VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV MATEMATIKY

Matematika 3
Sbírka úloh z pravděpodobnosti

Irena Hlavičková
Dana Hliněná

evropský sociální fond v ČR
EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Ústav matematiky FEKT VUT v Brně, 2015
http://www.umat.feec.vutbr.cz

Obsah

Úvod

Tento učební text by měl sloužit jako doplněk ke skriptům Matematika 3. Na začátku každé kapitoly jsou vždy zopakovány základní pojmy a vzorce vztahující se k probírané tématice. Vždy je předvedeno několik podrobně řešených příkladů, pak následují příklady pro samostatnou práci.
Při přípravě této sbírky jsme vycházely ze zkušeností z cvičení předmětu BMA3. Snažíme se upozorňovat na chyby, kterých se studenti často dopouštějí, a ukazovat cestu k jejich nápravě. Doufáme, že naše sbírka jednak pomůže studentům lépe pochopit probíranou látku a jednak že bude užitečným pomocníkem při přípravě na zkoušku.
Budeme čtenářům vděčné za jakékoli věcné připomínky, upozornění na případné chyby a náměty pro zdokonalení.

Autorky, březen 2015

1 Základy popisné statistiky

1.1 Základní pojmy a řešené příklady

Následující odstavec je tu víceméně ze zoufalství. Bohužel se ukazuje, že stále více studentů má potíže s pochopením vět typu: „Alespoň 3 mají vlastnost tu a tu.“ V příkladech z pravděpodobnosti a statistiky se podobné formulace vyskytují poměrně často, a proto stručně připomeneme základní názvosloví.

Slovní vyjádření rovností a nerovností

X je právě n	$X = n$,	tj. přesně n, ani více, ani méně
X je více než n	$X > n$	
X je alespoň n	$X \geq n$,	tj. n nebo více
X je méně než n	$X < n$	
X je nanejvýš n	$X \leq n$,	tj. n nebo méně

Příklad 1.1. V následující tabulce je seznam zahrádkářů, vždy jméno a rozloha pozemku v metrech čtverečních.

Beneš	1 200	Horáčková	500	Růžička	700
Černá	700	Kučera	800	Svobodová	1 000
Doležal	1 000	Moravčík	1 500	Veselý	1 800
Dvořák	900	Novák	900	Zeman	2 000
Fiala	1 100	Procházková	800	Žák	500

a) Kteří zahrádkáři mají pozemek o rozloze právě 1 000 m^2?

b) Kolik zahrádkářů má pozemek o rozloze alespoň 1 000 m^2?

c) Kolik zahrádkářů má pozemek o rozloze větší než 1 000 m^2?

d) Kteří zahrádkáři mají pozemek o rozloze menší než 700 m^2?

e) Kolik zahrádkářů má pozemek o rozloze nanejvýš 700 m^2?

Řešení. a) Pozemek o rozloze právě 1 000 m^2 mají Doležal a Svobodová.
b) Pozemek o rozloze alespoň 1 000 m^2 má 7 zahrádkářů (Beneš, Doležal, Fiala, Moravčík, Svobodová, Veselý a Zeman).
c) Pozemek o rozloze větší než 1 000 m^2 má 5 zahrádkářů (Beneš, Fiala, Moravčík, Veselý a Zeman).

d) Pozemek o rozloze menší než 700 m^2 mají Horáčková a Žák.
e) Pozemek o rozloze nanejvýš 700 m^2 mají 4 zahrádkáři (Černá, Horáčková, Růžička a Žák).

A teď už opravdu k popisné statistice:

Četnosti
Předpokládejme, že v souboru o rozsahu n může sledovaný znak x nabývat k různých hodnot (variant) $x_1, x_2, \dots, x_k$. **Četnost** varianty x_i je počet výskytů této hodnoty ve sledovaném souboru a označíme ji n_i, $i = 1, \dots, k$. Pak platí

$$n_1 + n_2 + \cdots + n_k = n.$$

Relativní četnost varianty x_i zavedeme jako

$$f_i = \frac{n_i}{n}\,.$$

Pro relativní četnosti platí

$$f_1 + \cdots + f_k = \frac{n_1}{n} + \cdots \frac{n_k}{n} = \frac{n_1 + \cdots + n_k}{n} = 1.$$

Relativní četnost se často vyjadřuje v procentech.
Kumulativní četnosti (opět absolutní nebo relativní) udávají, kolik jednotek má hodnotu znaku menší nebo rovnou vybrané variantě x_i.
Pro spojité znaky, které mohou nabýt jakékoli hodnoty z určitého intervalu, sestavujeme **intervalové rozdělení četností** – stanovujeme počty výskytů hodnot znaku, které náleží do předem vymezených intervalů. Počet intervalů, do kterých hodnoty rozdělíme, můžeme volit zhruba $\sqrt{n}$ nebo $1 + \log_2 n \doteq 1 + 3{,}3 \log n$.

Aritmetický průměr
Máme-li soubor rozsahu n a zjištěné hodnoty znaku jsou $x_1, \dots, x_n$, pak jejich aritmetický průměr je

$$\overline{x} = \frac{x_1 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Jestliže známe četnosti výskytu jednotlivých hodnot znaku, můžeme aritmetický průměr vypočítat jako

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot n_i = \sum_{i=1}^{k} x_i \cdot f_i. \tag{1.1}$$

Modus
Modus je hodnota, která se v souboru vyskytuje **nejčastěji**, označíme ji $\hat{x}$.

Medián
Medián je **prostřední hodnota** ze souboru uspořádaného podle velikosti. Označíme jej $\tilde{x}$ nebo též $\tilde{x}_{0,5}$.
Označíme-li prvky uspořádané podle velikosti jako $x_1, x_2, \ldots, x_n$ a počet prvků n je liché číslo, medián je prostřední hodnota,

$$\tilde{x} = x_{(n+1)/2}\,.$$

Je-li rozsah souboru n sudé číslo, je medián průměr ze dvou prostředních prvků,

$$\tilde{x} = \frac{1}{2}\left(x_{n/2} + x_{(n/2)+1}\right).$$

Kvantily
Pro $p \in (0,1)$ je kvantil $\tilde{x}_p$ neboli p-kvantil takové číslo, které odděluje nejmenších $p \cdot 100$ % hodnot statistického znaku od největších $(1-p) \cdot 100$ % hodnot. Speciálně 0,5-kvantil je medián, 0,25-kvantil je dolní kvartil a 0,75-kvantil je horní kvartil.

Aritmetický průměr, modus a medián patří mezi charakteristiky polohy. Často je ale důležité vědět, jak jsou hodnoty statistického souboru rozptýlené. Zajímá nás, jestli se znak pohybuje nejčastěji jen v určitém nevelkém intervalu, nebo zda je jeho rozpětí široké. K popisu těchto vlastností slouží charakteristiky variability:

Variační rozpětí
je rozdíl největší a nejmenší hodnoty znaku:

$$R = x_{\max} - x_{\min}.$$

Mezikvartilové rozpětí
je rozdíl třetího a prvního kvartilu:

$$\tilde{x}_{0,75} - \tilde{x}_{0,25}.$$

Rozptyl statistického znaku **v populaci** označíme σ^2 a definujeme jej jako

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2. \tag{1.2}$$

Rozptyl udává, jak se hodnoty statistického znaku průměrně liší od průměrné hodnoty, ovšem ve druhé mocnině. Výsledek je proto ve čtvercích použité měrné jednotky, což ztěžuje jeho interpretaci. Abychom se dostali zpátky na původní jednotky, rozptyl odmocníme, čímž získáme tzv. směrodatnou odchylku:

Směrodatná odchylka σ je odmocnina z rozptylu:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}.$$

Výběrový rozptyl
Jestliže nemáme k dispozici data pro celou populaci, ale pouze pro vybraný vzorek, pro odhad rozptylu se používá výběrový rozptyl, který značíme s^2 a je definován jako

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2. \tag{1.3}$$

Výběrová směrodatná odchylka se značí s a je to odmocnina z výběrového rozptylu,

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}. \tag{1.4}$$

Výpočet rozptylu je jednodušší, použijeme-li vztah

$$s^2 = \left(\frac{1}{n-1}\sum_{i=1}^{n}x_i^2\right) - \frac{n}{n-1}\,\overline{x}^2. \tag{1.5}$$

Příklad 1.2. 15 optických kabelů bylo testováno na pevnost v tahu. Byly zjištěny následující hodnoty (v N):

$$487, 530, 507, 499, 507, 498, 499, 515, 514, 514, 507, 488, 507, 516, 505.$$

Vypočítejte aritmetický průměr, výběrový rozptyl a výběrovou směrodatnou odchylku. Dále určete medián.

Řešení. Aritmetický průměr je

$$\overline{x} = \frac{1}{15}(487 + 530 + 507 + \cdots + 505) = 506{,}2.$$

Výběrový rozptyl bychom mohli počítat jako

$$s^2 = \frac{1}{14}((487 - 506{,}2)^2 + \cdots + (505 - 506{,}2)^2),$$

ale méně pracný je výpočet pomocí (1.5):

$$s^2 = \frac{1}{14}(487^2 + 530^2 + \cdots + 505^2) - \frac{15}{14}506{,}2^2 \doteq 124{,}03.$$

Výběrová směrodatná odchylka je pak

$$s = \sqrt{s^2} \doteq 11{,}14.$$

K určení mediánu je zapotřebí hodnoty uspořádat podle velikosti:

$$487, 488, 498, 499, 499, 505, 507, \underline{507}, 507, 507, 514, 514, 515, 516, 530.$$

Medián je prostřední, tj. v našem případě osmá, z nich:

$$\tilde{x} = 507.$$

Příklad 1.3. Máme záznamy o týdenních počtech dopravních nehod na určitém úseku dálnice během 20 týdnů:

$$6, 7, 6, 6, 5, 14, 6, 6, 7, 3, 7, 4, 6, 5, 3, 3, 8, 10, 6, 7.$$

Znázorněte četnosti a kumulativní četnosti pomocí histogramu a najděte modus.

Řešení. Tabulka četností je

Týdenní počet nehod x_i	3	4	5	6	7	8	10	14
Četnost n_i	3	1	2	7	4	1	1	1
Kumulativní četnost	3	4	6	13	17	18	19	20

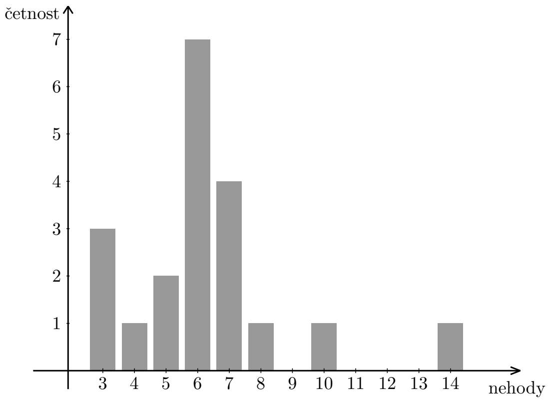
Příslušné histogramy vidíme na obrázcích 1.1 a 1.2.
Ještě zbývá určit modus. Nejčastěji došlo k 6 nehodám, a proto

$$\hat{x} = 6.$$

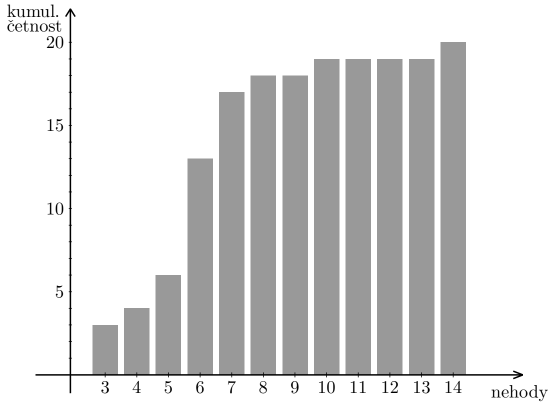
Příklad 1.4. V anketě byli studenti tázáni, kolik hodin týdně stráví přípravou na jistý předmět. Ankety se účastnilo 70 studentů, v následující tabulce jsou četnosti jejich odpovědí:

Počet hodin	1	2	3	4	5
n_i	21	29	17	2	1

Určete aritmetický průměr, modus a medián počtu hodin strávených přípravou.



