Matematická statistika

Tomáš Hobza

2. února 2011

Obsah

Literatura

[1] V. Rogalewicz: *Pravděpodobnost a statistika pro inženýry.* Skripta ČVUT-FEL, Praha, 2000.

[2] W. Sadowski: *Matematická statistika.* ALFA, Bratislava, 1975.

[3] J. Likeš, J. Machek: *Počet pravděpodobnosti.* SNTL, Praha, 1981.

[4] J. Likeš, J. Machek: *Matematická statistika.* SNTL, Praha, 1988.

[5] A. A. Svešnikov: *Sbírka úloh z teorie pravděpodobnosti, matematické statistiky a teorie náhodných funkcí.* SNTL, Praha, 1971.

[6] V. Dupač, M. Hušková: *Pravděpodobnost a matematická statistika.* UK - Nakladatelství Karolinum, Praha, 2003.

Kapitola 1

Úvod

Teorie pravděpodobnosti řeší následující problém: jistý jev má řadu různých následků (výsledků) a náhoda určuje, který z nich skutečně nastane. Teorie pravděpodobnosti dává návod, jak jednotlivé výsledky „předem“ kvantifikovat.

Matematická statistika řeší jistém smyslu opačnou situaci. Pozoruji nějaký jev, který mohl mít řadu příčin. Matematická statistika dává metody jak tyto příčiny třídit a kvantifikovat, např. kterou vybrat jako nejpravděpodobnější.

1.1 Klasická definice pravděpodobnosti

- uvažujme náhodný pokus, který může vykázat konečný počet n vzájemně se vylučujících výsledků (např. tři hody kostkou)
- předpokládáme, že všechny výsledky jsou stejně pravděpodobné (symetrie, homogenita)
- jestliže m z těchto výsledků má za následek realizaci jevu A (např. padnou dvě šestky) a zbylých $n - m$ výsledků tuto realizaci vylučuje, potom pokládáme pravděpodobnost jevu A rovnu

$$P(A) = \frac{m}{n} = \frac{\text{počet případů příznivých}}{\text{počet případů možných}}$$

(Laplace 1749-1829)

Nedokonalosti této definice:

a) jak spočítat pravděpodobnost na nesymetrické kostce?

b) co když mám nekonečně mnoho možností?

Nemůže být považována za definici v matematickém slova smyslu, ale v určitých případech může sloužit jako algoritmus výpočtu.

Příklad 1.1 Obarvená krychle je rozřezána na 1000 krychliček. Určete pravděpodobnost, že náhodně vybraná krychlička bude mít obarvené dvě stěny.

Celkový počet krychliček je $n = 1000$. Původní rychle měla 12 hran, na každé hraně 8 krychliček s dvěmi obarvenými stěnami. Takže počet případů příznivých je $m = 12 \cdot 8 = 96$. Hledaná pravděpodobnost p je tedy

$$p = \frac{96}{1000} = 0.096\,.$$

Příklad 1.2 Spočtěte pravděpodobnost, že při hodu dvěma kostkami bude součet výsledků a) 9, b) 10.

Součet 9 lze získat jako kombinaci výsledků 3 a 6 nebo 4 a 5. Součet 10 lze získat jako kombinaci výsledků 5 a 5 nebo 4 a 6. Takže na první pohled by se mohlo zdát, že $P(9) = P(10)$. Ale počet všech případů možných je $n = 36$ a případy příznivé jednotlivým součtům jsou

9: 3,6 6,3 4,5 5,4

10: 5,5 6,4 4,6

a tedy

$$P(9) = \frac{4}{36} = \frac{1}{9} \quad \text{a} \quad P(10) = \frac{3}{36} = \frac{1}{12}\,.$$

Pozor na to, co jsou „stejně možné“ výsledky!!!

Počty možných a příznivých výsledků se počítají pomocí kombinatorických metod.

1.1.1 Základní kombinatorické vzorce

Nechť je dána skupina A obsahující n různých prvků očíslovaných čísly $1, 2, \dots n$.

Uspořádaný výběr (variace) s opakováním: z A se vybírá k-krát po sobě po jednom prvku, vybraný prvek se vždy vrací. Počet všech různých k-tic $(i_1, \dots, i_k)$, které lze takto utvořit je

$$V'(k, n) = n^k\,.$$

Uspořádaný výběr (variace) bez opakování: z A se vybírá k-krát po jednom prvku, vybrané prvky se nevracejí. Počet všech různých k-tic $(i_1, \dots, i_k)$, které lze takto utvořit je

$$V(k, n) = n(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}\,.$$

Permutace: skupinu A lze uspořádat v posloupnosti $(i_1, \dots, i_n)$

$$P(n) = n!$$

způsoby.

Neuspořádaný výběr (kombinace) bez opakování: počet různých podmnožin o k-prvcích, které lze vybrat z A je

$$K(k,n) = \binom{n}{k} = \frac{n!}{k!(n-k)!}\,.$$

Neuspořádaný výběr (kombinace) s opakováním: počet neuspořádaných k-tic sestavených z n prvků tak, že se v ní každý vyskytuje nejvýše k-krát je

$$K'(k,n) = \binom{n+k-1}{k}\,.$$

Příklad 1.3 Náhodný pokus spočívá v 6 hodech kostkou. Kostka je pravidelná a hází se bez snahy o určitý výsledek. Jaká je pravděpodobnost, že nepadne ani jednou šestka?

Pro řešení lze použít např. variace s opakováním. Označme A jev, že nepadne šestka. Všech možných šestic $(i_1,\dots,i_6)$, kde i_j značí výsledek j-tého hodu je $n = 6^6$. Případů, kdy nepadne žádná šestka, je $m = 5^6$ a tedy

$$P(A) = \frac{5^6}{6^6} = \left(\frac{5}{6}\right)^6 = 0.3349\,.$$

Pravděpodobnost, že padne šestka alespoň jednou je potom

$$1 - P(A) = 0.6651\,.$$

Jaká je pravděpodobnost, že šestka padne právě jednou? Označme tento jev symbolem B. Příznivé jevu B jsou šestice $(i_1,\dots,i_6)$ ve kterých je právě na jedné pozici šestka a na ostaních jsou čísla od 1 do 5. Takových šestic je $m = 6\cdot 5^5$ (5^5 skupin tvaru $(6,i_2,\dots,i_6)$, $(i_1,6,i_3,\dots,i_6)$, $\dots(i_1,\dots,i_5,6)$) a tedy

$$P(B) = \frac{6\cdot 5^5}{6^6} = \left(\frac{5}{6}\right)^5 = 0.4019\,.$$

Jev, že šestka padne nejvýše jednou je sjednocením jevů jevů A a B a jeho pravděpodobnost je

$$P(A) + P(B) = 0.3349 + 0.4019 = 0.7368\,.$$

Konečně pravděpodobnost. že šestka padne více než jednou je

$$1 - 0.7368 = 0.2632\,.$$

Příklad 1.4 $2n$ sportovních mužstev bylo rozděleno do dvou podskupin. Určete pravděpodobnost toho, že se dvě nejsilnější mužstva dostanou

a) do různých podskupin

b) do jedné podskupiny.

Na problém můžeme nahlížet tak, že vybíráme mužstva do první podskupiny. Družstva druhé podskupiny pak už jsou určena jednoznačně. Počty případů příznivých a možných se pak dají vyjádřit

a)

$$\frac{\binom{2}{1}\binom{2n-2}{n-1}}{\binom{2n}{n}} = \frac{2\,\frac{(2n-2)!}{(n-1)!(n-1)!}}{\frac{(2n)!}{n!n!}} = \frac{n}{2n-1}$$

b)

$$\frac{\binom{2}{1}\binom{2n-2}{n-2}+\binom{2}{0}\binom{2n-2}{n}}{\binom{2n}{n}} = \frac{2\,\frac{(2n-2)!}{(n-2)!n!}}{\frac{(2n)!}{n!n!}} = \frac{n-1}{2n-1}\,.$$

Protože se jedná o komplementární jevy, měl by pro kontrolu součet výsledků dávat jedničku.

Cvičení 1.1 V bedně je 30 výrobků, z nichž 3 jsou vadné. Jaká je pravděpodobnost jevu A, že mezi 5 náhodně vybranými výrobky bude nejvýše jeden vadný? [0.936]

Cvičení 1.2 Vybíráme z pěti vstupenek po 10 Kč, tří vstupenek po 30 Kč a dvou vstupenek po 50 Kč. Vylosujeme-li náhodně 3 vstupenky, a jakou pravděpodobností bude jejich cena 70 Kč? [0.2917]

Cvičení 1.3 Na jedné poličce je náhodně rozestaveno 10 knih. Určete pravděpodobnost toho, že určité 3 knihy jsou postaveny vedle sebe. [1/15]

1.2 Geometrická pravděpodobnost

Geometrická pravděpodobnost představuje jisté zobecnění klasické definice, umožňuje totiž uvažovat nekonečně mnoho možných výsledků. Myšlenka je stejná, pouze „počet“ je nahrazen jinou veličinou.

Je-li výsledek stejně pravděpodobný v každém bodě nějakého geometrického objektu (pravděpodobnost je na tomto objektu rovnoměrně rozložena), pak pravděpodobnost nějakého jevu B vypočteme jako podíl velikosti (obsah, objem) množiny S_B všech bodů příznivých jevu B a velikosti celého objektu S,

$$P(B) = \frac{S_B}{S}\,.$$

Příklad 1.5 Dva vlaky mají přijet na vykládku k rampě. Každý z vlaků může přijet nezávisle na druhém kdykoliv během jednoho dne (každý čas příjezdu považujeme za stejně možný). S jakou pravděpodobností bude muset jeden z vlaků čekat, jestliže vykládka prvního vlaku trvá 2 hodiny a vykládka druhého trvá 3 hodiny?

Označme

T_1 čas příjezdu prvního vlaku

T_2 čas příjezdu druhého vlaku

B jev, že jeden z vlaků bude muset čekat

Časy příjezdů T_1, T_2 jsou v našem případě nezávislé, rovnoměrně rozložené na $\langle 0, 24\rangle$ a situaci tedy můžeme geometricky reprezentovat jako čtverec (viz. obr.).
První vlak bude muset čekat pokud

$$T_2 < T_1 < T_2 + 3$$

a druhý vlak bude čekat pokud

$$T_1 < T_2 < T_1 + 2\,.$$

Příslušné oblasti jsou graficky znázorněny na obrázku

Obr.

Z principu grafické pravděpodobnosti tedy vyplývá, že

$$P(B) = 1 - \frac{1}{24^2}\left(22 \cdot 11 + \frac{21^2}{2}\right) = 1 - \frac{1}{576}\left(242 + \frac{441}{2}\right) = 0.197\,.$$

Příklad 1.6 Určete pravděpodobnost, že kořeny kvadratické rovnice

$$x^2 + 2ax + b = 0$$

jsou reálné, jsou-li stejně možné hodnoty koeficientů a, b v obdélníku $|a| \leq n,\ |b| \leq m$.

Kořeny zadané rovnice budou reálné pokud její diskriminant bude větší nebo roven nule, tedy

$$D = 4a^2 = 4b \geq 0 \quad \text{neboli} \quad a^2 \geq b\,.$$

Situace je znázorněna na obrázku

Obr.

Z obrázku je vidět, že budeme muset rozdělit řešení na dva případy:

a) $\sqrt{m} < n$

V tomto případě je hledaná pravděpodobnost rovna

$$\frac{1}{4mn}\left[2mn + 2(n - \sqrt{m})m + 2\int_0^{\sqrt{m}} a^2\, da\right] = \frac{1}{4mn}\left[4mn - \frac{4}{3}m^{\frac{3}{2}}\right] = 1 - \frac{\sqrt{m}}{3n}\,.$$

b) $\sqrt{m} \geq n$

Pro tuto možnost dostáváme

$$\frac{1}{4mn}\left[2mn + 2\int_0^{n} a^2\, da\right] = \frac{1}{4mn}\left[2mn + \frac{2}{3}n^3\right] = \frac{1}{2} + \frac{n^2}{6m}\,.$$

Kapitola 2

Matematický model pravděpodobnosti

2.1 Náhodné jevy

Nyní budeme uvažovat obecný případ, kdy se nepředpokládá, že jednotlivé výsledky pokusu jsou stejně pravděpodobné. Dříve než budeme diskutovat pravděpodobnosti jevů, musíme probrat jevy samotné.

S každým pokusem nebo hrou je spojena množina všech možných výsledků. Označme symbolem Ω množinu všech možných, navzájem se vylučujících výsledků. Množinu Ω budeme nazývat *základní pravděpodobnostní prostor* (někdy také *jistý jev*) Libovolný možný výsledek značíme $\omega \in \Omega$, ω je v jistém smyslu nejjednodušší výsledek, budeme ho nazývat *elementární jev*.

Příklad 2.1 Uvažujme dva hody mincí. Potom je množina

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$$

všech možných navzájem se vylučujících výsledků tvořena čtyřmi elementárními jevy

$$\omega_1 = (H, H), \quad \omega_2 = (H, O), \quad \omega_3 = (O, H), \quad \omega_4 = (O, O),$$

kde symboly H a O značí, zda při hodu mincí padne hlava nebo orel.

Jevem budeme prozatím rozumět sjednocení jistých elementárních jevů. Jevy budeme značit velkými písmeny $A, B, C, \ldots$. Libovolný jev A je tedy podmnožinou množiny Ω ($A \subset \Omega$).

Příklad 2.2 Uvažujme situaci z předchozího příkladu a nechť A značí jev, že hlava padne v prvním hodu, B jev, že padne alespoň jedna hlava a C jev, že padnou alespoň tři hlavy.

Jev A je sjednocením dvou elementárních jevů a to

$$A = \{\omega_1, \omega_2\}.$$

Jev B se skládá z elementárních jevů

$$B = \{\omega_1, \omega_2, \omega_3\}$$

a jev C nemůže nastat. Takový jev budeme nazývat *nemožný jev* a značit $C = \emptyset$.

Jestliže nastane elementární jev ω a jestliže $\omega \in A$ budeme říkat, že *nastal jev* A.

Poznámka 2.1 a) jestliže nastane jev A neznamená to, že nemohl nasta také jiný jev. V předcházejícím příkladu, nastane-li $\omega_1 = (H, H)$, potom nastávají jevy A i B.

b) Ω se nazývá jistý jev, protože ať nastane jakýkoliv elementární jev ω, vždy $\omega \in \Omega$.

2.1.1 Operace s náhodnými jevy

Komplementární (opačný) jev A^C k jevu A se nazývá jev, který nastane právě tehdy, když nenastane jev A. Značíme

$$A^C = \Omega \setminus A\,.$$

Sjednocení jevů $A_1, \ldots, A_n$ se nazývá jev, který nastane právě tehdy, když nastane alespoň jeden z jevů $A_1, \ldots, A_n$. Značíme

$$A_1 \cup A_2 \cup \ldots \cup A_n \quad \text{nebo} \quad \bigcup_{j=1}^{n} A_j\,.$$

Průnikem jevů $A_1, \ldots, A_n$ nazýváme jev nastávající právě tehdy, když nastávají všechny jevy $A_1, \ldots, A_n$. Značíme

$$A_1 \cap A_2 \cap \ldots \cap A_n \quad \text{nebo} \quad \bigcap_{j=1}^{n} A_j\,.$$

Někdy budeme též používat zkrácený zápis $A \cap B = AB$.

Disjunktní (neslučitelné) nazýváme takové jevy A a B, které nemohou nastat současně, tj. jestliže $A \cap B = \emptyset$.
Pokud jsou jevy A a B disjunktní, budeme značit

$$A \cup B = A + B \quad \text{a} \quad \bigcup_{j=1}^{n} A_j = \sum_{j=1}^{n} A_j\,.$$

Říkáme, že jev A **implikuje jev** B (A má za následek B) jestliže jev B nastane vždy, když nastane jev A. Značíme

$$A \subset B\,.$$

Příklad 2.3 Označme jev A: dožiju se 70 let a jev B: dožiju se 40 let. Potom jev A implikuje jev B, $A \subset B$.

Věta 2.1 *Pro všechny jevy $A, B, C \subset \Omega$ platí:*

1. $A \subset A$

2. jestliže $A \subset B$ *a* $B \subset C$ *potom* $A \subset C$

3. $A \cap A = A, \quad A \cup A = A$

4. $A \cup B = B \cup A, \quad A \cap B = B \cap A$

5. $A \cap (B \cap C) = (A \cap B) \cap C, \quad A \cup (B \cup C) = (A \cup B) \cup C$

6. $\emptyset \subset A \subset \Omega$

7. $A \cap B \subset A \subset A \cup B$

8. $\emptyset \cap A = \emptyset, \quad \emptyset \cup A = A$

9. $\Omega \cap A = A, \quad \Omega \cup A = \Omega$

10. $(A^C)^C = A$

11. $(A \cup B)^C = A^C \cap B^C \quad a \quad (A \cap B)^C = A^C \cup B^C$ *(de Morganovy vztahy)*

12. $A \cup B = A + B \cap A^C$

13. $B = A \cap B + A^C \cap B = AB + A^C B$.

Důkaz. Dokažme například druhý de Morganův zákon $(A \cap B)^C = A^C \cup B^C$. Máme dokázat rovnost dvou množin, takže ukážeme dvě inkluze.
$\subset$:

$$\omega \in (A \cap B)^C \quad \Rightarrow \quad \omega \notin A \cap B \quad \Rightarrow \quad \omega \notin A \vee \omega \notin B \quad \Rightarrow$$
$$\Rightarrow \quad \omega \in A^C \vee \omega \in B^C \quad \Rightarrow \quad \omega \in A^C \cup B^C .$$

$\supset$: Tato inkluze se ukáže úplně stejně opačným postupem. □

Cvičení 2.1 Hodíme třikrát mincí. Vyjmenujte výsledky příznivé následujícím jevům:

A - padne právě jeden líc

B - padne nejvýš jednou rub

C - padne třikrát tatáž strana mince

D - padne alespoň dvakrát rub .

2.2 Axiomatická definice pravděpodobnossti

Připomeňme, že symbolem Ω jsem označili množinu všech možných vzájemně se vylučujících výsledků a její prvky ω jsme nazývali elementární jevy.

Definice 2.1 *Nechť je dána neprázdná množina* Ω *a systém* $\mathcal{A}$ *podmnožin množiny* Ω*, pro který platí*

1. $\Omega \in \mathcal{A}$

2. je-li $A \in \mathcal{A}$, potom $A^C = \Omega \setminus A \ \in \ \mathcal{A}$

3. jestliže $A_1, A_2, A_3, \ldots \in \mathcal{A}$, potom i $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Potom systém $\mathcal{A}$ nazveme σ-algebrou a prvky systému $\mathcal{A}$ jevy (náhodné jevy). Dvojici $(\Omega, \mathcal{A})$ nazýváme pozorovací prostor.

Zabývejme se teď základními vlastnostmi σ-algeber.

Tvrzení 2.1 *Nemožný jev $\emptyset$ je vždy prvkem σ-algebry $\mathcal{A}$, tj. $\emptyset \in \mathcal{A}$.*

Důkaz. Protože platí $\emptyset^C = \Omega \in \mathcal{A}$ dostáváme s použitím bodu b) z definice

$$\emptyset = (\emptyset^C)^C \in \mathcal{A}\,.$$

□

Tvrzení 2.2 *Jestliže $A_1, \ldots, A_n$ je konečná posloupnost jevů z $\mathcal{A}$, potom i*

$$\bigcup_{i=1}^{n} A_i \in \mathcal{A}\,.$$

Důkaz. Dodefinujme jevy

$$A_{n+1} = A_{n+2} = \ldots = \emptyset,$$

podle předchozího tvrzení víme, že $A_i \in \mathcal{A}$ i pro všechna $i \geq n+1$ a vlastnost 3) z Definice 2.1 implikuje

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}\,.$$

□

Tvrzení 2.3 *Jestliže $A_1, A_2, \ldots$ je posloupnost jevů z $\mathcal{A}$, potom i*

$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}\,.$$

Důkaz. Pro všechna $i \in I\!N$ platí podle bodu 2) z Definice 2.1

$$A_i^C \in \mathcal{A}$$

a tedy podle bodu 3) téže definice

$$\bigcup_{i=1}^{\infty} A_i^C \in \mathcal{A}\,.$$

Použitím de Morganova vztahu dostaneme

$$\left(\bigcup_{i=1}^{\infty} A_i^C\right) = \left(\bigcap_{i=1}^{\infty} A_i\right)^C \in \mathcal{A}$$

a podle bodu 2) tedy

$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}\,.$$

□

Poznámka 2.2 Volba systému $\mathcal{A}$ závisí na konkrétní situaci. Obecně nemusíme brát systém všech podmnožin množiny Ω.

Příklad 2.4 Pro hod kostkou je množina všech možných výsledků $\Omega = \{1, 2, 3, 4, 5, 6\}$ a za systém $\mathcal{A}$ můžeme vzít množinu všech podmnožin množiny Ω, $\mathcal{A} = \exp \Omega$. Pokud by nás ale například zajímalo pouze zda padne sudé nebo liché číslo, potom lze systém $\mathcal{A}$ volit jednodušeji

$$\mathcal{A} = \{\emptyset, \Omega, \{1, 3, 5\}, \{2, 4, 6\}\} .$$

Definice 2.2 *Jestliže $\Omega = \mathbb{R}$, pak nejmenší systém množin, který obsahuje všechny intervaly v $\mathbb{R}$ nazveme systémem (σ-algebrou) borelovských množin a značíme $\mathcal{B}(\mathbb{R})$ (zkráceně $\mathcal{B}$).*

Příklad 2.5 Jednodušší konečná σ-algebra na $\mathbb{R}$ může být generována konečným systémem intervalů pokrývajících $\mathbb{R}$. OBR:

Definice 2.3 *Nechť je dán pozorovací prostor $(\Omega, \mathcal{A})$. Funkci $\mathsf{P} : \mathcal{A} \to \mathbb{R}$ splňující podmínky*

1) $\mathsf{P}(\Omega) = 1$

2) $\mathsf{P}(A) \geq 0$ *pro všechny* $A \in \mathcal{A}$

3) $\mathsf{P}\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathsf{P}(A_k)$ *pro každou posl. disjunktních mn.* $A_1, A_2, \ldots \in \mathcal{A}$

nazýváme pravděpodobnostní mírou (pravděpodobností) na $\mathcal{A}$ a trojici $(\Omega, \mathcal{A}, \mathsf{P})$ pravděpodobnostním prostorem.

Věta 2.2 *Pravděpodobnost nemožného jevu je nulová,*

$$\mathsf{P}(\emptyset) = 0 .$$

Důkaz. Definujme jevy

$$A_1 = A_2 = A_3 = \ldots = \emptyset .$$

Protože všechny tyto jevy A_i jsou disjunktní, z bodu 3) Definice 2.3 plyne

$$\mathsf{P}(\emptyset) = \mathsf{P}\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathsf{P}(A_k) = \sum_{k=1}^{\infty} \mathsf{P}(\emptyset) ,$$

což může být splněno pouze tehdy, když $\mathsf{P}(\emptyset) = 0$. □

Věta 2.3 *Nechť $A_1, \ldots A_n \in \mathcal{A}$ jsou disjunktní jevy. Potom*

$$\mathsf{P}(A_1 + A_2 + \ldots + A_n) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \ldots + \mathsf{P}(A_n) .$$

Důkaz. Definujme jevy $A_{n+1} = A_{n+2} = \ldots = \emptyset$. Potom jsou všechny jevy $A_i, i \in \mathbb{N}$ disjunktní a platí

$$\mathsf{P}\left(\sum_{k=1}^{n} A_k\right) = \mathsf{P}\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathsf{P}(A_k) = \sum_{k=1}^{n} \mathsf{P}(A_k) ,$$

neboť podle předchozí věty platí $\mathsf{P}(\emptyset) = 0$. □

Věta 2.4 *Pro libovolné jevy A a B platí*

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)\,.$$

Důkaz. Pokud rozložíme jev $A \cup B$ na disjunktní sjednocení dvou jevů

$$A \cup B = A + A^C \cap B\,, \quad \text{(viz. Věta 2.1)}$$

podle předchozí věty dostaneme

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(A^C \cap B)\,. \tag{2.1}$$

Podobně rozložme jev B na $B = A \cap B + A^C \cap B$, potom platí

$$\mathsf{P}(B) = \mathsf{P}(A \cap B) + \mathsf{P}(A^C \cap B) \quad \text{neboli} \quad \mathsf{P}(A^C \cap B) = \mathsf{P}(B) - \mathsf{P}(A \cap B).$$

Dosazením do vztahu (2.1) dostáváme tvrzení věty. □

Důsledek. Pro libovolné jevy A a B platí

$$\mathsf{P}(A \cup B) \leq \mathsf{P}(A) + \mathsf{P}(B)\,.$$

Věta 2.5 *Jestliže $A \subset B$, potom*

$$\mathsf{P}(A) \leq \mathsf{P}(B)\,.$$

Důkaz. Použijeme opět rozklad $B = A \cap B + A^C \cap B$. Z předpokladů věty ovšem vyplývá, že $A \cap B = A$ a tedy

$$\mathsf{P}(B) = \mathsf{P}(A) + \mathsf{P}(A^C \cap B) \geq \mathsf{P}(A)\,.$$

□

Důsledek. Pro libovolný jev A platí

$$\mathsf{P}(A) \leq 1\,.$$

Důkaz. Protože každý jev splňuje $A \subset \Omega$, dostáváme $\mathsf{P}(A) \leq \mathsf{P}(\Omega) = 1$. □

Věta 2.6 *Pro každý jev A platí*

$$\mathsf{P}(A^C) = 1 - \mathsf{P}(A)\,.$$

Důkaz. Jistý jev můžeme rozložit na disjunktní sjednocení $\Omega = A + A^C$ a podle Věty 2.3 tedy platí

$$1 = P(A) + \mathsf{P}(A^C)\,.$$

□

Příklad 2.6 Tři hráči házejí postupně mincí. Vyhrává ten, komu padne dříve líc. určete pravděpodobnost výhry každého hráče.

Označme symbolem A_i jev, že první hráč vyhraje ve svém i-tém tahu a symbolem A jev, že vyhraje první hráč. Potom

$$\mathsf{P}(A) = \mathsf{P}\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathsf{P}(A_i) \tag{2.2}$$

protože jevy A_i jsou disjunktní (nemohou nastat současně). Dále platí

$$\mathsf{P}(A_1) = \frac{1}{2}, \quad \mathsf{P}(A_2) = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}, \quad \mathsf{P}(A_3) = \left(\frac{1}{2}\right)^7, \ldots$$

a obecně tedy můžeme psát

$$\mathsf{P}(A_i) = \left(\frac{1}{2}\right)^{3i-2}.$$

Dosazením do (2.2) dostáváme

$$\mathsf{P}(A) = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^{3i-2} = 4\sum_{i=1}^{\infty} \frac{1}{2^{3i}} = \frac{4}{7}\,.$$

Podobně pro jevy B_i, že vyhraje druhý hráč ve svém i-tém tahu a B, že vyhraje druhý hráč, platí

$$\mathsf{P}(B) = \sum_{i=1}^{\infty} \mathsf{P}(B_i) = \sum_{i=1}^{\infty} \frac{1}{2}\mathsf{P}(A_i) = \frac{2}{7}$$

a pravděpodobnost výhry třetího hráče tedy je 1/7.

Kapitola 3

Závislé a nezávislé jevy

Představme si, že už máme nějakou informaci o pokusu a zajímá nás pravděpodobnost nějakého jevu na základě této informace. Např.

- pravděpodobnost, že student dokončí studium, když dokončil první ročník
- pravděpodobnost, že student dokončí studium, když obhájil výzkumný úkol.

Je zřejmé, že odpovědi na tyto otázky nebudou stejné. Takovéto situace budeme schopni popisovat pomocí podmíněné pravděpodobnosti.

Definice 3.1 *Nechť jsou dány jevy $A, B \in \mathcal{A}$ takové, že $\mathsf{P}(B) > 0$. Podmíněnou pravděpodobností jevu A za podmínky, že nastal jev B, nazveme číslo*

$$\mathsf{P}(A|B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}\,. \tag{3.1}$$

Věta 3.1 *Je-li $\mathsf{P}(B) > 0$, pak $\mathsf{P}(.|B)$ je pravděpodobnost na pozorovacím prostoru $(\Omega, \mathcal{A})$.*

Důkaz. Na základě předpokladů věty a definice podmíněné pravděpodobnosti máme ukázat tři vlastnosti pravděpodobnostní míry.
1)

$$\mathsf{P}(\Omega|B) = \frac{\mathsf{P}(\Omega \cap B)}{\mathsf{P}(B)} = \frac{\mathsf{P}(B)}{\mathsf{P}(B)} = 1\,.$$

2) Pro každý jev $A \in \mathcal{A}$ platí

$$\mathsf{P}(A|B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)} \geq 0\,.$$

3) Pro libovolnou posloupnost $A_1, A_2, \ldots, A_n, \ldots$ disjunktních jevů z $\mathcal{A}$ platí

$$\mathsf{P}\left(\sum_{i=1}^{\infty} A_i \middle| B\right) = \frac{\mathsf{P}\left(\left(\sum_{i=1}^{\infty} A_i\right) \cap B\right)}{\mathsf{P}(B)} = \frac{\mathsf{P}\left(\sum_{i=1}^{\infty}(A_i \cap B)\right)}{\mathsf{P}(B)} =$$

$$\frac{\sum_{i=1}^{\infty} \mathsf{P}(A_i \cap B)}{\mathsf{P}(B)} = \sum_{i=1}^{\infty} \mathsf{P}(A_i|B)\,,$$

kde jsme v předposledním kroku využili toho, že i jevy $A_i \cap B,\ i \in \mathbb{N}$ jsou disjunktní.
□

Příklad 3.1 Na základě výzkumu dědičnosti barvy očí v jisté zemi, byla mužská populace rozdělena do čtyř skupin s procentuálním zastoupením uvedeným v tabulce.

četnost	otec	syn
5%	tmavooký	tmavooký
7.9%	tmavooký	světleoký
8.9%	světleoký	tmavooký
78.2%	světleoký	světleoký

Určete pravděpodobnost, že tmavooký otec bude mít tmavookého syna.

Označme jev A - syn bude tmavooký, a jev B - otec je tmavooký. Máme určit $\mathsf{P}(A|B)$, přitom

$$\mathsf{P}(A) = 0.139, \qquad \mathsf{P}(B) = 0.129, \qquad \mathsf{P}(A \cap B) = 0.05$$

a podle definice podmíněné pravděpodobnosti tedy

$$\mathsf{P}(A|B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)} = \frac{0.05}{0.129} \doteq 0.388\,,$$

což koresponduje s naším očekáváním.

Poznámka 3.1 V modelu $(\Omega, \mathcal{A}, \mathsf{P})$ vlastně nahrazujeme pravděpodobnostní míru P jinou mírou $\mathsf{P}(\cdot|B)$, která více odpovídá experimentu (pravděpodobnostních měr na $(\Omega, \mathcal{A})$ je nekonečně mnoho).

