhodný vektor pozorovaní $\mathbf{Y} = (Y_1, ..., Y_m)^T$ má distribučnú funkciu $F(\cdot, \theta^*)$ pre nejaké fixné, "skutočné", no neznáme $\theta^* \in \Theta$. Uvažujme "nulovú" hypotézu H_0, že skutočné θ^* patrí do množiny $\Theta_0 \subseteq \Theta$ a "alternatívnu" hypotézu H_1, že θ^* nepatrí do Θ_0, čo symbolicky značíme

$$H_0 : \theta \in \Theta_0 \text{ voči } H_1 : \theta \notin \Theta_0$$

Našim cieľom je skonštruovať procedúru (nazývanú štatistický test), ktorej vstupom budú realizácie náhodných premenných $Y_1, ..., Y_m$ a výsledkom bude buď "H_0 nezamietame", alebo "H_0 zamietame", s vlastnosťou:

Ak hypotéza H_0 platí, tak (z pohľadu pred získaním realizácií náhodných premenných $Y_1, ..., Y_m$) nastane výsledok "zamietame H_0" s vopred zvolenou pravdepodobnosťou α.

Pravdepodobnosť α, t.j. pravdepodobnosť, že H_0 zamietneme napriek tomu, že H_0 je platné, nazývame chyba 1. druhu daného testu a test s chybou 1. druhu rovnou α nazývame test "na hladine významnosti" α. Hodnota α sa najčastejšie volí $0,05$. V matematickej štatistike je predpis pre test so zadanou chybou 1. druhu α obvykle takýto:

Ak $g(Y_1, ..., Y_m) \in W_\alpha$, tak výsledok testu je "H_0 zamietame", inak je výsledok testu "H_0 nezamietame".

Pritom funkcia $g : \mathbb{R}^m \to \mathbb{R}$ a množina $W_\alpha \subset \mathbb{R}$ sú zvolené tak, aby platilo

$$P[g(Y_1, ..., Y_m) \in W_\alpha] = \alpha.$$

V takomto prípade nazývame množinu W_α kritickou oblasťou pre štatistiku (funkciu dát) $T = g(Y_1, ..., Y_m)$ a test $H_0 : \theta \in \Theta_0$ voči $H_1 : \theta \notin \Theta_0$ na hladine významnosti α.

Pre danú hypotézu H_0 je obvykle možné skonštruovať veľké množstvo rôznych testov na hladine významnosti α. Takéto testy sa pritom môžu podstatne líšiť a to tým, akú majú takzvanú chybu 2. druhu, t.j. s akou pravdepodobnosťou vrátia výsledok "H_0 nezamietame", ak H_0 neplatí. Chybami druhého druhu sa však nebudeme zaoberať.

Definícia 14.1 (*Kvantil*)**.** Nech X je náhodná premenná s distribučnou funkciou F. Nech G je kvantilová funkcia distribučnej funkcie F a nech $p \in (0,1)$. Potom číslo $G(p)$ nazveme p-kvantilom (alebo $100p$-percentným kvantilom) náhodnej premennej X. Pre náhodnú premennú X s rozdelením χ^2_k a t_k označujeme p-kvantil symbolom $\chi^2_k(p)$, resp. $t_k(p)$.

Ekvivalentná, ale názornejšia definícia kvantilu $\chi^2_k(p)$ je nasledovná: Nech náhodná premenná Z má rozdelenie χ^2_k. Potom $\chi^2_k(p)$ je to (jednoznačne určené) reálne číslo, pre ktoré platí $P[Z < \chi^2_k(p)] = p$. Podobne, ak náhodná premenná T má rozdelenie t_k, tak $t_k(p)$ je to (jednoznačne určené) reálne číslo, pre ktoré platí $P[T < t_k(p)] = p$.
Hodnoty $\chi^2_k(p)$ alebo $t_k(p)$ obvykle získame z tabuliek, alebo pomocou vhodného štatistického (numerického) programu. Tieto hodnoty nie je vo všeobecnosti jednoduché vypočítať.

14.2 Testovanie hypotéz o strednej hodnote a disperzii

Pre model $Y_1, ..., Y_m$ náhodného výberu z normálneho rozdelenia $N(\mu, \sigma^2)$ môžeme testovať hypotézy o strednej hodnote μ aj disperzii σ^2 pomocou nasledovnej vety.

Veta 14.1. Nech $Y_1, ..., Y_m$ je náhodný výber z rozdelenia $N(\mu, \sigma^2)$, kde $m \geq 2$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$, $\bar{Y}$ je výberový priemer a S^2 je výberový rozptyl náhodného výberu $Y_1, ..., Y_m$ a nech $\alpha \in (0,1)$.

Nech $\mu_0 \in \mathbb{R}$. Pre test hypotézy

$$H_0 : \mu = \mu_0 \text{ voči } H_1 : \mu \neq \mu_0$$

na hladine významnosti α, je množina

$$W_\alpha = (-\infty, -t_{m-1}(1-\alpha/2)) \cup (t_{m-1}(1-\alpha/2), \infty)$$

kritickou oblasťou štatistiky

$$\frac{\bar{Y} - \mu_0}{S}\sqrt{m}$$

Nech $\sigma_0^2 > 0$. Pre test hypotézy

$$H_0 : \sigma^2 = \sigma_0^2 \text{ voči } H_1 : \sigma^2 > \sigma_0^2$$

na hladine významnosti α, je množina

$$W_\alpha = \left(\chi^2_{m-1}(1-\alpha), \infty\right)$$

kritickou oblasťou štatistiky

$$\frac{(m-1)S^2}{\sigma_0^2}$$

Dôkaz. Veta je dôsledkom vety 12.2, definície štatistického testu a kvantilov. $\square$

14.3 Testovanie hypotézy o rozdiele stredných hodnôt

Pre dva nezávislé náhodné výbery z normálneho rozdelenia sa najčastejšie testuje hypotéza týkajúca sa rozdielu $\mu_1 - \mu_2$ stredných hodnôt, obzvlášť hypotéza $\mu_1 - \mu_2 = 0$ že náhodné výbery nevykazujú štatisticky významný rozdiel stredných hodnôt.