Obr. 1.1: Histogram četností z příkladu 1.3



Obr. 1.2: Histogram kumulativních četností

Řešení. Vidíme, že jednu hodinu tráví přípravou 21 studentů, dvě hodiny 29 studentů, atd., celkem je studentů 70. Průměrný počet hodin je tedy podle (1.1)

$$\overline{x} = \frac{1}{70}(1 \cdot 21 + 2 \cdot 29 + 3 \cdot 17 + 4 \cdot 2 + 5 \cdot 1) = \frac{143}{70} \doteq 2{,}04.$$

Modus je $\hat{x} = 2$, protože nejvíce studentů tráví přípravou 2 hodiny týdně.
Medián bude průměr 35. a 36. hodnoty v souboru seřazeném podle velikosti. V našem případě jsou obě hodnoty rovny dvěma (prvních 21 odpovědí je 1, následujících 29 je 2). Tedy

$$\tilde{x} = \frac{1}{2}(2 + 2) = 2.$$

Příklad 1.5. Na 50 místech v určité oblasti byla měřena koncentrace polétavého prachu v ovzduší (v μg/m^3). Byly zjištěny následující hodnoty (jsou již seřazené podle velikosti):

6	10	18	18	19	21	21	22	24	25
25	28	28	29	29	30	31	32	32	32
33	33	34	34	34	34	38	42	42	47
48	49	50	52	53	53	53	54	54	56
56	58	59	59	60	63	66	67	78	96

Najděte intervalové rozdělení četností a znázorněte je pomocí histogramu. Pomocí intervalových četností odhadněte aritmetický průměr. Dále najděte modální (nejčetnější) interval.

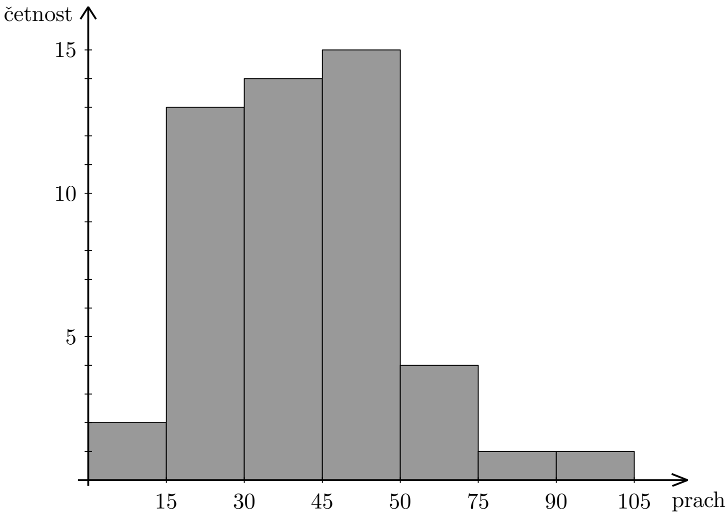
Řešení. Máme $n = 50$ hodnot, tedy počet intervalů by měl být kolem 7 ($\sqrt{50} \doteq 7$, $1 + 3{,}3 \log 50 \doteq 6{,}6$). Hodnoty jsou zhruba v rozpětí od 0 do 100, můžeme proto volit intervaly délky 15. Poznamenejme, že počet a délka intervalů nejsou určeny nijak exaktně, mohli jsme si je zvolit i jinak. Nicméně není dobré volit intervalů příliš málo ani příliš mnoho. Rozdělení intervalových četností v našem případě je

Interval	$\langle 0, 15)$	$\langle 15, 30)$	$\langle 30, 45)$	$\langle 45, 60)$	$\langle 60, 75)$	$\langle 75, 90)$	$\langle 90, 105)$
Četnost	2	13	14	15	4	1	1

Histogram intervalových četností je znázorněn na obrázku 1.3. Vidíme, že modální interval je $\langle 45, 60)$, v něm se naléžá nejvíce hodnot.
Pro odhad průměru můžeme použít středy intervalů:

$$\overline{x} \doteq \frac{1}{50}(7{,}5 \cdot 2 + 22{,}5 \cdot 13 + 37{,}5 \cdot 14 + 52{,}5 \cdot 15 + 67{,}5 \cdot 4 + 82{,}5 + 97{,}5) = 41{,}4.$$

Zdůrazněme, že se jedná pouze o odhad. Skutečná hodnota průměru, vypočítaná z přesných 50 hodnot, by vyšla 40,7.



Obr. 1.3: Histogram intervalových četností z příkladu 1.5

1.2 Příklady pro samostatnou práci

Příklad 1.6. Máme soubor hodnot

$$5, 1, 2, 1, 7, 5, 2, 5, 1, 1.$$

Vypočtěte aritmetický průměr, výběrový rozptyl a směrodatnou odchylku. Dále najděte modus a medián.

Výsledek: $\overline{x} = 3$; $s^2 \doteq 5{,}11$; $s \doteq 2{,}26$; modus $\hat{x} = 1$; medián $\tilde{x} = 2$

Příklad 1.7. Máme soubor hodnot

$$6, 3, 9, 4, 3, 6, 7, 15, 0, 7, 3, 15, 3, 2, 5, 1, 2, 8, 3, 6, 4, 3, 1, 4, 0.$$

Určete četnosti jednotlivých variant a pak pomocí těchto četností vypočítejte průměr, výběrový rozptyl a směrodatnou odchylku. Dále najděte modus a medián.

Výsledek: Tabulka četností viz níže; $\overline{x} = 4{,}8$; $s^2 \doteq 15{,}08$; $s \doteq 3{,}88$; modus $\hat{x} = 3$; medián $\tilde{x} = 4$

x_i	0	1	2	3	4	5	6	7	8	9	15
n_i	2	2	2	6	3	1	3	2	1	1	2

Příklad 1.8. Četnosti zkoumaného statistického znaku jsou v následující tabulce.

x_i	4	5	6	7	8
n_i	4	3	3	7	3

a) Určete relativní četnosti a absolutní i relativní kumulativní četnosti.
b) Vypočítejte průměr.
c) Najděte medián a horní a dolní kvartil.

Výsledek: a) viz níže; b) $\overline{x} = 6{,}1$; c) medián $\tilde{x} = 6{,}5$; dolní kvartil $\tilde{x}_{0,25} = 5$; horní kvartil $\tilde{x}_{0,75} = 7$

x_i	4	5	6	7	8
rel. četn. f_i	0,2	0,15	0,15	0,35	0,15
kumul. četn.	4	7	10	17	20
rel. kumul. četn.	0,2	0,35	0,5	0,85	1

Příklad 1.9. Známe rozdělení kumulativních četností zkoumaného znaku:

x_i	0	1	2	3	4	5	6	8	10
kumul. četnost	41	70	84	90	93	95	97	99	100

Části a) a b) se pokuste vyřešit, aniž byste hledali „obyčejné" četnosti jednotlivých variant.
a) Najděte medián a horní kvartil.
b) Kolikrát se ve statistickém souboru vyskytuje hodnota větší než 4?
c) Určete rozdělení četností.

Výsledek: a) medián $\tilde{x} = 1$; horní kvartil $\tilde{x}_{0,75} = 2$; b) sedmkrát; c) viz níže

x_i	0	1	2	3	4	5	6	8	10
n_i	41	29	14	6	3	2	2	2	1

Příklad 1.10. Testu inteligence se účastnilo 60 lidí. Intervalové četnosti jejich výsledků jsou v následující tabulce:

Interval	$\langle 90, 100)$	$\langle 100, 110)$	$\langle 110, 120)$	$\langle 120, 130)$	$\langle 130, 140)$
n_i	6	23	24	6	1

a) Kolika účastníkům testu bylo naměřeno IQ alespoň 110?
b) Kolika účastníkům testu bylo naměřeno IQ menší než 110?
c) Kolika účastníkům testu bylo naměřeno IQ nanejvýš 110?
d) Odhadněte průměrnou naměřenou hodnotu IQ.

Výsledek: a) 31; b) 29; c) z informací v tabulce nelze zjistit – nevíme, kolik lidí mělo IQ přesně 110; d) $\overline{x} \doteq 110{,}5$

2 Základy kombinatoriky

Tato kapitola bude trochu odlišná od ostatních kapitol naší sbírky. Na rozdíl od nich totiž přímo nenavazuje na teorii probíranou v předmětu Matematika 3, ale měla by sloužit spíše jako opakování a příprava pro studium tohoto předmětu. Kombinatorika se zabývá problémem, kolik máme možností, jak z nějakého výchozího souboru objektů sestavit skupinu s určitými vlastnostmi, a pro některé partie pravděpodobnosti je zcela nepostradatelná. Ve skriptech z Matematiky 3 se počítá s tím, že základy kombinatoriky jsou již studentům známy ze střední školy. Protože však zkušenost ukazuje, že mnohé věci jsou ve druhém ročníku vysoké školy už dávno zapomenuty, považujeme za vhodné některé základní pojmy připomenout. Budeme se vždy snažit na konkrétních příkladech vysvětlit, jak se ke vzorcům pro počty variací, permutací či kombinací vlastně dojde. Komu se výklad, určený hlavně těm studentům, pro které je toto odvětví matematiky obtížné, bude zdát příliš „polopatistický“, může jej samozřejmě přeskočit a věnovat se řešení příkladů samostatně.

2.1 Základní pojmy a řešené příklady

Příklad 2.1. Na vrchol kopce vedou čtyři cesty. Kolika způsoby si může turista, který chce na tento kopec vystoupit, naplánovat výlet, jestliže si pro výstup i sestup může vybrat kteroukoli z cest?

Řešení. Představme si, že po oněch čtyřech cestách vedou turistické značky: červená, modrá, zelená a žlutá. Turista může jít nahoru po kterékoli z nich. Má tedy čtyři možnosti, jak na kopec vystoupit. Dejme tomu, že si vybral červenou značku. Dolů pak může jít opět čtyřmi různými cestami. Máme tedy čtyři plány výletu, kdy se na kopec vystoupí po červené značce. Podobně pak čtyři plány začínající modrou, čtyři zelenou a čtyři žlutou značkou. Můžeme si je i vypsat (první písmeno označuje značku, po které šel turista nahoru, druhé značku pro sestup):

ČČ	MČ	ZČ	ŽČ
ČM	MM	ZM	ŽM
ČZ	MZ	ZZ	ŽZ
ČŽ	MŽ	ZŽ	ŽŽ

Počet možností pro naplánování výletu je tedy $4 \cdot 4 = \underline{\underline{16}}$.

Příklad 2.2. Řešme nyní stejnou úlohu, ale tentokrát turista nechce jít zpět stejnou cestou, kterou na kopec vyšel. Kolik má možností v tomto případě?

Řešení. Pro výstup si opět může vybrat kteroukoli ze čtyř značek. Jestliže si vybral červenou, má pro sestup už jen tři možnosti: modrou, zelenou nebo žlutou značku. Podobně je to pro případy, kdy si pro výstup zvolil některou jinou značku – pro každý z těchto případů má 3 možnosti cesty dolů. Možnosti si opět můžeme vypsat:

ČM	MČ	ZČ	ŽČ
ČZ	MZ	ZM	ŽM
ČŽ	MŽ	ZŽ	ŽZ

Celkem existuje $4 \cdot 3 = \underline{\underline{12}}$ variant výletu.

Jestliže postup předvedený v předchozích dvou příkladech zobecníme, dostaneme tzv. pravidlo součinu, které budeme později často používat i při řešení složitějších příkladů:

Pravidlo součinu
Jestliže objekt A můžeme vybrat m způsoby a po každém takovém výběru lze objekt B vybrat n způsoby, pak výběr uspořádané dvojice (A,B) lze uskutečnit $m \cdot n$ způsoby.

Poznamenejme, že čísla m a n nemusí být tak snadno určitelná jako u příkladů 2.1 a 2.2. Později uvidíme, že někdy budeme potřebovat složitější úvahu, abychom k nim dospěli. Nyní zkusíme pracovat s většími skupinami než s dvojicemi.

Příklad 2.3. U jistého typu automobilu si zákazník může zvolit vůz jedné z pěti barev. Automobil může a nemusí být vybaven airbagy, stejně tak může a nemusí mít klimatizaci. Kolika způsoby si může zákazník vybrat?

Řešení. Pro začátek z našich úvah vypustíme klimatizaci. Pokud budeme volit jen barvu a případné vybavení airbagy, máme celkem $5 \cdot 2 = 10$ možností. (Barvu můžeme zvolit 5 způsoby. Ke každé zvolené barvě máme ještě dvě možnosti, jak se rozhodnout: airbagy tam buď budou, nebo nebudou.) K libovolné z těchto 10 variant teď máme ještě dvě možnosti ohledně klimatizace. Celkem tedy existuje $5 \cdot 2 \cdot 2 = \underline{\underline{20}}$ možností pro výběr vybavení automobilu.