Věta 3.2 (Součinové pravidlo) *Pro každých $n+1$ jevů $A_0, A_1, \dots A_n$ z $\mathcal{A}$ takových, že $\mathsf{P}(A_0 A_1 \dots A_{n-1}) > 0$, platí*

$$\mathsf{P}(A_0 A_1 \dots A_n) = \mathsf{P}(A_0) \cdot \mathsf{P}(A_1|A_0) \cdot \mathsf{P}(A_2|A_0 A_1) \cdot \ldots \cdot \mathsf{P}(A_n|A_0 A_1 \dots A_{n-1})\,. \tag{3.2}$$

Důkaz. Protože

$$A_0 A_1 \dots A_{n-1} \subset A_0 A_1 \dots A_{n-2} \subset \dots \subset A_0\,,$$

platí

$$0 < \mathsf{P}(A_0 A_1 \dots A_{n-1}) \le \mathsf{P}(A_0 A_1 \dots A_{n-2}) \le \dots \le \mathsf{P}(A_0)$$

a všechny podmíněné pravděpodobnosti jsou tedy dobře definovány. Vztah (3.2) dokážeme matematickou indukcí.
a) Pro $n = 1$ máme

$$\mathsf{P}(A_0 A_1) = \mathsf{P}(A_0) \cdot \mathsf{P}(A_1|A_0),$$

což plyne přímo z definice podmíněné pravděpodobnosti.
b) Pro přechod $n \to n+1$ použijeme definici podmíněné pravděpodobnosti v prvním, a indukční předpoklad ve druhém kroku

$$\begin{aligned} \mathsf{P}(A_0 A_1 \dots A_{n+1}) &= \mathsf{P}(A_0 A_1 \dots A_n) \cdot \mathsf{P}(A_{n+1}|A_0 A_1 \dots A_n) \\ &= \mathsf{P}(A_0) \cdot \mathsf{P}(A_1|A_0) \cdot \ldots \cdot \mathsf{P}(A_n|A_0 A_1 \dots A_{n-1}) \cdot \mathsf{P}(A_{n+1}|A_0 \dots A_n)\,. \end{aligned}$$

□

Příklad 3.2 (Polya's urn scheme) V urně je b bílých a c černých koulí. Náhodně vybereme kouli, vrátíme ji zpět a přidáme dalších k koulí stejné barvy. Určete pravděpodobnost, že ve třech po sobě jdoucích pokusech vytáhneme vždy bílou kouli.

Nechť B_i značí jev, že v i-tém tahu vytáhneme bílou kouli. Hledaná pravděpodobnost potom je

$$\mathsf{P}(B_1B_2B_3) = \mathsf{P}(B_1) \cdot \mathsf{P}(B_2|B_1) \cdot \mathsf{P}(B_3|B_1B_2) = \frac{b}{b+c} \cdot \frac{b+k}{b+c+k} \cdot \frac{b+2k}{b+c+2k}\,.$$

Definice 3.2 *Konečnou množinu $\{B_1, \dots, B_n\}$ jevů z $\mathcal{A}$, pro kterou platí*

a) $B_i \cap B_j = \emptyset$ *pro* $i \neq j$

b) $\mathsf{P}(B_i) > 0\ \forall i \in \widehat{n}$

c) $\mathsf{P}\left(\sum_{i=1}^n B_i\right) = 1$,

nazveme úplný soubor jevů.

Věta 3.3 (O úplné pravděpodobnosti) *Nechť $\{B_1, \dots, B_n\}$ je úplný soubor jevů a nechť $A \in \mathcal{A}$. Potom platí*

$$\mathsf{P}(A) = \sum_{i=1}^n \mathsf{P}(A|B_i) \cdot \mathsf{P}(B_i)$$

(rozklad nepodmíněné pravděpodobnosti na podmíněné).

Důkaz. Podle Věty 2.1 platí pro libovolné jevy A, B rozklad $A = AB + AB^C$. Pokud jev B nahradíme jevem $\sum_{i=1}^n B_i$, platí

$$A = A\left(\sum_{i=1}^n B_i\right) + A\left(\sum_{i=1}^n B_i\right)^C .$$

Protože ale v našem případě platí

$$\mathsf{P}\left(\sum_{i=1}^n B_i\right)^C = 0,$$

dostáváme

$$\mathsf{P}(A) = \mathsf{P}\left(A\left(\sum_{i=1}^n B_i\right)\right) = \mathsf{P}\left(\sum_{i=1}^n AB_i\right) = \sum_{i=1}^n \mathsf{P}(AB_i) = \sum_{i=1}^n \mathsf{P}(A|B_i)\mathsf{P}(B_i)\,,$$

kde jsme využili nezávislosti jevů AB_i a definice podmíněné pravděpodobnosti. □

Příklad 3.3 Mějme krabici s n_1 lístky s číslem 1 a n_2 lístky s číslem 2. Vytáhneme náhodně lístek a podle výsledku se přesuneme do prvního nebo druhého osudí, která obsahují b_1 bílých a c_1 černých, resp. b_2 bílých a c_2 černých koulí. Jaká je pravděpodobnost, že vytáhneme bílou kouli?

Označme jevy B_1 - vytáhnu lístek s číslem 1, B_2 - vytáhnu lístek s číslem 2, A - vytáhnu bílou kouli. Protože jevy B_1, B_2 tvoří úplný systém jevů, můžeme použít tvrzení předchozí věty a dostáváme

$$\mathsf{P}(A) = \mathsf{P}(A|B_1)\mathsf{P}(B_1) + \mathsf{P}(A|B_2)\mathsf{P}(B_2) = \frac{b_1}{b_1+c_1} \cdot \frac{n_1}{n_1+n_2} + \frac{b_2}{b_2+c_2} \cdot \frac{n_2}{n_1+n_2} .$$

Poznamenejme jen, že postup přes všechny možné a příznivé výsledky by byl v tomto případě mnohem složitější.

Věta 3.4 (Bayesova věta) *Nechť* $\{B_1, \ldots, B_n\}$ *je úplný soubor jevů a nechť* $A \in \mathcal{A}$. *Potom pro každé* $i \in \widehat{n}$ *platí*

$$\mathsf{P}(B_i|A) = \frac{\mathsf{P}(A|B_i)\mathsf{P}(B_i)}{\sum_{i=1}^{n} \mathsf{P}(A|B_i)\mathsf{P}(B_i)} . \tag{3.3}$$

Důkaz. Použitím definice podmíněné pravděpodobnosti a věty o úplné pravděpodobnosti přímo dostaneme

$$\mathsf{P}(B_i|A) = \frac{\mathsf{P}(B_iA)}{\mathsf{P}(A)} = \frac{\mathsf{P}(AB_i)}{\mathsf{P}(A)} = \frac{\mathsf{P}(A|B_i)\mathsf{P}(B_i)}{\sum_{i=1}^{n} \mathsf{P}(A|B_i)\mathsf{P}(B_i)} .$$

□

Poznámka 3.2 Tato věta umožňuje zabývat se „opačným“ přístupem. Na konci pozorujeme jev A a ptáme se, jaká je pravděpodobnost, že nastala hypotéza B_i.

Příklad 3.4 Telegrafní zpráva se skládá ze symbolů „·“ a „-“. Při přenosu se zkreslí průměrně 2/5 teček a 1/3 čárek. Dále víme, že v signálech se symboly „·“ a „-“ vyskytují v poměru 5:3. Určete pravděpodobnost, že byl přijat vyslaný signál, jestliže a) byla přijata „·“, b) byla přijata „-“.

Označme A - jev, že byla přijata „·“, B - jev, že byla přijata „-“, C_1 - byla vyslána „·“ a C_2 - byla vyslána „-“. Jevy C_1, C_2 tvoří úplný soubor jevů a platí pro ně

$$\frac{\mathsf{P}(C_1)}{\mathsf{P}(C_2)} = \frac{5}{3}, \qquad \mathsf{P}(C_1) + \mathsf{P}(C_2) = 1$$

a tedy $\mathsf{P}(C_1) = 5/8$ a $\mathsf{P}(C_2) = 3/8$. Protože dále

$$\mathsf{P}(A|C_1) = \frac{3}{5} \quad \text{a} \quad \mathsf{P}(A|C_2) = \frac{1}{3}$$

dostáváme s pomocí Bayesovy věty

$$\mathsf{P}(C_1|A) = \frac{\mathsf{P}(A|C_1)\mathsf{P}(C_1)}{\mathsf{P}(A|C_1)\mathsf{P}(C_1) + \mathsf{P}(A|C_2)\mathsf{P}(C_2)} = \frac{\frac{3}{5} \cdot \frac{5}{8}}{\frac{3}{5} \cdot \frac{5}{8} + \frac{1}{3} \cdot \frac{3}{8}} = \frac{3}{4} .$$

Podobně pro bod b) dostáváme

$$\mathsf{P}(B|C_1) = \frac{2}{5} \quad \text{a} \quad \mathsf{P}(B|C_2) = \frac{2}{3}$$

a

$$\mathsf{P}(C_2|B) = \frac{\mathsf{P}(B|C_2)\mathsf{P}(C_2)}{\mathsf{P}(B|C_1)\mathsf{P}(C_1) + \mathsf{P}(B|C_2)\mathsf{P}(C_2)} = \frac{\frac{2}{3} \cdot \frac{3}{8}}{\frac{2}{5} \cdot \frac{5}{8} + \frac{2}{3} \cdot \frac{3}{8}} = \frac{1}{2} .$$

Cvičení 3.1 Tři střelci zasáhnou při každém výstřelu terč s pravděpodobnostmi po řadě 4/5, 3/4 a 2/3. Při současném výstřelu všech tří střelců nastaly dva zásahy. S jakou pravděpodobností se minul třetí střelec? [6/13]

Cvičení 3.2 Společnost 5 mužů a 10 žen se náhodně rozdělí na pět skupin po třech osobách. Určete pravděpodobnost, že v každé skupině bude jeden muž. [0.081]

Cvičení 3.3 Mějme 5 krabic a v každé z nich b bílých a c černých koulí. Z první krabice se náhodně vytáhne jedna koule a přemístí se do druhé. Pak se z druhé krabice náhodně jedna vybere a přemístí do třetí atd. Určete, jaká je po tomto přemísťování pravděpodobnost toho, že se z poslední krabice vytáhne bílá koule. [b/(b+c)]

Cvičení 3.4 Mějme tři krabice. V první krabici je 1 bílá, 2 černé a 3 modré koule, ve druhé 2 bílé, 1 černá a 1 modrá a ve třetí 4 bílé, 5 černých a 3 modré. Někdo náhodně vybere krabici, vytáhne z ní kouli a řekne nám, že je bílá. Z jaké krabice ji s největší pravděpodobností vytáhl? [pravděpodobnosti 1/6,1/2,1/3]

Cvičení 3.5 U elektrického spotřebiče se výrobní vada vyskytuje s pravděpodobností 0.1. U výrobků s touto vadou dochází v záruční době k poruše s pravděpodobností 0.5, výrobky bez vady mají pravděpodobnost poruchy v záruční době 0.01. Určete pravděpodobnost, že výrobek, který se v záruční době porouchal má uvedenou vadu.

Cvičení 3.6 Zamýšlíte koupit v autobazaru vůz jisté značky. Je ovšem známo, že 30% takových vozů má vadnou převodovku. Abyste získali více informací, najmete si mechanika, který je po projížďce schopen odhadnout stav vozu a jen s pravděpodobností 0,1 se zmýlí. Jaká je pravděpodobnost, že vůz, který zamýšlíte koupit, má vadnou převodovku

a) předtím, než si najmete mechanika

b) jestliže mechanik předpoví, že vůz je dobrý?

3.1 Nezávislé jevy

Definice 3.3 *Řekneme, že dva jevy A a B jsou nezávislé, jestliže*

$$\mathsf{P}(A \cap b) = \mathsf{P}(A) \cdot \mathsf{P}(B) .$$

Poznámka 3.3 Z definice podmíněné pravděpodobnosti víme, že $\mathsf{P}(A \cap B) = \mathsf{P}(A|B)\mathsf{P}(B)$, jsou=li jevy A a B nezávislé dostáváme

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

Neboli jsou-li jevy A a B nezávislé, znamená to, že výskyt jevu B neovlivní pravděpodobnost jevu A a naopak.

Příklad 3.5 Uvažujme hod kostkou a označme jevy A - padne sudé číslo a B - padne číslo nepřevyšující 2. Jsou jevy A a B nezávislé?

Jevy A, B a $A \cap B$ jsou reprezentovány výsledky

$$A = \{2, 4, 6\}, \quad B = \{1, 2\} \quad \text{a} \quad A \cap B = \{2\} .$$

To znamená, že

$$\mathsf{P}(A) = \frac{1}{2}, \quad \mathsf{P}(B) = \frac{1}{3} \quad \text{a} \quad \mathsf{P}(A \cap B) = \frac{1}{6}$$

a tedy $\mathsf{P}(A \cap B) = \mathsf{P}(A) \cdot \mathsf{P}(B)$, z čehož vyplývá, že jevy A a B jsou nezávislé. (poznamenejme jenom, že $\mathsf{P}(A|B) = 1/2$ a $\mathsf{P}(A) = 1/2$ a výskyt jevu B tedy neovlivnil pravděpodobnost jevu A.)

Příklad 3.6 Pozměňme trochu zadání předchozího příkladu a uvažujme hod kostkou a jevy A - padne sudé číslo a B - padne číslo nepřevyšující 3. Jsou jevy A a B nezávislé?

Jevy A, B a $A \cap B$ jsou reprezentovány výsledky

$$A = \{2, 4, 6\}, \quad B = \{1, 2, 3\} \quad \text{a} \quad A \cap B = \{2\}.$$

To znamená, že

$$\mathsf{P}(A) = \frac{1}{2}, \quad \mathsf{P}(B) = \frac{1}{2} \quad \text{a} \quad \mathsf{P}(A \cap B) = \frac{1}{6}$$

a tedy $\mathsf{P}(A \cap B) \neq \mathsf{P}(A) \cdot \mathsf{P}(B)$, z čehož vyplývá, že jevy A a B nejsou nezávislé. (poznamenejme, že $\mathsf{P}(A|B) = 1/3$ a $\mathsf{P}(A) = 1/2$ a výskyt jevu B tedy ovlivnil pravděpodobnost jevu A.)

Věta 3.5 *Nechť $A, B \in \mathcal{A}$, $\mathsf{P}(B) > 0$. Potom jsou jevy A a B nezávislé právě tehdy, když*

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

Důkaz. $\Rightarrow$: Z definice podmíněné pravděpodobnosti a nezávislosti jevů A a B dostáváme

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

$\Leftarrow$: Z předpokladu $\mathsf{P}(A|B) = \mathsf{P}(A)$ a definice podmíněné pravděpodobnosti vyplývá

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

a tedy jevy A a B jsou nezávislé. $\square$

Definice 3.4 *Řekneme, že jevy $A_1, A_2, \ldots, A_n$ jsou (sdruženě) nezávislé, jestliže*

$$\mathsf{P}(A_{i_1} \ldots A_{i_k}) = \mathsf{P}(A_{i_1}) \ldots \mathsf{P}(A_{i_k})$$

pro každý konečný výběr $\{i_1, \ldots, i_k\} \in \widehat{n}$.

Definice 3.5 *Jevy $A_1, \ldots, A_n$ se nazývají párově nezávislé, jestliže pro každé $i \neq j$ platí*

$$\mathsf{P}(A_i \cap A_j) = \mathsf{P}(A_i) \cdot \mathsf{P}(A_j).$$

Příklad 3.7 Předpokládejme náhodný pokus s množinou možných elementárních výsledků $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, které jsou všechny stejně pravděpodobné (např. hod kostkou ve tvaru čtyřstěnu). Definujme tři jevy

$$A = \{\omega_1, \omega_2\}, \quad B = \{\omega_1, \omega_3\}, \quad C = \{\omega_1, \omega_4\}.$$

Potom $\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = 1/2$ a protože $A \cap B = A \cap C = B \cap C = \{\omega_1\}$ dostáváme

$$\mathsf{P}(AB) = \mathsf{P}(AC) = \mathsf{P}(BC) = \frac{1}{4}$$

a tedy jevy A, B, C jsou párově nezávislé. Protože ale

$$\frac{1}{4} = \mathsf{P}(ABC) \neq \mathsf{P}(A)\mathsf{P}(B)\mathsf{P}(C) = \frac{1}{8},$$

jevy A, B, C nejsou sdruženě nezávislé.

Věta 3.6 *Jsou-li jevy A a B nezávislé, pak také jevy A a B^C jsou nezávislé.*

Důkaz. Zapišme jev A jako $A = AB + AB^C$. Potom $\mathsf{P}(A) = \mathsf{P}(AB) + \mathsf{P}(AB^C)$ a tedy

$$\mathsf{P}(AB^C) = \mathsf{P}(A) - \mathsf{P}(AB) = \mathsf{P}(A) - \mathsf{P}(A)\mathsf{P}(B) = \mathsf{P}(A)(1 - \mathsf{P}(B)) = \mathsf{P}(A)\mathsf{P}(B^C).$$

$\square$

Kapitola 4

Náhodné veličiny

V celé této kapitole budeme uvažovat pravděpodobnostní prostor $(\Omega, \mathcal{A}, \mathsf{P})$. Výsledkem pokusu ale nemusí být vždy číslo (tramvaj přijela-nepřijela, obrázky na kostce atd.). Abychom výsledky takových experimentů mohli matematicky zpracovávat, přiřazujeme každému výsledku číslo. To je základní princip náhodné veličiny, je to zobrazení $X : \Omega \to \mathbb{R}$.

Definice 4.1 *Náhodná veličina je zobrazení $X : \Omega \to \mathbb{R}$ takové, že pro každé $x \in \mathbb{R}$ platí*

$$X^{-1}((-\infty, x\rangle) = \{\omega \in \Omega | X(\omega) \leq x\} \in \mathcal{A}\,.$$

(tj. vzory všech intervalů $(-\infty, x\rangle$ jsou měřitelné (funkcí P) množiny)

Příklad 4.1 Předpokládejme, že házíme najednou dvěma kostkami. Potom množina všech možných výsledků má tvar

$$\Omega = \{(i,j) | i \in \widehat{6}, j \in \widehat{6} \text{ a } i \leq j\}\,.$$

Za σ-algebru $\mathcal{A}$ zvolme systém všech podmnožin množiny Ω a náhodná veličina X nechť vyjadřuje součet výsledků, tj.

$$X(i,j) = i + j\,.$$

Potom například

$$X^{-1}((-\infty, 1.2\rangle) = \{\omega \in \Omega | X(\omega) \leq 1.2\} = \emptyset\,,$$
$$\{\omega \in \Omega | X(\omega) \leq 3.1\} = \{(1,1), (1,2)\}\,,$$

protože $X(1,1) = 2$ a $X(1,2) = 3$ a

$$\{\omega \in \Omega | X(\omega) \leq 4\} = \{(1,1), (1,2), (1,3), (2,2)\}\,.$$

Nadále budeme v tomto textu používat následující označení:

$$\begin{aligned}
\{\omega \in \Omega | X(\omega) \leq x\} &\triangleq [X \leq x] \subset \Omega \\
\{\omega \in \Omega | X(\omega) < x\} &\triangleq [X < x] \\
\{\omega \in \Omega | X(\omega) \geq x\} &\triangleq [X \geq x] \\
\{\omega \in \Omega | X(\omega) > x\} &\triangleq [X > x] \\
\{\omega \in \Omega | X(\omega) = x\} &\triangleq [X = x] \\
\{\omega \in \Omega | X(\omega) \in S\} &\triangleq [X \in S]
\end{aligned}$$

kde S je podmnožina $\mathbb{R}$. Pokud bude $[X \in S] \in \mathcal{A}$ znamená to, že umíme určit $\mathsf{P}[X \in S]$.

Věta 4.1 *Nechť X, Y jsou náhodné veličiny. Potom*

1) $X+Y$ je také náhodná veličina

2) kX je náhodná veličina pro všechna $k \in \mathbb{R}$

3) XY je náhodná veličina

4) je-li navíc $[Y=0]=\emptyset$, pak X/Y je náhodná veličina.

4.1 Distribuční funkce náhodných veličin

Jednou z možností, jak popsat náhodnou veličina je distribuční funkce.

Definice 4.2 *Nechť X je náhodná veličina. Funkci $F_X : \mathbb{R} \to \langle 0,1\rangle$ definovanou pro všechna $x \in \mathbb{R}$ vztahem*

$$F_X(x) = \mathsf{P}[X \leq x]\,, \tag{4.1}$$

nazýváme distribuční funkce náhodné veličiny X.

Příklad 4.2 Uvažujme tři hody mincí. Množina všech možných výsledků je

$$\Omega = \{(H,H,H),(H,H,O),(H,O,H),(O,H,H),H,O,O),(O,H,O),(O,O,H),(O,O,O)\}\,,$$

kde H značí, že padla hlava a O, že padl orel. Nechť náhodná veličina X vyjadřuje počet hlav, které padly. Tedy např. $X(H,H,H)=3, X(H,O,O)=1,\ldots$ Distribuční funkce náhodné veličiny X bude mít tvar

pro $x<0$	$[X \leq x]=\emptyset$	a $F_X(x)=0$
pro $0 \leq x < 1$	$[X \leq x]=\{(O,O,O)\}$	a $F_X(x)=\frac{1}{8}$
pro $1 \leq x < 2$	$[X \leq x]=\{(O,O,O),(O,O,H),(O,H,O),(H,O,O)\}$	a $F_X(x)=\frac{1}{2}$
pro $2 \leq x < 3$	$[X \leq x]=\Omega \setminus \{(H,H,H)\}$	a $F_X(x)=\frac{7}{8}$
pro $x \geq 3$	$[X \leq x]=\Omega$	a $F_X(x)=1$,

viz. následující obrázek.

Příklad 4.3 Uvažujme náhodnou veličinu X, která popisuje polohu velké ručičky hodin. Tu můžeme udávat pomocí úhlu, o který se posunula od počátku. Pravděpodobnostní model popisující tuto situaci potom bude

$$\Omega = \langle 0,2\pi), \quad \mathcal{A}=\mathcal{B}(\langle 0,2\pi)), \quad \mathsf{P}(A)=\frac{1}{2\pi}\int_0^{2\pi} I_A(x)\,dx\,,$$

kde $I_A(x)$ je charakteristická funkce množiny, tj.

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

Distribuční funkce náhodné veličiny X má potom tvar

$$F_X(x) = \mathsf{P}[X \leq x] = \begin{cases} 0 & \text{pro } x < 0, \\ \dfrac{x}{2\pi} & \text{pro } 0 \leq x < 2\pi, \\ 1 & \text{pro } x \geq 2\pi. \end{cases}$$

Následující věta shrne základní vlastnosti distribuční funkce.

Věta 4.2 *Nechť X je náhodná veličina a $F_X(x)$ její distribuční funkce. Potom*

1) $0 \leq F_X(x) \leq 1$ *pro všechna* $x \in \mathbb{R}$

2) jestliže $x_1 < x_2$*, potom* $F_X(x_1) \leq F_X(x_2)$ *(tj.* F_X *je neklesající)*

3) $\lim_{x \to -\infty} F_X(x) = 0$

4) $\lim_{x \to \infty} F_X(x) = 1$

5) pro všechna $a \in \mathbb{R}$ *platí* $\lim_{x \to a+} F_X(x) = F_X(a)$ *(tj.* F_X *je zprava spojitá)*

6) F_X *má nejvýše spočetně mnoho bodů nespojitosti.*

Důkaz. ad 1) Pro všechna $x \in \mathbb{R}$ platí $0 \leq \mathsf{P}[X \leq x] \leq 1$ a tvrzení vyplývá z definice distribuční funkce.

ad 2) Náhodný jev $[X \leq x_2]$ se dá zapsat jako disjunktní sjednocení

$$[X \leq x_2] = [X \leq x_1] + [x_1 < X \leq x_2]$$

a tedy

$$F_X(x_2) = \mathsf{P}[X \leq x_2] = \mathsf{P}[X \leq x_1] + \mathsf{P}[x_1 < X \leq x_2] \geq \mathsf{P}[X \leq x_1] = F_X(x_1)\,.$$

ad 4) Podle definice limity máme dokázat

$$(\forall \varepsilon > 0)(\exists x_\varepsilon \in \mathbb{R})(\forall x > x_\varepsilon)(F_X(x) > 1 - \varepsilon)\,.$$

Zvolme $\varepsilon > 0$ pevně. Protože X je náhodná veličina a $X(\omega)$ je konečná pro všechna $\omega \in \Omega$, lze rozložit Ω na

$$\Omega = [X \leq 0] + \sum_{k=1}^{\infty} [k - 1 < X \leq k]$$

a po aplikaci pravděpodobnosti na tuto rovnost dostaneme

$$1 = \mathsf{P}[X \leq 0] + \sum_{k=1}^{\infty} \mathsf{P}[k - 1 < X \leq k]\,.$$

To ale znamená, že řada nezáporných členů na pravé straně je konvergentní a tedy existuje $N_\varepsilon \in \mathbb{N}$ tak, že pro všechna $n > N_\varepsilon$

$$\mathsf{P}[X \leq 0] + \sum_{k=1}^{n} \mathsf{P}[k - 1 < X \leq k] > 1 - \varepsilon\,,$$

neboli pro všechna $n > N_\varepsilon$
$$\mathsf{P}[X \le n] > 1 - \varepsilon\,.$$
Zvolme $x_\varepsilon = N_\varepsilon + 1$, pro tuto hodnotu tedy platí $\mathsf{P}[X \le x_\varepsilon] > 1 - \varepsilon$ a aplikací bodu 2) dostáváme pro všechna $x > x_\varepsilon$
$$F_X(x) = \mathsf{P}[X \le x] \ge \mathsf{P}[X \le x_\varepsilon] > 1 - \varepsilon\,.$$

ad 3) V tomto případě máme dokázat, že
$$(\forall \varepsilon > 0)(\exists x_\varepsilon \in \mathbb{R})(\forall x < x_\varepsilon)(F_X(x) < \varepsilon)\,.$$
Důkaz se provede podobně, jako v předchozím bodě. Zvolme $\varepsilon > 0$ pevně a rozložme Ω na
$$\Omega = [X > 0] + \sum_{k=0}^{\infty} [-(k+1) < X \le -k]\,.$$
Aplikací pravděpodobnosti na tuto rovnost dostaneme
$$1 = \mathsf{P}[X > 0] + \sum_{k=0}^{\infty} \mathsf{P}[-(k+1) < X \le -k]\,.$$
To ale znamená, že existuje $N_\varepsilon \in \mathbb{N}$ tak, že pro všechna $n > N_\varepsilon$
$$\mathsf{P}[X > 0] + \sum_{k=0}^{n} \mathsf{P}[-(k+1) < X \le -k] > 1 - \varepsilon\,,$$
což je ekvivalentní výrazu
$$\mathsf{P}[X \le -(n+1)] < \varepsilon\,.$$
Zvolme $x_\varepsilon = -(N_\varepsilon + 2)$, pro tuto hodnotu tedy platí $\mathsf{P}[X \le x_\varepsilon] < \varepsilon$ a aplikací bodu 2) dostáváme pro všechna $x < x_\varepsilon$
$$F_X(x) = \mathsf{P}[X \le x] \le \mathsf{P}[X \le x_\varepsilon] < \varepsilon\,.$$

ad 5) Pro důkaz tohoto tvrzení věty rozložíme na disjunktní sjednocení jev
$$[a < X \le a+1] = \sum_{n=1}^{\infty} \left[a + \frac{1}{n+1} < X \le a + \frac{1}{n}\right]\,.$$
Platí tedy
$$\mathsf{P}[a < X \le a+1] = \sum_{n=1}^{N} \mathsf{P}\left[a + \frac{1}{n+1} < X \le a + \frac{1}{n}\right] + Z_N\,,$$
kde Z_N je zbytek konvergující k nule pro $N \to \infty$. Poslední výraz lze přepsat do tvaru
$$F_X(a+1) - F_X(a) = \mathsf{P}[a + \frac{1}{N+1} < X \le a+1] + Z_N = F_X(a+1) - F_X\left(a + \frac{1}{N+1}\right)\,,$$
ze kterého vyplývá, že
$$F_X\left(a + \frac{1}{N+1}\right) = F_X(a) + Z_N \quad \text{a tedy} \quad \lim_{N\to\infty} F_X\left(a + \frac{1}{N+1}\right) = F_X(a)\,.$$

Stejná hodnota limity pro $x \to a_+$ vyplývá z monotonie distribuční funkce F_X.

ad 6) Označme S_n množinu bodů, kde má funkce $F_X(x)$ skok velikosti větší nebo rovné $1/n$. Množina všech bodů nespojitosti S funkce F_X se dá potom vyjádřit jako

$$S = \bigcup_{n=1}^{\infty} S_n \,.$$

Pro každé n je množina S_n konečná, jinak by totiž hodnoty funkce F_X přesáhly jedničku, z čehož vyplývá, že množina S je nejvýše spočetná. $\square$

Poznámka 4.1 Jestliže nějaká reálná funkce splňuje vlastnosti 1),...,5) z předchozí věty, pak je to distribuční funkce nějaké náhodné veličiny.

Poznámka 4.2 Někdy se definuje

$$F_X(x) = \mathsf{P}[X < x] \,.$$

Taková distribuční funkce je potom spojitá zleva, ostatní vlastnosti zůstávají v platnosti.

Příklad 4.4 Přístroj obsahuje 5 tranzistorů. Předpokládáme, že došlo k poruše způsobené jedním tranzistorem. Opravář prohlíží jeden po druhém bez vracení. Určete distribuční funkci náhodné veličiny X označující počet potřebných zkoušek.

Náhodná veličina X může nabývat hodnot z množiny $\{1, 2, 3, 4, 5\}$. Pravděpodobnosti jednotlivých hodnot jsou

$$\mathsf{P}[X = 1] = \frac{1}{5}, \quad \mathsf{P}[X = 2] = \frac{4}{5} \cdot \frac{1}{4} = \frac{1}{5}, \quad \mathsf{P}[X = 3] = \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{5},$$

$$\mathsf{P}[X = 4] = \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{5}, \quad \mathsf{P}[X = 5] = \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{5}\,,$$

z čehož vyplývají příslušné hodnoty distribuční funkce:

$$\begin{aligned}
F_X(x) &= 0 && \text{pro } x < 1,\\
F_X(x) &= \mathsf{P}[X = 1] = \tfrac{1}{5} && \text{pro } 1 \le x < 2,\\
F_X(x) &= \mathsf{P}[X = 1] + \mathsf{P}[X = 2] = \tfrac{2}{5} && \text{pro } 2 \le x < 3,\\
F_X(x) &= \mathsf{P}[X = 1] + \mathsf{P}[X = 2] + \mathsf{P}[X = 3] = \tfrac{3}{5} && \text{pro } 3 \le x < 4,\\
F_X(x) &= \mathsf{P}[X = 1] + \ldots + \mathsf{P}[X = 4] = \tfrac{4}{5} && \text{pro } 4 \le x < 5,\\
F_X(x) &= \mathsf{P}[X = 1] + \ldots + \mathsf{P}[X = 5] = 1 && \text{pro } x \ge 5.
\end{aligned}$$

OBR.