Veta 14.2. Uvažujme rovnaké predpoklady a značenie ako vo vete 12.6.
Nech $\alpha \in (0, 1)$ a $\Delta_0 \in \mathbb{R}$. Pre test hypotézy

$$H_0 : \mu_1 - \mu_2 = \Delta_0 \text{ voči } H_1 : \mu_1 - \mu_2 \neq \Delta_0$$

na hladine významnosti α, je množina

$$W_\alpha = (-\infty, -t_{m-2}(1 - \alpha/2)) \cup (t_{m-2}(1 - \alpha/2), \infty)$$

kritickou oblasťou štatistiky

$$T = \frac{\bar{Y}_1 - \bar{Y}_2 - \Delta_0}{S\sqrt{\frac{1}{m_1} + \frac{1}{m_2}}}.$$

Dôkaz. Veta je priamym dôsledkom vety 12.6, definície štatistického testu a kvantilov. □

14.4 Testovanie hypotézy o sklone regresnej priamky

Ak modelujeme pozorovania $Y_1, ..., Y_m$ lineárnym regresným modelom priamkou (definícia 12.5), môžeme v princípe testovať mnoho hypotéz týkajúcich sa parametrov a, b, alebo disperzie σ^2 odchýlok. Avšak obvykle sa testuje hypotéza o sklone regresnej priamky, t.j. o parametri a. Obzvlášť často sa testuje konkrétne hypotéza $a = 0$, t.j. hypotéza, že pozorovania Y nevykazujú štatisticky významný rastúci (alebo klesajúci) trend pri zmene hodnôt x.

Veta 14.3. Uvažujme rovnaké predpoklady a značenie ako vo vete 12.3.
Nech $\alpha \in (0, 1)$ a $a_0 \in \mathbb{R}$. Pre test hypotézy

$$H_0 : a = a_0 \text{ voči } H_1 : a \neq a_0$$

na hladine významnosti α, je množina

$$W_\alpha = (-\infty, -t_{m-2}(1 - \alpha/2)) \cup (t_{m-2}(1 - \alpha/2), \infty)$$

kritickou oblasťou štatistiky

$$\frac{\hat{a} - a_0}{S}\sqrt{\sum_{i=1}^{m} x_i^2 - m\bar{x}^2}.$$

Dôkaz. Veta je dôsledkom vety 12.3, definície štatistického testu a kvantilov. □

14.5 Cvičenia

Úloha 14.1. V roku 1882 vykonal Michelson 23 nezávislých meraní rýchlosti svetla, pričom priemer nameraných hodnôt bol $\bar{x} = 299756,2$ km/s a výberový rozptyl meraní vyšiel $S^2 = 11473,54$. Môžeme predpokladať, že merania zodpovedali realizáciám nezávislých náhodných premenných s normálnym rozdelením $N(\mu, \sigma^2)$. Na hladine významnosti $\alpha = 0,05$ testujte hypotézu, že Michelsonove merania neboli zaťažené výchylkou strednej hodnoty, t.j. testujte, že $\mu = 299792,5$ km/s, čo je "presná" rýchlosť svetla určená modernými metódami. Vieme, že $t_{22}(0,975) = 2,074$.

Príklad 14.1. Z veľkého súboru rezistorov rovnakého typu a nominálnej hodnoty sme vybrali 16 kusov. Na základe dlhodobých skúseností môžeme predpokladať, že v základnom súbore majú hodnoty odporu rezistorov v $k\Omega$ rozdelenie $N(\mu, \sigma^2)$, avšak μ a σ^2 sú neznáme. Výberový priemer odporu vybratých rezistorov je $9,3$ $k\Omega$ a výberový rozptyl $6,25$ $(k\Omega)^2$.
a) Určte realizáciu 99-percentného intervalu spoľahlivosti pre parameter μ.
b) Určte realizáciu 99-percentného intervalu spoľahlivosti pre parameter σ^2.
c) Na hladine významnosti $\alpha = 0,1$ testujte hypotézu, že $\mu = 10$ $k\Omega$.
d) Na hladine významnosti $\alpha = 0,1$ testujte hypotézu, že $\sigma^2 = 4$ $(k\Omega)^2$.
Potrebné kvantily sú nasledovné: $t_{15}(0,995) = 2,95$, $t_{15}(0,95) = 1,75$, $\chi^2_{15}(0,995) = 32,80$, $\chi^2_{15}(0,95) = 25,00$, $\chi^2_{15}(0,005) = 4,60$, $\chi^2_{15}(0,05) = 7,26$.

Literatúra

[1] N.Alon, J.H.Spencer: The Probabilistic Method, Wiley-Interscience, 2008

[2] T.M.Cover, J.A.Thomas: Elements of Information Theory, Wiley-Interscience, 2006

[3] M.Mitzenmacher, E.Upfal: Probability and Computing, Cambridge University Press, 2005

[4] R.Motwani, P.Raghavan: Randomized Algorithms, Cambridge University Press, 1995

[5] S.M.Ross: Simulation, Academic Press, 2012