Příklad 2.4. V prvním semestru prvního ročníku studenti skládají čtyři zkoušky (BMA1, BMTD, BFY1 a BEL1[1]). U každé zkoušky může být student hodnocen stupněm A až F. Kolika různých celkových výsledků („vysvědčení“) může dosáhnout?

Řešení. Pro zkoušku z BMA1 má 6 možností výsledku. Ke každému z nich pak pro zkoušku z BMTD zase 6. Tím máme $6 \cdot 6$ možností pro první dvě zkoušky. Ke každé z nich je 6 možností pro zkoušku z BFY1. To dělá $6 \cdot 6 \cdot 6$ možností a ke každé z nich pak ještě je 6 možností pro výsledek zkoušky z BEL1. Výsledných možností pro celé studentovo „vysvědčení“ je tedy $6 \cdot 6 \cdot 6 \cdot 6 = 6^4 = \underline{\underline{1296}}$.

[1] Podle studijního plánu pro školní rok 2007/2008.

V tomto příkladu jsme sestavovali uspořádané čtveřice, udávající postupně známky z matematiky, materiálů, fyziky a elektrotechniky, ze šestiprvkové množiny možných výsledků {A, B, C, D, E, F}. Jednalo se o tzv. variace s opakováním.

Variace s opakováním
Variace s opakováním jsou uspořádané k-tice vybírané z n prvkové množiny, přičemž jednotlivé prvky se v k-ticích mohou opakovat.
Protože každý z k prvků můžeme vybrat n způsoby, počet těchto variací je

$$\underbrace{n \cdots n}_{k\ krát} = n^k.$$

S variacemi s opakováním jsme se setkali už v příkladu 2.1, kde jsme sestavovali uspořádané dvojice ze čtyřprvkové množiny {Č, M, Z, Ž}. Výsledek tohoto příkladu můžete porovnat s výše uvedeným vzorcem.

Příklad 2.5. Nový děkan jisté fakulty si vybírá 3 proděkany: pro studium, pro vědu a výzkum a pro vnější vztahy. V úvahu připadá celkem 10 lidí, se kterými je zadobře. Kolik je možností pro nové vedení fakulty?

Řešení. Jako studijního proděkana si může vybrat kohokoli z 10. K němu pak jako vědeckého proděkana už jen někoho ze zbývajících 9 lidí. Tím máme $10 \cdot 9$ možností pro první dva proděkany a ke každé z nich je pak ještě 8 možností pro volbu třetího. Celkem je $10 \cdot 9 \cdot 8 = \underline{\underline{720}}$ možností.

Tentokrát jsme z množiny deseti lidí sestavovali uspořádané trojice proděkanů, ve kterých se prvky nemohly opakovat. Jak již asi tušíte, byly to tzv. variace bez opakování.

Variace bez opakování
Variace bez opakování jsou uspořádané k-tice vybírané z n prvkové množiny, přičemž jednotlivé prvky se v k-ticích nesmí opakovat.
Protože první prvek můžeme vybrat n způsoby, druhý $n-1$ způsoby atd., počet těchto variací je

$$\underbrace{n \cdot (n-1) \cdots (n-k+1)}_{celkem\ násobíme\ k\ čísel}$$

Všímavý čtenář jistě rozeznal, že na variace bez opakování byl již příklad 2.2.

Příklad 2.6. Pět přátel si koupilo lístky do kina. Kolika způsoby se mohou na vyhrazených pěti sedadlech rozesadit?

Řešení. Můžeme v podstatě zopakovat úvahu z příkladu s proděkany. Na první sedadlo si může sednout kterýkoli z pěti kamarádů, na druhé pak už jen někdo ze zbylých čtyř, na třetí pak ze tří, atd. Celkem se mohou usadit $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = \underline{\underline{120}}$ způsoby.

Nyní jsme zkoumali, kolika způsoby se dá – co do pořadí – přeskupit 5 objektů (v našem případě návštěvníků kina). Též by se dalo říct, že jsme sestavovali uspořádané pětice z pěti prvků. Tím se dostáváme k tzv. permutacím. Dříve, než je budeme definovat, připomeňme pro jistotu pojem, který s permutacemi úzce souvisí.

Faktoriál
Je-li n přirozené číslo, pak symbolem $n!$ (čteno „n faktoriál") rozumíme součin všech přirozených čísel od 1 do n:

$$n! = n \cdot (n-1) \cdots 1$$

Speciálně se definuje

$$0! = 1.$$

Výsledek předchozího příkladu se tedy mohl zapsat jako 5!.

Permutace
Permutace jsou uspořádané n-tice vybírané z n prvkové množiny, přičemž jednotlivé prvky se nesmí opakovat.
Počet permutací z n prvků (neboli počet všech možných přeskupení n prvků) je

$$n \cdot (n-1) \cdots 1 = n!$$

Nyní zkusíme o něco obtížnější variantu příkladu s návštěvníky kina.

Příklad 2.7. Sedm přátel, čtyři chlapci a tři dívky, si koupilo lístky do kina. Chtějí se rozesadit tak, aby chlapci a dívky seděli střídavě. Kolik mají možností, jak to udělat?

Řešení. Na počet možných rozesazení se dá přijít více způsoby. Můžeme si například říct, že první sedadlo musí obsadit některý ze čtyř chlapců. Ke každé z těchto čtyř možností máme 3 možnosti, jak obsadit druhé sedadlo některou z dívek. Ke každé z výsledných $4 \cdot 3$ možností pro první dvě sedadla máme 3 možnosti, jak obsadit třetí sedadlo některým ze zbývajících chlapců, atd. Tímto způsobem dostáváme výsledek ve tvaru $4 \cdot 3 \cdot 3 \cdot 2 \cdot 2 \cdot 1 \cdot 1 = \underline{\underline{144}}$.
K témuž výsledku se ale můžeme dostat i tak, že si řekneme, že určitá 4 sedadla jsou vyhrazena chlapcům. Ti je mohou obsadit $4! = 24$ způsoby. Ke každému z těchto 24 způsobů rozesazení chlapců máme 3! rozesazení dívek na 3 sedadla, která jsou jim vyhrazena. To dává celkem $4! \cdot 3! = 24 \cdot 6 = \underline{\underline{144}}$ možností.

Druhý z postupů předvedených v tomto příkladu použijeme nyní.

Příklad 2.8. Kolik různých slov se dá sestavit ze všech písmen slova ANANAS? (Slovem zde rozumíme libovolné přeskupení zadaných písmen; význam takovéto „slovo" mít nemusí.)

Řešení. Máme k dispozici 6 písmen, z nichž některá se opakují. Kdyby všechna písmena byla navzájem různá, odpověď na otázku by byla jednoduchá. Proto budeme pro začátek považovat jednotlivá A a N za různá písmena – slovo by pak mohlo vypadat např. ANᴀɴaS. Máme 6! možností, jak tyto znaky přeskupit. Ovšem některá z těchto přeskupení jsou ve skutečnosti totožná. Např. ASɴaNᴀ je totéž slovo jako aSNᴀɴA. Všech 6! přeskupení proto můžeme rozdělit do několika (zatím nevíme kolika – tento počet je právě úkolem zjistit) skupin, které se skládají vždy ze stejných slov. Kdybychom to např. brali podle abecedy, v první skupině by byla slova AᴀaNɴS, ᴀAaɴNS, ..., ve druhé AᴀaNSɴ, aAᴀɴSN, ..., atd.
Zkoumejme nyní, kolik prvků má první skupina (AAANNS). Je to podobný úkol jako v příkladu 2.7. Pro rozmístění tří různých písmen A na prvních třech místech slova máme 3! možností. Ke každé z těchto možností jsou pak 2 (nebo, chceme-li, 2!) možnosti, jak uspořádat dvě různá písmena N na dalších dvou místech. Poslední S je pak umístěno pevně. Celkem tedy je $3! \cdot 2!$ slov (se zatím rozlišitelnými A a N), která dávají slovo AAANNS. Stejným způsobem bychom dospěli k tomu, že v každé z ostatních skupin je také $3! \cdot 2!$ slov. Když situaci zrekapitulujeme, vidíme, že máme celkem 6! slov rozdělených do skupin. V každé skupině je $3! \cdot 2!$ slov. Počet skupin neboli počet různých slov sestavitelných z písmen ANANAS je tedy

$$\frac{6!}{3! \cdot 2!} = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = \underline{\underline{60}}.$$

Obecně, zkoumáme-li přeskupení n předmětů, z nich některé se opakují, mluvíme o permutacích s opakováním. Ke vzorci pro jejich celkový počet bychom dospěli stejnou úvahou jako v předchozím příkladu.

Permutace s opakováním
Permutace s opakováním jsou permutace sestavované z n prvkové množiny, ve které je n_1 prvků prvního druhu, n_2 prvků druhého druhu, ..., n_k prvků k-tého druhu, kde $n_1 + n_2 + \cdots + n_k = n$. Jejich počet je

$$\frac{n!}{n_1! \cdot n_2! \cdots n_k!}.$$

Příklad 2.9. Máme balíček 32 různých karet. Hráč dostane 5 karet. Kolik je možností pro tuto pětici? (Nezáleží na tom, v jakém pořadí byly karty rozdány. Záleží pouze na tom, které karty hráč dostal.)

Řešení. Kdyby na pořadí karet záleželo, byli bychom ve stejné situaci jako v příkladu 2.5. Šlo by o variace bez opakování. První karta by mohla být kterákoli z 32, k ní druhá pak kterákoli z 31, atd. Celkem bychom měli $32 \cdot 31 \cdot 30 \cdot 29 \cdot 28 = 24\,165\,120$ možností. Jenže zde na pořadí karet nezáleží. Uspořádaná pětice $K\heartsuit, Q\clubsuit, 10\heartsuit, 8\diamondsuit, 7\spadesuit$ hráči vyjde nastejno jako $8\diamondsuit, Q\clubsuit, 7\spadesuit, K\heartsuit, 10\heartsuit$. Všech 24 165 120 možností proto můžeme rozdělit do (opět zatím neznámého počtu) skupin. Každá skupina se skládá z pětic karet, které se liší pouze pořadím. Pět karet můžeme přeskupit 5! způsoby, každá skupina má proto 5! prvků. Počet skupin neboli počet všech různých pětic karet sestavitelných z 32 karet je proto

$$\frac{32 \cdot 31 \cdot 30 \cdot 29 \cdot 28}{5!}\,.$$

S tímto výsledkem můžeme ještě chvíli „čarovat“. Zlomek rozšíříme výrazem 27!. Po úpravě pak dostaneme výsledek v kompaktnějším tvaru:

$$\frac{32 \cdot 31 \cdot 30 \cdot 29 \cdot 28}{5!} = \frac{32 \cdot 31 \cdot 30 \cdot 29 \cdot 28 \cdot 27!}{5! \cdot 27!} = \frac{32!}{5! \cdot 27!}$$

Jiný způsob řešení: Představme si, že máme karty nějak uspořádané, např. $A\heartsuit, A\diamondsuit, A\spadesuit, A\clubsuit, K\heartsuit, K\diamondsuit, \ldots, 7\heartsuit, 7\diamondsuit, 7\spadesuit, 7\clubsuit$. Každá pětice karet se teď dá interpretovat jako 32-členná posloupnost pěti jedniček a dvaceti sedmi nul. Jedničky jsou na místech vybraných karet, nuly na všech ostatních. Např. posloupnost $(1, 0, 1, 0, 0, 1, 0, \ldots, 0, 1, 1)$ by znamenala, že hráč dostal $A\heartsuit, A\spadesuit, K\diamondsuit, 7\spadesuit$ a $7\clubsuit$. Zadanou otázku teď můžeme převést na problém, kolika způsoby lze navzájem přeskupit pět jedniček a 27 nul. Na tuto otázku už ale odpověď známe. Jde o něco podobného jako v příkladu s ananasem a výsledek je

$$\frac{32!}{5! \cdot 27!}\,.$$

Oběma způsoby jsme dospěli k témuž výsledku, který po vyčíslení dá $\underline{\underline{201\,376}}$ možných pětic sestavitelných z 32 karet.