Příklad 4.5 Náhodná veličina X je zadána distribuční funkcí

$$F_X(x) = \begin{cases} 0 & \text{pro } x \le 0,\\ 1 - e^{-2x} & \text{pro} x > 0. \end{cases}$$

Najděte $\mathsf{P}[1 < X \leq 2], \mathsf{P}[-1 < X \leq 2], \mathsf{P}[X = 1]$.

$$\mathsf{P}[1 < X \leq 2] = \mathsf{P}[X \leq 2] - \mathsf{P}[X \leq 1] = F_X(2) - F_X(1) = 1 - e^{-4} - 1 + e^{-2} = e^{-2} - e^{-4}\,.$$

$$\mathsf{P}[-1 < X \leq 2] = F_X(2) - F_X(-1) = 1 - e^{-4} - 0 = 1 - e^{-4}\,.$$

Všimněte si, že $\mathsf{P}[-1 < X \leq 2] = \mathsf{P}[0 < X \leq 2]$ z čehož vyplývá $\mathsf{P}[-1 < X \leq 0] = 0$.

$$\mathsf{P}[X = 1] = \mathsf{P}[X \leq 1] - \mathsf{P}[X < 1] = F_X(1) - \lim_{t \to 1_-} F_X(t) = F_X(1) - F_X(1) = 0\,.$$

Poznámka 4.3 Z příkladu je vidět, že pravděpodobnost určité hodnoty bude nenulová pouze tehdy, když distribuční funkce bude mít v tomto bodě skok.

4.2 Mnohorozměrné distribuční funkce

Často měříme více veličin najednou, např. teplota, tlak, vlhkost, nebo míry 90-60-90, atd. Každé číslo je náhodná veličina samo o sobě, abychom si ale udělali celkovou představu, musíme znát všechna. Jednotlivé veličiny mohou být nezávislé, mohou ale mít i silnou vazbu, např. úhlopříčka čtverce a jeho plocha, průměrný roční příjem a výše důchodu. Budeme chtít získávat souhrnné informace o obou(všech) proměnných, informace o jedné z veličin, informace o jedné z nich za podmínky na druhou (ostatní). Toto vše budeme schopni popisovat pomocí mnohorozměrné distribuční funkce.

Vektor náhodných veličin $Z = (X_1, X_2, \ldots, X_n)$ budeme nazývat náhodný vektor. Pro jednoduchost budeme většinu pojmů definovat pro náhodné vektory o dvou složkách, případné rozšíření bude zřejmé.

Definice 4.3 *Nechť X, Y jsou náhodné veličiny. Pak jejich sdružená distribuční funkce je pro všechny $(x, y) \in \mathbb{R}^2$ definována předpisem*

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

Poznámka 4.4 Sdružené distribuční funkce mají podobné vlastnosti jako distribuční funkce jedné proměnné. Proto vyslovíme jen některé věty a bez důkazů, které jsou podobné jako v případě jedné proměnné.

Věta 4.3 *Nechť $F_{X,Y}$ je sdružená distribuční funkce náhodného vektoru (X, Y). Potom pro všechna $x_1 \leq x_2$ a $y_1 \leq y_2$ platí*

$$F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2)\,.$$

Věta 4.4 *Nechť $F_{X,Y}$ je sdružená distribuční funkce náhodného vektoru (X, Y). Potom*

$$\lim_{x \to -\infty} F_{X,Y}(x, y) = 0 \quad \textit{pro všechna} \quad y \in \mathbb{R}\,,$$

$$\lim_{y \to -\infty} F_{X,Y}(x, y) = 0 \quad \textit{pro všechna} \quad x \in \mathbb{R}\,.$$

Věta 4.5 *Nechť $F_{X,Y}$ je sdružená distribuční funkce náhodného vektoru (X, Y). Potom*

$$\lim_{x \to \infty} F_{X,Y}(x, y) = F_Y(y) \quad \textit{pro všechna} \quad y \in \mathbb{R}\,,$$

$$\lim_{y \to \infty} F_{X,Y}(x, y) = F_X(x) \quad \textit{pro všechna} \quad x \in \mathbb{R}\,.$$

Poznámka 4.5 Distribuční funkce F_X, F_Y z předchozí věty budeme nazývat *marginální distribuční funkce* náhodného vektoru (X,Y). Obecněji, pokud $Z = (X_1, X_2, \ldots, X_n)$ je náhodný vektor, vektor $Z' = (X_{i_1}, X_{i_2}, \ldots, X_{i_k})$, pro $k < n$, nazýváme *marginální náhodný vektor*, jeho rozdělení marginální rozdělení a speciálně náhodnou veličinu X_i nazýváme *marginální náhodná veličina.*

Definice 4.4 *Náhodné veličiny X a Y nazýváme (statisticky) nezávislé, jestliže jsou nezávislé jevy*

$$[a < X \leq b] \quad a \quad [c < Y \leq d]$$

pro libovolné dvojice $(a,b), (c,d)$, kde $a, b, c, d \in \overline{\mathbb{R}}$ a $a \leq b, c \leq d$.

Věta 4.6 *Náhodné veličiny X, Y jsou nezávislé právě tehdy, když pro každou dvojici $(x,y) \in \mathbb{R}^2$ platí rovnost*

$$F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)\,.$$

(hodnota sdružené distribuční funkce je rovna součinu hodnot marginálních distribučních funkcí.)

Důkaz. $\Rightarrow$: Nechť jsou X, Y nezávislé, tj. pro všechna $a, b, c, d \in \overline{\mathbb{R}}$, $a \leq b, c \leq d$ platí

$$\mathsf{P}([a < X \leq b][c < Y \leq d]) = \mathsf{P}[a < X \leq b] \cdot \mathsf{P}[c < Y \leq d]\,.$$

Položme $a = -\infty, b = x, c = -\infty, d = y$. Potom pro všechna $(x,y) \in \mathbb{R}^2$ platí

$$\begin{aligned} F_{X,Y}(x,y) &= \mathsf{P}([-\infty < X \leq x][-\infty < Y \leq y]) \\ &= \mathsf{P}[-\infty < X \leq x] \cdot \mathsf{P}[-\infty < Y \leq y] = F_X(x) \cdot F_Y(y)\,. \end{aligned}$$

$\Leftarrow$:

$$\begin{aligned} &\mathsf{P}([a < X \leq b][c < Y \leq d]) = F_{X,Y}(b,d) - F_{X,Y}(b,c) - F_{X,Y}(a,d) + F_{X,Y}(a,c) \\ &= F_X(b)F_Y(d) - F_X(b)F_Y(c) - F_X(a)F_Y(d) + F_X(a)F_Y(c) \\ &= (F_X(b) - F_X(a))(F_Y(d) - F_Y(c)) = \mathsf{P}[a < X \leq b] \cdot \mathsf{P}[c < Y \leq d]\,. \end{aligned}$$

$\square$

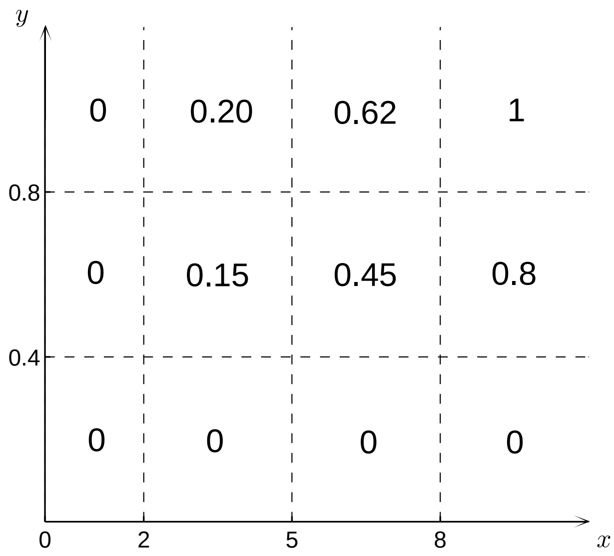
Příklad 4.6 Náhodný vektor (X,Y) je zadán pravděpodobnostmi

$Y \setminus X$	2	5	8
0.4	0.15	0.30	0.35
0.8	0.05	0.12	0.03

Určete sdruženou a příslušné marginální distribuční funkce a rozhodněte, zda jsou náhodné veličiny X, Y nezávislé.

Hodnoty sdružené distribuční funkce jsou vidět na následujícím Obrázku 4.1. Marginální distribuční funkce mají tedy tvar

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x,y) = \begin{cases} 0 & \text{pro } x < 2, \\ 0.2 & \text{pro } 2 \leq x < 5, \\ 0.62 & \text{pro } 5 \leq x < 8, \\ 1 & \text{pro } x \geq 8, \end{cases}$$



Obrázek 4.1: Hodnoty distribuční funkce $F_{X,Y}(x,y)$.

a

$$F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x,y) = \begin{cases} 0 & \text{pro } y < 0.4, \\ 0.8 & \text{pro } 0.4 \leq y < 0.8, \\ 1 & \text{pro } y \geq 0.8. \end{cases}$$

Protože například pro $x = 3$ a $y = 0.5$ neplatí $F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)$, veličiny X a Y nejsou nezávislé.

Kapitola 5

Diskrétní náhodné veličiny

Definice 5.1 *Říkáme, že náhodná veličina X má diskrétní rozdělení, jestliže existuje konečná nebo spočetná množina reálných čísel $H = \{x_1, x_2, \ldots, x_n, \ldots\}$ taková, že pro všechna $x_j \in H$*

$$\mathsf{P}[X = x_j] > 0 \quad a \quad \sum_{x_j \in H} \mathsf{P}[X = x_j] = 1\,.$$

Poznámka 5.1 Jinými slovy, obor hodnot náhodné veličiny X je nejvýše spočetná množina.

Definice 5.2 *Funkci $P(x) = \mathsf{P}[X = x]$ definovanou pro všechna $x \in \mathbb{R}$ nazveme pravděpodobnostní funkcí náhodné veličiny X.*

Poznámka 5.2 V bodech, kde je $P(x)$ nenulová budeme její hodnoty značit

$$P(x_i) = \mathsf{P}[X = x_i] \triangleq p_i\,.$$

Poznámka 5.3 Diskrétní náhodné veličiny se obvykle zadávají pravděpodobnostními funkcemi ve formě tabulky, např.

x_i	0	1	2	3
p_i	0.25	0.3	0.1	0.35

Distribuční funkce diskrétní náhodné veličiny má tvar

$$F_X(x) = \sum_{j: x_j \le x} \mathsf{P}[X = x_j] = \sum_{j: x_j \le x} p_j \quad \text{pro} \quad x \in \mathbb{R}\,.$$

Z toho tedy vyplývá, že distribuční funkce $F_X(x)$ má skoky v bodech $x_1, x_2, \ldots$ a je konstantní v intervalech $\langle x_j, x_{j+1})$, $j = 1, 2, \ldots$. Navíc velikost skoku v bodě x_j je rovna $\mathsf{P}[X = x_j]$. Můžeme tedy psát

$$\mathsf{P}[X = a] = P(a), \qquad \mathsf{P}[X < a] = F(a) - P(a), \qquad \mathsf{P}[X \le a] = F(a),$$

$$\mathsf{P}[X > a] = 1 - F(a), \qquad \mathsf{P}[X \ge a] = 1 - F(a) + P(a)\,.$$

Obrázek

Poznámka 5.4 Pokud jsou oborem hodnot náhodné veličiny X nezáporná celá čísla $\{0, 1, 2, \ldots\}$, je možno zapsat distribuční funkci ve tvaru

$$F_X(x) = \sum_{n=0}^{[x]} \mathsf{P}[X = n]\,,$$

kde symbol $[x]$ označuje celou část z čísla x.

Cvičení 5.1 Automobil má na trase 4 semafory. Ty dávají s pravděpodobností 2/3 zelenou nebo s pravděpodobností 1/3 červenou (pracují nezávisle). Určete rozdělení pravděpodobnosti náhodné veličiny udávající počet semaforů, které automobil pojede před prvním zastavením.

Cvičení 5.2 Terč je rozdělen na tři oblasti očíslované 1,2,3. Za zásah oblasti 1 získává střelec 10 bodů, za oblast 2 získává 5 bodů a za oblast 3 dostává -1 bod. Pravděpodobnosti zásahu střelce jsou po řadě 0.5, 0.3, 0.2. Určete rozdělení pravděpodobnosti náhodného počtu bodů získaného při dvou výstřelech.

5.1 Příklady diskrétních rozdělení

5.1.1 Alternativní rozdělení (Bernoulliovo)

Toto rozdělení slouží pro popis náhodných situací majících dva možné výsledky (ano $\times$ ne, správně $\times$ špatně). Například hod mincí, kostkou (když nás zajímá pouze sudost, lichost). Obvykle se těmto výsledkům přiřazují hodnoty 0 (reprezentující neúspěch s předpokládanou pravděpodobností $1 - p$) a 1 (úspěch s pravděpodobností p).

Definice 5.3 *Řekneme, že náhodná veličina X má alternativní rozdělení s parametrem $p \in (0, 1)$, jestliže platí*

$$\mathsf{P}[X = 1] = p, \qquad \mathsf{P}[X = 0] = 1 - p\,.$$

Poznámka 5.5 Pro vyjádření toho, že náhodná veličina X má alternativní rozdělení budeme používat označení

$$X \sim Be(p)\,.$$

Pravděpodobnostní funkce alternativního rozdělení se dá obecně zapsat ve tvaru

$$P(x) = \mathsf{P}[X = x] = \begin{cases} p^x(1-p)^{1-x} & \text{pro } x \in \{0, 1\}, \\ 0 & \text{jinak.} \end{cases}$$

5.1.2 Binomické rozdělení

Uvažujme náhodnou veličinu X vyjadřující počet výskytů jevu A v n nezávislých pokusech. Jestliže pravděpodobnost jevu A je pokaždé rovna p, pravděpodobnost, že jev A nastane z n pokusů právě k-krát je

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

Protože náhodná veličina X může nabývat hodnot $0, 1, \ldots n$ a podle binomické věty platí

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

jedná se o rozdělení pravděpodobnosti.

Definice 5.4 *Řekneme, že náhodná veličina X má binomické rozdělení s parametry $n \in I\!N$ a $p \in (0, 1)$, jestliže*

$$\mathsf{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k} \quad \textit{pro} \quad k = 0, 1, \ldots, n\,.$$

Značíme $X \sim Bi(n, p)$.

Poznámka 5.6 Platí, jestliže $X_j, j \in \widehat{n}$ jsou nezávislé náhodné veličiny s alternativním rozdělením s parametrem p potom náhodná veličina definovaná předpisem

$$X \triangleq X_1 + X_2 + \ldots + X_n$$

má binomické rozdělení s parametry n a p.

Příklad 5.1 Uvažujme šachovou partii, ve které jsou dva stejně silní soupeři (pravděpodobnost výhry i prohry je 1/2). Rozhodněte, zda je pravděpodobnější

a) vyhrát 3 partie ze 4, nebo 5 partií z 8,

b) vyhrát alespoň 3 partie ze 4, nebo alespoň 5 partií z 8.

Zaměřme se na jednoho hráče a označme X náhodnou veličinu vyjadřující počet jeho výher ze 4 partií a Y počet výher z 8 partií. Potom
a)

$$\mathsf{P}[X = 3] = \binom{4}{3} \left(\frac{1}{2}\right)^3 \frac{1}{2} = \frac{1}{4}$$

a

$$\mathsf{P}[Y = 5] = \binom{8}{5} \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^3 = \frac{7}{32}$$

b)

$$\mathsf{P}([X = 3] \cup [X = 4]) = \mathsf{P}[X = 3] + \mathsf{P}[X = 4] = \frac{1}{4} + \binom{4}{4} \left(\frac{1}{2}\right)^4 = \frac{5}{16}$$

a

$$\mathsf{P}([Y = 5] \cup [Y = 6] \cup [Y = 7] \cup [Y = 8]) = \sum_{k=5}^{8} \binom{8}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{8-k} = \frac{93}{256}\,.$$

5.1.3 Pascalovo (geometrické) rozdělení

Definice 5.5 *Řekneme, že náhodná veličina X má Pascalovo rozdělení s parametrem $p \in (0,1)$, jestliže*

$$\mathsf{P}[X = k] = p \cdot (1-p)^k \quad \textit{pro} \quad k = 0, 1, 2, \dots .$$

Značíme $X \sim Pa(p)$.

Poznámka 5.7 Uvažujme nekonečnou posloupnost Bernoulliovských pokusů (jev A nastává s pravděpodobností p, nenastává s pravděpodobností $1-p$). Náhodná veličina $X \sim Pa(p)$ potom vyjadřuje počet pokusů před prvním výskytem jevu A.

Ukažme, že se skutečně jedná o pravděpodobnostní rozdělení.

$$\sum_{k=0}^{\infty} \mathsf{P}[X = k] = \sum_{k=0}^{\infty} p(1-p)^k = \frac{p}{1-(1-p)} = 1 .$$

Pro Pascalovo rozdělení je také možno jednoduše vyjádřit distribuční funkci

$$F_X(x) = \sum_{k=0}^{[x]} p(1-p)^k = p \cdot \frac{1-(1-p)^{[x]+1}}{1-(1-p)} = 1-(1-p)^{[x]+1}, \quad x \in \mathbb{R}.$$

Příklad 5.2 Uvažujme začátek hry člověče nezlob se, kdy hráč nasadí figurku do hry až mu padne šestka. Jaká je pravděpodobnost, že hráč nasadí ve čtvrtém, resp. nejpozději ve čtvrtém kole?

Označme X počet pokusů, než padne šestka. Potom $X \sim Pa(1/6)$ a platí

$$\mathsf{P}[X = 3] = \frac{1}{6} \cdot \left(1 - \frac{1}{6}\right)^3 = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^3 \doteq 0.0965$$

a

$$\mathsf{P}[X \leq 3] = F_X(3) = 1 - \left(\frac{5}{6}\right)^4 \doteq 0.5177 .$$

5.1.4 Poissonovo rozdělení

Definice 5.6 *Řekneme, že náhodná veličina X má Poissonovo rozdělení s parametrem $\lambda > 0$, jestliže*

$$\mathsf{P}[X = k] = \begin{cases} e^{-\lambda} \dfrac{\lambda^k}{k!} & \textit{pro } k = 0, 1, 2, \dots \\ 0 & \textit{jinak.} \end{cases}$$

Značení $X \sim Po(\lambda)$.

Poznámka 5.8 Ověřme, že se jedná o rozdělení pravděpodobnosti,

$$\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1.$$

Poznámka 5.9 Poissonovo rozdělení bylo odvozeno při popisu radioaktivního rozpadu. Použití:

a) Při popisu počtu událostí, které nastanou během časového intervalu (λ - průměrný počet událostí za časový interval). Např. Počet signálů v telefonní ústředně.
Při popisu počtu částic v jednotce plochy nebo objemu (λ - průměrný počet částic v jednotce). Např. počet částic v zorném poli mikroskopu.

b) Aproximace binomického rozdělení
Uvažujme posloupnost náhodných veličin $X_1, X_2, X_3, \ldots$, kde $X_i \sim Bi(n, p_n)$. Nechť $np_n = \lambda$, kde $\lambda > 0$ je parametr. Předpoklad říká, že relativní počet „příznivých" výsledků mezi n pokusy by „se měl pohybovat kolem čísla λ". Mluví pouze o dlouhodobém chování, nikoli o jedné realizaci pokusu.
Není problém spočítat rozdělení X_i pro malá i, pro i velká ($i = 10000$) je již výpočet problematický. a právě tady nám pomůže Poissonovo rozdělení.

Věta 5.1 *Nechť $(X_n)_{n=1}^{\infty}$ je posloupnost náhodných veličin taková, že $X_n \sim Bi(n, p_n)$. Nechť*

$$\lim_{n\to\infty} p_n = 0 \quad a \quad \lim_{n\to\infty} np_n = \lambda, \quad \lambda \in I\!R^+.$$

Potom platí

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

(neboli $\lim_{n\to\infty} \mathsf{P}[X_n = k] = \mathsf{P}[X = k]$, kde $X \sim Po(\lambda)$).

Důkaz. Rozepišme člen za limitou do tvaru

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!}\,\frac{p_n^k}{(1-p_n)^k}\,(1-p_n)^n.$$

Limitu výrazu na pravé straně budeme zkoumat po částech. Protože podle předpokladů $np_n \to \lambda$, můžeme psát $p_n = \lambda/n + z_n$, přičemž $nz_n \to 0$. Potom

$$(1-p_n)^n = \left(1 - \frac{\lambda + nz_n}{n}\right)^n = \left[\left(1 + \frac{1}{\frac{n}{-(\lambda+nz_n)}}\right)^{\frac{n}{-(\lambda+nz_n)}}\right]^{-(\lambda+nz_n)} \longrightarrow e^{-\lambda},$$

pro $n \to \infty$, kde jsme využili faktu, že $(1 + 1/p_n)^{p_n} \to e$ pro $|p_n| \to \infty$.

Dále

$$\left(\frac{p_n}{1-p_n}\right)^k = \left(\frac{\lambda/n + z_n}{1 - \lambda/n - z_n}\right)^k = \left(\frac{\lambda + nz_n}{n - \lambda - nz_n}\right)^k,$$

kde

$$(\lambda + nz_n)^k \to \lambda^k$$

a všechny zbývající členy mají limitu

$$\frac{n(n-1)\ldots(n-k+1)}{k!\,(n-\lambda-nz_n)^k} \longrightarrow \frac{1}{k!},$$

protože se jedná o limitu typu polynom stupně k lomeno polynom stupně k. Celkem tedy dostáváme tvrzení věty. $\square$

Poznámka 5.10 Pro malá p a velká n tedy můžeme aproximovat rozdělení $Bi(n,p)$ rozdělením $Po(np)$. Tato aproximace se často doporučuje pro $p < 0.1$ a $n > 30$. Např. hodnoty pravděpodobnosti náhodné veličiny $Y \sim Bi(1000, 0.0015)$ můžeme aproximovat pravděpodobnostmi náhodné veličiny s rozdělením $Po(1.5)$. Pro ilustraci zjednodušení výpočtu vyjádřeme

$$\mathsf{P}[Y=5] = \binom{1000}{5}(0.0015)^5(0.9985)^{995} \doteq e^{-1.5}\frac{(1.5)^5}{5!}.$$

Příklad 5.3 Pravděpodobnost závady na jeden kilometr vodiče je 0.01. Kabel je složen ze 100 vodičů, funkční bude pokud je 99 vodičů v pořádku. Jaká je pravděpodobnost, že kabel délky jeden kilometr bude fungovat?

Označme X počet chybných vodičů, tedy $X \sim Bi(100, 0.01)$ a A jev, že kabel bude fungovat. Potom

$$\mathsf{P}(A) = \mathsf{P}[X \le 1] = \mathsf{P}[X=0]+\mathsf{P}[X=1] = \binom{100}{0}0.01^0 \cdot 0.99^{100} + \binom{100}{1}0.01^1 \cdot 0.99^{99} = 0.736\,.$$

Protože je n velké a p malé, můžeme také příslušné pravděpodobnosti aproximovat pomocí Poissonova rozdělení s parametrem $\lambda = np = 1$,

$$\mathsf{P}[X=0]+\mathsf{P}[X=1] = e^{-1}\frac{1}{0!} + e^{-1}\frac{1}{1!} = 2e^{-1} = 0.736\,.$$

Příklad 5.4 Pokusy se zjistilo, že radioaktivní látka vyzářila během 7.5 sekundy průměrně 3.87 α-částice. Určete pravděpodobnost toho, že za 1 sekundu vyzáří tato látka alespoň 1 α-částici.

Označme X počet částic vyzářených za 1 sekundu. Protože průměrný počet částic vyzářených za 1 sekundu je $\lambda = 3.87/7.5 = 0.516$, můžeme náhodnou veličinu X popisovat pomocí rozdělení $Po(0.516)$ a tedy

$$\mathsf{P}[X \ge 1] = 1 - \mathsf{P}[X=0] = 1 - e^{-0.516}\frac{0.516^0}{0!} \doteq 0.403\,.$$

Cvičení 5.3 Určete pravděpodobnost, že mezi 200 výrobky jsou alespoň 4 zmetky, jestliže zmetky tvoří průměrně 1% produkce. [0.143]

Cvičení 5.4 Během hodiny přijde na telefonní ústřednu průměrně 60 výzev. Jaká je pravděpodobnost toho, že během půl minuty, na kterou se manipulantka vzdálila, nebude nikdo volat? [0.61]

Cvičení 5.5 Korektura 500 stránek obsahuje 1500 tiskových chyb. Určete pravděpodobnost toho, že na namátkou vybrané stránce jsou alespoň 3 chyby. [0.577]

5.2 Diskrétní náhodné vektory

Definice 5.7 *Řekneme, že náhodný vektor* $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ *je diskrétní, jestliže existuje konečná nebo spočetná množina* n*-tic* $\{(x_{11}, x_{21}, \ldots, x_{n1}), (x_{12}, x_{22}, \ldots, x_{n2}), \ldots\}$ *taková, že*

$$P(x_{1i}, x_{2i}, \ldots, x_{ni}) = \mathsf{P}[X_1 = x_{1i}, X_2 = x_{2i}, \ldots, X_n = x_{ni}] > 0 \quad \textit{pro všechna} \quad i = 1, 2, \ldots$$

a
$$\sum_i P(x_{1i}, x_{2i}, \ldots, x_{ni}) = 1\,.$$
Funkci $P(x_1, x_2, \ldots, x_n)$ nazýváme pravděpodobnostní funkcí náhodného vektoru $\boldsymbol{X}$.

5.2.1 Multinomické rozdělení

Definice 5.8 *Řekneme, že náhodný vektor $\boldsymbol{X} = (X_1, X_2, \ldots, X_k)$ má multinomické rozdělení s parametry $n, p_1, \ldots, p_k$, kde $n \in I\!N$, $0 < p_j < 1$, $j \in \widehat{k}$ a $\sum_{j=1}^k p_j = 1$, jestliže*
$$\mathsf{P}[X_1 = x_1, \ldots, X_k = x_k] = \frac{n!}{x_1!x_2!\ldots x_k!}\, p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$
pro $x_j = 0, 1, \ldots n, j \in \widehat{k}$, a $\sum_{j=1}^k x_j = n$ a je nulová jinak.

Poznámka 5.11 Název tohoto rozdělení pochází z toho, že $\mathsf{P}[X_1 = x_1, \ldots, X_k = x_k]$ je obecný člen rozvoje $(p_1 + p_2 + \ldots + p_k)^n$. Multinomické rozdělení je vlastně zobecněním binomického: uvažujme n nezávislých pokusů, z nichž v každém nastane právě jeden z k vzájemně disjunktních jevů $A_1, A_2, \ldots, A_k$. Přitom, předpokládáme, že pravděpodobnost, že v i-tém pokusu nastane jev A_j je p_j a $\sum_{j=1}^k p_j = 1$. Binomické rozdělení je tedy speciálním případem multinomického pro $k = 2$.

Příklad 5.5 U výrobku měříme znak Z. Výrobek je dobrý, jestliže $T_D \leq Z \leq T_H$. Uvažujme jevy $A_1 : Z < T_D$, $A_2 : T_D \leq Z \leq T_H$ a $A_3 : Z > T_H$. Za normálních podmínek je známo, že
$$\mathsf{P}(A_j) = p_j, \quad j = 1, 2, 3 \quad \text{přičemž} \quad \sum_{j=1}^3 p_j = 1\,.$$
Potom pravděpodobnost, že z n výrobků jich bude mít x_1 $Z < T_D$, x_2 hodnotu $T_D \leq Z \leq T_H$ a x_3 hodnotu $Z > T_H$ je rovna
$$\begin{aligned}\mathsf{P}[X_1 = x_1, X_2 = x_2, X_3 = x_3] &= \frac{n!}{x_1!x_2!x_3!}\, p_1^{x_1} p_2^{x_2} p_3^{x_3} \\ &= \frac{n!}{x_1!x_2!(n - x_1 - x_2)!}\, p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{n - x_1 - x_2}\,.\end{aligned}$$
Např. pro $p_1 = 0.02, p_2 = 0.95, p_3 = 0.03$ a $n = 6$
$$\mathsf{P}[X_1 = 1, X_2 = 4, X_3 = 1] = \frac{6!}{1!4!1!}\, 0.02 \cdot 0.95^4 \cdot 0.03 = 0.015\,.$$

5.2.2 Vlastnosti diskrétních náhodných vektorů

Uvažujme diskrétní náhodný vektor (X_1, X_2). Potom sdružená distribuční funkce tohoto vektoru je dána předpisem
$$F_{X_1,X_2}(x_1, x_2) = \sum_{t_1 \leq x_1} \sum_{t_2 \leq x_2} \mathsf{P}[X_1 = t_1, X_2 = t_2]\,.$$

Věta 5.2 *Je-li* $\mathsf{P}[X_1 = x_1, X_2 = x_2]$ *sdružená pravděpodobnostní funkce náhodného vektoru* X_1, X_2*, potom pro marginální pravděpodobnostní funkce náhodných veličin* X_1 *a* X_2 *platí*

$$\mathsf{P}[X_1 = x_1] = \sum_{x_2} \mathsf{P}[X_1 = x_1, X_2 = x_2] \quad \textit{pro všechna} \quad x_1 \in \mathbb{R}$$

a

$$\mathsf{P}[X_2 = x_2] = \sum_{x_1} \mathsf{P}[X_1 = x_1, X_2 = x_2] \quad \textit{pro všechna} \quad x_2 \in \mathbb{R}\,.$$

Důkaz. Pro všechna $x_1 \in \mathbb{R}$ platí

$$\begin{aligned} F_{X_1}(x_1) &= \lim_{x_2 \to \infty} F_{X_1,X_2}(x_1, x_2) = \lim_{x_2 \to \infty} \sum_{t_1 \leq x_1} \sum_{t_2 \leq x_2} \mathsf{P}[X_1 = t_1, X_2 = t_2] \\ &= \sum_{t_1 \leq x_1} \left(\sum_{t_2} \mathsf{P}[X_1 = t_1, X_2 = t_2] \right), \end{aligned}$$

ale podle definice platí také pro všechna $x_1 \in \mathbb{R}$

$$F_{X_1}(x_1) = \sum_{t_1 \leq x_1} \mathsf{P}[X_1 = t_1]\,.$$

Protože obě sumy pro $t_1 \leq x_1$ mají stejné hodnoty pro všechna $x_1 \in \mathbb{R}$, musí být shodné i jejich argumenty, tedy

$$\mathsf{P}[X_1 = t_1] = \sum_{t_2} \mathsf{P}[X_1 = t_1, X_2 = t_2] \quad \text{pro všechna} \quad t_1 \in \mathbb{R}\,.$$

Stejně se ukáže i druhá část tvrzení. □

Věta 5.3 *Dvě diskrétní náhodné veličiny* X, Y *jsou nezávislé právě tehdy, když*

$$\mathsf{P}[X = x, Y = y] = \mathsf{P}[X = x] \cdot \mathsf{P}[Y = y] \quad \textit{pro všechny} \quad (x, y) \in \mathbb{R}^2\,.$$

Důkaz. Podle Věty 4.6 víme, že dvě náhodné veličiny jsou nezávislé právě tehdy, když

$$F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y) \quad \forall (x, y) \in \mathbb{R}^2\,.$$

Tvrzení této věty dostaneme přímo přechodem ke skokům distribuční funkce. □

Příklad 5.6 Mějme dvě diskrétní náhodné veličiny zadané sdruženou pravděpodobnostní funkcí $\mathsf{P}[X_1 = x_1, X_2 = x_2]$ ve formě tabulky

$X_1 \backslash X_2$	0	1	2
0	0.42	0.12	0.06
1	0.28	0.08	0.04

Určete marginální pravděpodobnostní funkce a rozhodněte, zda jsou náhodné veličiny X_1, X_2 nezávislé.