Tento příklad byl velmi důležitý. Otázka, kolika způsoby lze vybrat z n-prvkové množiny k-prvkovou skupinu, ve které nezáleží na pořadí, ale pouze na tom, které prvky byly vybrány, se vyskytuje velmi často. V právě předvedeném příkladu jsme odvodili vztah pro výpočet počtu pětiprvkových podmnožin z dvaatřicetiprvkové množiny. V následujícím rámečku je situace popsána obecně.

Kombinace, kombinační čísla

Kombinace k-té třídy z n prvků jsou k-prvkové skupiny vybírané z n-prvkové množiny, přičemž nezáleží na pořadí, v jakém prvky byly vybrány.
Jinými slovy, jsou to k-prvkové podmnožiny z množiny mající n prvků.
Počet takovýchto kombinací je

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

Symbol $\binom{n}{k}$ čteme „n nad k“ a nazýváme jej kombinační číslo.

U karet ještě chvíli zůstaneme a zkusíme o něco složitější příklady. Výpočty kombinačních čísel pro jistotu v prvních několika případech rozepíšeme.

Příklad 2.10. Kolika způsoby můžeme z balíčku 52 karet (4×13 karet v barvách $\heartsuit, \diamondsuit, \spadesuit, \clubsuit$ o hodnotách 2, 3, …, K, A) vybrat pětici karet, v které jsou právě dvě esa?

Řešení. Nejprve se podíváme, kolika způsoby lze z balíčku vybrat dvojici es. Esa máme celkem 4, z nich vybíráme dvojici, nezáleží nám na tom, v jakém pořadí jsme karty vybrali, ale jen na tom, z kterých es se dvojice skládá. Počet všech možných dvojic je proto

$$\binom{4}{2} = \frac{4!}{2! \cdot 2!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 2 \cdot 1} = 6.$$

Jestliže už je nějaká dvojice es vybrána, musíme k ní – pro doplnění do pětice – vybrat ještě tři karty jiné než esa. Takovýchto karet máme 48, vybíráme z nich trojici. Pro tento výběr je počet možností

$$\binom{48}{3} = \frac{48!}{3! \cdot 45!} = \frac{48 \cdot 47 \cdot 46 \cdot 45!}{3 \cdot 2 \cdot 1 \cdot 45!} = 17\,296.$$

Nyní použijeme pravidlo součinu – dvojici es lze vybrat $\binom{4}{2}$ způsoby, ke každé takovéto dvojici lze zbývající trojici doplnit $\binom{48}{3}$ způsoby, pětici složenou ze dvou es a tří „ne-es“ můžeme proto sestavit

$$\binom{4}{2} \cdot \binom{48}{3} = 6 \cdot 17\,296 = \underline{\underline{103\,776}}$$

způsoby.

Kombinační čísla se dají snadno vypočítat pomocí tzv. *Pascalova trojúhelníka.* Tímto názvem označujeme tabulku trojúhelníkového tvaru

$$\begin{array}{lccccccccccccc}
n=0 & & & & & & & 1 & & & & & & \\
n=1 & & & & & & 1 & & 1 & & & & & \\
n=2 & & & & & 1 & & 2 & & 1 & & & & \\
n=3 & & & & 1 & & 3 & & 3 & & 1 & & & \\
n=4 & & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
n=5 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
n=6 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}$$

Každý řádek zde obsahuje všechna kombinační čísla pro totéž n. Podle známého vzorce, $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$, dostaneme sečtením dvou sousedních kombinačních čísel některého řádku to kombinační číslo, jež stojí v dalším řádku pod mezerou mezi nimi. Pascalův trojúhelník tak může sloužit k dosti rychlému výpočtu kombinačních čísel.

Příklad 2.11. Kolika způsoby můžeme z balíčku 52 karet vybrat 6 karet tak, aby nanejvýš jedna z vybraných karet byla piková?

Řešení. Má-li být ve vybrané šestici nanejvýš jedna piková karta, znamená to, že šestice obsahuje buď právě jednu takovou kartu, nebo žádnou. Celkový počet proto dostaneme tak, že sečteme počet šestic skládajících se z jedné pikové karty (takovýchto karet máme celkem k dispozici 13) a pěti jiných karet (jiných karet je celkem 39) a počet šestic složených ze samých karet jiných než pikových. Jednotlivé počty určíme podobně jako v předchozím příkladu, celkový výsledek je pak

$$\underline{\underline{13 \cdot \binom{39}{5} + \binom{39}{6}}}.$$

Postup použitý v tomto příkladu můžeme zobecnit. Často se podaří objekty, jejichž počet chceme zjistit, rozdělit do více skupin, jejichž počty prvků umíme určit. Celkový počet objektů je pak roven součtu počtů prvků jednotlivých skupin. Musíme však dát pozor na to, aby jednotlivé skupiny byly disjunktní, tj. aby žádný objekt nespadal do více skupin současně. Jako varování by nám mohl posloužit následující příklad.

Příklad 2.12. Jistá vysoká škola promíjí přijímací zkoušku uchazečům, kteří maturovali z matematiky nebo z fyziky a dosáhli alespoň z jednoho z těchto předmětů klasifikace výborně. Z celkového počtu 1000 přihlášených na tuto školu maturovalo z matematiky na výbornou 100 uchazečů a z fyziky 80. Kolik uchazečů bude přijato bez přijímací zkoušky?
Pokus o řešení: Někdo by si třeba mohl říct, že výsledek je

$$100 + 80 = 180.$$

Tento výsledek však vůbec nemusí být správný. Uchazeči, kteří maturovali za jedna z matematiky i z fyziky, jsou zde započítáni dvakrát. Na základě informací, které máme, nejsme schopni počet lidí přijatých bez přijímaček přesně určit. Kdybychom ale navíc věděli, že 30 uchazečů maturovalo na výbornou z obou předmětů, úloha by již řešitelná byla. Počet přijatých bez přijímací zkoušky by byl

$$100 + 80 - 30 = 150.$$

Vzali jsme 100 lidí, kteří mají jedničku z matematiky, a k nim jsme přidali z těch, kdo mají jedničku z fyziky, jen ty, kteří ještě nebyli započítáni v „matematicích“.
Obecně je řešení našeho příkladu

$$\underline{\underline{100 + 80 - m}}, \quad \text{kde } m \text{ je počet uchazečů s jedničkou z obou předmětů.}$$

Příklad 2.13. Kolik lze sestavit pěticiferných čísel z číslic $1, 2, 3$, takových že na začátku nebo na konci je jednička?

Řešení. Čísel, která začínají jedničkou, je 3^4 (na prvním místě je napevno jednička, na druhém až pátém místě může být kterákoli ze tří možností). Podobně máme 3^4 čísel, která jedničkou končí. Abychom získali počet čísel, která mají na začátku nebo na konci jedničku, nestačí vzít $2 \cdot 3^4$, protože čísla jedničkou začínající i končící by byla započítána dvakrát. Takovýchto čísel je 3^3 (na prvním a pátém místě jedničky, na zbylých třech místech cokoli ze tří možností). Výsledek je tedy

$$3^4 + 3^4 - 3^3 = 81 + 81 - 27 = \underline{\underline{135}}.$$

Jiný způsob řešení Příklad můžeme řešit též tak, že zjistíme, kolik čísel nevyhovuje požadované podmínce, a tento počet pak odečteme od celkového počtu možností. Celkem můžeme z číslic $1, 2, 3$ sestavit 3^5 pěticiferných čísel. Počet čísel, která jedničku na začátku ani na konci nemají, je $2 \cdot 3 \cdot 3 \cdot 3 \cdot 2$, protože první a pátou cifru můžeme vybrat pouze ze dvou možností, zatímco všechny ostatní ze tří. Výsledek je pak

$$3^5 - 2 \cdot 3^3 \cdot 2 = 243 - 108 = 135$$

Druhý předvedený způsob řešení – určení počtu možností, které zadané podmínce nevyhovují, a jeho odečtení od počtu všech možností – je v některých případech výhodnější než přímé hledání počtu vyhovujících možností. Často se používá v příkladech typu „kolik existuje skupin obsahujících aspoň jeden prvek s danou vlastností".
Na závěr si uvedeme příklad, který jsme díky vhodnému přeformulování schopni vyřešit mnohem snadněji.

Příklad 2.14. Kolik lze sestavit trojciferných čísel, takových že číslice na místě jednotek je menší než číslice na místě stovek, a ta je menší než číslice na místě desítek ?

Řešení. Stačí si uvědomit, že když máme k dispozici tři různé číslice (od 0 po 9), tak hledané trojciferné číslo je už jednoznačně dané – nejmenší číslice bude na místě jednotek, prostřední na místě stovek a největší na místě desítek. Proto úlohu můžeme přeformulovat následovně: Kolik je trojprvkových podmnožin množiny $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$? Hledaný počet je $\binom{10}{3} = 120$.

2.2 Příklady pro samostatnou práci

Příklad 2.15. V zábavném testu v novinách je úkolem přiřadit k literárním dílům jejich autory. Celkem je zadáno 10 děl a 10 autorů. Kolika způsoby může být tento test vyplněn, jestliže ke každému dílu nějakého autora přiřadíme?

Výsledek: 10!

Příklad 2.16. U přijímací zkoušky z matematiky je 20 otázek a u každé z nich je 5 možností odpovědi. Kolika způsoby může uchazeč test vyplnit, jestliže
a) u každé otázky zaškrtne právě jednu odpověď?
b) u každé otázky buď zaškrtne jednu odpověď, nebo otázku nechá nezodpovězenou?

Výsledek: a) 5^{20}, b) 6^{20}

Příklad 2.17. Kolika způsoby je možné rozměnit 100 Kč, jestliže jsou k dispozici 50, 20, 10 a 5 - koruny? Bankovek a mincí je neomezený počet.

Výsledek: 49

Příklad 2.18. Finále sportovní soutěže se účastní 12 závodníků. Nikdo není diskvalifikován, na žádném místě nemohou skončit dva nebo více závodníků současně.
a) Kolik je možných výsledků soutěže?
b) Kolik je možných výsledků takových, že loňský vítěz neobhájí své umístění? (Předpokládáme, že loňský vítěz se účastní.)
c) Kolik je možných výsledků takových, že první tři místa budou obsazena stejnými závodníky jako minule, i když třeba v jiném pořadí?

Výsledek: a) $12!$, b) $12! - 11! = 11 \cdot 11!$, c) $3! \cdot 9!$

Příklad 2.19. U stánku prodávají trička v šesti velikostech, čtyřech různých barvách a buď s jedním ze čtyř různých obrázků, nebo bez obrázku. Kolik různých typů triček mohou maximálně mít?

Výsledek: $6 \cdot 4 \cdot 5 = 120$

Příklad 2.20. Kolik existuje čtyřpísmenných slov skládajících se z 26 písmen abecedy (bez háčků a čárek), která obsahují písmeno x?

Výsledek: $26^4 - 25^4$

Příklad 2.21. V mateřské školce je 10 dětí. Když jdou na procházku, mají vytvořit 5 dvojic. Kolika způsoby to můžou udělat? Zajímá nás jen to, kdo je s kým ve dvojici, pořadí dvojic ne.

Výsledek: 945

Příklad 2.22. Do tanečních chodí 20 chlapců a 30 dívek.
a) Kolika způsoby z nich lze vytvořit jeden taneční pár?
b) Kolika způsoby mohou naráz vytvořit 20 tanečních párů?

Výsledek: a) $20 \cdot 30$, b) $30 \cdot 29 \cdots 12 \cdot 11$

Příklad 2.23. Otylka se rozhodla, že bude dvakrát týdně chodit na aerobic.
a) Kolika způsoby si může vybrat dvojici dnů (mimo víkend)?
b) Kolika způsoby si může vybrat dvojici dnů (opět mimo víkend), jestliže nechce chodit dva dny po sobě?

Výsledek: a) $\binom{5}{2} = 10$, b) $\binom{5}{2} - 4 = 6$

Příklad 2.24. Osm přátel, čtyři chlapci a čtyři dívky, si koupilo lístky do kina. Kolika způsoby se mohou rozesadit, jestliže
a) chlapci a dívky chtějí sedět střídavě?
b) všechny dívky chtějí sedět pospolu? (Chlapci v jednom bloku sedět nemusí.)
c) Adam a Eva chtějí sedět vedle sebe?

Výsledek: a) $2 \cdot 4! \cdot 4!$, b) $5 \cdot 4! \cdot 4!$, c) $7 \cdot 2 \cdot 6!$

Příklad 2.25. Kolika způsoby můžeme ze třídy, v níž je 15 chlapců a 17 dívek, vybrat sedmičlennou skupinu tak, aby v ní byli
a) tři chlapci a čtyři dívky,
b) aspoň jeden chlapec,
c) aspoň jeden chlapec a aspoň jedna dívka,
d) aspoň šest chlapců?

Výsledek: a) $\binom{15}{3} \cdot \binom{17}{4}$, b) $\binom{32}{7} - \binom{17}{7}$, c) $\binom{32}{7} - \binom{17}{7} - \binom{15}{7}$, d) $\binom{15}{6} \cdot 17 + \binom{15}{7}$

Příklad 2.26. Kolika způsoby se může 30 studentů rozdělit ke třem učitelům, jestliže
a) každý učitel jich může přijmout právě 10?
b) každý učitel může přijmout neomezené množství?

Výsledek: a) $\binom{30}{10} \cdot \binom{20}{10} \cdot \binom{10}{10} = 30!/(10! \cdot 10! \cdot 10!)$, b) 3^{30}

Příklad 2.27. Pět závodníků, Adam, Bořek, Cyril, Dušan a Emil, běželo 100m. Kolik je takových pořadí, kdy Adam doběhl dříve než Bořek?

Výsledek: 60

3 Klasická a podmíněná pravděpodobnost

3.1 Základní pojmy a řešené příklady

Před uvedením formální definice pravděpodobnostního prostoru začneme několika příklady pro osvětlení problematiky. V každém z nich se jedná o dobře známý náhodný proces s jedním z několika (předem známých) možných výsledků. Naše chápání náhodných jevů

- *Hod mincí* – má 2 možné výsledky hlava / orel, neboli 1/0.
  Každý padá se zhruba stejnou četností (říkáme „s pravděpodobností $\frac{1}{2}$“).

- *Hod kostkou* – má 6 možných výsledků $1, 2, 3, 4, 5, 6$.
  Opět každý padá se stejnou četností (říkáme „s pravděpodobností $\frac{1}{6}$“).

- *Zamíchání karet* – předpokládáme, že každé možné zamíchání karet lze asi stejně dobře očekávat. Zde je však všech možností nesrovnatelně více – $32! \doteq 2{,}6 \cdot 10^{35}$, což vůbec nejsme schopni ani vypsat.

  Tento přirozený příklad zároveň přináší zajímavou filozofickou otázku: Jak můžeme tvrdit, že každé možné pořadí karet nastane stejně pravděpodobně, když se patrně za celou dobu existence lidstva a karet skutečně objeví jen nepatrný zlomek všech zamíchání? Toto lze seriózně rozřešit pouze použitím přesného formálního matematického modelu pravděpodobnosti.

- *Tah sportky* – očekáváme, že každé příští tažené číslo bude „stejně pravděpodobné“ ze všech čísel zbylých v osudí. Celý tah je však ve výsledku neuspořádaným výběrem, takže počet všech možností je $\binom{49}{6} = 13983816$. I to je příliš vysoké číslo pro snadnou představu.

A teď si pro takové příklady uvedeme příslušný matematický aparát:
Označme Ω množinu všech možných výsledků pokusu, který provádíme.
Náhodný jev A je jakákoli podmnožina množiny Ω, $A \subseteq \Omega$.

Klasická pravděpodobnost
Předpokládejme, že pokus má n možných výsledků (množina Ω má n prvků) a že všechny výsledky jsou stejně pravděpodobné. Dále předpokládejme, že z těchto n výsledků jich je m příznivých jevu A (neboli množina A má m prvků, $|A| = m$). Klasická pravděpodobnost jevu A se definuje jako podíl počtu příznivých výsledků ku počtu všech možných výsledků:

$$P(A) = \frac{m}{n}\,.$$

Opačný jev (doplněk) k jevu A: $\overline{A} = \Omega - A$ (obsahuje všechny možné výsledky, které neodpovídají jevu A).

$$P(\overline{A}) = 1 - P(A)$$

Příklad 3.1. Jaká je pravděpodobnost toho, že při hodu klasickou šestistěnnou kostkou padne číslo 6?

Řešení. Všech možností na jedné kostce je 6 a jen jedna z nich je, že padne číslo 6, tedy $P(\text{padne šestka}) = \frac{1}{6}$.

Příklad 3.2. Číslice 1, 2, 3, 4 jsou napsané na čtyřech kartách. Náhodně vybereme dvě karty a naskládáme je vedle sebe v tom pořadí, v jakém jsme je vybrali. Vypočítejte pravděpodobnost toho, že takto vzniklé dvouciferné číslo bude liché.

Řešení. Všech možností je $4 \cdot 3$, protože první kartu vybíráme ze čtyř karet a druhou už jenom ze tří. Liché číslo má na místě jednotek liché číslo, v našem případě to může být 1 nebo 3. Na místo desítek pak můžeme použít v obou případech 2, 4 nebo zbylé liché číslo (21,31,41,13,23,43), příznivých případů je proto 6. Tedy $P(\text{vybereme liché dvouciferné číslo}) = \frac{6}{4\cdot 3} = \frac{6}{12} = \frac{1}{2}$.

Příklad 3.3. Jaká je pravděpodobnost toho, že při jednom hodu klasickou šestistěnnou kostkou padne číslo menší než 6?

Řešení. Jestliže má padnout číslo menší než 6, musí padnout 1, 2, 3, 4 nebo 5. Takže všech možností je 6, příznivých je 5, pravděpodobnost je tedy $P = \frac{5}{6}$. Úlohu jsme mohli vyřešit i jako pravděpodobnost doplňkového jevu, který v našem případě je, že padne šestka. Tento jev má pravděpodobnost $\frac{1}{6}$, tedy pravděpodobnost našeho jevu je $P = 1 - \frac{1}{6} = \frac{5}{6}$.

Příklad 3.4. Jaká je pravděpodobnost toho, že při hodu třemi klasickými šestistěnnými kostkami padne součet menší než 17?

Řešení. Zatímco v předchozím příkladu jsme výhodu doplňkového jevu nezaznamenali, teď je rozhodně nejjednodušší doplňkový jev použít. Doplňkový jev zřejmě je, že součet bude alespoň 17. Víme, že maximální součet je $3 \cdot 6 = 18$. Proto si stačí všimnout jen dvou možných součtů, které jsou navíc velice jednoduché. Součet 18 padne, jen když na všech kostkách padne šestka (jediná možnost 6,6,6) a součet 17 padne, jen když na dvou

kostkách padnou šestky a na jedné pětka (tři možnosti 5,6,6; 6,5,6; 6,6,5). Všech možností na třech kostkách je $6 \cdot 6 \cdot 6 = 216$, počet příznivých možností je 1+3, tedy pravděpodobnost doplňkového jevu je $P' = \frac{4}{216}$. Pravděpodobnost našeho jevu je $P = 1 - P' = 1 - \frac{4}{216} = \frac{212}{216}$.
Nesprávný pokus o řešení: Někdo by možná mohl příklad řešit tímto způsobem: Házíme-li třemi kostkami, může padnout součet 3 (padly tři jedničky), 4 (dvě jedničky a dvojka), ..., 18 (tři šestky). Tj. máme 16 možností, jak může náš „pokus" dopadnout. Z toho správných – menších než 17 – je 14. Pravděpodobnost toho, že padne součet menší než 17, je proto $\frac{14}{16}$. Rozmyslete si, proč je tento postup špatný.

Příklad 3.5. Házíme třemi klasickými šestistěnnými kostkami. Zjistěte, jestli je pravděpodobnější, že padne součet 9, nebo součet 10.

Nesprávný pokus o řešení: Jedna z možností (bohužel nesprávná) může vypadat takto:

$$10 = 6 + 3 + 1 = 6 + 2 + 2 = 5 + 4 + 1 = 5 + 3 + 2 = 4 + 4 + 2 = 4 + 3 + 3$$

a

$$9 = 6 + 2 + 1 = 5 + 3 + 1 = 5 + 2 + 2 = 4 + 4 + 1 = 4 + 3 + 2 = 3 + 3 + 3.$$

Jelikož počty rozkladů jsou u obou čísel stejné, může někdo nabýt dojmu, že pravděpodobnosti uvedených součtů jsou stejné. Všímavý čtenář (obzvlášť jestliže se zamyslel nad předchozím příkladem) ví, že například součet $6 + 3 + 1$ má větší šanci než například součet $4 + 4 + 2$ nebo $3 + 3 + 3$. Proč?

Řešení. Zatímco možnost, že padnou tři trojky, je jediná, máme tři možnosti, že padnou dvě čtyřky a jedna dvojka, šest možností, že padne jedna šestka, jedna trojka a jedna jednička. Když teď spočítáme skutečné počty možností pro jednotlivé součty, zjistíme, že pro součet 10 jich je víc (zkuste si to sami ověřit). Proto pravděpodobnost, že padne součet 10, je vyšší než pravděpodobnost součtu 9.

Příklad 3.6. Jaká je pravděpodobnost, že při hodu třemi klasickými šestistěnnými kostkami padne aspoň jedna šestka?

Řešení. Všech možností na třech kostkách je $6 \cdot 6 \cdot 6 = 216$. Jestliže má padnout aspoň jedna šestka, musíme spočítat pravděpodobnosti toho, že padne jedna, dvě nebo tři šestky, a protože to jsou jevy neslučitelné, výsledná pravděpodobnost bude součtem těchto tří pravděpodobností. Jednodušší však je spočítat pravděpodobnost doplňkového jevu, tedy pravděpodobnost toho, že nepadne ani jedna šestka. To znamená, že na každé kostce máme 5 možností (to, že padne šestka, nemůže nastat), a proto počet příznivých možností je $5 \cdot 5 \cdot 5 = 125$. Pravděpodobnost doplňkového jevu je

$$P(\text{nepadne ani jedna šestka}) = \frac{125}{216},$$

a

$$P(\text{padne aspoň jedna šestka}) = 1 - \frac{125}{216} = \frac{91}{216} \doteq 0{,}421.$$

Příklad 3.7. Ve Sportce hráč tipuje 6 čísel ze 49. Potom je 6 čísel vylosováno. Sázející získává výhru v 1. pořadí, jestliže vyjde všech 6 čísel, která vsadil. Výhru ve 2. pořadí získá, jestliže uhodl 5 čísel, ve 3. pořadí, jestliže uhodl 4 čísla, a nejnižší je výhra ve 4. pořadí pro 3 uhodnutá čísla. Jaká je pravděpodobnost, že sázející
a) získá výhru v 1. pořadí,
b) získá výhru ve 4. pořadí,
c) uhodne aspoň jedno číslo?

Řešení. a) Všech možností, jak mohlo být vylosováno 6 čísel ze 49, je $\binom{49}{6}$. Příznivá je pouze jediná možnost – když byla vylosována přesně ta šestice, kterou hráč vsadil. Proto

$$P = \frac{1}{\binom{49}{6}} = \frac{1}{13\,983\,816} \doteq 7 \cdot 10^{-8} \quad \text{(s výhrou moc nepočítejte...)}$$

b) Počet všech možností zůstává $\binom{49}{6}$. Nyní musíme zjistit, kolik je možností pro vylosované šestice, tak aby se skládaly ze tří vsazených čísel a tří čísel, která hráč nevsadil. Trojici uhodnutých čísel můžeme vybrat $\binom{6}{3}$ způsoby. Ke každému z nich ještě musíme doplnit trojici „špatných“ čísel. Tato čísla vybíráme ze 43 čísel, která hráč nevsadil, a máme pro ně $\binom{43}{3}$ možností. Proto můžeme šestici skládající se ze tří vsazených a tří nevsazených čísel sestavit $\binom{6}{3} \cdot \binom{43}{3}$ způsoby. Celkem pak

$$P = \frac{\binom{6}{3} \cdot \binom{43}{3}}{\binom{49}{6}} \doteq 0{,}018 \quad \text{(taky žádná sláva...)}$$

c) Zde bude nejlepší vypočítat nejprve pravděpodobnost jevu opačného, tj. toho, že neuhodne vůbec žádné číslo. Celá šestice vylosovaných čísel tedy musí být sestavena pouze ze 43 čísel, která hráč nevsadil. Na to máme $\binom{43}{6}$ možností. Pravděpodobnost, že tipující uhodne aspoň jedno číslo, je tedy

$$P = 1 - P(\text{neuhodne nic}) = 1 - \frac{\binom{43}{6}}{\binom{49}{6}} \doteq 1 - 0{,}436 = 0{,}564.$$

Nyní si objasníme další důležité pojmy:

Průnik jevů A, B: $A \cap B$ (obsahuje všechny výsledky, které odpovídají oběma jevům A, B)

Sjednocení jevů A a B: $A \cup B$ (obsahuje všechny výsledky, které odpovídají alespoň jednomu z jevů A, B). Pro pravděpodobnost sjednocení jevů obecně platí

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

Pro jevy, které nemohou nastat současně (**neslučitelné jevy**, jejich průnik je prázdný), platí

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

Oba vztahy plynou ze známé vlastnosti zjednocení a průniku množin

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

Příklad 3.8. Jaká je pravděpodobnost toho, že při hodu klasickou šestistěnnou kostkou padne buď číslo 3, nebo 6?