Podle Věty 5.2 platí pro pravděpodobnostní funkci náhodné veličiny X_1

$$\mathsf{P}[X_1 = 0] = \sum_{k=0}^{2} \mathsf{P}[X_1 = 0, X_2 = k] = 0.42 + 0.12 + 0.06 = 0.6$$

a

$$\mathsf{P}[X_1 = 1] = \sum_{k=0}^{2} \mathsf{P}[X_1 = 1, X_2 = k] = 0.28 + 0.08 + 0.04 = 0.4\,.$$

Podobně dostaneme hodnoty pravděpodobnostní funkce náhodné veličiny X_2,

$$\mathsf{P}[X_2 = 0] = 0.7, \quad \mathsf{P}[X_2 = 1] = 0.2, \quad \mathsf{P}[X_2 = 2] = 0.1\,.$$

Jak je vidět, hodnoty marginálních pravděpodobnostních funkcí získáme součtem příslušných řádků nebo sloupců a je výhodné je zapisovat přímo do tabulky

$X_1 \backslash X_2$	0	1	2	P_{X_1}
0	0.42	0.12	0.06	**0.6**
1	0.28	0.08	0.04	**0.4**
P_{X_2}	**0.7**	**0.2**	**0.1**	1

Z této tabulky je snadno vidět, že pro všechna $i \in \{0, 1\}$ a $j \in \{0, 1, 2\}$ platí

$$\mathsf{P}[X_1 = i, X_2 = j] = \mathsf{P}[X_1 = i] \cdot \mathsf{P}[X_2 = j]$$

a tedy veličiny X_1, X_2 jsou nezávislé.

Definice 5.9 *Nechť (X_1, X_2) je diskrétní náhodný vektor se sdruženou pravděpodobnostní funkcí $\mathsf{P}[X_1 = x_1, X_2 = x_2]$. Jestliže $\mathsf{P}[X_2 = x_2] > 0$, definujeme podmíněnou pravděpodobnostní funkci náhodné veličiny X_1 za podmínky, že veličina X_2 nabyla hodnoty x_2, výrazem*

$$\mathsf{P}[X_1 = x_1 | X_2 = x_2] = \frac{\mathsf{P}[X_1 = x_1, X_2 = x_2]}{\mathsf{P}[X_2 = x_2]} = \frac{\mathsf{P}[X_1 = x_1, X_2 = x_2]}{\sum_{x_1} \mathsf{P}[X_1 = x_1, X_2 = x_2]}\,.$$

Příklad 5.7 Házíme dvěma kostkami. Najděte rozdělení náhodné veličiny označující maximum z počtu ok padlých na obou kostkách, jestliže víme, že minimum nabylo hodnotu 3.

Označme X maximum a Y minimum počtu ok na dvou kostkách. Y nabývá hodnotu 3, jestliže na první kostce padla hodnota z množiny $\{4, 5, 6\}$ a na druhé z množiny $\{3\}$, nebo na první $\{3\}$ a na druhé $\{4, 5, 6\}$ nebo na obou kostkách $\{3\}$. Celkem je to 7 možností z 36, takže $\mathsf{P}[Y = 3] = 7/36$. Pokud $Y = 3$, může náhodná veličina X nabývat hodnot $\{3, 4, 5, 6\}$ a

$$\mathsf{P}[X = 3, Y = 3] = \frac{1}{36} \quad \text{a} \quad \mathsf{P}[X = i, Y = 3] = \frac{2}{36}, \quad \text{pro} \quad i = 4, 5, 6\,.$$

Podle definice podmíněné pravděpodobnosti tedy platí

$$\mathsf{P}[X = 3 | Y = 3] = \frac{1/36}{7/36} = \frac{1}{7} \quad \text{a} \quad \mathsf{P}[X = i | Y = 3] = \frac{2/36}{7/36} = \frac{2}{7}, \quad \text{pro} \quad i = 4, 5, 6\,.$$

Kapitola 6

Absolutně spojité náhodné veličiny

Diskrétní a spojité náhodné veličiny tvoří jen dvě malé disjunktní třídy pravděpodobnostních distribucí. Jednodušeji se s nimi pracuje, pokud neznáme teorii míry a Lebesgueův integrál. Musíme ale poznamenat, že existuje mnoho distribucí, které nejsou ani spojité ani diskrétní.

Definice 6.1 *Řekneme, že náhodná veličina X má absolutně spojité rozdělení (ASR), jestliže existuje nezáporná reálná funkce $f_X(x)$ taková, že pro každé $x \in \mathbb{R}$ platí*

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt\,.$$

Funkci $f_X(x)$ nazýváme (pravděpodobnostní) hustota náhodné veličiny X.

Příklad 6.1 Mějme náhodnou veličinu X zadanou pomocí své distribuční funkce

$$F_X(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ \frac{x}{2} & \text{pro } 0 < x \leq 2, \\ 1 & \text{pro } x \geq 2. \end{cases}$$

Tato náhodná veličina X má absolutně spojité rozdělení, neboť když vezmeme funkci

$$f_X(x) = \begin{cases} 0 & \text{pro } x \leq 0 \vee x \geq 2, \\ \frac{1}{2} & \text{pro } 0 < x < 2, \end{cases} \qquad \text{platí} \quad F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt\,.$$

Obr.

Věta 6.1 *Nechť má náhodná veličina X absolutně spojité rozdělení. Potom ve všech bodech, kde existuje derivace distribuční funkce F_X platí*

$$f_X(x) = F'(x) = \frac{dF}{dx}(x)\,.$$

Důkaz. Důkaz plyne z vlastností integrálu a derivace. □

Věta 6.2 *Nechť má náhodná veličina X absolutně spojité rozdělení. Potom pro každou množinu $A \in \mathcal{B}(\mathbb{R})$ platí*

$$\mathsf{P}[X \in A] = \int_A f_X(x)\,dx\,.$$

Obr.

Důkaz. Tato věta se dokazuje pomocí vlastností σ-algeber, naznačme důkaz alespoň pro interval.

$$\mathsf{P}[a < X \leq b] = F_X(b) - F_X(a) = \int_{-\infty}^{b} f_X(x)\,dx - \int_{-\infty}^{a} f_X(x)\,dx = \int_{a}^{b} f_X(x)\,dx\,.$$

$\square$

Poznámka 6.1 Tato vlastnost je analogická vlastnosti diskrétních veličin, pro které platí

$$\mathsf{P}[X \in A] = \sum_{j:x_j \in A} \mathsf{P}[X = x_j]\,.$$

Pravděpodobnostní funkce tedy odpovídá ve spojitém případě hustotě pravděpodobnosti.

Poznámka 6.2 Má-li náhodná veličina X ASR, platí

$$\mathsf{P}[X = a] = 0, \qquad \forall a \in \mathbb{R}\,,$$

což vyplývá např. z toho, že $\int_a^a f_X(x)\,dx = 0$ nebo $\mathsf{P}[X = a] = F_X(a) - \lim_{x\to a_-} F_X(x) = 0$, protože F_X je spojitá. To tedy znamená, že

$$\mathsf{P}[X < a] = \mathsf{P}[X \leq a]$$

a

$$\mathsf{P}[a < X < b] = \mathsf{P}[a \leq X < b] = \mathsf{P}[a < X \leq b] = \mathsf{P}[a \leq X \leq b]\,.$$

Poznámka 6.3 Dvě základní vlastnosti každé hustoty pravděpodobnosti jsou

a)

$$f_X(x) \geq 0 \quad \text{pro všechna} \quad x \in \mathbb{R}$$

b)

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \lim_{x\to\infty} F_X(x) = 1\,.$$

Příklad 6.2 Náhodná veličina X má hustotu

$$f_X(x) = \frac{a}{1+x^2} \qquad x \in \mathbb{R}\,.$$

Určete konstantu a, distribuční funkci F_X a $\mathsf{P}[-1 < X < 1]$.

Protože

$$\int_{-\infty}^{\infty} \frac{a}{1+x^2}\,dx = a[\operatorname{arctg}(x)]_{-\infty}^{\infty} = \pi a\,,$$

konstanta $a = 1/\pi$. Dále pro všechna $x \in \mathbb{R}$

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\pi}\frac{1}{1+t^2}\,dt = \frac{1}{\pi}[\operatorname{arctg}(t)]_{-\infty}^{t} = \frac{1}{\pi}\left[\operatorname{arctg}(x) + \frac{\pi}{2}\right]$$

a

$$\mathsf{P}[-1 < X < 1] = \int_{-1}^{1} \frac{1}{\pi}\frac{1}{1+x^2}\,dx = F_X(1) - F_X(-1) = \frac{1}{2}\,.$$

Příklad 6.3 Z bodu A do bodu B přenášíme signál složený z „+1“ a „-1“. Signál je ale zašuměn, tj. příjemce dostává signál plus šum. Víme, že šum je náhodná veličina s hustotou pravděpodobnosti

$$f_X(x) = \begin{cases} \frac{1}{2}\,e^{-x} & \text{pro } x > 0, \\ \frac{1}{2}\,e^{x} & \text{pro } x \leq 0. \end{cases}$$

Signál S je v místě B detekován následovně. Pokud $S > d$ detekujeme symbol „+1“, je-li $S \leq d$ detekujeme symbol „-1“. Stanovte optimální rozhodovací mez d a pravděpodobnost chyby, jestliže víme, že průměrně posíláme 60% „+1“ a 40% „-1“.

Cvičení 6.1 Náhodná veličina X má hustotu

$$f_X(x) = ax^2 e^{-x} \qquad x \geq 0, .$$

Určete konstantu a, distribuční funkci F_X a $\mathsf{P}[0 < X < 1]$. $[a = 1/2,\ ,1 - 5/(2e)]$

Cvičení 6.2 Distribuční funkce zdaněných ročních příjmů je

$$F_X(x) = \begin{cases} 1 - \left(\dfrac{c}{x}\right)^{\alpha} & \text{pro } x \geq c, \\ 0 & \text{jinak,} \end{cases}$$

kde $\alpha > 0$. Určete hodnotu ročního příjmu, kterou náhodně zvolený daňový poplatník překročí s pravděpodobností 0.5. $[c \cdot 2^{1/\alpha}]$

6.1 Absolutně spojité náhodné vektory

Definice 6.2 *Říkáme, že náhodné veličiny $X_1, X_2, \dots, X_n$ mají sdružené absolutně spojité rozdělení (SASR), jestliže existuje nezáporná funkce $f_{X_1,\dots,X_n}(x_1, \dots, x_n)$ taková, že*

$$F_{X_1,\dots,X_n}(x_1, \dots, x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f_{X_1,\dots,X_n}(t_1, \dots, t_n)\,dt_1 \dots dt_n$$

pro všechny n-tice $(x_1, \dots, x_n) \in I\!R^n$. Funkci $f_{X_1,\dots,X_n}$ nazýváme sdružená hustota pravděpodobnosti náhodných veličin $X_1, X_2, \dots, X_n$.

Poznámka 6.4 Podobně jako u jedné náhodné veličiny platí

a)

$$f_{X_1,\dots,X_n}(x_1, \dots, x_n) = \frac{\partial^n F_{X_1,\dots,X_n}(x_1, \dots, x_n)}{\partial x_1 \dots \partial x_n},$$

tam kde derivace existuje.

b)

$$\mathsf{P}[a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2, \dots, a_n < X_n \leq b_n] =$$
$$\int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_n}^{b_n} f_{X_1,\dots,X_n}(x_1, \dots, x_n)\,dx_1 \dots dx_n\,.$$

Věta 6.3 *Mají-li náhodné veličiny X_1, X_2, X_3 SASR, potom také X_1, X_3 mají SASR a platí*

$$f_{X_1,X_3}(x_1, x_3) = \int_{-\infty}^{\infty} f_{X_1,X_2,X_3}(x_1, x_2, x_3)\, dx_2 \,.$$

Poznámka 6.5 Tvrzení je opět analogické diskrétnímu případu, kde platí

$$\mathsf{P}[X_1 = x_1] = \sum_{x_2} \mathsf{P}[X_1 = x_1, X_2 = x_2] \,.$$

Důkaz. Použitím Fubiniho věty a záměny integrálu a limity ve třetím kroku (všechny předpoklady jsou v našem případě splněny) dostaneme pro všechna $(x_1, x_3) \in I\!R^2$

$$\begin{aligned}
F_{X_1,X_3}(x_1, x_3) &= \lim_{x_2 \to \infty} F_{X_1,X_2,X_3}(x_1, x_2, x_3) \\
&= \lim_{x_2 \to \infty} \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \int_{-\infty}^{x_3} f_{X_1,X_2,X_3}(t_1, t_2, t_3)\, dt_1 dt_2 dt_3 \\
&= \int_{-\infty}^{x_1} \int_{-\infty}^{x_3} \left(\lim_{x_2 \to \infty} \int_{-\infty}^{x_2} f_{X_1,X_2,X_3}(t_1, t_2, t_3)\, dt_2 \right) dt_1 dt_3 \\
&= \int_{-\infty}^{x_1} \int_{-\infty}^{x_3} \left(\int_{-\infty}^{\infty} f_{X_1,X_2,X_3}(t_1, t_2, t_3)\, dt_2 \right) dt_1 dt_3 \,,
\end{aligned}$$

z čehož pomocí definice SASR plyne tvrzení věty. □

Příklad 6.4 Náhodné veličiny X, Y jsou zadány pomocí sdružené hustoty pravděpodobnosti

$$f_{X,Y}(x, y) = \begin{cases} e^{-(x+y)} & \text{pro } \min\{x, y\} \geq 0, \\ 0 & \text{jinak.} \end{cases}$$

Určete marginální hustotu $f_Y(y)$ a $\mathsf{P}[X > 2Y]$.

Podle předchozí věty

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx = \begin{cases} \int_0^{\infty} e^{-(x+y)}\, dx = e^{-y} & \text{pro } y \geq 0, \\ 0 & \text{pro } y < 0. \end{cases}$$

Dále

$$\mathsf{P}[X > 2Y] = \iint_{x > 2y} f_{X,Y}(x, y)\, dx dy = \int_0^{\infty} \left(\int_{2y}^{\infty} e^{-(x+y)}\, dx \right) dy = \int_0^{\infty} e^{-3y}\, dy = \frac{1}{3} \,.$$

Příklad 6.5 Náhodné veličiny X, Y jsou zadány pomocí sdružené hustoty pravděpodobnosti

$$f_{X,Y}(x, y) = \begin{cases} \frac{2}{a^2} & \text{pro } 0 \leq x < y < a, \\ 0 & \text{jinak.} \end{cases}$$

Určete marginální hustoty $f_X(x)$ a $f_Y(y)$.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy = \begin{cases} \int_x^a \frac{2}{a^2}\, dy = \frac{2}{a^2}\,(a - x) & \text{pro } 0 \leq x < a, \\ 0 & \text{jinak,} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx = \begin{cases} \int_0^y \frac{2}{a^2}\, dx = \frac{2}{a^2}\, y & \text{pro } 0 \leq y < a, \\ 0 & \text{jinak.} \end{cases}$$

Věta 6.4 *Nechť náhodné veličiny $X_1, \dots, X_n$ mají SASR. Potom $X_1, \dots, X_n$ jsou nezávislé právě tehdy, když*

$$f_{X_1,\dots,X_n}(x_1, \dots, x_n) = \prod_{j=1}^{n} f_{X_j}(x_j) \quad \textit{pro všechny} \quad (x_1, \dots, x_n) \in \mathbb{R}^n .$$

Důkaz. Ukažme, že tvrzení platí pro $n = 2$, tedy

$$X, Y \quad \text{jsou nezávislé} \quad \Leftrightarrow \quad f_{X,Y}(x, y) = f_X(x) f_Y(y),\ \forall (x, y) \in \mathbb{R}^2 .$$

$\Leftarrow$: Pro všechna $(x, y) \in \mathbb{R}^2$ platí

$$\begin{aligned} F_{X,Y}(x, y) &= \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t, s)\, dt ds = \int_{-\infty}^{x} \int_{-\infty}^{y} f_X(t) f_Y(s)\, dt ds \\ &= \left(\int_{-\infty}^{x} f_X(t)\, dt \right) \left(\int_{-\infty}^{y} f_Y(s)\, ds \right) = F_X(x) F_Y(y), \end{aligned}$$

z čehož podle Věty 4.6 vyplývá nezávislost náhodných veličin X, Y.
$\Rightarrow$: Opačným postupem dostaneme pro všechna $(x, y) \in \mathbb{R}^2$

$$F_{X,Y}(x, y) = F_X(x) F_Y(y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_X(t) f_Y(s)\, dt ds,$$

a podle definice SASR tedy

$$f_{X,Y}(x, y) = f_X(x) f_Y(y),\ \forall (x, y) \in \mathbb{R}^2 .$$

□

Příklad 6.6 Náhodné veličiny X_1, X_2, X_3 jsou zadány sdruženou hustotou

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = \begin{cases} 2x_1(x_2 + x_3) & \text{pro } 0 < x_j < 1, j \in \widehat{3}, \\ 0 & \text{jinak.} \end{cases}$$

Rozhodněte, zda jsou nezávislé.

Spočtěme marginální hustoty pravděpodobnosti jednotlivých veličin. Pro $0 < x_1 < 1$ platí

$$f_{X_1}(x_1) = \int_0^1 \int_0^1 2x_1(x_2 + x_3)\, dx_2 dx_3 = 2x_1 \int_0^1 \frac{1}{2} + x_3\, dx_3 = 2x_1 .$$

Pro $0 < x_2 < 1$ platí

$$f_{X_2}(x_2) = \int_0^1 \int_0^1 2x_1(x_2 + x_3)\, dx_1 dx_3 = \left(\int_0^1 2x_1\, dx_1 \right) \int_0^1 (x_2 + x_3)\, dx_3 = x_2 + \frac{1}{2}$$

a symetricky, pro $0 < x_3 < 1$

$$f_{X_3}(x_3) = x_3 + \frac{1}{2} .$$

Celkem tedy

$$f_{X_1,X_2,X_3} \neq f_{X_1} f_{X_2} f_{X_3}$$

a veličiny X_1, X_2, X_3 nejsou nezávislé.

Definice 6.3 *Nechť X, Y jsou náhodné veličiny. Potom podmíněnou distribuční funkcí náhodné veličiny X při dané hodnotě $Y = y$ definujeme jako*

$$F_{X|Y}(x|y) = \lim_{\varepsilon \to 0_+} \mathsf{P}[X \le x \,|\, y - \varepsilon < Y \le y + \varepsilon]$$

za předpokladu, že limita existuje. Pokud navíc existuje funkce $f_{X|Y}(x|y) \ge 0$ taková, že

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} f_{X|Y}(t|y)\,dt \qquad \forall x \in I\!R\,,$$

nazýváme ji podmíněnou hustotou náhodné veličiny X za podmínky $Y = y$.

Poznámka 6.6 Definice podmíněného rozdělení je pro spojité náhodné veličiny komplikovanější než pro diskrétní. Vztah

$$\mathsf{P}[X \le x|Y = y] = \frac{\mathsf{P}[X \le x, Y = y]}{\mathsf{P}[Y = y]}$$

totiž nelze použít, neboť pro spojitou náhodnou veličinu platí $\mathsf{P}[Y = y] = 0$ pro všechna $y \in I\!R$.

Věta 6.5 *Nechť náhodné veličiny X, Y mají SASR. Potom v každém bodě $(x, y) \in I\!R^2$, kde je*

a) *$f_{X,Y}(x, y)$ spojitá,*

b) *$f_Y(y)$ spojitá a $f_Y(y) > 0$,*

existuje podmíněná hustotu náhodné veličiny X za podmínky $Y = y$ a platí

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}\,.$$

Příklad 6.7 Náhodné veličiny X, Y mají pro $a > 2$ sdruženou hustotu

$$f_{X,Y}(x, y) = \begin{cases} c(1 + x + y)^{-a} & \text{pro } x > 0, y > 0, \\ 0 & \text{jinak.} \end{cases}$$

Stanovte konstantu c, marginální hustotu f_X a podmíněnou hustotu $f_{Y|X}$.

$$\begin{aligned} 1 &= \int_0^\infty \int_0^\infty c(1 + x + y)^{-a}\,dxdy = c \int_0^\infty \left[\frac{(1 + x + y)^{-a+1}}{-a + 1} \right]_0^\infty dy \\ &= \frac{c}{1 - a} \int_0^\infty (1 + y)^{-a+1}\,dy = \frac{c}{(1 - a)(2 - a)} \end{aligned}$$

a konstanta $c = (1 - a)(2 - a)$. Pro $x > 0$ platí

$$f_X(x) = \int_0^\infty c(1 + x + y)^{-a}\,dy = \frac{c}{1 - a} \left[(1 + x + y)^{-a+1}\right]_0^\infty = (a - 2)(1 + x)^{-a+1}$$

a podmíněná hustota má pro $x > 0, y > 0$ tvar

$$f_{Y|X}(y|x) = (a - 1)(1 + x)^{a-1}(1 + x + y)^{-a}\,.$$

6.2 Funkce náhodných veličin

V řadě problémů nás zajímá rozdělení pravděpodobnosti funkce $Y = h(X)$, pokud známe rozdělení náhodné veličiny X.

Příklad 6.8 X je diskrétní náhodná veličina zadaná tabulkou

x_i	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	π
p_i	0.1	0.3	0.2	0.1	0.3

Určete rozdělení náhodné veličiny $Y = \sin X$.

Pokud do tabulky doplníme hodnoty $\sin x_i$

x_i	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	π
$\sin x_i$	0	$\frac{\sqrt{2}}{2}$	1	$\frac{\sqrt{2}}{2}$	0
p_i	0.1	0.3	0.2	0.1	0.3

vidíme, že náhodná veličina Y má rozdělení reprezentované tabulkou

y_i	0	$\frac{\sqrt{2}}{2}$	1
p_i	0.4	0.4	0.2

neboť například

$$\mathsf{P}\left[Y = \frac{\sqrt{2}}{2}\right] = \mathsf{P}\left(\left[X = \frac{\pi}{4}\right] \cup \left[X = \frac{3\pi}{4}\right]\right) = \mathsf{P}\left[X = \frac{\pi}{4}\right] + \mathsf{P}\left[X = \frac{3\pi}{4}\right] = 0.4\,.$$

Poznámka 6.7 Po transformaci se může rozdělení náhodné veličiny hodně změnit, dokonce ze spojité se může stát diskrétní (ne naopak). Protože pro diskrétní náhodné veličiny se dá nové rozdělení získat přímým výpočtem, omezíme se dále pouze na spojité náhodné veličiny. Ukážeme dvě metody:

- výpočet distribuční funkce transformované náhodné veličiny
- vzorec pro přímý výpočet hustoty transformované náhodné veličiny.

Pokud je například funkce h ostře rostoucí a existuje tedy inverzní funkce h^{-1}, můžeme psát

$$\begin{aligned} F_Y(y) &= \mathsf{P}[Y \le y] = \mathsf{P}[h(X) \le y] = \mathsf{P}[X \in \{t \in I\!R \,|\, h(t) \le y\}] \\ &= \mathsf{P}[X \le h^{-1}(y)] = F_X(h^{-1}(y))\,. \end{aligned}$$

Poslední výraz už umíme určit, protože známe rozdělení náhodné veličiny X.

Příklad 6.9 Náhodná veličina X je dána hustotou pravděpodobnosti

$$f_X(x) = \begin{cases} \frac{1}{5} & \text{pro } x \in \langle -2, 3 \rangle, \\ 0 & \text{jinak.} \end{cases}$$

Najděte hustotu pravděpodobnosti náhodné veličiny $Y = X^2$.

Vyřešme problém nejdříve pro obecnou náhodnou veličinu X.

$$\begin{aligned} F_Y(y) &= \mathsf{P}[Y \le y] = \mathsf{P}[X^2 \le y] \\ &= \begin{cases} 0 & \text{pro } y < 0, \\ \mathsf{P}[-\sqrt{y} \le X \le \sqrt{y}] = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\,dx = F_X(\sqrt{y}) - F_X(-\sqrt{y}) & \text{pro } y \ge 0 \end{cases} \end{aligned}$$

a hustota pravděpodobnosti náhodné veličiny Y má tedy tvar

$$f_Y(y) = (F_Y(y))' = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}),$$

pro $y > 0$. V našem konkrétním případě vypadá situace následovně

$$F_Y(y) = \left\{ \begin{array}{lll} 0 & y < 0 & 0 \\ \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{5}\,dx = \frac{2}{5}\sqrt{y} & y \in \langle 0, 4) & \frac{1}{5\sqrt{y}} \\ \int_{-2}^{\sqrt{y}} \frac{1}{5}\,dx = \frac{1}{5}(\sqrt{y}+2) & y \in \langle 4, 9\rangle & \frac{1}{10\sqrt{y}} \\ \int -2^3 \frac{1}{5}\,dx = 1 & y > 9 & 0 \end{array} \right\} = f_Y(y)\,.$$

Věta 6.6 *Nechť X je spojitá náhodná veličina, $h : \mathbb{R} \to \mathbb{R}$ je ryze monotónní funkce na množině $X(\Omega)$ a h^{-1} je diferencovatelná. Potom náhodná veličina $Y = h(X)$ má hustotu*

$$f_Y(y) = f_X\left(h^{-1}(y)\right) \cdot \left| \frac{dh^{-1}}{dy}(y) \right| .$$

Důkaz. Pro rostoucí funkci $h(x)$ dostaneme

$$F_Y(y) = \mathsf{P}[Y \le y] = \mathsf{P}[h(X) \le y] = \mathsf{P}[X \le h^{-1}(y)] = F_X(h^{-1}(y))$$

a tedy

$$f_Y(y) = f_X(h^{-1}(y)) \cdot \frac{dh^{-1}}{dy}(y)\,.$$

Podobně pro klesající

$$F_Y(y) = \mathsf{P}[h(X) \le y] = 1 - \mathsf{P}[h(x) \ge y] = 1 - \mathsf{P}[X \le h^{-1}(y)] = 1 - F_X(h^{-1}(y))$$

a

$$f_Y(y) = -f_X(h^{-1}(y)) \cdot \frac{dh^{-1}}{dy}(y)\,.$$

Protože derivace klesající funkce je záporná, můžeme oba obdržené výrazy zapsat v kýženém tvaru pomocí absolutní hodnoty. $\square$

Příklad 6.10 Nechť náhodná veličina X má hustotu $f_X(x) \ge 0$ pro $x > 0$ a $f_X(x) = 0$ pro $x \le 0$. Určete rozdělení náhodně veličiny $Y = a \cdot \ln X, a \ne 0$.

Ve značení věty máme

$$h(x) = a \cdot \ln x \quad \text{a} \quad h^{-1}(y) = e^{\frac{y}{a}}$$

a z jejího tvrzení vyplývá

$$f_Y(y) = f_X(e^{\frac{y}{a}}) e^{\frac{y}{a}} \frac{1}{|a|}\,.$$

Např. pro $c > 0$, $f_X(x) = 1/c$ pro $x \in (0, c)$ dostáváme, že náhodná veličina $Y = -\ln X$ má hustotu

$$f_Y(y) = \begin{cases} \frac{1}{c}\, e^{-y} & \text{pro } y > -\ln c, \\ 0 & \text{jinak.} \end{cases}$$

Příklad 6.11 Náhodná veličina X má hustotu $f_X(x) = 1/\pi$ pro $x \in (-\pi/2, \pi/2)$. Určete rozdělení náhodné veličiny $Y = \operatorname{tg} X$.

Protože $h^{-1}(y) = \operatorname{arctg}(y)$, dostáváme

$$f_Y(y) = \frac{1}{\pi} \cdot \frac{1}{1+y^2}, \qquad y \in (-\infty, \infty)\,.$$

Cvičení 6.3 Náhodná veličina X má hustotu $f_X(x) = x/8$ pro $0 \le x < 4$. Určete hustotu pravděpodobnosti náhodné veličiny Y=2X+4. $[(y-4)/32, y \in \langle 4, 12\rangle]$

Cvičení 6.4 Rychlost částice plynu je náhodná veličina V s hustotu

$$f_V(v) = av^2 e^{-bv}, \qquad v > 0,$$

kde konstanta $b > 0$ a její hodnota závisí na teplotě plynu a hmotnosti částice. Určete konstantu a a pravděpodobnostní rozdělení kinetické energie částice $W = mV^2/2$.

Cvičení 6.5 X_1, X_2 jsou nezávislé náhodné veličiny s hustotami

$$f_{X_j}(x_j) = 2e^{-2x_j}, \qquad x_j \ge 0,\ j = 1, 2\,.$$

Určete hustotu pravděpodobnosti náhodné veličiny $Y = X_1/X_2$. $[1/(1+y)^2, y > 0]$

6.3 Funkce více náhodných veličin

Nechť náhodný vektor $(X_1, \ldots, X_n)$ má sdruženou hustotu pravděpodobnosti $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ a nechť $h : I\!R^n \to I\!R$. Potom distribuční funkce náhodné veličiny $Y = h(X_1, \ldots, X_n)$ je

$$F_Y(y) = \mathsf{P}[Y \le y] = \int \!\!\cdots\!\! \int_{\{x_1,\ldots,x_n \in I\!R^n \,|\, h(x_1,\ldots,x_n) \le y\}} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n\,.$$

Příklad 6.12 Najděte rozdělení součtu $Y = X_1 + X_2$, pokud jsou náhodné veličiny X_1, X_2 nezávislé.

Užitím Fubiniovy věty a nezávislosti dostáváme

$$\begin{aligned} F_Y(y) &= \mathsf{P}[X_1 + X_2 \le y] = \iint_{x_1+x_2 \le y} f_{X_1,X_2}(x_1, x_2)\, dx_1 dx_2 \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{y-x_1} f_{X_1,X_2}(x_1, x_2)\, dx_2 \right) dx_1 = \int_{-\infty}^{\infty} f_{X_1}(x_1) \left(\int_{-\infty}^{y-x_1} f_{X_2}(x_2)\, dx_2 \right) dx_1 \\ &= \int_{-\infty}^{\infty} f_{X_1}(x_1) \cdot F_{X_2}(y - x_1)\, dx_1 \end{aligned}$$

a tedy

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(y - x_1)\, dx_1\,.$$

Symetricky, pokud použijeme opačné řezy ve Fubiniově větě, dostaneme

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_2}(x_2) f_{X_1}(y - x_2)\, dx_2 \,.$$

Věta 6.7 *Nechť náhodný vektor $X_1, \dots, X_n$ má SASR a $(Y_1, \dots, Y_n) = h(X_1, \dots, X_n)$, kde $h : \mathbb{R}^n \to \mathbb{R}^n$ je spojité bijektivní zobrazení na otevřené množině G takové, že $\int \cdots \int_G f_{X_1,\dots,X_n}\, dx_1 \dots dx_n = 1$. Nechť dále inverzní zobrazení $(x_1, \dots, x_n) = h^{-1}(y_1, \dots, y_n)$ je spojité, diferencovatelné a na množině $h(G)$ splňuje podmínku*

$$J = \begin{vmatrix} \frac{\partial h_1^{-1}}{\partial y_1} & \dots & \frac{\partial h_1^{-1}}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_n^{-1}}{\partial y_1} & \dots & \frac{\partial h_n^{-1}}{\partial y_n} \end{vmatrix} \neq 0 \,.$$

Potom i náhodný vektor $(Y_1, \dots, Y_n)$ má SASR a platí

$$f_{(Y_1,\dots,Y_n)}(y_1, \dots, y_n) = f_{X_1,\dots,X_n}(h_1^{-1}(y_1, \dots, y_n), \dots, h_n^{-1}(y_1, \dots, y_n) \cdot |J| \,.$$

Příklad 6.13 Nechť náhodné veličiny X_1, X_2 mají sdruženou hustotu

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} 4x_1x_2 & \text{pro } 0 < x_1 < 1, 0 < x_2 < 1, \\ 0 & \text{jinak.} \end{cases}$$

Najděte sdruženou hustotu pravděpodobnosti veličin X_1^2, X_2^2.

Ve značení věty máme $(Y_1, Y_2) = h(X_1, X_2) = (X_1^2, X_2^2)$, tj. $X_1 = \sqrt{Y_1}$, $X_2 = \sqrt{Y_2}$ a $h^{-1}(y_1, y_2) = (\sqrt{y_1}, \sqrt{y_2})$ a hodnota jakobiánu tedy je

$$J = \begin{vmatrix} \frac{1}{2\sqrt{y_1}} & 0 \\ 0 & \frac{1}{2\sqrt{y_2}} \end{vmatrix} = \frac{1}{4\sqrt{y_1y_2}} \,,$$

jež je na dané oblasti nenulová. Podle tvrzení věty tedy

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} 4\sqrt{y_1}\sqrt{y_2} \cdot \frac{1}{4\sqrt{y_1y_2}} = 1 & \text{pro } 0 < y_1 < 1, 0 < y_2 < 1, \\ 0 & \text{jinak.} \end{cases}$$

Příklad 6.14 Nechť jsou náhodné veličiny X_1, X_2 nezávislé a nechť platí $f_{X_2}(x_2) = 0$ pro všechna $x \leq 0$. Určete hustotu pravděpodobnosti náhodné veličiny $Y = X_1/X_2$.

Označme $Y_1 = X_1/X_2, Y_2 = X_2$. Inverzní funkce jsou $X_2 = Y_2$, $X_1 = Y_1Y_2$ a pro hodnotu jakobiánu platí

$$J = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} = y_2 \neq 0 \quad \text{pro} \quad x_2 > 0 \,.$$

Podle věty má sdružená hustota tvar

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(y_1y_2, y_2)y_2 = f_{X_1}(y_1y_2) f_{X_2}(y_2) y_2$$

a marginální hustota je potom

$$f_{Y_1}(y_1) = \int_0^{\infty} f_{X_1}(y_1y_2) f_{X_2}(y_2) y_2 \, dy_2 \,.$$

Příklad 6.15 Nechť náhodné veličiny X_1, X_2, X_3 mají sdruženou hustotu

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = \begin{cases} e^{(-x_1+x_2+x_3)} & \text{pro } x_1 > 0, x_2 > 0, x_3 > 0, \\ 0 & \text{jinak.} \end{cases}$$

Najděte hustotu náhodné veličiny $Y = (X_1 + X_2 + X_3)/2$.

Označme

$$Y_1 = \frac{X_1 + X_2 + X_3}{3}, \quad Y_2 = X_2, \quad Y_3 = X_3\,.$$

Inverzní funkce tedy jsou $X_1 = 3Y_1 - Y_2 - Y_3, X_2 = Y_2, X_3 = Y_3$ a hodnota jakobiánu transformace je

$$J = \begin{vmatrix} 3 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 3\,.$$

Sdružená hustota pravděpodobnosti tedy je

$$f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) = \begin{cases} 3e^{-3y_1} & \text{pro } 3y_1 - y_2 - y_3 > 0, y_2 > 0, y_3 > 0, \\ 0 & \text{jinak} \end{cases}$$

a postupným vyintegrováním proměnných y_2, y_3 dostáváme

$$f_{Y_1,Y_2}(y_1,y_2) = \begin{cases} \int_0^{3y_1-y_2} 3e^{-3y_1}\, dy_3 = 3(3y_1 - y_2)e^{-3y_1} & \text{pro } 3y_1 - y_2 > 0, y_2 > 0, \\ 0 & \text{jinak,} \end{cases}$$

$$f_{Y_1}(y_1) = \begin{cases} \int_0^{3y_1} 3(3y_1 - y_2)e^{-3y_1} dy_2 = \frac{27}{2} y_1^2 e^{-3y_1} & \text{pro } y_1 > 0, y_2 > 0, \\ 0 & \text{jinak.} \end{cases}$$

6.4 Příklady spojitých náhodných veličin

6.4.1 Rovnoměrné rozdělení s parametry $a < b,\ a, b \in I\!R$

Definice 6.4 *Řekneme, že náhodná veličina X má rovnoměrné rozdělení s parametry $a < b,\ a, b \in I\!R$, jestliže její hustota má tvar*

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{pro } x \in (a,b), \\ 0 & \text{jinak.} \end{cases}$$

Značíme $X \sim U(a,b)$.

Obr.

6.4.2 Gamma rozdělení s parametry $\alpha, \beta > 0$

Definice 6.5 *Řekneme, že náhodná veličina X má Gamma rozdělení s parametry $\alpha, \beta > 0$, jestliže její hustota pravděpodobnosti má tvar*

$$f_X(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)} \cdot \dfrac{1}{\beta^\alpha} x^{\alpha-1} e^{\frac{x}{\beta}} & \text{pro } x > 0, \\ 0 & \text{jinak.} \end{cases}$$

Značíme $X \sim Gamma(\alpha, beta)$.

Poznámka 6.8 Připomeňme definici Gamma funkce

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x}\, dx, \qquad p > 0$$

a její základní vlastnosti

$$\Gamma(1) = 1,\ \Gamma(p+1) = p\Gamma(p) \quad \text{a tedy} \quad \Gamma(p+1) = p! \quad \text{a} \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}\,.$$

6.4.3 Normální (Gaussovo) rozdělení s parametry $\mu \in \mathbb{R}, \sigma^2 > 0$

Definice 6.6 *Řekneme, že náhodná veličina* X *má normální (Gaussovo) rozdělení s parametry* $\mu \in \mathbb{R}, \sigma^2 > 0$*, jestliže její hustota pravděpodobnosti má tvar*

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

Značíme $X \sim N(\mu, \sigma^2)$.

Obr.

Ověřme, že se jedná o hustotu pravděpodobnosti

$$\begin{aligned} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx &= \left| \begin{array}{c} \frac{x-\mu}{\sigma} = t \\ \frac{dx}{\sigma} = dt \end{array} \right| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\, dt = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{t^2}{2}}\, dt \\ &= \left| \begin{array}{c} \frac{t^2}{2} = z \\ tdt = dz \end{array} \right| = \frac{1}{\sqrt{\pi}} \int_0^{\infty} z^{-\frac{1}{2}} e^{-z}\, dz = \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) = 1\,. \end{aligned}$$

Normální rozdělení s parametry $\mu = 0$ a $\sigma^2 = 1$ bude mít výsadní postavení, budeme ho nazývat standardní normální rozdělení a zavedeme pro jeho hustotu a distribuční funkci označení

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt\,.$$

OBR.

Tvrzení 6.1 *Pro distribuční funkci* $\Phi(x)$ *standardní normální náhodné veličiny platí*

$$\Phi(x) = 1 - \Phi(-x) \quad \textit{pro všechna} \quad x \in \mathbb{R}\,.$$

Důkaz.

$$\begin{aligned} \Phi(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt = |t = -y| = -\frac{1}{\sqrt{2\pi}} \int_{\infty}^{-x} e^{-\frac{y^2}{2}}\, dy = \frac{1}{\sqrt{2\pi}} \int_{-x}^{\infty} e^{-\frac{y^2}{2}}\, dy \\ &= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, dy - \int_{-\infty}^{-x} e^{-\frac{y^2}{2}}\, dy \right) = 1 - \Phi(x)\,. \end{aligned}$$

□

Věta 6.8 *Jestliže $X \sim N(\mu, \sigma^2)$, potom*

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

Důkaz.

$$F_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\, dt = \left| \begin{array}{l} \frac{t-\mu}{\sigma} = z \\ \frac{dt}{\sigma} = dz \end{array} \right| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{z^2}{2}}\, dz = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

□

Poznámka 6.9 Protože integrály potřebné k vyčíslení hodnot distribuční funkce normálního rozdělení lze počítat pouze numericky, tyto hodnoty jsou uvedeny ve statistických tabulkách. Z předchozích vět vyplývá, že stačí tabelovat pouze hodnoty distribuční funkce standardního normálního rozdělení pro $x > 0$.

Uvažujme například náhodnou veličinu $X \sim N(2, 16)$ a spočtěme pravděpodobnost $\mathsf{P}[X \in (1,3)]$.

$$\begin{aligned} \mathsf{P}[X \in (1,3)] &= F_X(3) - F_X(1) = \Phi\left(\frac{3-2}{4}\right) - \Phi\left(\frac{1-2}{4}\right) \\ &= \Phi\left(\frac{1}{4}\right) - \Phi\left(\frac{-1}{4}\right) = 2\Phi\left(\frac{1}{4}\right) - 1 \doteq 0.197\,. \end{aligned}$$

Věta 6.9 *Nechť $X \sim N(\mu, \sigma^2)$, potom platí*

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

Důkaz. Označme $Y = aX + b$. To znamená $X = (Y - b)/a$ a aplikací Věty 6.6 o transformaci dostaneme

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

což je hustota pravděpodobnosti náhodné veličiny s rozdělením $N(a\mu + b, a^2\sigma^2)$. □

Důsledek. Nechť $X \sim N(\mu, \sigma^2)$. Potom

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

Důkaz. V předchozí větě položíme $a = 1/\sigma, b = -\mu/\sigma$ a dostaneme

$$\frac{X-\mu}{\sigma} \sim N\left(\frac{\mu}{\sigma} - \frac{\mu}{\sigma}, \frac{1}{\sigma^2}\sigma^2\right) = N(0,1)\,.$$

□

Příklad 6.16 Nechť $X \sim N(\mu, \sigma^2)$. Určete $\mathsf{P}[|X - \mu| \leq k\sigma]$ pro $k > 0$.

Pro všechna $k > 0$ platí

$$\mathsf{P}[|X-\mu| \leq k\sigma] = \mathsf{P}[\mu - k\sigma \leq X \leq \mu + k\sigma] = F_X(\mu+k\sigma) - F_X(\mu-k\sigma) = \Phi(k) - \Phi(-k)$$

a pro vybrané hodnoty k tedy dostáváme

$$\mathsf{P}[|X-\mu| \le \sigma] = \Phi(1) - \Phi(-1) = 0.6827\,,$$

$$\mathsf{P}[|X-\mu| \le 2\sigma] = \Phi(2) - \Phi(-2) = 0.9555\,,$$

$$\mathsf{P}[|X-\mu| \le 3\sigma] = \Phi(3) - \Phi(-3) = 0.9973\,.$$

Poslednímu výrazu se někdy říká *pravidlo 3σ:* hodnoty náhodné veličiny $X \sim N(\mu, \sigma^2)$ budou téměř jistě (s pravděpodobností 99.73%) ležet v intervalu $(\mu - 3\sigma, \mu + 3\sigma)$.

Cvičení 6.6 Délka skoků sportovce Petra měřená v cm má normální rozdělení $N(\mu_1, \sigma_1^2)$, kde $\mu_1 = 690$ a $\sigma_1 = 10$. Délka skoků sportovce Pavla měřená v cm má také normální rozdělení $N(\mu_2, \sigma_2^2)$, kde $\mu_2 = 705$ a $\sigma_2 = 15$. Na závody se kvalifikuje ten, kdo ze dvou skoků alespoň jednou skočí více než 700 cm.

a) S jakou pravděpodobností se oba kvalifikují na závody?

b) S jakou pravděpodobností se kvalifikuje Pavel, ale Petr ne?

6.4.4 Logaritmicko-normální rozdělení s parametry $\mu \in \mathbb{R}, \sigma^2 > 0$

Definice 6.7 *Řekneme, že náhodná veličina X má logaritmicko-normální rozdělení s parametry $\mu \in \mathbb{R}, \sigma^2 > 0$, jestliže její hustota pravděpodobnosti má tvar*

$$f_X(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \qquad x > 0\,.$$

Značíme $X \sim LN(\mu, \sigma^2)$.

OBR.

Poznámka 6.10 Uvažujeme-li náhodnou veličinu $Y = \ln X$, kde $X \sim LN(\mu, \sigma^2)$, potom $Y \sim N(\mu, \sigma^2)$, z čehož také pochází název logaritmicko-normálního rozdělení. Že to opravdu platí, lze jednoduše ukázat pomocí Věty 6.6 o transformaci náhodné veličiny monotónní funkcí. Naopak, pokud $Y \sim N(\mu, \sigma^2)$, potom $X = e^Y \sim LN(\mu, \sigma^2)$.

Toto rozdělení se používá například při popisu částic sypkých materiálů nebo v teorii spolehlivosti.

6.4.5 Exponenciální rozdělení s parametry $\mu \in \mathbb{R}, \theta > 0$

Definice 6.8 *Řekneme, že náhodná veličina X má exponenciální rozdělení s parametry $\mu \in \mathbb{R}, \theta > 0$, jestliže její hustota pravděpodobnosti má tvar*

$$f_X(x) = \begin{cases} \frac{1}{\theta}\, e^{-\frac{x-\mu}{\theta}} & \textit{pro } x > \mu, \\ 0 & \textit{jinak.} \end{cases}$$

Značíme $X \sim Exp(\mu, \theta)$.

Pro exponenciální rozdělení není problém vyjádřit explicitně distribuční funkci. Platí totiž pro všechna $x > \mu$

$$F_X(x) = \int_\mu^x \frac{1}{\theta}\, e^{-\frac{t-\mu}{\theta}}\, dt = \left[-e^{-\frac{t-\mu}{\theta}}\right]_\mu^x = 1 - e^{-\frac{x-\mu}{\theta}}\,.$$

Poznámka 6.11 Poznamenejme, že exponenciální rozdělení je speciální případ Gamma rozdělení, neboť $Exp(0,\theta) = Exp(\theta) = Gamma(1,\theta)$.

Poznámka 6.12 Exponenciální rozdělení je také někdy nazýváno rozdělení *„bez paměti"*, pro $X \sim Exp(\theta)$ totiž platí

$$\mathsf{P}[X > a + x \,|\, X > a] = \mathsf{P}[X > x]\,.$$

Tento vztah se dá interpretovat například následujícími dvěma způsoby:

- informace o tom, že událost nenastala po dobu a hodin, nemění pravděpodobnost výskytu události v dalších hodinách,
- nechť X značí dobu do poruchy. Pravděpodobnost, že zařízení, které pracovalo bez poruchy a hodin, bude pracovat ještě alespoň x hodin, je rovna pravděpodobnosti, že zařízení, které nebylo v provozu, bude pracovat alespoň x hodin.

Ukažme si, že uvedený vztah skutečně platí. Pro všechny $a > 0, x > 0$,

$$\begin{aligned}\mathsf{P}[X > a + x \,|\, X > a] &= \frac{\mathsf{P}[X > a + x, X > a]}{\mathsf{P}[X > a]} = \frac{1 - \mathsf{P}[X \le a + x]}{1 - \mathsf{P}[X \le a]} \\ &= \frac{e^{-\frac{a+x}{\theta}}}{e^{-\frac{a}{\theta}}} = e^{-\frac{x}{\theta}} = 1 - F_X(x) = \mathsf{P}[X > x]\,.\end{aligned}$$

Exponenciální rozdělení nachází široké uplatnění v teorii spolehlivosti a životnosti (elektronické součástky), v teorii hromadné obsluhy atd.

Cvičení 6.7 Ve velké počítačové síti se průměrně přihlašuje 25 uživatelů za hodinu. Určete pravděpodobnost, že

a) se nikdo nepřihlásí během intervalu délky 6 minut

b) do dalšího přihlášení uběhnou 2-3 minuty

c) určete délku časového intervalu, aby pravděpodobnost, že se nikdo nepřihlásí, byla 0.9.

V následujících třech odstavcích se zmíníme o rozděleních důležitých pro praktické aplikace v matematické statistice.

6.4.6 Pearsonovo χ^2 rozdělení s n stupni volnosti, $n \in I\!N$

Definice 6.9 *Řekneme, že náhodná veličina Y má Pearsonovo χ^2 rozdělení s n stupni volnosti, $n \in I\!N$, jestliže její hustota pravděpodobnosti má tvar*

$$f_Y(y) = \begin{cases} \frac{1}{2\Gamma\left(\frac{1}{2}\right)} \left(\frac{y}{2}\right)^{\frac{n}{2}-1} e^{-\frac{y}{2}} & \text{pro } y > 0, \\ 0 & \text{jinak.} \end{cases}$$

Značíme $Y \sim \chi^2(n)$.

Poznámka 6.13 Platí následující tvrzení. Nechť $X_1, \ldots, X_n$ jsou nezávislé náhodné veličiny, $X_i \sim N(0,1), \forall i \in \widehat{n}$, a definujme náhodnou veličinu

$$Y = X_1^2 + X_2^2 + \ldots + X_n^2 .$$

Potom $Y \sim \chi^2(n)$, tj. Y má χ^2 rozdělení s n stupni volnosti.

χ^2 rozdělení je opět speciálním případem Gamma rozdělení, $\chi^2(n) = Gamma(n/2, 2)$, a ukazuje se jako velmi důležité v testech dobré shody a testování hypotéz.

6.4.7 Studentovo rozdělení s n stupni volnosti, $n \in I\!N$

Definice 6.10 *Nechť X, Y jsou nezávislé náhodné veličiny, $X \sim N(0,1), Y \sim \chi^2(n)$. Potom říkáme, že náhodná veličina*

$$T = \frac{X}{\sqrt{\frac{Y}{n}}}$$

má Studentovo rozdělení s n stupni volnosti. Značíme $T \sim t(n)$.

6.4.8 Fisherovo rozdělení (F-rozdělení)

Definice 6.11 *Nechť X, Y jsou nezávislé náhodné veličiny, $X \sim \chi^2(m), Y \sim \chi^2(n)$. Potom říkáme, že náhodná veličina*

$$Z = \frac{\frac{X}{m}}{\frac{Y}{n}}$$

má Fisherovo rozdělení s m, n stupni volnosti. Značíme $Z \sim F(m, n)$.

Kapitola 7

Charakteristiky náhodných veličin

7.1 Střední hodnota

Definice 7.1 *Nechť náhodná veličina X má diskrétní rozdělení. Potom střední hodnotou náhodné veličiny X nazýváme číslo*

$$\mathsf{E}\, X = \sum_i x_i \mathsf{P}[X = x_i] = \sum_i x_i p_i \,,$$

pokud příslušná řada absolutně konverguje.

Definice 7.2 *Nechť má náhodná veličina X absolutně spojité rozdělení s hustotou f_X. Potom její střední hodnotou rozumíme číslo*

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

pokud integrál absolutně konverguje.

Věta 7.1 *Nechť je dána náhodná veličina X, zobrazení $h : \mathbb{R} \to \mathbb{R}$ a nechť $Y = h(X)$ je náhodná veličina. Potom*

a) má-li X diskrétní rozdělení, platí

$$\mathsf{E}\, Y = \sum_i h(x_i) p_i, \quad \textit{pokud } \mathsf{E}\, Y \textit{ existuje,}$$

b) má-li X absolutně spojité rozdělení, platí

$$\mathsf{E}\, Y = \int_{-\infty}^{\infty} h(x) f_X(x)\, dx, \quad \textit{pokud } \mathsf{E}\, Y \textit{ existuje.}$$

Důkaz. a) Pro diskrétní rozdělení je důkaz triviální. Pro hodnoty x_i, kterých může nabývat náhodná veličina X, platí $h(x_i) = y_i$ a pro získání pravděpodobnostního rozdělení náhodné veličiny Y stačí sečíst hodnoty p_i, p_j pro taková i, j, kde $h(x_i) = h(x_j)$.

b) Dokažme tvrzení pro ostře rostoucí funkci h na $I\!R$. S využitím věty o transformaci a faktu, že mimo interval $(h(-\infty), h(\infty))$ je hustota f_Y nulová, dostaneme

$$\begin{aligned}\int_{-\infty}^{\infty} h(x) f_X(x)\, dx &= \begin{bmatrix} y = h(x) & x = h^{-1}(y) \\ dx = \frac{dh^{-1}(y)}{dy}\, dy & \end{bmatrix} = \int_{h(-\infty)}^{h(\infty)} y f_X(h^{-1}(y)) \frac{dh^{-1}(y)}{dy}\, dy = \\ &= \int_{h(-\infty)}^{h(\infty)} y f_Y(y)\, dy = \int_{-\infty}^{\infty} y f_Y(y)\, dy = \mathsf{E}\, Y\,.\end{aligned}$$

□

Poznámka 7.1 a) Jsou-li X, Y diskrétní náhodné veličiny a $h : I\!R^2 \to I\!R$, potom

$$\mathsf{E}\, h(X,Y) = \sum_{i,j} h(x_i, y_j) \mathsf{P}[X = x_i, Y = y_j] = \sum_{i,j} h(x_i, y_j) p_{ij}\,,$$

pokud řada absolutně konverguje.

b) Jsou-li X, Y absolutně spojité náhodné veličiny a $h : I\!R^2 \to I\!R$, potom

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

pokud integrál absolutně konverguje.

c) Je-li $\boldsymbol{X} = (X_1, \ldots, X_n)$ náhodný vektor, potom definujeme jeho střední hodnotu po složkách, tj.

$$\mathsf{E}\, \boldsymbol{X} = \mathsf{E}\,(X_1, \ldots, X_n) = (\,\mathsf{E}\, X_1, \ldots, \mathsf{E}\, X_n)\,.$$

Věta 7.2 *Nechť mají náhodné veličiny X, Y diskrétní nebo absolutně spojité rozdělení a nechť existují $\mathsf{E}\, X$, $\mathsf{E}\, Y$. Potom existuje i $\mathsf{E}\,(X+Y)$ a platí*

$$\mathsf{E}\,(X+Y) = \mathsf{E}\, X + \mathsf{E}\, Y\,.$$

Důkaz. Označme $Z = X + Y$, tj. $Z = h(X,Y)$. V diskrétním případě

$$\begin{aligned}\mathsf{E}\,(X+Y) &= \sum_{i,j} (x_i + y_j) \mathsf{P}[X = x_i, Y = y_j] = \sum_i x_i \sum_j \mathsf{P}[X = x_i, Y = y_j] \\ &+ \sum_j y_j \sum_i \mathsf{P}[X = x_i, Y = y_j] = \sum_i x_i \mathsf{P}[X = x_i] + \sum_j y_j \mathsf{P}[Y = y_j] \\ &= \mathsf{E}\, X + \mathsf{E}\, Y\,.\end{aligned}$$

Pro spojitý případ

$$\begin{aligned}\mathsf{E}\,(X+Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) f_{X,Y}(x,y)\, dx dy = \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy \right) dx \\ &+ \int_{-\infty}^{\infty} y \left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx \right) dy = \int_{-\infty}^{\infty} f_X(x)\, dx + \int_{-\infty}^{\infty} f_Y(y)\, dy \\ &= \mathsf{E}\, X + \mathsf{E}\, Y\,.\end{aligned}$$

□

Věta 7.3 *Pro diskrétní nebo absolutně spojitou náhodnou veličinu X mající $\mathsf{E}\,X$ a $c \in \mathbb{R}$ platí*
$$\mathsf{E}\,(cX) = c\,\mathsf{E}\,X\,.$$

Důkaz. Ukažme například pro diskrétní
$$\mathsf{E}\,(cX) = \sum_i cx_ip_i = c\sum_i x_ip_i = c\,\mathsf{E}\,X\,.$$
□

Tvrzení 7.1 *Pro libovolné $a \in \mathbb{R}$ platí $\mathsf{E}\,a = a$.*

Důkaz. Konstantu a si můžeme představit jako náhodnou veličinu X, pro kterou platí $\mathsf{P}[X = a] = 1$. Potom $\mathsf{E}\,X = a\,\mathsf{P}[X = a] = a$. □

Důsledek. Pro diskrétní nebo absolutně spojitou náhodnou veličinu X a konstanty $a, c \in \mathbb{R}$ platí
$$\mathsf{E}\,(a + cX) = a + c\,\mathsf{E}\,X\,.$$

Věta 7.4 *Nechť X, Y jsou nezávislé náhodné veličiny mající diskrétní nebo absolutně spojité rozdělení. Potom*
$$\mathsf{E}\,(XY) = \mathsf{E}\,X \cdot \mathsf{E}\,Y\,.$$

Důkaz. Dokažme tvrzení například pro spojité náhodné veličiny. S využitím nezávislosti a Fubiniovy věty dostaneme
$$\begin{aligned}\mathsf{E}\,(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)\,dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\,dxdy \\ &= \left(\int_{-\infty}^{\infty} xf_X(x)\,dx\right)\left(\int_{-\infty}^{\infty} yf_Y(y)\,dy\right) = \mathsf{E}\,X \cdot \mathsf{E}\,Y\,.\end{aligned}$$
□

7.2 Momenty náhodných veličin

Definice 7.3 *Nechť X je náhodná veličina a $k \in \mathbb{N}$. Potom k-tým obecným momentem náhodné veličiny X nazýváme číslo*
$$\mu'_k = \mathsf{E}\,(X^k)\,,$$
pokud existuje. Dále k-tým centrálním momentem náhodné veličiny X nazýváme číslo
$$\mu_k = \mathsf{E}\,(X - \mathsf{E}\,X)^k$$
pokud pravá strana existuje.

Poznámka 7.2 Speciálně
$$\begin{array}{ll}\mu'_0 = 1 & \mu_0 = 1 \\ \mu'_1 = \mathsf{E}\,X & \mu_1 = \mathsf{E}\,(X - \mathsf{E}\,X) = 0 \\ \mu'_2 = \mathsf{E}\,X^2 & \mu_2 = \mathsf{E}\,(X - \mathsf{E}\,X)^2 = \mu'_2 - (\mu'_1)^2\end{array}$$
a obecně platí, že každý centrální moment se dá vyjádřit pomocí obecných.

Definice 7.4 *Druhý centrální moment μ_2 se nazývá rozptyl náhodné veličiny X a značí se*

$$\mathsf{Var}\,X = \mathsf{E}\,(X - \mathsf{E}\,X)^2\,.$$

Směrodatnou odchylkou náhodné veličiny X rozumíme číslo

$$\mathsf{s.d.}\,X = \sqrt{\mathsf{Var}\,X}\,.$$

Poznámka 7.3 $\mathsf{E}\,X$ udává hodnotu, kolem které náhodná veličina kolísá, $\mathsf{Var}\,X$ udává velikost tohoto kolísání.

Věta 7.5 *Nechť X je náhodná veličina pro kterou existuje $\mathsf{E}\,X$, $\mathsf{E}\,X^2$. Potom*

$$\mathsf{Var}\,X = \mathsf{E}\,X^2 - (\,\mathsf{E}\,X)^2\,.$$

Důkaz. Podle pravidel pro počítání se střední hodnotou platí

$$\begin{aligned}\mathsf{Var}\,X &= \mathsf{E}\,(X - \mathsf{E}\,X)^2 = \mathsf{E}\,(X^2 - 2X\,\mathsf{E}\,X + (\,\mathsf{E}\,X)^2) = \mathsf{E}\,X^2 - \mathsf{E}\,(2X\,\mathsf{E}\,X) + \mathsf{E}\,(\,\mathsf{E}\,X)^2 \\ &= \mathsf{E}\,X^2 - 2\,\mathsf{E}\,X(\,\mathsf{E}\,X) + (\,\mathsf{E}\,X)^2 = \mathsf{E}\,X^2 - (\,\mathsf{E}\,X)^2\,.\end{aligned}$$

$\square$

Věta 7.6 *Nechť X je náhodná veličina a $a, b \in I\!R$. Potom*

$$\mathsf{Var}\,(aX + b) = a^2\,\mathsf{Var}\,X\,.$$

Důkaz. Pomocí stejných pravidel jako v předchozím důkazu dostaneme

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

$\square$

Věta 7.7 *Nechť X, Y jsou nezávislé náhodné veličiny. Potom platí*

$$\mathsf{Var}\,(X + Y) = \mathsf{Var}\,X + \mathsf{Var}\,Y\,.$$

Důkaz. Opět podobným způsobem dostaneme v tomto případě

$$\begin{aligned}\mathsf{Var}\,(X + Y) &= \mathsf{E}\,(X + Y - \mathsf{E}\,(X + Y))^2 = \mathsf{E}\,(X - \mathsf{E}\,X + Y - \mathsf{E}\,Y)^2 \\ &= \mathsf{E}\,(X - \mathsf{E}\,X)^2 + \mathsf{E}\,(Y - \mathsf{E}\,Y)^2 + \mathsf{E}\,(2(X - \mathsf{E}\,X)(Y - \mathsf{E}\,Y)) \\ &= \mathsf{Var}\,X + \mathsf{Var}\,Y\,,\end{aligned}$$

protože pro nezávislé náhodné veličiny platí

$$\begin{aligned}\mathsf{E}\,(2(X - \mathsf{E}\,X)(Y - \mathsf{E}\,Y)) &= 2\,\mathsf{E}\,(XY - X\,\mathsf{E}\,Y - Y\,\mathsf{E}\,X + \mathsf{E}\,X\,\mathsf{E}\,Y) \\ &= 2(\,\mathsf{E}\,(XY) - \mathsf{E}\,X\,\mathsf{E}\,Y) = 0\,.\end{aligned}$$

$\square$

Definice 7.5 *Číslo*

$$\alpha_3(X) = \frac{\mu_3(X)}{[\mathsf{Var}\, X]^{\frac{3}{2}}}\,,$$

pokud existuje, nazýváme koeficient šikmosti náhodné veličiny X a číslo

$$\alpha_4(X) = \frac{\mu_4(X)}{(\mathsf{Var}\, X)^2} - 3\,,$$

pokud existuje, nazýváme koeficient špičatosti náhodné veličiny X.

Poznámka 7.4 Koeficient šikmosti měří míru asymetrie, naklonění hustoty náhodné veličiny na jednu stranu. Pokud

- $\alpha_3(X) < 0$ hustota náhodné veličiny je vychýlená vlevo od $\mathsf{E}\, X$,
- $\alpha_3(X) > 0$ hustota náhodné veličiny je vychýlená vpravo od $\mathsf{E}\, X$,
- $\alpha_3(X) = 0$ hustota náhodné veličiny je symetrická vůči $\mathsf{E}\, X$.

Koeficient špičatosti udává, zda je hustota daného rozdělení „plošší“ nebo „špičatější“ než hustota normálního rozdělení.

Obrázky

Poznamenejme jenom, že pokud $X \sim N(\mu, \sigma^2)$, potom $\alpha_3(X) = \alpha_4(X) = 0$.

Příklad 7.1 Uvažujme náhodnou veličinu $X \sim Bi(n, p)$ a určeme její střední hodnotu a rozptyl.

Podle definice střední hodnoty pro diskrétní rozdělení platí

$$\begin{aligned}
\mathsf{E}\, X &= \sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=1}^{n} \frac{n!}{(n-i)!(i-1)!} p^i (1-p)^{n-i} \\
&= np \sum_{i=1}^{n} \frac{(n-1)!}{(n-i)!(i-1)!} p^{i-1} (1-p)^{n-i} = np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-j-1} \\
&= np(p + (1-p))^{n-1} = np\,.
\end{aligned}$$

Podobně,

$$\mathsf{E}\, X^2 = \sum_{i=0}^{n} i^2 \mathsf{P}[X = i] = \sum_{i=0}^{n} i(i-1) \mathsf{P}[X = i] + \sum_{i=0}^{n} i \mathsf{P}[X = i] = n(n-1)p^2 + np\,,$$

kde první suma se sečte stejným postupem jako předchozí střední hodnota a druhá suma je přímo již spočítaná střední hodnota náhodné veličiny X. Pro rozptyl tedy platí

$$\mathsf{Var}\, X = \mathsf{E}\, X^2 - (\mathsf{E}\, X)^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)\,.$$

Příklad 7.2 Uvažujme náhodnou veličinu $X \sim N(\mu, \sigma^2)$ a určeme její střední hodnotu a rozptyl.