Řešení. Pravděpodobnost, že padne šestka, je $\frac{1}{6}$ a pravděpodobnost toho, že padne trojka, je také $\frac{1}{6}$. Jelikož tyto dva jevy jsou neslučitelné, můžeme použít součtovou větu a dostaneme výslednou pravděpodobnost: $P = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.

Podmíněná pravděpodobnost
Víme, že určitý jev A nastal, a zkoumáme pravděpodobnost jevu B za této podmínky. Podmíněná pravděpodobnost jevu B za podmínky, že nastal jev A, je

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

Podmíněnou pravděpodobnost ovšem nemusíme vždy počítat podle tohoto vzorce, někdy je lepší přímý výpočet. Často spíše naopak počítáme pravděpodobnost průniku jako

$$P(A \cap B) = P(A) \cdot P(B/A).$$

Zřejmě

$$P(A \cap B) = P(B \cap A) = P(B) \cdot P(A/B),$$

a proto

$$P(A) \cdot P(B/A) = P(B) \cdot P(A/B),$$

což se někdy může hodit.

Nezávislost jevů
Intuitivně: Jevy A, B jsou nezávislé, jestliže to, že nastal jev A, nijak neovlivní pravděpodobnost toho, že nastane jev B, a naopak neboli jestliže

$$P(B/A) = P(B) \quad \text{a} \quad P(A/B) = P(A).$$

Začneme několika jednoduchými ukázkami.

- Nezávislé jevy:
  dva hody toutéž kostkou za sebou,
  hod kostkou a současné zamíchání karet,
  dva různé tahy sportky.

- Závislé jevy:
  vrchní a spodní číslo padlé při jednom hodu kostky,
  výběr první a druhé karty ze zamíchaného balíčku,
  dvě akumulační kurzové sázky obsahující stejné utkání.

Obecně definujeme, že jevy A, B jsou nezávislé, jestliže

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

Příklad 3.9. Hodíme dvěma stejnými kostkami najednou. Jaká je pravděpodobnost, že nám padne 4 a 5?

Řešení. Hody obou kostek jsou nezávislé, neboť v poctivé situaci mezi nimi není fyzická vazba. Pravděpodobnost, že na první kostce padne 4, je $\frac{1}{6}$, a obdobně je $\frac{1}{6}$ pravděpodobnost toho, že na druhé padne 5. Obě možnosti najednou padnou s pravděpodobností po vynásobení $\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$.
Musíme si však dát pozor na to, že kostky jsou stejné a my nemáme určeno, na které padne 4. Proto se úloha vlastně rozpadá na dva disjunktní jevy: „na první padne 4 a na druhé 5“ a „na první 5 a na druhé 4“, a proto výsledná pravděpodobnost je součtem pravděpodobností obou možností $P = \frac{1}{36} + \frac{1}{36} = \frac{1}{18}$.

Příklad 3.10. Vytáhneme-li z karetní hry o 32 kartách (8 karet z každé barvy) postupně dvě karty, jaká je pravděpodobnost, že to budou dva králi?

Řešení. Máme za úkol vypočítat pravděpodobnost jevu $K_1 \cap K_2$, kde jev K_1 znamená, že první vytažená karta je král, a K_2 znamená, že druhá vytažená karta je král.
K řešení nejprve přistoupíme z hlediska klasické pravděpodobnosti:
Všech možností pro postupné vytažení dvou karet je $32 \cdot 31$ a počet možností pro vytažení dvou králů je $4 \cdot 3$. Tedy

$$P(K_1 \cap K_2) = \frac{4 \cdot 3}{32 \cdot 31} = \frac{3}{8 \cdot 31}\,.$$

Nesprávný pokus o řešení Studenti často dělají následující chybu. Řeknou si: „Mám spočítat pravděpodobnost průniku dvou jevů. Na to je přece pěkný vzorec $P(A \cap B) = P(A) \cdot P(B)$!“ Zkusíme to tedy tímto způsobem. Určíme pravděpodobnosti jednotlivých jevů, v našem případě K_1 a K_2. Pro pravděpodobnost K_1 evidentně platí $P(K_1) = \frac{4}{32} = \frac{1}{8}$. Pravděpodobnost K_2 je také $\frac{1}{8}$. (Tohle někdo vidí na první pohled a někdo ne. Těm druhým je určen příklad 3.11.) Je tedy

$$P(K_1 \cap K_2) = \frac{1}{8} \cdot \frac{1}{8} = \frac{1}{64} \text{ ???}$$

Už jsme viděli, že ne. Problém je v tom, že jevy K_1 a K_2 **nejsou nezávislé**. Pravděpodobnost toho, že jako druhý bude vytažen král, je ovlivněna tím, jaká karta byla vytažena jako první. A proto $P(K_1 \cap K_2)$ jako $P(K_1) \cdot P(K_2)$ počítat nemůžeme.

Správné řešení pomocí součinu Násobit přímo pravděpodobnosti jevů K_1 a K_2 jsme nemohli. Můžeme však použít vztah

$$P(K_1 \cap K_2) = P(K_1) \cdot P(K_2/K_1).$$

$P(K_1)$ už známe, zbývá určit $P(K_2/K_1)$, tj. pravděpodobnost toho, že jako druhý bude vytažen král, víme-li, že první karta byla král. Po vytažení první karty nám v balíčku zůstalo $32 - 1$ karet. Protože první karta byla král, počet králů se také snížil o jednoho. Počet příznivých možností je proto $4 - 1$. Tedy dostáváme

$$P(K_2/K_1) = \frac{3}{31}\,.$$

Dohromady pak máme

$$P(K_1 \cap K_2) = P(K_1) \cdot P(K_2/K_1) = \frac{4}{32} \cdot \frac{3}{31} = \frac{3}{8 \cdot 31}\,.$$

Vidíme, že jsme došli ke stejnému výsledku jako předtím, když jsme uvažovali počty možností pro celé dvojice karet.

V následujícím příkladu se mimo jiné vrátíme k výpočtu $P(K_2)$. Mohli by si jej přečíst i ti, kterým bylo hned jasné, že je to $\frac{1}{8}$.

Příklad 3.11. Z karetní hry o 32 kartách (8 karet z každé barvy) vytáhneme postupně dvě karty.
a) Jaká je pravděpodobnost, že druhá karta bude král?
b) Jako druhou kartu jsme vytáhli krále. Jaká je pravděpodobnost, že král byl vytažen i jako první karta?

Řešení. a) Mohly nastat dva případy: první karta král buď byla (nastal jev K_1), nebo nebyla (nastal jev $\overline{K_1}$). Pravděpodobnost K_2 tedy můžeme vypočítat jako součet pravděpodobností jevů $K_1 \cap K_2$ (první i druhá karta byla král) a $\overline{K_1} \cap K_2$ (první karta král nebyla a druhá byla). Pravděpodobnosti těchto průniků vypočteme podobně jako v předchozím příkladu. Celkem máme

$$\begin{aligned} P(K_2) &= P(K_1 \cap K_2) + P(\overline{K_1} \cap K_2) = P(K_1) \cdot P(K_2/K_1) + P(\overline{K_1}) \cdot P(K_2/\overline{K_1}) = \\ &= \frac{4}{32} \cdot \frac{3}{31} + \frac{28}{32} \cdot \frac{4}{31} = \frac{4}{32}\left(\frac{3}{31} + \frac{28}{31}\right) = \frac{1}{8}\,. \end{aligned}$$

b) Zde se po nás chce pravděpodobnost toho, že jako první byl vytažen král za podmínky, že jako druhý byl vytažen král, neboli $P(K_1/K_2)$. Výpočet lze provést dvěma způsoby. Můžeme například čistě dosadit do vzorce pro podmíněnou pravděpodobnost:

$$P(K_1/K_2) = \frac{P(K_1 \cap K_2)}{P(K_2)} = \frac{\frac{3}{8 \cdot 31}}{\frac{1}{8}} = \frac{3}{31}\,.$$

Ve druhém způsobu použijeme klasickou pravděpodobnost. Víme, že druhý byl vytažen král. Počet všech možných dvojic karet, kde král je na druhém místě, dostaneme jako součet počtů dvojic, kde král je na obou místech, a dvojic, kde první karta je jiná než král. Celkem to je $4 \cdot 3 + 28 \cdot 4$ dvojic. Příznivé možnosti jsou ty, kde král je na prvním místě, a takovýchto možností je $4 \cdot 3$. Dohromady pak

$$P(K_1/K_2) = \frac{4 \cdot 3}{4 \cdot 3 + 28 \cdot 4} = \frac{3}{3 + 28} = \frac{3}{31}\,.$$

V části a) tohoto příkladu jsme situaci rozdělili na dva možné případy. Pro každý z nich jsme pravděpodobnost spočítali zvlášť a příslušné pravděpodobnosti jsme pak sečetli. Někdy je možných výchozích situací víc než dvě. V takovém případě se používá věta o úplné pravděpodobnosti. V části b) jsme počítali pravděpodobnost jedné z výchozích možností, když už jsme znali celkový výsledek pokusu. Obecně se v takovémto případě může použít Bayesův vzorec.

Věta o úplné pravděpodobnosti
Vezměme systém navzájem disjunktních podmnožin (tzv. hypotéz) $H_1, \ldots, H_k$ množiny Ω, takový že $H_1 \cup \cdots \cup H_k = \Omega$. Pak

$$P(A) = P(H_1) \cdot P(A/H_1) + \cdots + P(H_k) \cdot P(A/H_k).$$

Bayesův vzorec

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{P(A)} = \frac{P(H_i) \cdot P(A/H_i)}{P(H_1) \cdot P(A/H_1) + \cdots + P(H_k) \cdot P(A/H_k)}\,.$$

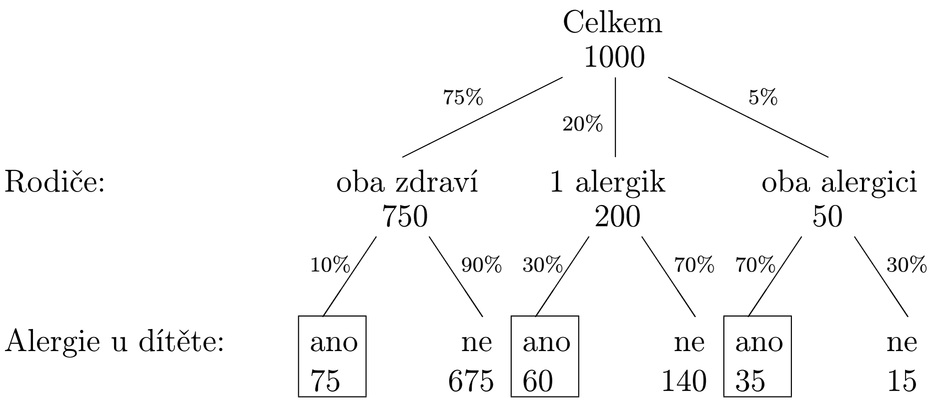
Příklad 3.12. Pravděpodobnost, že dítě bude trpět určitou alergií, je 0,7, jsou-li oba jeho rodiče alergici, 0,3, je-li jen jeden z rodičů alergik, a 0,1, jestliže žádný z rodičů alergií netrpí. Mezi rodiči dětí, které nyní zkoumáme, je 75 % párů, kde alergií netrpí nikdo, 20 % párů, kde má alergii jeden, a 5 % párů, v nichž mají alergii oba.
a) Jaká je pravděpodobnost, že náhodně vybrané dítě bude mít alergii?
b) Je-li dítě alergické, jaká je pravděpodobnost, že ani jeden z jeho rodičů alergií netrpí?

Řešení. Toto je typická ukázka příkladu na úplnou pravděpodobnost (část a)) a Bayesův vzorec (část b)). Než však začneme dosazovat do nějakých vzorců, pokusme se příklad vyřešit „selským rozumem“.
a) Zde by nám stačilo zjistit, jaká část (kolik procent) všech zkoumaných dětí má alergii. Představme si, že bylo vyšetřeno celkem 1000 dětí. Podle zadání z nich 750 mělo oba rodiče zdravé, 200 mělo alergického jednoho z rodičů a 50 mělo oba rodiče alergiky. V zadání bylo řečeno, že například pravděpodobnost, že dítě zdravých rodičů bude mít alergii, je 0,1. To znamená, že můžeme očekávat, že ze 750 dětí se zdravými rodiči bude mít alergii 10 %, tj. 75. Podobně můžeme čekat, že 30 % z 200 dětí s jedním alergickým rodičem bude mít

alergii. To dělá 60 dětí. Ještě zbývá 70 % z 50 dětí s oběma rodiči alergickými, tj. 35 dětí. Celkem tedy můžeme očekávat, že z 1000 dětí jich alergii bude mít 75+60+35 = 170. Když teď k problému přistoupíme z hlediska klasické pravděpodobnosti, tak pravděpodobnost, že náhodně vybereme dítě, které má alergii, bude

$$P = \frac{170}{1000} = 0{,}17.$$

Je zřejmé, že kdybychom si zvolili jiný výchozí počet dětí než 1000, dospěli bychom k témuž výsledku.
Celou předcházející úvahu si můžeme i graficky znázornit pomocí stromu:



Nyní se vraťme k tomu, jak jsme vlastně ke 170 alergickým dětem dospěli. Brali jsme 10 % ze sedmdesáti pěti procent celku, 30 % z dvaceti procent celku a 70 % z pěti procent celku. Dohromady jsme tedy měli

$$0{,}75 \cdot 0{,}1 + 0{,}2 \cdot 0{,}3 + 0{,}05 \cdot 0{,}7 = 0{,}17, \text{ tj. } 17\%.$$

b) Když už jsme zdolali část a), je řešení snadné. Zůstaňme nejprve u konkrétního počtu 1000 dětí. Již jsme zjistili, že se dá očekávat, že z tohoto počtu bude alergických dětí 170. Nyní víme, že bylo vybráno dítě alergické. Muselo to tedy být některé z oněch 170 – zúžil se nám počet všech možností, které jsme měli na výběr. Ptáme se, jaká je pravděpodobnost, že rodiče vybraného dítěte jsou oba zdraví. Alergických dětí se zdravými rodiči je 75 – viz strom. Pravděpodobnost, že oba rodiče vybraného alergického dítěte jsou zdraví, je tedy

$$P = \frac{75}{170} \doteq 0{,}441.$$

Kdybychom nechtěli brát v úvahu nějaký konkrétní počet, můžeme problém řešit takto: Víme, že z celku (ať už jakéhokoli) je 17 % dětí alergických. Potřebujeme zjistit, jakou část (kolik procent) z tohoto nového celku tvoří alergické děti se zdravými rodiči. Takovýchto dětí je 10 % ze sedmdesáti pěti procent původního celku neboli 7,5 % původního celku. Výsledek je tedy opět

$$P = \frac{0{,}75 \cdot 0{,}1}{0{,}17} = \frac{0{,}075}{0{,}17} \doteq 0{,}441.$$

Řešení pomocí vzorců

a) Otázka byla, jaká je pravděpodobnost, že náhodně vybrané dítě bude alergické. Když se nad tím člověk zamyslí, řekne si nejspíš něco ve smyslu: „To přece záleží na tom, jaké má rodiče!“ Máme několik možností, které mohly nastat, a při každé z nich je pravděpodobnost zkoumaného jevu jiná. Toto je klasická situace pro použití vzorce pro úplnou pravděpodobnost. Možné výchozí situace bereme jako jednotlivé hypotézy $H_1, H_2, \ldots$. V našem případě to bude, zda jsou oba rodiče zdraví, jeden alergický, nebo oba alergičtí. Nyní si všechny zkoumané jevy označíme a sepíšeme jejich pravděpodobnosti.

A	...	náhodně vybrané dítě je alergické		
H_1	...	rodiče jsou oba zdraví	$P(H_1) = 0{,}75$	$P(A/H_1) = 0{,}1$
H_2	...	jeden z rodičů je alergik	$P(H_2) = 0{,}2$	$P(A/H_2) = 0{,}3$
H_3	...	oba rodiče jsou alergici	$P(H_3) = 0{,}05$	$P(A/H_3) = 0{,}7$

Dosadíme-li nyní do vzorce pro úplnou pravděpodobnost, dostáváme

$$\begin{aligned} P(A) &= P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + P(H_3) \cdot P(A/H_3) = \\ &= 0{,}75 \cdot 0{,}1 + 0{,}2 \cdot 0{,}3 + 0{,}05 \cdot 0{,}7 = 0{,}17. \end{aligned}$$

Porovnejte s předchozím („selským“) řešením!

b) V tomto případě si člověk možná řekne: „To je divná otázka, jaksi pozpátku... Jó, kdyby to bylo naopak – když má dítě oba rodiče zdravé, tak jaká je pravděpodobnost, že bude alergické – tak to bych věděl!“ Tohle je typická situace pro použití Bayesova vzorce. Známe už výsledek pokusu a chceme zjistit pravděpodobnost jedné z možných výchozích situací. Při našem označení jevů máme vypočítat pravděpodobnost hypotézy H_1 za podmínky, že nastal jev A. Použitím Bayesova vzorce dostáváme

$$P(H_1/A) = \frac{P(H_1) \cdot P(A/H_1)}{P(A)} = \frac{0{,}75 \cdot 0{,}1}{0{,}17} \doteq 0{,}441.$$

Opět porovnejte se „selským“ řešením!

Příklad 3.13. Zuzanka má dva pytlíky s kuličkami. V prvním je pět bílých a šest zelených kuliček a ve druhém jsou čtyři bílé a sedm zelených kuliček. Označme U_1−výběr kuliček z prvního pytlíku, U_2−výběr ze druhého pytlíku, B−výběr bílé kuličky, Z−výběr zelené kuličky. Víme, že výběr z prvního a druhého pytlíku je stejně pravděpodobný. Vypočítejte $P(B/U_1), P(B/U_2), P(Z/U_1), P(Z/U_2), P(B \cap U_1), P(B \cap U_2), P(B)$.

Řešení. Výpočet prvních čtyř pravděpodobností je jednoduchý. Je třeba si uvědomit, co vlastně například $P(B/U_1)$ znamená. Je to pravděpodobnost toho, že za předpokladu výběru z prvního pytlíku vytáhneme bílou kuličku. Představme si to tak, že už máme ruku v prvním pytlíku a ptáme se jen na to, jaká je pravděpodobnost, že vytáhneme bílou kuličku. V prvním pytlíku je 11 kuliček, z toho pět bílých, a tedy

$$P(B/U_1) = \frac{5}{11}.$$

Podobně spočítáme, že

$$P(B/U_2) = \frac{4}{11},$$

$$P(Z/U_1) = \frac{6}{11} = 1 - P(B/U_1),$$
$$P(Z/U_2) = \frac{7}{11} = 1 - P(B/U_2).$$

Další jev $(B \cap U_1)$ znamená, že vybírám bílou kuličku z prvního pytlíku, tedy že si vyberu správný pytlík (první) a pak z něj ještě musím vytáhnout bílou kuličku. Protože výběr z prvního a z druhého pytlíku je stejně pravdepodobný a tyto dva jevy jsou neslučitelné, je pravděpodobnost toho, že vyberu kuličku z prvního pytlíku $P(U_1) = \frac{1}{2}$. A teď, když už s pravděpodobností $\frac{1}{2}$ je naše ruka v prvním pytlíku, stačí vytáhnout bílou kuličku, a to lze s pravděpodobností $\frac{5}{11}$. Proto

$$P(B \cap U_1) = \frac{1}{2} \cdot \frac{5}{11} = \frac{5}{22} \doteq 0{,}227$$

a podobně

$$P(B \cap U_2) = \frac{1}{2} \cdot \frac{4}{11} = \frac{4}{22} \doteq 0{,}182.$$

Všimněte si, že jsme vlastně použili vzorec pro výpočet průniku dvou jevů pomocí podmíněné pravděpodobnosti. Zkuste si rozmyslet, jestli se ke druhému výsledku nedalo dojít jiným způsobem. Zůstává spočítat poslední pravděpodobnost. K tomu si stačí uvědomit, že bílou kuličku můžeme vybrat z prvního nebo druhého pytlíku a tyto jevy jsou neslučitelné, proto

$$P(B) = P(B \cap U_1) + P(B \cap U_2) = \frac{4+5}{22} \doteq 0{,}409.$$

Na závěr se ještě vraťme k poslednímu výsledku. Když sečteme bílé kuličky z obou pytlíků, dostaneme číslo 9, a počet všech kuliček je 22. Častý dotaz studentů je, proč jsme tu pravděpodobnost nespočítali hned jako podíl počtu bílých kuliček ku počtu všech kuliček. Proč? Z jediného důvodu, není to obecný postup. V tomto konkrétním příkladu je to možné a korektní řešení. Zkuste si ale vypočítat tuto úlohu s jedinou úpravou – předpokládejte, že pravděpodobnost výběru z prvního pytlíku je například dvakrát větší než výběr z druhého pytlíku. Pak se ještě zamyslete nad tím, kdy si výpočet poslední pravděpodobnosti můžeme zjednodušit a kdy to nesmíme udělat.

Příklad 3.14. Studenti jsou rozděleni do tří stejně početných skupin. Jistého testu se zúčastnilo 40 % studentů z první skupiny, 30 % studentů ze druhé skupiny a 30 % studentů ze třetí skupiny. Víme, že studenti z první skupiny mívají na testech úspěšnost 60 %, studenti ze druhé skupiny mívají úspěšnost 35 % a ze třetí skupiny 70 %. Učitel si náhodně vybere jeden test a začne jej opravovat.
a) Jaká je pravděpodobnost, že opravuje úspěšný test?
b) Opravuje úspěšný test. Jaká je pravděpodobnost, že ho psal student z první skupiny?

Řešení. (Upozorňujeme, že trochu lehkomyslné – pokuste se najít ono podezřelé místo. Pozorně si pak přečtěte i další příklad.)
a) Úspěšný test mohl napsat student z první, druhé nebo třetí skupiny a tyto tři jevy jsou neslučitelné. Pak si stačí uvědomit, že pravděpodobnost, že opravujeme úspěšný test, který psal student z první skupiny, je $P(U \cap S_1) = 0{,}4 \cdot 0{,}6$. Podobně spočítáme další

dvě pravděpodobnosti ($P(U \cap S_2), P(U \cap S_3)$). Pravděpodobnost $P(U)$ napsání úspěšného testu je tedy

$$P(U) = 0{,}4 \cdot 0{,}6 + 0{,}3 \cdot 0{,}35 + 0{,}3 \cdot 0{,}7 = 0{,}555.$$

b) Potřebujeme vypočítat pravděpodobnost toho, že test psal student z první skupiny, za přepokladu, že vybraný test je úspěšný. Můžeme použít Bayesův vzorec a využijeme výsledek části a).

$$P(S_1/U) = \frac{P(S_1 \cap U)}{P(U)},$$

$$P(S_1/U) = \frac{0{,}4 \cdot 0{,}6}{0{,}555} \doteq 0{,}4332.$$

Následující příklad je na první pohled stejný, ale pozor, je mírně zákeřný.

Příklad 3.15. Tentokrát se testu zúčastnilo 60 % studentů z první skupiny, 50 % studentů ze druhé skupiny a 70 % studentů ze třetí skupiny. Víme, že studenti z první skupiny mívají na testech úspěšnost 60 %, studenti ze druhé skupiny mívají úspěšnost 35 % a ze třetí skupiny 70 %. Učitel si opět náhodně vybere jeden test.
a) Jaká je pravděpodobnost, že opravuje úspěšný test?
b) Test je úspěšný. Jaká je pravděpodobnost, že ho psal student z první skupiny?