Podle definice střední hodnoty pro spojité rozdělení platí

$$
\begin{aligned}
\mathsf{E}\,X &= \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}(x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx + \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\mu e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \mu\,,
\end{aligned}
$$

neboť druhý integrál (po vytknutí konstanty μ) je integrálem z hustoty pravděpodobnosti a jeho hodnota je tedy 1 a první integrál je po zavedení substituce $(x-\mu)/\sigma = t$ roven

$$
\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} te^{-\frac{t^2}{2}}\,dt = \frac{\sigma}{\sqrt{2\pi}}\left[\int_{-\infty}^{0} te^{-\frac{t^2}{2}}\,dt + \int_{0}^{\infty} te^{-\frac{t^2}{2}}\,dt\right] = 0\,,
$$

protože se jedná o antisymetrický integrál. Dále,

$$
\begin{aligned}
\mathsf{E}\,X^2 &= \int_{-\infty}^{\infty} x^2 f_X(x)\,dx = \int_{-\infty}^{\infty}(x-\mu+\mu)^2 f_X(x)\,dx \\
&= \int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)\,dx + 2\mu\int_{-\infty}^{\infty}(x-\mu)f_X(x)\,dx + \mu^2\int_{-\infty}^{\infty} f_X(x)\,dx\,. \quad (7.1)
\end{aligned}
$$

Ze znalosti střední hodnoty $\mathsf{E}\,X = \mu$ dostaneme přímo, že prostřední integrál je nulový a z vlastností hustoty vyplyne, že poslední integrál je roven jedné. Zbývá tedy vyčíslit první integrál, tj.

$$
\begin{aligned}
&\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{(x-\mu)^2}{\sigma^2}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \left|\begin{array}{c} \frac{x-\mu}{\sigma} = t \\ \frac{dx}{\sigma} = dt \end{array}\right| = \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}}\,dt \\
&= \left|\begin{array}{c} \frac{t^2}{2} = s \\ tdt = ds \end{array}\right| = \frac{2\sigma^2}{\sqrt{2\pi}}\int_{0}^{\infty}\sqrt{2s}e^{-s}\,ds = \frac{2\sigma^2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right) = \frac{2\sigma^2}{\sqrt{\pi}}\frac{1}{2}\Gamma\left(\frac{1}{2}\right)\frac{\sigma^2}{\sqrt{\pi}}\sqrt{\pi} = \sigma^2\,.
\end{aligned}
$$

Dosazením do vztahu (7.1) dostaneme

$$
\mathsf{E}\,X^2 = \sigma^2 + \mu^2
$$

a rozptyl má tedy hodnotu

$$
\mathsf{Var}\,X = \mathsf{E}\,X^2 - (\mathsf{E}\,X)^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2\,.
$$

Střední hodnotu náhodného vektoru jsme definovali po složkách, $\mathsf{E}\,(X_1,\ldots,X_n) = (\mathsf{E}\,X_1,\ldots,\mathsf{E}\,X_n)$. Podobně můžeme definovat i další charakteristiky jako jsou rozptyl nebo momenty. Budeme ale popisovat chování jednotlivých složek, což bude postačovat v případě nezávislých náhodných veličin. Pokud ale budou složky náhodného vektoru závislé, budou nás zajímat i charakteristiky popisující jejich interakci. Jednou z nich je kovariance náhodných veličin. Pro jednoduchost budeme všechny pojmy definovat pro dvourozměrné vektory.

Definice 7.6 *Kovarianci náhodných veličin X, Y definujeme jako*

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

pokud střední hodnota existuje.

Jako jednoduchý důsledek této definice vidíme, že platí

$$\mathsf{Cov}\,(X,X) = \mathsf{Var}\,X\,.$$

Věta 7.8 *Nechť X, Y jsou náhodné veličiny. Potom platí*

$$\mathsf{Cov}\,(X,Y) = \mathsf{E}\,(XY) - (\,\mathsf{E}\,X)(\,\mathsf{E}\,Y)\,,$$

pokud příslušné střední hodnoty existují.

Důkaz. Je ponechán studentům. □

Definice 7.7 *Nechť X, Y jsou náhodné veličiny. Číslo*

$$\varrho(X,Y) = \frac{\mathsf{Cov}\,(X,Y)}{\mathsf{s.d.}\,X\ \mathsf{s.d.}\,Y} = \frac{\mathsf{Cov}\,(X,Y)}{\sqrt{\mathsf{Var}\,X\ \mathsf{Var}\,Y}}\,,$$

pokud existuje, nazýváme koeficient korelace náhodných veličin X, Y.

Věta 7.9 *Nechť X, Y jsou náhodné veličiny. Potom platí*

$$-1 \le \varrho(X,Y) \le 1\,.$$

Přitom rovnost

$$\varrho(X,Y) = 1 \text{ nastává} \quad \Leftrightarrow \quad (\exists\beta > 0)(Y - \,\mathsf{E}\,Y = \beta(X - \,\mathsf{E}\,X))\,,$$

$$\varrho(X,Y) = -1 \text{ nastává} \quad \Leftrightarrow \quad (\exists\gamma < 0)(Y - \,\mathsf{E}\,Y = \gamma(X - \,\mathsf{E}\,X))\,.$$

Rovnost $\varrho(X,Y) = 1$ tedy vyjadřuje lineární závislost náhodných veličin X, Y s kladnou směrnicí, rovnost $\varrho(X,Y) = -1$ lineární závislost se zápornou směrnicí.

Lemma 7.1 *(Schwarzova nerovnost) Nechť X, Y jsou náhodné veličiny. Potom platí*

$$(\,\mathsf{E}\,(XY))^2 \le \,\mathsf{E}\,(X^2)\,\mathsf{E}\,(Y^2)\ ,.$$

Důkaz. Definujme náhodnou veličinu $Z = aX - Y$, $a \in \mathbb{R}$. Potom pro všechna $a \in \mathbb{R}$ platí

$$0 \le \,\mathsf{E}\,Z^2 = \,\mathsf{E}\,(aX - Y)^2 = \,\mathsf{E}\,(a^2X^2 - 2aXY + Y^2) = a^2\,\mathsf{E}\,X^2 - 2a\,\mathsf{E}\,(XY) + \,\mathsf{E}\,Y^2\,.$$

Dostali jsme tedy kvadratickou nerovnici v neznámé a, která platí pro všechna $a \in \mathbb{R}$. To ale znamená, že její diskriminant musí být menší nebo roven 0, tj.

$$4(\,\mathsf{E}\,(XY))^2 - 4\,\mathsf{E}\,X^2\,\mathsf{E}\,Y^2 \le 0$$

což je ekvivalentní dokazovanému tvrzení. □

Důkaz věty 7.9. Pokud použijeme Schwarzovu nerovnost pro náhodné veličiny $X - \,\mathsf{E}\,X, Y - \,\mathsf{E}\,Y$, dostaneme

$$[\,\mathsf{E}\,(X - \,\mathsf{E}\,X)(Y - \,\mathsf{E}\,Y)]^2 \le \,\mathsf{E}\,(X - \,\mathsf{E}\,X)^2\,\mathsf{E}\,(Y - \,\mathsf{E}\,Y)^2$$

což je ekvivalentní zápisu

$$\mathsf{Cov}^2(X,Y) \leq \mathsf{Var}\,X\,\mathsf{Var}\,Y$$

a tedy

$$\varrho^2(X,Y) \leq 1\,.$$

Z ekvivalence uvedené v druhé části tvrzení dokážeme pouze jednu implikaci a to zprava doleva. Předpokládejme tedy, že $Y - \mathsf{E}\,Y = \beta(X - \mathsf{E}\,X)$. To znamená, že $Y = \beta X + \mathsf{E}\,Y - \beta\,\mathsf{E}\,X$ a Y se dá vyjádřit ve tvaru $Y = \beta X + C$, kde C je konstanta. Potom

$$\begin{aligned} \mathsf{Cov}\,(X,Y) &= \mathsf{E}\,(X - \mathsf{E}\,X)(\beta X + C - \mathsf{E}\,(\beta X + C)) \\ &= \mathsf{E}\,(\beta X^2 - 2\beta X\,\mathsf{E}\,X + \beta(\mathsf{E}\,X)^2) = \beta\,\mathsf{E}\,(X - \mathsf{E}\,X)^2 = \beta\,\mathsf{Var}\,X\,. \end{aligned}$$

Protože dále podle pravidel pro počítání s rozptylem platí $\mathsf{Var}\,Y = \beta^2\,\mathsf{Var}\,X$ dostáváme pro koeficient korelace

$$\varrho(X,Y) = \frac{\beta\,\mathsf{Var}\,X}{\sqrt{\beta^2\,\mathsf{Var}^2 X}} = \begin{cases} +1 & \text{pro } \beta > 0, \\ -1 & \text{pro } \beta < 0. \end{cases}$$

□

Definice 7.8 *Nechť X, Y jsou náhodné veličiny. Pokud $\varrho(X,Y) = 0$, pak se náhodné veličiny X, Y nazývají nekorelované.*

Věta 7.10 *Nezávislé náhodné veličiny X, Y jsou nekorelované.*

Důkaz. Z nezávislosti náhodných veličin vyplývá

$$\mathsf{Cov}\,(X,Y) = \mathsf{E}\,(XY) - \mathsf{E}\,X\,\mathsf{E}\,Y = \mathsf{E}\,X\,\mathsf{E}\,Y - \mathsf{E}\,X\,\mathsf{E}\,Y = 0$$

a tedy i $\varrho(X,Y) = 0$. □

Příklad 7.3 Abychom ilustrovali, že předchozí tvrzení neplatí v opačném směru uvažujme následující příklad. Nechť náhodné veličiny X, Y mají rovnoměrné rozdělení na kruhu $K = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 \leq 1\}$, tj.

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi} & \text{pro } (x,y) \in K, \\ 0 & \text{jinak.} \end{cases}$$

Úkolem je rozhodnout, zda jsou náhodné veličiny X, Y nezávislé a nekorelované.

Marginální hustoty mají tvar

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,dy = \frac{2}{\pi}\sqrt{1-x^2}, \quad x \in \langle -1, 1 \rangle$$

a ze symetrie i

$$f_Y(y) = \frac{2}{\pi}\sqrt{1-y^2}, \quad y \in \langle -1, 1 \rangle\,.$$

Protože tedy platí $f_{X,Y}(x,y) \neq f_X(x) f_Y(y)$, veličiny X, Y nejsou nezávislé. Spočtěme jejich kovarianci,

$$\mathsf{E}\,X = \int_{-1}^{1} \frac{2}{\pi} x\sqrt{1-x^2}\,dx = \frac{2}{\pi}\left[\int_{-1}^{0} x\sqrt{1-x^2}\,dx + \int_{0}^{1} x\sqrt{1-x^2}\,dx\right] = 0$$

a za symetrie i $\mathsf{E}\,Y = 0$. Dále,

$$\mathsf{E}\,(XY) = \iint_K xy\frac{1}{\pi}\,dxdy = \frac{1}{\pi}\int_{-1}^{1}\left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} xy\,dy\right)dx = 0\,,$$

kovariance má tudíž hodnotu $\mathsf{Cov}\,(X,Y) = \mathsf{E}\,(XY) - \mathsf{E}\,X\,\mathsf{E}\,Y = 0$ a veličiny X, Y jsou nekorelované.

Poznámka 7.5 Koeficient korelace je vlastně kovariance normovaná na interval $\langle -1, 1\rangle$, což umožňuje lepší srovnání, a vyjadřuje lineární závislost.

Velká absolutní hodnota $\varrho(X,Y)$ vyjadřuje velkou míru lineární závislosti náhodných veličin X, Y. Nízká hodnota $\varrho(X,Y)$ říká, že veličiny X, Y jsou „téměř nekorelované“, což může znamenat, že jsou nezávislé nebo, že jejich závislost je odlišná od lineární.

Vysoká hodnota $\varrho(X,Y)$ navíc znamená, že se hodnoty obou veličin vyvíjejí podobně, nemusí ale mezi nimi existovat příčinný vztah!

Definice 7.9 *Nechť $\boldsymbol{X} = (X_1, \dots, X_n)$ je náhodný vektor. Matici*

$$\Sigma = (\sigma_{ij})_{i,j=1}^{n}, \quad \textit{kde} \quad \sigma_{ij} = \mathsf{Cov}\,(X_i, X_j)\,,$$

nazveme kovariační maticí náhodného vektoru $\boldsymbol{X}$.

Poznámka 7.6 Z definice vyplývá. že je matice Σ symetrická, na diagonále má rozptyly $\sigma_{ii} = \mathsf{Var}\,X_i$ náhodných veličin $X_1, \dots, X_n$ a pokud jsou tyto veličiny nezávislé je Σ diagonální matice.

Cvičení 7.1 Nechť má náhodná veličina X rovnoměrné rozdělení na intervalu (a, b), tj. $X \sim U(a, b)$. Určete střední hodnotu $\mathsf{E}\,X$ a rozptyl $\mathsf{Var}\,X$.

Cvičení 7.2 Spojitá náhodná veličina má trojúhelníkové rozdělení na intervalu $\langle -a, a\rangle$ (hustota je rovnoramenný trojúhelník se základnou $\langle -a, a\rangle$, mimo je nulová). Zapište matematicky tuto hustotu a spočtěte $\mathsf{E}\,X$, $\mathsf{Var}\,X$.

Cvičení 7.3 Váha tělesa se může se stejnou pravděpodobností rovnat každému celému číslu od 1 do 10. Jsou k dispozici tři sady závaží o hmotnostech

a) 1,2,2,5,10

b) 1,2,3,4,10

c) 1,1,2,5,10 .

Závaží se dává jen na jednu misku dvoumiskových vah a používá se co nejméně závaží. Určete při které sadě je průměrný počet závaží nutných k vyvážení minimální.

Cvičení 7.4 Firma získá z každého prodaného výrobku 100 Kč. Za výměnu během záruční doby zaplatí 300 Kč. Životnost výrobku v letech má normální rozdělení $N(3, 1)$. Jakou záruční dobu v měsících má firma stanovit, aby střední (průměrný) zisk byl alespoň 60 Kč?

Cvičení 7.5 Určete hodnotu konstanty c a koeficient korelace diskrétních náhodných veličin X, Y zadaných sdruženou pravděpodobnostní funkcí

$$P[X = x, Y = y] = c\,(x + y), \qquad \text{pro} \qquad x = 1, 2, 3 \quad \text{a} \quad y = 1, 2, 3.$$

[-1/23]

Cvičení 7.6 Náhodný vektor (X, Y) má hustotu

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{6}\left(\frac{x}{2} + \frac{y}{3}\right) & x \in (0, 2) \text{ a } y \in (0, 3) \\ 0 & \text{jinak} . \end{cases}$$

Určete marginální hustoty pravděpodobnosti f_X, f_Y a spočtěte koeficient korelace $\varrho(X, Y)$ náhodných veličin X, Y. [-1/11]

7.3 Kvantily náhodných veličin

Definice 7.10 *Nechť X je náhodná veličina s distribuční funkcí F_X a nechť $\alpha \in (0, 1)$. Potom bod x_α nazýváme α-kvantilem náhodné veličiny X, právě když platí*

$$x_\alpha = \inf\{x : F_X(x) \geq \alpha\} .$$

Pokud je F_X ostře rostoucí a spojitá, potom je x_α takový bod z $\mathbb{R}$, že

$$F_X(x_\alpha) = \alpha, \quad \textit{tj.} \quad x_\alpha = F_X^{-1}(\alpha) .$$

Obrázky

Poznámka 7.7 Některé kvantily mají vlastní názvy:

$x_{0.5}$ - medián (50% dat leží vlevo od této hodnoty, 50% dat vpravo)

$x_{0.25}$ - dolní kvartil

$x_{0.75}$ - horní kvartil

$x_{0.1}$ - dolní decil

$(x_{0.75} - x_{0,25})$ - mezikvartilové rozpětí (délka intervalu, kde leží 50% nejčastějších hodnot) .

Příklad 7.4 Uvažujme náhodnou veličinu X s hustotou pravděpodobnosti

$$f_X(x) = \frac{1}{\pi} \frac{\theta}{\theta^2 + (x - \mu)^2}, \quad \theta > 0, \mu \in \mathbb{R}, x \in \mathbb{R} .$$

Jedná se o takzvané Cauchyovo rozdělení. Není těžké ověřit, že neexistuje střední hodnota náhodné veličiny X, neboť integrál

$$\mathsf{E}\,X = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\theta x}{\theta^2 + (x - \mu)^2}\, dx$$

neexistuje. Tuto charakteristiku tedy nemůžeme použít. Distribuční funkci, ale vyjádřit lze. Pro všechna $x \in \mathbb{R}$ platí

$$
\begin{aligned}
F_X(x) &= \frac{1}{\pi}\int_{-\infty}^{x} \frac{\theta}{\theta^2+(t-\mu)^2}\,dt = \left| \begin{array}{c} \frac{t-\mu}{\theta} = y \\ \frac{1}{\theta}dt = dy \end{array} \right| = \frac{1}{\pi}\int_{-\infty}^{\frac{x-\mu}{\theta}} \frac{1}{1+y^2}\,dy \\
&= \frac{1}{\pi}\operatorname{arctg}\frac{x-\mu}{\theta} + \frac{1}{2}\,.
\end{aligned}
$$

α-kvantil tedy určíme z rovnosti

$$
F_X(x_\alpha) = \alpha, \quad \text{tj.} \quad \frac{1}{\pi}\operatorname{arctg}\frac{x_\alpha-\mu}{\theta} + \frac{1}{2} = \alpha
$$

a tedy

$$
x_\alpha = \theta \operatorname{tg}\left(\pi\left(\alpha - \frac{1}{2}\right)\right) + \mu\,.
$$

Speciálně pro medián platí $x_{0.5} = \mu$.

7.4 Charakteristická a momentová funkce

Kapitola 8

Limitní věty

8.1 Zákon velkých čísel

Dosud jsme se zabývali jednotlivými náhodnými veličinami. V této kapitole budeme studovat posloupnosti náhodných veličin. Příklady takových posloupností zajímavých prakticky i teoreticky:

- Nechť Z_n označuje počet výskytů jistého jevu A v n nezávislých pokusech, přičemž pravděpodobnost jevu A v jednotlivých pokusech je p. Označme $X_n = Z_n/n$, X_n tedy vyjadřuje relativní četnost jevu A v n pokusech. Jak se chová X_n pro velká n?
- Uvažujme n nezávislých měření $X_1, \ldots, X_n$ a jejich průměr $\overline{X}_n = 1/n \sum_{i=1}^{n} X_i$. Má rozdělení náhodné veličiny $\overline{X}_n$ pro velká n nějaké zajímavé vlastnosti? Závisí na rozdělení veličin X_i?

Definice 8.1 *Nechť $(X_n)_{n=1}^{\infty}$ je posloupnost náhodných veličin a $c \in \mathbb{R}$. Říkáme, že posloupnost $(X_n)_{n=1}^{\infty}$ konverguje k c podle pravděpodobnosti, jestliže pro každé $\varepsilon > 0$ platí*

$$\lim_{n\to\infty} \mathsf{P}[|X_n - c| \geq \varepsilon] = 0\,.$$

Značíme $X_n \xrightarrow{P} c$.

Příklad 8.1 Jsou dány náhodné veličiny $X_n \sim N(\mu, \sigma^2/n)$ Dokažte, že $X_n \xrightarrow{P} \mu$.

Podle definice máme ukázat, že pro každé $\varepsilon > 0$

$$\mathsf{P}[|X_n - \mu| \geq \varepsilon] \to 0\,,$$

což je ekvivalentní tvrzení, že pro každé $\varepsilon > 0$

$$\mathsf{P}[|X_n - \mu| < \varepsilon] \to 1\,.$$

Vyjádřeme tuto pravděpodobnost. Pro každé $\varepsilon > 0$ platí

$$\begin{aligned}
\mathsf{P}[|X_n - \mu| < \varepsilon] &= \mathsf{P}[\mu - \varepsilon < X_n < \mu + \varepsilon] = \int_{\mu-\varepsilon}^{\mu+\varepsilon} \frac{1}{\sqrt{2\pi\frac{\sigma^2}{n}}}\, e^{-\frac{n(x-\mu)^2}{2\sigma^2}}\, dx \\
&= \left| \begin{array}{c} \frac{x-\mu}{\sigma/\sqrt{n}} = t \\ \frac{1}{\sigma/\sqrt{n}} dx = dt \end{array} \right| = \int_{-\frac{\sqrt{n}\varepsilon}{\sigma}}^{\frac{\sqrt{n}\varepsilon}{\sigma}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t^2}{2}}\, dt \xrightarrow{n\to\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t^2}{2}}\, dt = 1\,.
\end{aligned}$$

Věta 8.1 *(Čebyševova nerovnost) Nechť X je náhodná veličina se střední hodnotou $\mathsf{E}\,X$ a konečným rozptylem $\mathsf{Var}\,X$. Potom pro každé $\varepsilon > 0$ platí*

$$\mathsf{P}[|X - \mathsf{E}\,X| \geq \varepsilon] \leq \frac{\mathsf{Var}\,X}{\varepsilon^2}\,.$$

Důkaz. Proveďme důkaz pro spojité náhodné veličiny. Uvažujme nejprve náhodnou veličinu Y s konečným druhým momentem $\mathsf{E}\,Y^2$. Potom pro každé $\varepsilon > 0$ platí

$$\begin{aligned}\mathsf{E}\,(Y^2) &= \int_{-\infty}^{\infty} y^2 f_Y(y)\,dy = \int_{|y|<\varepsilon} y^2 f_Y(y)\,dy + \int_{|y|\geq\varepsilon} y^2 f_Y(y)\,dy \\ &\geq \int_{|y|\geq\varepsilon} y^2 f_Y(y)\,dy \geq \varepsilon^2 \int_{|y|\geq\varepsilon} f_Y(y)\,dy = \varepsilon^2 \mathsf{P}[|Y| \geq \varepsilon]\,.\end{aligned}$$

Aplikací této nerovnosti na náhodnou veličinu $Y = X - \mathsf{E}\,X$ dostaneme

$$\mathsf{P}[|X - \mathsf{E}\,X| \geq \varepsilon] \leq \frac{\mathsf{E}\,(X - \mathsf{E}\,X)^2}{\varepsilon^2} = \frac{\mathsf{Var}\,X}{\varepsilon^2}\,.$$

□

Věta 8.2 *(Zákon velkých čísel) Nechť $(X_n)_{n=1}^{\infty}$ je posloupnost nezávislých, stejně rozdělených náhodných veličin, pro které existuje $\mu = \mathsf{E}\,X_i$ a $\sigma^2 = \mathsf{E}\,(X_i - \mu)^2$, $i \in \mathbb{N}$. Potom*

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu\,.$$

Důkaz. K důkazu použijeme Čebyševovy nerovnosti. Spočtěme nejdříve střední hodnotu a rozptyl náhodné veličiny $\overline{X}_n$. Podle pravidel pro počítání se střední hodnotou

$$\mathsf{E}\,\overline{X}_n = \mathsf{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\,X_i = \mu$$

a s využitím nezávislosti dále

$$\mathsf{Var}\,\overline{X}_n = \mathsf{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathsf{Var}\,X_i = \frac{\sigma^2}{n}\,.$$

Dosazením do Čebyševovy nerovnosti dostaneme, že pro všechna $\varepsilon > 0$ platí

$$0 \leq \mathsf{P}[|\overline{X}_n - \mu| \geq \varepsilon] \leq \frac{\sigma^2}{n\varepsilon^2} \longrightarrow 0\,,$$

což je podle definice konvergence podle pravděpodobnosti ekvivalentní dokazovanému tvrzení. □

Poznámka 8.1 Speciálním případem předchozí věty je Bernoulliho věta (1713). Nechť jsou dány nezávislé náhodné veličiny $X_n \sim Be(p)$ a označme $S_n = \sum_{i=1}^{n} X_i$. S_n tedy vyjadřuje počet úspěchů v n Bernoulliovských pokusech. Potom platí

$$\frac{S_n}{n} \xrightarrow{P} p\,.$$

Toto tvrzení vyplývá např. přímo z předcházející věty užitím faktu, že střední hodnota náhodné veličiny $X \sim Be(p)$ je p.

Příklad 8.2 S jakou pravděpodobností můžeme tvrdit, že při 1000 hodech mincí padne orel 400-600 krát?

Nechť veličiny $X_1, \dots, X_{1000}$ vyjadřují , zda v i-tém hodu padl orel (hodnota 1) nebo panna (hodnota 0). Veličiny X_i jsou nezávislé a platí pro ně

$$\mathsf{E}\, X_i = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0 = \frac{1}{2}, \quad \mathsf{E}\, X_i^2 = \frac{1}{2} \cdot 1^2 + \frac{1}{2} \cdot 0^2 = \frac{1}{2} \quad \text{a} \quad \mathsf{Var}\, X_i = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}\,.$$

Označme dále $Y = \sum_{i=1}^{1000} X_i$ počet orlů v 1000 hodech mincí. Pro náhodnou veličinu Y platí

$$\mathsf{E}\, Y = \sum_{i=1}^{1000} \mathsf{E}\, X_i = 500, \quad \mathsf{Var}\, Y = \sum_{i=1}^{1000} \mathsf{Var}\, X_i = 250\,.$$

Podle Čebyševovy nerovnosti platí pro všechna $\varepsilon > 0$ nerovnost

$$\mathsf{P}[|Y - \mathsf{E}\, Y| \geq \varepsilon] \leq \frac{\mathsf{Var}\, Y}{\varepsilon^2}\,.$$

V našem konkrétním případě dostaneme

$$\mathsf{P}[|Y - 500| \geq 100] \leq \frac{250}{100^2}\,,$$

tedy

$$\mathsf{P}[|Y - 500| < 100] \geq 1 - 0,025 = 0,975\,.$$

S pravděpodobností vyšší než 0,975 můžeme tedy říci, že v 1000 hodech mincí padne orel 400-600 krát.

8.2 Centrální limitní věta

V různých úlohách matematické statistiky nás zajímá rozdělení součtu nebo průměru n nezávislých náhodných veličin. Někdy je však stanovení přesného rozdělení obtížné. Centrální limitní věta říká, že za jistých podmínek lze toto rozdělení aproximovat normálním.

Definice 8.2 *Nechť $(X_n)_{n=1}^{\infty}$ je posloupnost náhodných veličin s distribučními funkcemi F_{X_n} a nechť X je náhodná veličina s distribuční funkcí F_X. Řekneme, že posloupnost $(X_n)_{n=1}^{\infty}$ konverguje k náhodné veličině X, jestliže*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

ve všech bodech spojitosti funkce F_X. Značíme $X_n \xrightarrow{L} X$ (in law).

Poznámka 8.2 Řekneme tedy, že $(X_n)_{n=1}^{\infty}$ konverguje v distribuci k rozdělení $N(\mu, \sigma^2)$, jestliže

$$\lim_{n\to\infty} F_{X_n}(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

pro všechna $x \in I\!\!R$, budeme značit $X_n \xrightarrow{L} N(\mu, \sigma^2)$.

Věta 8.3 *(CLV - Lindeberg-Lévy) Nechť* $(X_n)_{n=1}^{\infty}$ *je posloupnost nezávislých, stejně rozdělených náhodných veličin se společnou střední hodnotou* μ *a konečným rozptylem* σ^2. *Potom platí*

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow{L} N(0,1)\,.$$

Poznámka 8.3 Pokud označíme $Y_n = \sum_{i=1}^{n} X_i$, platí $\mathsf{E}\, Y_n = n\mu$, $\mathsf{Var}\, Y_n = n\sigma^2$ a tvrzení lze zapsat v lépe pamatovatelném tvaru

$$\frac{Y_n - \mathsf{E}\, Y_n}{\sqrt{\mathsf{Var}\, Y_n}} \xrightarrow{L} N(0,1)\,.$$

Další možností je přepsat tvrzení pomocí průměru $\overline{X}_n$ do tvaru

$$\sqrt{n}\,\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow{L} N(0,1)\,.$$

Věta 8.4 *(Moivre-Laplace 1718) Nechť* $(X_n)_{n=1}^{\infty}$ *je posloupnost nezávislých, stejně rozdělených náhodných veličin,* $X_i \sim Be(p)$, *pro všechna* $i \in I\!N$. *Potom platí*

$$\frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}} \xrightarrow{L} N(0,1)\,.$$

Poznámka 8.4 Pokud jsou náhodné veličiny $X_i \sim Be(p)$ nezávislé, platí $Y_n = \sum_{i=1}^{n} X_i \sim Bi(n,p)$ a podle předchozí věty

$$\lim_{n\to\infty} \mathsf{P}\left[\frac{Y_n - np}{\sqrt{np(1-p)}} \le x\right] = \Phi(x), \quad \text{pro všechna} \quad x \in I\!R\,.$$

Této verze centrální limitní věty se dá výhodně využít pro aproximaci binomického rozdělení pomocí normálního. Taková aproximace se doporučuje pro $np(1-p) \ge 9$.

Příklad 8.3 Loď má nosnost 5000 kg. Váha cestujících je náhodná veličina se střední hodnotou 70 kg a směrodatnou odchylkou 20 kg. Kolik cestujících můžeme nalodit, aby pravděpodobnost přetížení člunu byla menší než 0.001?

Označme X_i hmotnost i-tého cestujícího a $X = \sum_{i=1}^{n} X_i$ hmotnost všech cestujících. Potom

$$\mathsf{E}\, X = \sum_{i=1}^{n} \mathsf{E}\, X_i = 70n \quad \text{a} \quad \mathsf{Var}\, X = \sum_{i=1}^{n} \mathsf{Var}\, X_i = 400n\,.$$

(za předpokladu, že hmotnosti jednotlivých cestujících jsou nezávislé). Zajímá nás hodnota n tak, aby

$$\mathsf{P}[X \ge 5000] < 0.001\,.$$

Pokud označíme U náhodnou veličinu s rozdělením $N(0,1)$, dostaneme pomocí CLV aproximaci

$$\begin{aligned}\mathsf{P}[X \ge 5000] &= \mathsf{P}\left[\frac{X - 70n}{20\sqrt{n}} \ge \frac{5000 - 70n}{20\sqrt{n}}\right] \doteq \mathsf{P}\left[U \ge \frac{5000 - 70n}{20\sqrt{n}}\right] \\ &= 1 - \Phi\left(\frac{5000 - 70n}{20\sqrt{n}}\right) < 0.001\,,\end{aligned}$$

jinak zapsáno

$$\Phi\left(\frac{5000-70n}{20\sqrt{n}}\right) > 0.999\,.$$

Po nalezení příslušné hodnoty funkce Φ v tabulkách dostaneme nerovnost

$$\frac{5000-70n}{20\sqrt{n}} > 3.09$$

neboli

$$70n + 61.8\sqrt{n} - 5000 < 0\,.$$

Vyřešením této kvadratické nerovnice v proměnné $\sqrt{n}$ najdeme jeden přípustný kořen $\sqrt{n_1} = 8.02$ a závěr je, že můžeme nalodit maximálně 64 cestujících.

Příklad 8.4 S jakou pravděpodobností můžeme tvrdit, že při 1000 hodech mincí padne orel 450-550 krát?

V označení zavedeném v příkladu 8.2 máme najít pravděpodobnost $\mathsf{P}[450 \le Y \le 550]$. V tomto případě je hodnota $np(1-p) = 250$ dostatečně velká, abychom mohli použít aproximaci pomocí centrální limitní věty a s pomocí náhodné veličiny $U \sim N(0,1)$ dostáváme

$$\begin{aligned} \mathsf{P}[450 \le Y \le 550] &= \mathsf{P}\left[\frac{-50}{\sqrt{250}} \le \frac{Y-500}{\sqrt{250}} \le \frac{50}{\sqrt{250}}\right] \doteq \mathsf{P}\left[\frac{-50}{\sqrt{250}} \le U \le \frac{50}{\sqrt{250}}\right] \\ &= 2\Phi\left(\frac{50}{\sqrt{250}}\right) - 1 = 0.9984\,. \end{aligned}$$

Podobně bychom například spočítali

$$\mathsf{P}[480 \le Y \le 520] \doteq 2\Phi\left(\frac{20}{\sqrt{250}}\right) - 1 = 0.7941\,.$$

Pro ilustraci, přesná hodnota spočtená mnohem náročněji pomocí binomického rozdělení je v tomto případě 0.7939.

Příklad 8.5 Házíme symetrickou kostkou. S jakou pravděpodobností padne v 600 hodech více než 110 šestek?

Označme Y počet šestek ze 600 hodů. Y má tedy binomické rozdělení s parametry $n = 600$ a $p = 1/6$ a hodnota $np(1-p) = 250/3$ nás opravňuje použít aproximaci normálním rozdělením. Již známým postupem dostaneme

$$\begin{aligned} \mathsf{P}[Y \ge 110] &= \mathsf{P}\left[\frac{Y-100}{\sqrt{250/3}} \ge \frac{10}{\sqrt{250/3}}\right] \doteq \mathsf{P}\left[U \ge \frac{10}{\sqrt{250/3}}\right] \\ &= 1 - \Phi\left(\frac{10}{\sqrt{250/3}}\right) = 1 - 0.863 = 0.137\,. \end{aligned}$$

Příklad 8.6 Před volbami je v populaci 52% příznivců koalice. Jaká je pravděpodobnost, že průzkum veřejného mínění o rozsahu $n = 1500$ respondentů ukáže nesprávně převahu opozice?

Označme X počet příznivců koalice ve výběru. Pokud byl výběr proveden nezávisle, má X binomické rozdělení s parametry $n = 1500$ a $p = 0.52$ a $np = 780$, $np(1-p) = 374.4$. Použitím aproximace dostáváme, že hledaná hodnota pravděpodobnosti je přibližně

$$\mathsf{P}[X \leq 749] = \mathsf{P}\left[\frac{X - 780}{\sqrt{374.4}} \leq \frac{-31}{\sqrt{374.4}}\right] \doteq \Phi(-1.602) = 1 - \Phi(1.602) = 0.055\,.$$

Příklad 8.7 Uvažujme nezávislé náhodné veličiny $X_i \sim U(0,1)$ rovnoměrně rozdělené na intervalu $(0,1)$. Protože pro toto rozdělení platí

$$\mathsf{E}\,X_i = \frac{1}{2}, \quad \mathsf{E}\,X_i^2 = \int_0^1 x^2\,dx = \frac{1}{3} \quad \text{a} \quad \mathsf{Var}\,X_i = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

dostaneme použitím centrální limitní věty

$$\frac{\sum_{i=1}^n X_i - \frac{n}{2}}{\sqrt{\frac{n}{12}}} \xrightarrow{L} N(0,1)\,.$$

Tohoto limitního vztahu se dá využít pro sestrojení jednoduchého generátoru náhodných čísel z rozdělení $N(0,1)$. Například pro $n = 12$ podle předchozího vztahu $U = \sum_{i=1}^{12} X_i - 6$ má přibližně standardní normální rozdělení. Stačí tedy nagenerovat 12 čísel z rozdělení $U(0,1)$ a jejich sečtením a odečtením 6 dostaneme realizaci náhodného čísla s rozdělením blízkým standardnímu normálnímu.

Kapitola 9

Statistika

Zjednodušeně lze popsat úlohu matematické statistiky takto: pro danou reálnou situaci vytvoříme model a na základě experimentálních dat pak odhadujeme hodnoty volných parametrů modelu, testujeme hypotézy o těchto parametrech a ověřujeme shodu modelu se skutečností (v jistém smyslu opačný postup oproti pravděpodobnosti). Kroky:

- odhad tvaru rozdělení (volba modelu) - vychází z intuitivní představy nebo zkušenosti. Pomůckou může být např. histogram. **Obr.**
- odhad parametrů rozdělení (modelu)
  a) bodový odhad - předpokládáme parametrický model $X \sim F_X(x, \theta)$ s neznámou hodnotou parametru θ. Úkolem je najít funkci $\widehat{\theta}(\boldsymbol{X})$ vektoru pozorování $\boldsymbol{X}$, která odhaduje nejpravděpodobnější hodnotu parametru θ.
  b) intervalový odhad - předpokládáme stejný model, ale hledáme interval, kde leží hodnota parametru θ s jistou pravděpodobností. Tj. hledáme funkce $\underline{\theta}(\boldsymbol{X}), \overline{\theta}(\boldsymbol{X})$ takové, že $\mathsf{P}[\theta \in (\underline{\theta}(\boldsymbol{X}), \overline{\theta}(\boldsymbol{X}))] \geq 1 - \alpha$ pro $\alpha \in (0, 1)$.
- ověření správnosti modelu
  a) testování hypotéz (parametrické testy) - opět uvažujeme parametrický model a nějakou hypotézu o parametru θ. Na základě dat se snažíme rozhodnout, zda je nebo není možné zamítnout tuto hypotézu. Například hypotézou může být, že správná hodnota parametru θ je 5, $H_0 : \theta = 5$. Na základě dat zkoumáme, zda $\mathsf{P}[H_0 \text{ platí}] \geq 1 - \alpha$. Pokud ne, zamítáme hypotézu H_0.
  b) testy dobré shody - v tomto případě bude otázkou přímo tvar modelu nebo pravděpodobnostního rozdělení. Např. otázka: řídí se počet zákazníků Poissonovým rozdělením?

Symboly Ω, ω, X budou mít ve statistice poněkud odlišný význam. Ω bude představovat nějakou populaci, ω bude individuum, element a zobrazení $X : \Omega \to \mathbb{R}$ bude vyjadřovat určitou vlastnost populace Ω.

Definice 9.1 *n-tici nezávislých náhodných veličin $X_1, \ldots, X_n$, stejně rozdělených s distribuční funkcí F, nazýváme náhodný výběr z rozdělení F.*

Poznámka 9.1 Příklady výběrů:

- provedeme n-krát fyzikální měření, experiment, nezávisle na sobě, stejnou technologií
- z kamionu rajčat náhodně vybereme n z nich a otestujeme obsah dusičnanů v nich

Poznámka 9.2 Výsledný model bude záviset na velikosti výběru. Nejlepší by zřejmě bylo provést test pro všechny elementy. To ale může být drahé nebo i nemožné (destruktivní test). Dalším problémem je, jak zajistit nezávislost (např. při výzkumech veřejného mínění).

Poznámka 9.3 Jakmile je pokus hotov, náhodný výběr se stává vektorem naměřených hodnot (pevných čísel). Je to vlastně jedna realizace, budeme ji značit malými písmeny. $x_1, \dots, x_n$ tedy bude značit jednu konkrétní realizaci náhodného výběru $X_1, \dots, X_n$.

Z realizací náhodných výběru potom počítáme charakteristiky umožňující jisté shrnutí výsledků, případně provedení závěrů o vlastnostech rozdělení, které studujeme. Např. u n fyzikálních měření může jít o stanovení jedné konkrétní hodnoty. Opakování má zmenšit možnou chybu výsledku, v tomto případě bude vhodnou charakteristikou průměr naměřených hodnot.

Definice 9.2 *Nechť $X_1, \dots, X_n$ je náhodný výběr. Statistikou rozumíme každou funkci náhodného výběru $X_1, \dots, X_n$, která nezávisí na konkrétních hodnotách parametrů příslušného rozdělení.*

Poznámka 9.4 Příklady nejužívanějších statistik:

- výběrový průměr
$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \,.$$
Jedná se o protějšek střední hodnoty. Pokud $\mathsf{E}\, X_i = \mu$ a $\mathsf{Var}\, X_i = \sigma^2$, potom $\mathsf{E}\, \overline{X}_n = \mu$ a $\mathsf{Var}\, \overline{X}_n = \sigma^2/n$.
- výběrový rozptyl
$$s_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2 \,,$$
protějšek rozptylu. Ukážeme, že $\mathsf{E}\, s_n^2 = \sigma^2$.
- výběrová směrodatná odchylka
$$s_n = \left[\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2 \right]^{\frac{1}{2}} .$$
- výběrový r-tý moment
$$m_r = \frac{1}{n} \sum_{i=1}^{n} X_i^r, \quad r = 1, 2, \dots$$

9.1 Odhady parametrů

9.1.1 Bodové odhady

V této kapitole budeme uvažovat náhodný výběr z rozdělení patřícího do parametrické rodiny distribucí $\{F_X(x,\theta)|\theta \in \Theta\}$, kde $\Theta \subset \mathbb{R}^d$ je množina všech možných hodnot parametru θ(může se tedy jednak i o vektor).

Definice 9.3 *Bodový odhad parametru θ je jakákoliv funkce $\widehat{\theta}_n(X_1,\dots,X_n)$ náhodného výběru $X_1,\dots,X_n$, jejíž funkční předpis nezávisí na θ.*

Poznámka 9.5 Bodový odhad je tedy statistika. Protože je to funkce náhodných veličin, je sám o sobě náhodná veličina s nějakým rozdělením.

Definice 9.4 *Odhad $\widehat{\theta}_n(X_1,\dots,X_n)$ parametru θ se nazývá nestranný, jestliže*

$$\mathsf{E}\,\widehat{\theta}_n(X_1,\dots,X_n) = \theta \quad \textit{pro všechna} \quad \theta \in \Theta\,.$$

Definice 9.5 *Odhad $\widehat{\theta}_n(X_1,\dots,X_n)$ parametru θ se nazývá konzistentní, jestliže pro $n \to \infty$*

$$\widehat{\theta}_n(X_1,\dots,X_n) \xrightarrow{P} \theta \quad \textit{pro všechna} \quad \theta \in \Theta\,.$$

Poznámka 9.6 Nestrannost znamená, že odhad není zatížen systematickou chybou. Konzistence znamená, že volbou dostatečně velkého n lze učinit chybu libovolně malou.

Obr.

Konzistence odhadu $\widehat{\theta}_n(X_1,\dots,X_n)$ je ekvivalentní současné platnosti dvojice limitních vztahů

$$\mathsf{E}\,\widehat{\theta}_n(X_1,\dots,X_n) \to \theta \quad \text{a} \quad \mathsf{Var}\,\widehat{\theta}_n(X_1,\dots,X_n) \to 0\,,$$

pro $n \to \infty$.

Věta 9.1 *Nechť $X_1,\dots,X_n$ je náhodný výběr z rozdělení s distribuční funkcí $F_{\mu,\sigma^2}(x)$, kde $\mu = \mathsf{E}\,X_1, \sigma^2 = \mathsf{Var}\,X_1$. Potom výběrový průměr $\overline{X}_n$ je nestranný a konzistentní odhad parametru μ a výběrový rozptyl s_n^2 je nestranný a konzistentní odhad parametru σ^2.*

Důkaz. Pro výběrový průměr platí

$$\mathsf{E}\,\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\,X_i = \mu$$

a je tedy nestranným odhadem parametru μ a podle zákona velkých čísel $\overline{X}_n \xrightarrow{P} \mu$ a je tedy i konzistentním odhadem parametru μ.

Zabývejme se nyní nestranností výběrového rozptylu.

$$
\begin{aligned}
\mathsf{E}\, s_n^2 &= \frac{1}{n-1}\,\mathsf{E}\sum_{i=1}^{n}(X_i-\mu-(\overline{X}_n-\mu))^2\\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}\mathsf{E}\,(X_i-\mu)^2-2\,\mathsf{E}\sum_{i=1}^{n}(X_i-\mu)(\overline{X}_n-\mu)+n\,\mathsf{E}\,(\overline{X}_n-\mu)^2\right)\\
&= \frac{1}{n-1}\left(n\,\mathsf{Var}\,X_1-2n\,\mathsf{E}\,(\overline{X}_n-\mu)^2+n\,\mathsf{E}\,(\overline{X}_n-\mu)^2\right)\\
&= \frac{1}{n-1}\left(n\sigma^2-n\,\mathsf{E}\,(\overline{X}_n-\mu)^2\right)=\frac{1}{n-1}\left(n\sigma^2-n\frac{\sigma^2}{n}\right)=\sigma^2\,,
\end{aligned}
$$

kde jsme využili toho, že $\mathsf{Var}\,\overline{X}_n=\sigma^2/n$. □

Poznámka 9.7 Situace se může změnit, pokud budeme znát hodnotu jednoho z parametrů. Např. pokud známe hodnotu $\mu=c$, potom výběrový rozptyl s_n^2 nebude nestranným odhadem parametru σ^2. Nestranným odhadem bude statistika

$$\sigma_n^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-c)^2\,.$$

Poznámka 9.8 Často existuje více nestranných odhadů. Například pro náhodný výběr z normálního rozdělení $N(\mu,\sigma^2)$ jsou výběrový průměr a výběrový medián nestrannými odhady parametru μ. V takovém případě se snažíme najít takový nestranný odhad, který má nejmenší rozptyl.

Definice 9.6 *Odhad $\widehat{\theta}_n(X_1,\dots,X_n)$ se nazývá nejlepší nestranný odhad parametru θ, pokud je nestranný a pro každý jiný nestranný odhad $\widehat{\theta}_n^0(X_1,\dots,X_n)$ platí*

$$\mathsf{Var}\,(\widehat{\theta}_n^0)\geq\mathsf{Var}\,(\widehat{\theta}_n)\quad\textit{pro všechna}\quad\theta\in\Theta.$$

Poznámka 9.9 Existuje dolní mez, pod kterou nemůže rozptyl žádného nestranného odhadu klesnout (Rao-Cramerova mez). Pokud najdeme odhad s touto mezní hodnotou rozptylu, je to nejlepší nestranný odhad.

Platí tvrzení: je-li $X_1,\dots,X_n$ náhodný výběr z binomického, Poissonova, exponenciálního nebo normálního rozdělení je výběrový průměr $\overline{X}_n$ nejlepší nestranný odhad $\mathsf{E}\,X$. Jde-li o výběr z normálního rozdělení $N(\mu,\sigma^2)$, je výběrový rozptyl s_n^2 nejlepší nestranný odhad parametru σ^2.

9.1.2 Metoda momentů

Tato metoda je výpočetně jednoduchá a rychlá, její odhady však někdy nemívají příliš dobré vlastnosti. Často se používá pro získání prvních odhadů, které se následně použijí ve složitějších metodách (MLE, iterační řešení rovnic).

Myšlenka metody je následující. Nechť $X_1,\dots,X_n$ je náhodný výběr z rozdělení s distribuční funkcí $F(x,\theta_1,\dots,\theta_k)$ a nechť $X\sim F(x,\theta_1,\dots,\theta_k)$. Definovali jsme r-tý obecný moment jako

$$\mu_r'=\mathsf{E}\,X^r\,.$$

Tyto momenty umíme spočítat, jsou funkcemi $\theta_1, \ldots, \theta_k$. Dále r-tý výběrový moment je definován předpisem

$$m_r = \frac{1}{n} \sum_{i=1}^{n} X_i^r \,.$$

Věty o zákonech velkých čísel implikují, že $m_r \xrightarrow{P} \mu'_r$, takže metoda momentů spočívá v porovnání prvních k momentů

$$\mu'_j = m_j \quad \text{pro} \quad j \in \widehat{k} \,.$$

Máme tedy k rovnic pro k neznámých.

Postup v praxi:

1. Napočítáme teoretické momenty
$$\mu_j = \mu_j(\theta_1, \ldots, \theta_k) = \mathsf{E}\, X^j, \quad j \in \widehat{k} \,.$$
2. Spočítáme výběrové momenty $m_j, j \in \widehat{k}$.
3. Ze soustavy rovnic
$$\mu_j(\theta_1, \ldots, \theta_k) = m_j, \quad j \in \widehat{k} \,,$$
vyjádříme odhady
$$\widehat{\theta}_j = \widehat{\theta}_j(m_1, \ldots, m_k) \,.$$
Jsou to vlastně funkce náhodného výběru $X_1, \ldots, X_n$.

Příklad 9.1 Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozdělení jako náhodná veličina X, kde

$$f_X(x, \theta_1, \theta_2) = \begin{cases} \frac{1}{\theta_2} & \text{pro } \theta_1 \leq x \leq \theta_1 + \theta_2, \\ 0 & \text{jinak.} \end{cases}$$

Na základě náhodného výběru odhadněte parametry θ_1, θ_2.

Vyjádřeme první dva obecné momenty náhodné veličiny X,

$$\mathsf{E}\, X = \int_{\theta_1}^{\theta_1 + \theta_2} \frac{x}{\theta_2} \, dx = \theta_1 + \frac{\theta_2}{2}$$

a

$$\mathsf{E}\, X^2 = \int_{\theta_1}^{\theta_1 + \theta_2} \frac{x^2}{\theta_2} \, dx = \theta_1^2 + \theta_1\theta_2 + \frac{\theta_2^2}{3} \,.$$

Soustava rovnic má tedy tvar

$$\begin{aligned} \theta_1 + \frac{\theta_2}{2} &= m_1 \\ \theta_1^2 + \theta_1\theta_2 + \frac{\theta_2^2}{3} &= m_2 \end{aligned}$$

a tedy

$$\theta_1 = m_1 - \frac{\theta_2}{2}$$

a dosazením do druhé rovnice dostaneme

$$\theta_2 = \sqrt{12(m_2 - m_1^2)}\,.$$

Získali jsme tudíž odhady

$$\widehat{\theta}_2 = \sqrt{12}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}_n^2\right)^{\frac{1}{2}} \quad \text{a} \quad \widehat{\theta}_1 = \overline{X}_n - \frac{\widehat{\theta}_2}{2}\,.$$

Příklad 9.2 Najděte odhad parametrů binomického rozdělení $Bi(n,p)$ na základě náhodného výběru $X_1,\ldots,X_n$.

Víme, že pro binomickou náhodnou veličinu platí $\mathsf{E}\,X = np$, $\mathsf{Var}\,X = np(1-p)$ a tedy $\mathsf{E}\,X^2 = \mathsf{Var}\,X + (\mathsf{E}\,X)^2 = np - np^2 + n^2p^2$. Soustava rovnic tedy bude mít tvar

$$\begin{aligned} np &= m_1 \\ np - np^2 + n^2p^2 &= m_2\,. \end{aligned}$$

Z první rovnice dostaneme $n = m_1/p$ a dosazením do druhé vyjádříme p ve tvaru $p = 1 + m_1 - m_2/m_1$. Celkově mají tedy odhady metodou momentů tvar

$$\widehat{p} = 1 + m_1 - \frac{m_2}{m_1} \quad \text{a} \quad \widehat{n} = \frac{m_1^2}{m_1 + m_1^2 - m_2}\,.$$

Cvičení 9.1 Nechť $X_1,\ldots,X_n$ je náhodný výběr (tzn. jsou to nezávislé stejně rozdělené náhodné veličiny) z rozdělení s hustotou pravděpodobnosti

$$f_{X_i}(x_i) = \begin{cases} \frac{1}{\theta} e^{-\frac{x_i-\mu}{\theta}} & \text{pro } x_i > \mu, \\ 0 & \text{jinak,} \end{cases}$$

kde $\theta > 0, \mu \in I\!R$. Metodou momentů nalezněte odhady parametrů θ, μ.

9.1.3 Maximálně věrohodné odhady

Začněme příkladem.

Příklad 9.3 Nechť X_1, X_2, X_3, X_4 je náhodný výběr z Bernoulliova (alternativního) rozdělení s parametrem p (neznámým). Máme za úkol odhadnout hodnotu parametru p, jestliže jsme realizací výběru získali data $0, 0, 1, 0$.

Spočtěme sdruženou pravděpodobnost tohoto výsledku:

$$P(p) = \mathsf{P}[X_1 = X_2 = X_4 = 0, X_3 = 1] = p(1-p)^3\,,$$

Princip metody spočívá v tom, že za odhad parametru p vezmeme hodnotu, pro kterou je získaná realizace nejpravděpodobnější. Najděme maximum funkce $P(p)$,

$$\frac{dP}{dp}(p) = (1-p)^3 - 3p(1-p)^2 = (1-p)^2(1-4p) = 0\,.$$

Stacionární body tedy jsou $p_{1,2} = 1, p_3 = 1/4$, přičemž maximum je nabýváno v bodě $p_3 = 1/4$. Dostáváme tedy odhad $\widehat{p} = 1/4$, což je hodnota, která se dala předvídat přímo ze zadání.

Ukázaný postup nyní zobecníme.

Definice 9.7 *Sdruženou hustotu pravděpodobnosti náhodného výběru* $X_1, \dots, X_n$ *uvažovanou jako funkci parametru* θ *při dané realizaci* $(X_1, \dots, X_n) = (x_1, \dots, x_n)$ *nazýváme věrohodnostní funkcí a značíme*

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta)\,,$$

pro absolutně spojité rozdělení, případně

$$L(\theta) = \prod_{i=1}^{n} \mathsf{P}_\theta[X_i = x_i]$$

pro diskrétní rozdělení.

Definice 9.8 *Nechť* $X_1, \dots, X_n$ *je náhodný výběr z rozdělení s hustotou* $f(x, \theta), \theta \in \Theta$. *Odhad* $\widehat{\theta}(X_1, \dots, X_n)$ *nazýváme maximálně věrohodný odhad (MLE) parametru* θ*, jestliže*

$$L(\widehat{\theta}(x_1, \dots, x_n)) \geq L(\theta) \quad \textit{pro všechna} \quad \theta \in \Theta\,.$$

Postup pro získání odhadu:
Máme realizaci náhodného výběru $x_1, \dots, x_n$, předpokládáme, že z rozdělení s hustotou $f_X(x, \theta_1, \dots, \theta_k)$.

1. Sestavíme věrohodnostní funkci (sdruženou hustotu) $L(\theta_1, \dots, \theta_k)$.

2. Maximalizujeme $L(\theta_1, \dots, \theta_k)$ přes všechny možné hodnoty parametrů $\theta_1, \dots, \theta_k$. Často je výhodné maximalizovat funkci $\ln L(\theta_1, \dots, \theta_k)$. MLE tedy získáme jako řešení soustavy k-rovnic

$$\frac{\partial \ln L(\theta_1, \dots, \theta_k)}{\partial \theta_j} = 0, \qquad j = 1, \dots, k\,.$$

   Uvedené rovnice se nazývají věrohodnostní rovnice.

Poznámka 9.10 Vlastnosti MLE:

- často jsou výpočetně náročné, soustava se řeší numericky
- nemusí být nestranné, ale jsou konzistentní
- funkce $g(\widehat{\theta})$, kde $\widehat{\theta}$ je maximálně věrohodný odhad, je odhadem funkce $g(\theta)$ (např. odhad směrodatné odchylky $\widehat{\sigma}_n = \sqrt{\widehat{\sigma}_n^2}$)
- za jistých podmínek známe asymptotické rozdělení MLE odhadů.

Příklad 9.4 Běžný typ biologického experimentu je počítání mikroorganismů v políčcích čtvercové sítě, kterou je rozděleno zorné pole mikroskopu. V prvních třech řádcích následující tabulky jsou uvedeny k - počty mikroorganismů v políčku, n_k - počty políček, kde bylo pozorováno k mikroorganismů a n_k/n - relativní četnosti, kde $n = 118$ je celkový počet políček čtvercové sítě.

k	0	1	2	3	4	5	6	7
n_k	5	19	26	26	21	13	5	3
n_k/n	0.042	0.161	0.220	0.220	0.178	0.110	0.042	0.025
$\widehat{p}_k$	0.052	0.153	0.227	0.224	0.166	0.980	0.049	0.020

Naším úkolem je odhadnout hodnotu parametru λ Poissonova rozdělení, kterým by bylo nejlepší situaci modelovat.

Začněme nejdříve obecně a uvažujme realizaci náhodného výběru $x_1, \ldots, x_n$ z Poissonova rozdělení s parametrem λ. Věrohodnostní funkce má tvar

$$L(\lambda) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$$

a její logaritmus je

$$\ln L(\lambda) = \sum_{i=1}^{n} \ln\left(e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}\right) = \sum_{i=1}^{n} (-\lambda + x_i \ln\lambda - \ln(x_i!)) = -n\lambda + \ln\lambda \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \ln(x_i!)\,.$$

Následnou derivací získáme rovnici

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = -n + \frac{1}{\lambda} \sum_{i=1}^{n} x_i = 0$$

a jako její řešení maximálně věrohodný odhad

$$\widehat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}_n\,.$$

V našem konkrétním případě bylo $n = 118$ pozorování a

$$\sum_{i=1}^{118} x_i = 19 + 2 \cdot 26 + 3 \cdot 26 + 4 \cdot 21 + 5 \cdot 13 + 6 \cdot 5 + 7 \cdot 3 = 349$$

a dostáváme tudíž odhad parametru λ,

$$\widehat{\lambda} = \frac{349}{118} = 2.96\,.$$

Teoretické pravděpodobnosti spočtené s odhadnutou hodnotou parametru jsou pro srovnání s relativními četnostmi uvedeny v posledním řádku předchozí tabulky.

Příklad 9.5 Nechť $x_1, \ldots, x_n$ je realizace náhodného výběru z rozdělení $N(\mu, \sigma^2)$. Najděte maximálně věrohodné odhady parametrů μ, σ^2.

Sestavme věrohodnostní funkci

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

a její logaritmus

$$\ln L(\mu, \sigma^2) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln\sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\,.$$

Derivacemi podle parametrů μ, σ^2 získáme soustavu věrohodnostních rovnic

$$\begin{aligned}
\frac{\partial \ln L(\mu,\sigma^2)}{\partial \mu} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i - n\mu\right) = 0 \\
\frac{\partial \ln L(\mu,\sigma^2)}{\partial \sigma^2} &= -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0\,.
\end{aligned}$$

Protože $\sigma^2 > 0$ dostaneme z první rovnice

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}_n\,.$$

Dosazením do druhé rovnice soustavy pak získáme

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x}_n)^2\,.$$

Pro normální rozdělení jsme tedy získali maximálně věrohodné odhady

$$\widehat{\mu} = \overline{x}_n \quad \text{a} \quad \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x}_n)^2\,.$$

Cvičení 9.2 Nechť $X_1,\dots,X_n$ je náhodný výběr (tzn. jsou to nezávislé stejně rozdělené náhodné veličiny) z rozdělení s hustotou pravděpodobnosti

$$f_{X_i}(x_i) = (\theta+1)x^{\theta}, \quad \text{pro} \quad 0 \le x \le 1\,.$$

Najděte maximálně věrohodný odhad parametru θ.

9.1.4 Intervalové odhady

Místo konkrétního bodového odhadu parametru θ nás někdy zajímá interval, ve kterém bude hodnota parametru ležet s určitou velkou pravděpodobností $1-\alpha$ ($\alpha > 0$ malé).

Hledáme tedy na základě náhodného výběru dvě statistiky, zn. $\underline{\theta}(X_1,\dots,X_n)$ a $\overline{\theta}(X_1,\dots,X_n)$ tak, aby platilo

$$\mathsf{P}[\theta \le \underline{\theta}(X_1,\dots,X_n)] \le \frac{\alpha}{2} \quad \text{a} \quad \mathsf{P}[\theta \ge \overline{\theta}(X_1,\dots,X_n)] \le \frac{\alpha}{2}$$

a tedy

$$\mathsf{P}[\underline{\theta}(X_1,\dots,X_n) \le \theta \le \overline{\theta}(X_1,\dots,X_n)] \ge 1-\alpha\,. \tag{9.1}$$

Definice 9.9 *Dvojice statistik* $(\underline{\theta},\overline{\theta})$ *splňující (9.1) se nazývá* $100\cdot(1-\alpha)\%$ *interval spolehlivosti. Statistika* $\underline{\theta}(X_1,\dots,X_n)$ *se nazývá dolní mez, statistika* $\overline{\theta}(X_1,\dots,X_n)$ *horní mez intervalu spolehlivosti. Číslo* $1-\alpha$ *se nazývá koeficient spolehlivosti.*

Poznámka 9.11 Nejčastěji se používají hodnoty $\alpha = 0.05, \alpha = 0.01$, kdy mluvíme o 95% intervalu spolehlivosti respektive o 99% intervalu spolehlivosti.

Interpretace 99% intervalu spolehlivosti: skutečná hodnota parametru θ bude s 99% pravděpodobností ležet uvnitř intervalu $(\underline{\theta},\overline{\theta})$. V praxi ale máme konkrétní realizaci $(x_1,\dots,x_n)$ náhodného výběru, v tomto případě si význam intervalu spolehlivosti můžeme představit takto: při 100 opakováních pokusu, bude v 99 případech hodnota parametru θ ležet uvnitř spočteného intervalu spolehlivosti.

Poznámka 9.12 Někdy nás zajímá jen horní nebo dolní mez, potom konstruujeme statistiky tak, aby

$$\mathsf{P}[\underline{\theta}(X_1,\dots,X_n) \geq \theta] \leq \alpha \quad \text{nebo} \quad \mathsf{P}[\theta \geq \overline{\theta}(X_1,\dots,X_n)] \leq \alpha$$

neboli

$$\mathsf{P}[\theta \geq \underline{\theta}(X_1,\dots,X_n)] \geq 1-\alpha \quad \text{a} \quad \mathsf{P}[\theta \leq \overline{\theta}(X_1,\dots,X_n)] \geq 1-\alpha$$

a hovoříme o dolním resp. horním intervalovém odhadu nebo obecně o jednostranných intervalech spolehlivosti.

Obecně je konstrukce intervalových odhadů poměrně složitá, omezíme se na normální rozdělení.

Věta 9.2 *Nechť $X_1,\dots,X_n$ je náhodný výběr z rozdělení $N(\mu,\sigma^2)$ a známe hodnotu σ^2. Potom platí*

$$\mathsf{P}\left[\overline{X}_n - u_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}} < \mu < \overline{X}_n + u_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha\,,$$

kde u_β je β-kvantil standardního normálního rozdělení $N(0,1)$.

Poznámka 9.13 To tedy znamená, že

$$\left(\overline{X}_n - u_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}, \overline{X}_n + u_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}\right)$$

je $100\cdot(1-\alpha)\%$ intervalem spolehlivosti pro parametr μ.

Důkaz. Pro výběrový průměr $\overline{X}_n = 1/n\sum_{i=1}^n X_i$ platí $\mathsf{E}\,\overline{X}_n = \mu$ a $\mathsf{Var}\,\overline{X}_n = \sigma^2/n$, pomocí momentové funkce se dá ukázat, že $\overline{X}_n \sim N(\mu,\sigma^2/n)$. To znamená, že

$$Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

a platí tedy

$$\mathsf{P}[u_{\alpha/2} < Z < u_{1-\alpha/2}] = \mathsf{P}[Z < u_{1-\alpha/2}] - \mathsf{P}[Z < u_{\alpha/2}] = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1-\alpha\,.$$

Využijeme-li dále toho, že ze symetrie rozdělení $N(0,1)$ plyne $u_{\alpha/2} = -u_{1-\alpha/2}$, dostáváme konečně

$$\mathsf{P}\left[-u_{1-\alpha/2} < \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} < u_{1-\alpha/2}\right] = \mathsf{P}\left[-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \overline{X}_n - \mu < u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$
$$\mathsf{P}\left[\overline{X}_n - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X}_n + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha\,.$$

□

Další intervaly spolehlivosti pro normální rozdělení již nebudeme odvozovat, ale uvedeme si je v souhrnné tabulce.

Tabulka

Příklad 9.6 Při kontrolních zkouškách 16 žárovek byl stanoven odhad střední hodnoty $\overline{x}_{16} = 3000$ hodin a směrodatné odchylky $\sigma_{16} = 20$ hodin jejich životnosti. Za předpokladu, že životnost žárovky je náhodná veličina s normálním rozdělením $N(\mu, \sigma^2)$ určete 90% interval spolehlivosti pro parametry μ a σ.

Pro dosazení do vzorců pro intervaly spolehlivosti budeme potřebovat výběrovou směrodatnou odchylku s_{16}. Protože obecně platí $s_n^2 = n/(n-1)\sigma_n^2$, dostaneme $s_{16} = 80/\sqrt{15}$.

V tabulce najdeme interval spolehlivosti pro parametr μ a tento případ ve tvaru

$$\left(\overline{X}_n - t_{1-\alpha/2}(n-1)\,\frac{s_n}{\sqrt{n}}, \overline{X}_n + t_{1-\alpha/2}(n-1)\,\frac{s_n}{\sqrt{n}}\right).$$

Po dosazení za $\alpha = 0.1, n = 16$ a $t_{0.95}(15) = 1.753$

$$\left(3000 - 1.753\frac{20}{\sqrt{15}}, 3000 + 1.753\frac{20}{\sqrt{15}}\right) = (2990.9, 3009.1)$$

a parametr μ tedy s pravděpodobností 90% leží v intervalu $(2990.9, 3009.1)$. Pro dolní interval spolehlivosti (střední životnost větší než?) dostaneme

$$\overline{x}_{16} - t_{0.9}(15)\frac{s_{16}}{\sqrt{16}} = 3000 - 1.3406\frac{20}{\sqrt{15}} = 2993.3$$

a s pravděpodobností 90% je $\mu > 2993.3$.

Pro parametr σ^2 má interval spolehlivosti tvar

$$\left(\frac{(n-1)s_n^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)s_n^2}{\chi^2_{\alpha/2}(n-1)}\right)$$

a dosazením hodnot $\chi_{0.95}(15) = 24.996, \chi_{0.05}(15) = 4.6009$ dostaneme interval $(256, 1391)$. Hodnota parametru σ tedy leží s pravděpodobností 90% v intervalu $(16, 37.3)$. Podobně pro horní interval spolehlivosti (rozptyl menší než?)

$$\frac{15 s_{16}^2}{\chi^2_{0.1}(15)} = \frac{16 \cdot 400}{5.2293} = 1223.8\,,$$

z čehož vyplývá, že s 90% pravděpodobností je $\sigma < 34.9$.

Příklad 9.7 Hloubka moře se měří přístrojem, jehož systematická chyba je rovna nule a náhodné chyby mají normální rozdělení se směrodatnou odchylkou $\sigma = 20\ m$. Kolik nezávislých měření je třeba provést, aby se s pravděpodobností 95% stanovila hloubka s chybou menší než $10\ m$?

Označme výsledky jednotlivých měření $Y_j, j = 1, \dots, n$. Ze zadání víme, že $Y_j = h + X_j$. kde h je skutečná hloubka moře a $X_j \sim N(0, \sigma^2)$ je náhodná chyba měření. Zprůměrováním měření dostaneme

$$\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_j = \frac{1}{n}\sum_{i=1}^{n}(h + X_j) = h + \frac{1}{n}\sum_{i=1}^{n} X_j\,,$$

kde $1/n \sum_{i=1}^{n} X_j$ vyjadřuje chybu, jaké se dopustíme po zprůměrování měření.

Pro interval spolehlivosti pro parametr μ platí

$$\mathsf{P}\left[\overline{X}_n - u_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}} < \mu < \overline{X}_n + u_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha\,,$$

což můžeme přepsat do tvaru ($\mu = 0$)

$$\mathsf{P}\left[|\overline{X}_n - \mu| < u_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}\right] = \mathsf{P}\left[|\overline{X}_n| < u_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha\,.$$

My bychom chtěli určit počet měření n tak, aby chyba, které se dopustíme, byla s pravděpodobností 95% menší než 10 m, tzn.

$$\mathsf{P}[|\overline{X}_n| < 10] = 0.95\,.$$

Z předchozího vztahu vyplývá, že pro takové n musí platit

$$10 = u_{0.975}\frac{20}{\sqrt{n}}$$

a tedy

$$\sqrt{n} = 2\cdot u_{0.975} = 2\cdot 1.96 \doteq 4\,.$$

Závěr zní, že abychom s pravděpodobností 95% určili hloubku s přesností na 10 m, musíme provést minimálně 16 měření.

9.2 Testování hypotéz

Na základě náhodného výběru z populace budeme konstruovat testy, pomocí kterých zamítneme nebo nezamítneme nějakou hypotézu o populaci (populace je popsána nějakým rozdělením pravděpodobnosti s parametrem θ, otázka může být $\theta = 3$?). Například:

- Příjem dodávky zboží. Dodávku přijmeme, pokud procento zmetků p v ní bude menší než nějaká hranice (např. 0.05). Nemůžeme ale prověřit všechny výrobky. Provedeme náhodný výběr a otestujeme. Na základě výsledků se máme rozhodnout, zda dodávku přijmout či ne. Testujeme vlastně nulovou hypotézu $H_0 : p \leq 0.05$ proti alternativě $H_1 : p \geq 0.05$.

- Zlepšení technologického nebo výrobního procesu. Představme si, že byl navržen nový způsob povrchové úpravy materiálu proti korozi. Než se zavede do výroby je třeba rozhodnout, zda je skutečně lepší než ten současný. Naměří se vzorky n_1 a n_2 pozorování pro starou a novou metodu. Zajímá nás například procento povrchu zasaženého korozí a ptáme se, jestli střední hodnota povrchu zasaženého korozí je při nové metodě menší nebo ne. Testujeme tedy nulovou hypotézu $H_0 : \mu_2 < \mu_1$ proti alternativě $H_1 : \mu_2 \geq \mu_1$.

9.2.1 Matematický popis

Předpokládáme, že máme náhodný výběr $X_1, \ldots, X_n$ z rozdělení, které je prvkem rodiny distribucí $\{F_\theta \,|\, \theta \in \Theta\}$, kde $\Theta \subset I\!R^d$ je množina všech možných hodnot parametru θ. Dále předpokládáme, že o parametru θ existují dvě konkurující si hypotézy.

$$\text{Nulová hypotéza} \quad H_0 : \theta \in \Theta_0$$

reprezentovaná množinou parametrů $\Theta_0 \subset \Theta$ a

$$\text{alternativní hypotéza} \quad H_1 : \theta \in \Theta_1$$

kde $\Theta_1 = \Theta \setminus \Theta_0$.

Testem nulové hypotézy H_0 proti alternativní hypotéze H_1 rozumíme rozhodovací proces založený na náhodném výběru $X_1, \ldots, X_n$, na jehož základě *zamítneme* nebo *nezamítneme* platnost hypotézy H_0. Naše rozhodnutí ovšem nemusí být správné. Celkem mohou nastat následující čtyři případy

a) platí H_0 a naše rozhodnutí je nezamítnout H_0,

b) platí H_0 a naše rozhodnutí je zamítnout H_0,

c) platí H_1 a naše rozhodnutí je nezamítnout H_0,

d) platí H_1 a naše rozhodnutí je zamítnout H_0.

V případech a) a d) je naše rozhodnutí správné. V případě b) říkáme, že se dopouštíme *chyby prvního druhu*, v případě c) říkáme, že se dopouštíme *chyby druhého druhu.*

Postavení hypotéz tedy není symetrické. Za nulovou hypotézu volíme tu, jejíž neoprávněné zamítnutí, tj. chyba prvního druhu, je závažnější. Můžeme totiž kontrolovat velikost pouze jedné z uvedených chyb., volba padla na chybu prvního druhu. Velikost chyby prvního druhu tedy stanovíme na rozumné úrovni (obvykle $\alpha = 0.05, \alpha = 0.01$), chyba druhého druhu potom může být vysoká.

Při této volbě uděláme chybu s pravděpodobností nejvýše α, pokud hypotézu H_0 zamítneme. Buďto totiž platí H_1 a udělali jsme dobře nebo platí H_0 a my jsme udělali chybu prvního druhu ale maximálně s pravděpodobností α. Pokud bychom ale hypotézu H_0 přijali, může být pravděpodobnost chyby rozhodnutí velká.

Volíme tedy mezi dvěma výsledky testu:

a) na základě testu zamítáme hypotézu H_0 (s možnou malou chybou α),

b) na základě testu nezamítáme H_0 (ale ani nepřijímáme, přijetí H_0, by mohlo být chybné).

Parametrický prostor jsme rozdělili na dvě části Θ_0, Θ_1, příslušné hypotézám H_0, H_1. Stejně se rozpadne i výběrový prostor $\mathcal{X}^n$ $((X_1, \ldots, X_n) \in \mathcal{X}^n)$ na body, kdy zamítáme H_0 a nezamítáme H_0.

Množinu $W \subset \mathcal{X}^n$ bodů z $\mathcal{X}^n$, kterým je přiřazeno rozhodnutí zamítnout H_0, nazýváme *kritickým oborem* pro testování H_0. Je-li tedy realizace náhodného výběru $(x_1, \ldots, x_n) \in$

W, zamítáme H_0. Volba množiny W se řídí především požadavkem, aby pravděpodobnost chyby prvního druhu nepřesáhla α, tj.

$$\mathsf{P}_\theta\left[(X_1, \ldots, X_n) \in W\right] \leq \alpha \quad \text{pro všechna} \quad \theta \in \Theta_0 \,.$$

Číslo α se nazývá *hladina významnosti testu*. Celkově tedy budeme říkat, že testujeme nulovou hypotézu H_0 oproti alternativě H_1 na hladině významnosti α.

9.2.2 Testy hypotéz $\times$ intervalové odhady

Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozdělení s distribuční funkcí $F(x, \theta)$ a mějme testovat hypotézu

$$H_0 : \theta = \theta_0 \quad \text{proti} \quad H_1 : \theta \neq \theta_0 \,.$$

Nechť dále $(\underline{\theta}, \overline{\theta})$ je $100(1-\alpha)\%$ interval spolehlivosti pro parametr θ na základě náhodného výběru $X_1, \ldots, X_n$ (tj. $\mathsf{P}[\theta \in (\underline{\theta}, \overline{\theta})] \geq 1-\alpha$). Potom

$$W = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \,|\, \theta_0 \notin (\underline{\theta}(x_1, \ldots, x_n), \overline{\theta}(x_1, \ldots, x_n))\}$$

je kritický obor testu H_0 proti H_1 na hladině významnosti α, neboť za platnosti nulové hypotézy, tedy $\theta = \theta_0$, platí

$$\mathsf{P}[(X_1, \ldots, X_n) \in W] = \mathsf{P}[\theta_0 \notin (\underline{\theta}, \overline{\theta})] \leq \alpha \,.$$

Jinými slovy hypotézu H_0 zamítáme, pokud $\theta_0 \notin (\underline{\theta}, \overline{\theta})$.

Podobně máme-li testovat $H_0 : \theta = \theta_0$ proti (jednostranné alternativě) $H_1 : \theta > \theta_0$ a máme dolní intervalový odhad $\underline{\theta}$ ($\mathsf{P}[\theta \geq \underline{\theta}] \geq 1-\alpha$), je kritickým oborem na hladině α množina

$$W = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \,|\, \theta_0 < \underline{\theta}(x_1, \ldots, x_n)\} \,.$$

Hypotézu H_0 proti jednostranné alternativě H_1 tedy na hladině významnosti α zamítáme, pokud $\theta_0 < \underline{\theta}(x_1, \ldots, x_n)$.

Příklad 9.8 Podle jisté teorie je podíl železa v určité chemické sloučenině 12.1%. Byl proveden rozbor 9 vzorků s následujícími výsledky

$$11.7,\ 12.2,\ 10.9,\ 11.4.\ 11.3,\ 12.0,\ 11.1,\ 10.7,\ 11.6 \,.$$

Testujte platnost teoretické hypotézy na hladině významnosti 5% (předpokládáme normální rozdělení dat s parametry μ, σ^2).

Máme tedy testovat

$$H_0 : mu = 12.1 \quad \text{proti} \quad H_1 : \mu \neq 12.1 \,.$$

Najděme příslušný oboustranný interval spolehlivosti pro parametr μ, který má tvar

$$\left(\overline{X}_n - t_{1-\alpha/2}(n-1)\,\frac{s_n}{\sqrt{n}}, \overline{X}_n + t_{1-\alpha/2}(n-1)\,\frac{s_n}{\sqrt{n}}\right) \,.$$

Podle zadání je

$$\overline{X}_n = \frac{1}{9}\sum_{i=1}^{9} X_i = 11.43, \quad s_n^2 = \frac{1}{8}\sum_{i=1}^{9}(X_i - 11.43)^2 = 0.24,$$

$$s_n = 0.49 \quad \text{a} \quad t_{0.975}(8) = 2.306 \,.$$

Po dosazení dostaneme interval spolehlivosti $(11.05, 11.81)$. Protože $12.1 \notin (11.05, 11.81)$ na hladině významnosti 5% zamítáme hypotézu H_0, že obsah železa je 12.1%.

9.2.3 Testy o parametrech normálního rozdělení

1. Testy o střední hodnotě

Uvažujeme náhodný výběr $X_1, \dots, X_n$ z rozdělení $N(\mu, \sigma^2)$ a úkolem je otestovat hypotézu

$$H_0 : \mu = \mu_0 \quad \text{proti} \quad H_1 : \mu \neq \mu_0,\ H_1' : \mu > \mu_0,\ H_1'' : \mu < \mu_0 \,.$$

Odvození jednoho jednoduchého testu si ukážeme na příkladu, další testy potom uvedeme v tabulce.

Příklad 9.9 Pevnost vyráběného ocelového drátu je náhodná veličina s rozdělením $N(\mu_0, \sigma^2)$, kde $\mu_0 = 1250, \sigma = 150$. Novým postupem bylo vyrobeno 25 vzorků a stanoven průměr $\overline{X}_{25} = 1312$, předpokládáme stejné σ. Na hladině významnosti 0.05 ověřte, zda je nový postup lepší.

Máme tedy otestovat hypotézu

$$H_0 : \mu = \mu_0 \quad \text{proti} \quad H_1 : \mu > \mu_0$$

(při novém postupu je drát pevnější). Víme, že $\overline{X}_n$ má má rozdělení $N(\mu, \sigma^2/n)$. Čím větší bude rozdíl $\overline{X}_n - \mu_0$, tím spíš bude nový postup lepší. Pokud rozdíl normujeme za předpokladu platnosti H_0, tj. $\mu = \mu_0$, dostaneme testovací statistiku

$$Z_n = \frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \,,$$

která má za předpokladu H_0 standardní normální rozdělení. Za zmíněného předpokladu tedy platí

$$\mathsf{P}[Z_n \geq u_{0.95}] = 0.05 \,.$$

Pokud hodnota Z_n překročí $u_{0.95}$ (jev, který měl za předpokladu H_0 nastat s malou pravděpodobností), bude to říkat, že Z_n asi nemá rozdělení $N(0,1)$, což znamená, že neplatí H_0. Hypotézu H_0 tedy zamítneme, pokud

$$Z_n \geq u_{0.95} \,.$$

Podle zadání je

$$Z_n = \frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}} = \frac{1312 - 1250}{30} = 2.067$$

a příslušný kvantil má hodnotu $u_{0.95} = 1.64$. Protože $Z_n = 2.067 > u_{0.95} = 1.64$ nulovou hypotézu na hladině významnosti 0.05 zamítáme a statisticky tedy bylo prokázáno, že nový výrobní postup dává lepší výsledky.

Další testy o parametrech normálního rozdělení jsou uvedeny v tabulce.

tabulka

Příklad 9.10 Pro kontrolu nastavení měřícího přístroje bylo provedeno 10 měření zkušebního vzorku se správnou hodnotou $\mu_0 = 15.2$. Výsledky měření jsou

$$15.21,\ 15.19,\ 15.16.\ 15.26,\ 15.22,\ 15.23,\ 15.26,\ 15.23,\ 15.29, 15.23\,.$$

Lze považovat na hladině významnosti 0.05 naměřené odchylky za náhodné chyby nebo je podezření na systematickou chybu? (předpokládáme, že chyby mají normální rozdělení)

Máme tedy otestovat platnost hypotézy

$$H_0 : \mu = \mu_0 \quad \text{proti} \quad H_1 : \mu \neq \mu_0,$$

pro kterou najdeme v tabulce test ve tvaru

$$\frac{\sqrt{n}}{s_n}\,|\overline{X}_n - \mu_0| \geq t_{1-\alpha/2}(n-1).$$

Ze zadaných dat dostaneme $\overline{X}_n = 15.228, s_n = 0.03706$ a po dosazení do testu

$$\frac{\sqrt{10}}{0.03706}\,|15.228 - 15.2| = 2.3893 > 2.2622 = t_{0.975}(9)\,.$$

Nulovou hypotézu na hladině 0.05 zamítáme, tzn. usuzujeme na přítomnost systematické chyby.

2. Testy o rozptylu

Příklad 9.11 Při přesném seřízení automatického obráběcího stroje nepřekračuje rozptyl délky součástky 300 μm^2. Bylo změřeno 15 součástek a výběrový rozptyl činil $s_n^2 = 580 \mu m^2$. Je tím na hladině významnosti 0.05 prokázáno, že seřízení stroje není v pořádku? (předpokládáme normální rozdělení délek součástek).

Úkolem je otestovat hypotézu

$$H_0 : \sigma^2 \leq 300 \quad \text{proti} \quad H_1 : \sigma^2 > 300\,.$$

Tvrdíme, že z tabulky lze použít test pro hypotézu $H_0 : \sigma^2 = \sigma_0^2$ proti $H_1 : \sigma^2 > \sigma_0^2$ ve tvaru

$$\frac{(n-1)s_n^2}{\sigma_0^2} \geq \chi^2_{1-\alpha}(n-1)\,. \tag{9.2}$$

Zdůvodnění je následující. Podobně jako v příkladu 9.9 se dá ukázat, že

$$\mathsf{P}\left[\frac{(n-1)s_n^2}{\sigma^2} \geq \chi^2_{1-\alpha}(n-1)\right] = \alpha\,.$$

Pro hypotézu $H_0 : \sigma^2 = \sigma_0^2$ z tohoto vztahu dostáváme test (9.2). Teď máme nulovou hypotézu ve tvaru $H_0 : \sigma \leq \sigma_0$ a pro všechna $\sigma \leq \sigma_0$ platí

$$\left[\frac{(n-1)s_n^2}{\sigma_0^2} \geq \chi^2_{1-\alpha}(n-1)\right] \subset \left[\frac{(n-1)s_n^2}{\sigma^2} \geq \chi^2_{1-\alpha}(n-1)\right],$$

z čehož následně pro všechna $\sigma \leq \sigma_0$ vyplývá

$$\mathsf{P}\left[\frac{(n-1)s_n^2}{\sigma_0^2} \geq \chi^2_{1-\alpha}(n-1)\right] \leq \mathsf{P}\left[\frac{(n-1)s_n^2}{\sigma^2} \geq \chi^2_{1-\alpha}(n-1)\right] = \alpha\,.$$

Jedná se tedy také o test na hladině významnosti α.

Dosazením konkrétních hodnot z tohoto příkladu dostaneme

$$\frac{14 \cdot 580}{300} = 27.07 \geq 23.7 = \chi^2_{0.95}(14)$$

a hypotézu H_0 na hladině významnosti 0.05 zamítáme. Stroj je tedy nutné znovu seřídit.

3. Testy rovnosti středních hodnot dvou normálních rozdělení

Uvažujeme dva nezávislé náhodné výběry $X_1, \dots, X_{n_1}$ a $Y_1, \dots, Y_{n_2}$ z rozdělení $N(\mu_1, \sigma^2)$ a $N(\mu_2, \sigma^2)$. Úkolem je otestovat hypotézu

$$H_0 : \mu_1 = \mu_2 \quad \text{proti} \quad H_1 : \mu_1 \neq \mu_2 \,.$$

Poznámka 9.14 Všimněme si, že pro použití testu uvedeného pro tuto hypotézu v tabulce je důležité splnit předpoklady o rovnosti rozptylů obou normálních rozdělení. Pokud tato informace není explicitně dána, musíme rovnost rozptylů otestovat (viz. následující sekce).

Příklad 9.12 Dvě nádrže obsahují chlorovanou vodu. Bylo odebráno $n_1 = 25$ vzorků z první nádrže a $n_2 = 10$ z druhé s těmito výsledky

$$\overline{X}_{n_1} = 34.48, \quad s^2_{n_1} = 1.7482,$$

$$\overline{Y}_{n_2} = 35.59, \quad s^2_{n_2} = 1.7121.$$

Předpokládáme, že jde o výběry z normálního rozdělení se stejným rozptylem ($N(\mu_1, \sigma^2)$, $N(\mu_2, \sigma^2)$). Lze na hladině 0.05 tvrdit, že obsah chloru je v obou nádržích stejný?

Máme tedy otestovat

$$H_0 : \mu_1 = \mu_2 \quad \text{proti} \quad H_1 : \mu_1 \neq \mu_2 \,.$$

Podle vzorců uvedených v tabulce je

$$(s^*)^2 = \frac{1}{25 + 10 - 2}\left[(25-1)s^2_{n_1} + (10-1)s^2_{n_2}\right] \doteq 1.7384,$$

$s^* = 1.3185$ a hodnota testovací statistiky je

$$\sqrt{\frac{n_1 + n_2}{n_1 + n_2}}\frac{1}{s^*}|\overline{X}_{n_1} - \overline{Y}n_2| \doteq 2.25\,.$$

Protože hodnota příslušného kvantilu je $t_{0.975}(33) = 2.035$, nulovou hypotézu na hladině významnosti 0.05 zamítáme.

4. Testy rovnosti rozptylů dvou normálních rozdělení

Uvažujeme dva nezávislé náhodné výběry $X_1, \dots, X_{n_1}$ a $Y_1, \dots, Y_{n_2}$ z rozdělení $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_2^2)$. Úkolem je otestovat hypotézu

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{proti} \quad H_1 : \sigma_1^2 \neq \sigma_2^2 \,.$$

Příklad 9.13 U 16 mužů a 13 žen byly změřeny hodnoty diastolického tlaku s těmito výsledky

$$\overline{X}_M = 77.375, \quad s_M^2 = 69.7167,$$
$$\overline{Y}_Z = 71.0769, \quad s_Z^2 = 85.0769.$$

Na hladině 0.05 ověřte hypotézu, že muži mají stejný diastolický tlak jako ženy. (předpokládáme normální rozdělení $N(\mu_M, \sigma_M^2)$ a $N(\mu_Z, \sigma_Z^2)$)

Ověřme nejdříve hypotézu o rovnosti rozptylů

$$H_0 : \sigma_M^2 = \sigma_Z^2 \quad \text{proti} \quad H_1 : \sigma_M^2 \neq \sigma_Z^2 \,.$$

Pro testovací statistiku platí

$$\frac{s_M^2}{s_Z^2} = \frac{69.7167}{85.0769} = 0.8195 \in (0.34, 3.17),$$

neboť $F_{0.025}(15, 12) = 1/F_{0.975}(12, 15) = 1/2.97 \doteq 0.34$ a $F_{0.975}(15, 12) = 3.17$. Hypotézu H_0 nezamítáme, rozptyly jednotlivých výběrů se tudíž statisticky neliší a můžeme použít test pro rovnost středních hodnot,

$$H_0 : \mu_M = \mu_Z \quad \text{proti} \quad H_1 : \mu_M \neq \mu_Z \,.$$

V tomto případě je

$$(s^*)^2 = \frac{1}{27}(15 \cdot 69.1767 + 12 \cdot 85.0769) = 76.54,$$

$s^* = 8.75$ a

$$T_n = \sqrt{\frac{16 \cdot 13}{29}} \frac{1}{8.75} |77.375 - 71.0769| = 1.928 < 2.0518 = t_{0.975}(27) \,.$$

Nulovou hypotézu tedy nelze na hladině významnosti 0.05 zamítnout.

Pokud bychom testovali hypotézu

$$H_0 : \mu_M = \mu_Z \quad \text{proti} \quad H_1 : \mu_M > \mu_Z \,,$$

test by byl

$$T_n > t_{1-\alpha}(n_1 + n_2 - 2) \,.$$

Protože $t_{0.975}(27) = 1.703 < 1.928 = T_n$, nulovou hypotézu H_0 oproti alternativě H_1 v tomto případě na hladině 0.05 zamítáme.

Cvičení 9.3 Automatická plnící linka plní lahve tekutým mýdlem. Náhodný výběr 20 lahví dal výběrový rozptyl obsahu $s_n^2 = 0.0153$. Pokud rozptyl obsahu přesáhne hodnotu 0.01, nepřijatelné množství lahví bude přeplněno nebo nedoplněno. Můžeme na hladině 0.05 tvrdit, že výrobce má problém s plněním lahví?

Cvičení 9.4 Výrobce tvrdí, že jeho lék je účinný v 75% případů. Nemocnice zaznamenala úspěšnost léku u 136 ze 200 pacientů, kteří jím byli léčeni. Je mezi tvrzením výrobce a zjištěnou úspěšností léku statisticky významný rozdíl?

9.2.4 χ^2-test dobré shody

Představme si, že máme náhodný výběr $X_1, \ldots, X_n$ z neznámého rozdělení pravděpodobnosti s distribuční funkcí F a zajímá nás, zda tato data pocházejí z určitého konkrétního rozdělení s distribuční funkcí F_0, tzn. chtěli bychom otestovat hypotézu

$$H_0 : F = F_0 \quad \text{proti} \quad H_1 : F \neq F_0 \,.$$

Je třeba najít nějaké míry neshody mezi empirickou distribuční funkcí danou pozorovanými daty a distribuční funkcí F_0. Za tímto účelem rozdělíme pozorovací prostor na l disjunktních tříd (např. u spojitého rozdělení to budou disjunktní intervaly pokrývající celý pozorovací prostor). Označme

n počet pozorování

n_i počet pozorování $X_1, \ldots, X_n$ ležících v i-té třídě

p_i pravděpodobnost, že náhodná veličina X padne do i-té třídy vypočtenou podle teoretického rozdělení F_0 .

Pearsonova χ^2-statistika, vyjadřující zmíněnou míru neshody, je potom definována předpisem

$$S_n = \sum_{i=1}^{l} \frac{(n_i - np_i)^2}{np_i} \,.$$

Dá se ukázat, že pro $n \to \infty$ platí $S_n \xrightarrow{L} \chi^2(l - r - 1)$, kde l je počet tříd na které je rozdělen pozorovací prostor a r je počet parametrů, které bylo nutné odhadnout z daného výběru.

Poznámka 9.15 Poznamenejme jenom, že definovaná statistika se dá interpretovat následujícím způsobem

$$S_n = \sum_{i=1}^{l} \frac{(\text{pozorováno v } i\text{-té třídě} - \text{teoreticky v } i\text{-té třídě})^2}{\text{teoreticky v } i\text{-té třídě}} \,.$$

Test hypotézy H_0 tedy bude: zamítáme hypotézu H_0 na hladině významnosti α pokud

$$S_n = \sum_{i=1}^{l} \frac{(n_i - np_i)^2}{np_i} > \chi_{1-\alpha}(l - r - 1) \,,$$

tedy pokud hodnota statistiky S_n překročí hodnoty $1-\alpha$-kvantilu χ^2 rozdělení s $(l-r-1)$ stupni volnosti.

Příklad 9.14 Při 120 hodech kostkou jsme získali následující výsledky

výsledek	1	2	3	4	5	6
četnost n_i	15	16	25	31	15	14

Je možno označit tuto kostku na hladině 0.05 za falešnou?

Máme tedy otestovat hypotézu, zda dané výsledky pocházejí z rovnoměrného rozdělení na množině $\{1, 2, 3, 4, 5, 6\}$, tj.

$$H_0 : F = U\{1, 2, 3, 4, 5, 6\} \quad \text{proti} \quad H_1 : F \neq U\{1, 2, 3, 4, 5, 6\}\,.$$

V tomto případě máme šest tříd odpovídajících jednotlivým výsledkům hodů. Pro teoretické pravděpodobnosti platí $p_i = 1/6, i = 1, \ldots, 6$ a teoretické počty jednotlivých výsledků ze 120 hodů jsou $np_i = 20$. Po dosazení dostaneme hodnotu testovací statistiky

$$S_n = \sum_{i=1}^{l} \frac{(n_i - 20)^2}{20} = 10.8\,.$$

Protože máme $l = 6$ tříd a nemuseli jsme odhadovat žádný parametr rozdělení ($r = 0$), porovnáme hodnotu statistiky s kvantilem $\chi^2_{0.95}(5) = 11.07 > S_n$. Na základě výsledků pokusu tedy nelze zamítnout hypotézu, že je kostka pravidelná.

Příklad 9.15 Nechť je dán generátor náhodných čísel z rovnoměrného rozdělení $U(0, 1)$ a naším úkolem je ověřit jeho kvalitu. Vygenerujeme 1000 náhodných čísel a interval $(0, 1)$ na deset intervalů stejné délky. Napozorované četnosti náhodných čísel v jednotlivých intervalech jsou uvedeny v tabulce

i-tý interval	1	2	3	4	5	6	7	8	9	10
četnost n_i	91	92	97	99	108	103	102	110	99	99

Je možné na základě výsledků na hladině významnosti 0.05 prokázat špatnou kvalitu generátoru?

Nulová hypotéza má tedy tvar

$$H_0 : F = U(0, 1) \quad \text{proti} \quad H_1 : F \neq U(0, 1)$$

a

$$n = 1000, \quad p_i = 0.1 \quad \text{a} \quad np_i = 100\,.$$

Po dosazení dostaneme

$$S_n = \frac{1}{100} \sum_{i=1}^{10} (n_i - 100)^2 = 3.33 < 16.919 = \chi^2_{0.95}(9)\,.$$

Na hladině významnosti 0.05 tedy nelze zamítnout hypotézu, že generátor je rovnoměrný.

Cvičení 9.5 Počet defektů na tištěných spojích má hypoteticky Poissonovo rozdělení. Bylo prozkoumáno 60 tištěných desek a bylo pozorováno

Počet chyb	0	1	2	3
Četnosti	32	15	9	4

Na hladině významnosti 5% otestujte, zda se skutečně jedná o Poissonovo rozdělení.