Pokus o řešení (bohužel nevydařený)
a) Podobně jako v předchozím příkladu spočítáme pravděpodobnost úspěšného testu:

$$P(U) = 0{,}6 \cdot 0{,}6 + 0{,}5 \cdot 0{,}35 + 0{,}7.0{,}7 = 1{,}025.$$

Výsledek nemůže být správný, možné hodnoty pravděpodobnosti jsou vždy z intervalu $\langle 0, 1 \rangle$. Kde jsme udělali chybu? Je vůbec předešlý příklad správně vyřešen? Jaký je rozdíl mezi tímto a předešlým příkladem?
A konečně správné řešení Všimněte si počtů procent zúčastněných studentů z jednotlivých skupin. Jejich součet v prvním příkladu je 100 %, ve druhém je součet 180 %. (Pozor! V zadání chyba není.) Chyba nastala v řešení druhého příkladu při určení pravděpodobností $P(U \cap S_1), P(U \cap S_2), P(U \cap S_3)$, přesněji v určení pravděpodobností $P(S_1), P(S_2), P(S_3)$. Už v prvním příkladu jsme tyto pravděpodobnosti určili poněkud unáhleně, ovšem tam jsme měli štěstí. Nyní jsme například za pravděpodobnost jevu, že student je z první skupiny, brali číslo 0,6, a to byla chyba. Totiž součet pravděpodobností neslučitelných jevů (student může patřit jenom do jedné ze tří skupin) musí být 1. V našem nesprávném řešení je tento součet $0{,}6 + 0{,}5 + 0{,}7 = 1{,}8 > 1$. Proto potřebujeme tyto pravděpodobnosti znormovat, upravit je tak, aby poměr zůstal stejný, ale součet byl roven 1. Proto $P(S_1) = \frac{0{,}6}{1{,}8}, P(S_2) = \frac{0{,}5}{1{,}8}, P(S_3) = \frac{0{,}7}{1{,}8}$. Takže jaký je rozdíl mezi prvním a druhým příkladem už víme a zřejmě je už jasné i to, že první příklad jsme řešili správně (i když možná víceméně náhodou). Oba příklady jsme uvedli záměrně, aby si studenti uvědomili, že není možné bez přemýšlení použít vzoreček. A teď už to správné řešení:
a)

$$P(U) = \frac{0,6}{1,8} \cdot 0{,}6 + \frac{0,5}{1,8} \cdot 0{,}35 + \frac{0,7}{1,8} \cdot 0{,}7 \doteq 0{,}569.$$

b)

$$P(S_1/U) = \frac{P(S_1 \cap U)}{P(U)},$$

$$P(S_1/U) = \frac{\frac{0{,}6}{1{,}8} \cdot 0{,}6}{P(U)} \doteq 0{,}351.$$

Vyzkoušejte si část b) spočítat s nesprávnými hodnotami $P(U \cap S_1), P(U \cap S_2), P(U \cap S_3)$ a výsledek porovnejte s naším. Porovnání je moc zajímavé, tak si to zkuste!

3.2 Příklady pro samostatnou práci

Přímý výpočet pravděpodobnosti

Příklad 3.16. Hloupý Honza si má vybrat princeznu. Předvedli mu 6 zamaskovaných ježibab, 3 zamaskované komorné a 2 princezny.
Jaká je pravděpodobnost, že Honza vybere princeznu?

Výsledek: $\frac{2}{11}$

Příklad 3.17. Házíme třikrát šestistěnnou kostkou. Jaká je pravděpodobnost, že součet bude
a) sudé číslo,
b) prvočíslo,
c) číslo dělitelné 3,
d) osm,
e) větší než 12?
Výsledek: a) 0.5, b) $\frac{73}{216}$, c) $\frac{1}{3}$, d) $\frac{21}{216}$, e)$\frac{7}{27}$.

Příklad 3.18. Házíme třemi šestistěnnými kostkami. Který ze součtů je pravděpodobnější, 11, nebo 12?
Výsledek: Pravděpodobnější je součet 11.

Příklad 3.19. Házíme třikrát šestistěnnou kostkou. Jaká je pravděpodobnost, že součet bude 9, když víme, že na při prvním hodu padla dvojka?
Výsledek: $\frac{1}{6}$

Příklad 3.20. Zuzanka si hraje s písmenky A, A, A, E, I, K, M, M, T, T. Jaká je pravděpodobnost, že při náhodném uspořádaní písmen dostane slovo MATEMATIKA?
Výsledek: $\frac{1}{151\,200}$

Příklad 3.21. Číslice 1, 2, 3, 4, 5 jsou napsané na pěti kartách. Náhodně vybereme tři karty a naskládáme je vedle sebe v tom pořadí, v jakém jsme je vybrali. Vypočítejte pravděpodobnost toho, že takto vzniklé trojciferné číslo bude sudé.
Výsledek: 0,4

Příklad 3.22. Ze 32 hracích karet vybíráme dvakrát po sobě po jedné kartě. Zjistěte
a) jaká je pravděpodobnost, že obě karty jsou esa, když jsme první kartu nevrátili,
b) jaká je pravděpodobnost, že obě karty jsou esa, když jsme první kartu vrátili,
c) jaká je pravděpodobnost, že obě karty jsou stejné barvy, když jsme první kartu nevrátili,
d) jaká je pravděpodobnost, že obě karty jsou stejné barvy, když jsme první kartu vrátili.
Výsledek: a) přibližně 0,012, b) přibližně 0,015625, c) přibližně 0,2258, d) přibližně 0,25.

Příklad 3.23. Ze 32 hracích karet (z toho 4 jsou esa) vybíráme postupně po jedné kartě, žádnou kartu nevrátíme zpátky. Zjistěte, jaká je pravděpodobnost, že
a) pátá karta bude eso,
b) první eso bude vytaženo jako pátá karta v pořadí.
Výsledek: a) $\frac{1}{8} = 0{,}125$, b) přibližně 0,081

Příklad 3.24. Krychli, která má nabarvené všechny stěny, rozřežeme na 1000 malých krychlí stejných rozměrů. Všechny pomícháme a náhodně vybereme jednu malou krychli. Vypočítejte, jaká je pravděpodobnost, že krychle bude mít
a) jednu zabarvenou stěnu,
b) dvě zabarvené stěny,
c) tři zabarvené stěny,
d) všechny stěny nezabarvené.

Výsledek: a) přibližně 0,384, b) přibližně 0,096, c) přibližně 0,008, d) přibližně 0,512.

Příklad 3.25. Vypočtěte pravděpodobnost, že ve Sportce sázející
a) vyhraje 2. cenu
b) vyhraje, ale nanejvýš 3. cenu

Výsledek: a) $\frac{\binom{6}{5}\cdot\binom{43}{1}}{\binom{49}{6}} \doteq 1{,}84 \cdot 10^{-5}$, b) $\frac{\binom{6}{4}\cdot\binom{43}{2}}{\binom{49}{6}} + \frac{\binom{6}{3}\cdot\binom{43}{3}}{\binom{49}{6}} \doteq 1{,}86 \cdot 10^{-2}$

Příklad 3.26. Jaký počet uhodnutých čísel ve Sportce je nejpravděpodobnější?
Výsledek: Intuice napovídá a výpočet to potvrzuje, že 0 – příslušná pravděpodobnost je přibližně 0,436, pro jedno číslo 0,413, ostatní už jsou daleko menší.

Příklad 3.27. 40 procent myší se stává minutu po podání léku vysoce agresivními. Určete pravděpodobnost toho, že pokud 15 myším bude podán lék, tak minutu po podání budou agresivní nanejvýš 3 z nich.
Výsledek: přibližně 0,0905019.

Příklad 3.28. Test sestává z 8 otázek, u každé z nich je z variant a,b,c správná právě jedna. Jeden student se na test nepřipravil, místo toho háže kostkou u každé otázky. Pokud padne 1 nebo 2, zatrhne variantu a, pokud padne 3 nebo 4, variantu b, pokud 5 nebo 6, variantu c. Určete pravděpodobnost, že bude po opravení testu mít aspoň šest otázek správně.
Výsledek: $\frac{129}{3^8} \doteq 0{,}01966$

Příklad 3.29. Vykonal se experiment, který spočíval v křížení bílého a fialového hrachu, přičemž se předpokládalo, že pokusné rostliny kříženy dosud nebyly. Podle pravidel dědičnosti lze očekávat, že $\frac{3}{4}$ nových potomků rozkvetou fialově a $\frac{1}{4}$ bíle. Vzklíčilo 10 rostlin. Určete pravděpodobnost toho, že
a) ani jedna rostlina nerozkvete bíle,
b) fialově rozkvetou alespoň tři,
c) fialově rozkvete alespoň 6, ale nanejvýš 8 rostlin.
Výsledek: a) přibližně 0,0563, b) přibližně 0,9996, c) přibližně 0,6779.

Příklad 3.30. Dva sportovní střelci střílejí nezávisle na sobě na stejný cíl, každý jednu střelu. Pravděpodobnost, že cíl zasáhne první střelec, je 0,9 a druhý 0,8. Vypočítejte pravděpodobnost, že cíl nezasáhne ani jeden střelec.
Výsledek: 0,002.

Příklad 3.31. Mezi 10 kvalitních výrobků bylo omylem přimícháno 5 zmetků (máme 15 výrobků). Náhodně vybereme tři výrobky. Určete pravděpodobnost toho, že
a) všechny budou kvalitní,
b) právě jeden bude chybný,
c) minimálně jeden bude chybný.

Výsledek: a) přibližně 0,264, b) přibližně 0,538, c) přibližně 0,736.

Příklad 3.32. Jaká je pravděpodobnost toho, že tři náhodně vybraní lidé nemají narozeniny ve stejný den i měsíc (29. únor neuvažujeme).
Výsledek: přibližně 0,9917.

Příklad 3.33. V pytlíku máme 8 kuliček – 3 červené a 5 modrých. Náhodně vybereme dvě kuličky (naráz). Co je pravděpodobnější: že vybereme dvě modré, modrou a červenou, nebo dvě červené kuličky?
Výsledek: Nejpravděpodobnější je výběr modré a červené kuličky.

Příklad 3.34. Ke zkoušce je zadáno 30 otázek. Student si vylosuje dvě. Zkoušku udělá, jestliže obě dvě umí. Lojza se chystá na zkoušku, ale nehodlá se předřít.
a) Kolik otázek se musí naučit, aby měl nadpoloviční šanci, že u zkoušky uspěje?
b) Kolik by se musel naučit, aby pravděpodobnost úspěchu nebyla menší než 90%?
Výsledek: a) 22 b) 29

Příklad 3.35. Ke zkoušce je zadáno 30 otázek, po deseti z každého ze tří okruhů, z nichž se předmět skládal. Student si vylosuje z každého okruhu jednu otázku. Zkoušku udělá, jestliže umí všechny tři.
Kolik otázek se musí student naučit z každého okruhu, aby měl nadpoloviční šanci, že u zkoušky uspěje?
Výsledek: 8

Podmíněná pravděpodobnost, nezávislost jevů

Příklad 3.36. Házíme klasickou šestistěnnou kostkou.
a) Víme, že padlo sudé číslo. Jaká je pravděpodobnost, že padla dvojka?
b) Víme, že padla dvojka. Jaká je pravděpodobnost, že padlo sudé číslo?
Výsledek: a) $\frac{1}{3}$ b) 1

Příklad 3.37. Jistý předmět učí tři profesoři: Hodný, Zlý a Střední. Profesor Hodný učí stejný počet studentů jako profesor Střední. Profesor Zlý učí poloviční počet studentů než profesor Hodný. Profesor Hodný vyhazuje 20 % studentů, profesor Střední 30 % a profesor Zlý 60 %.
a) Jaká je pravděpodobnost, že náhodně vybraný student neudělá zkoušku?
b) Jestliže byl student od zkoušky vyhozen, jaká je pravděpodobnost, že ji dělal u profesora Zlého?
Výsledek: a) 0,32 b) 0,375

Příklad 3.38. Přibližně jedno z 10000 dětí se narodí s určitou vrozenou vadou (která se projeví až později). Na včasné odhalení této vady lékaři vyvinuli test, který správně funguje s pravděpodobností 0,99, pokud dítě tuto vadu opravdu má. Jestliže je dítě zdravé, je zde pravděpodobnost 0,05, že bude chybně prohlášeno za nemocné.

a) Jaká je pravděpodobnost, že vyšetřované dítě bude označeno jako nemocné?
b) Dítě bylo označeno za nemocné. Jaká je pravděpodobnost, že nemocné opravdu je?
Výsledek: a) 0,050094 b) $0{,}0001 \cdot 0{,}99/0{,}050094 \doteq 0{,}002$

Příklad 3.39. Pravděpodobnost, že se tenistovi podaří první podání, je 0,6. U druhého podání má pst úspěchu 0,9. (Druhé podání dává jedině v případě, že se mu první nepodařilo.)
a) Tento „pokus“ má tři možné výsledky: tenista může uspět buď při prvním podání, nebo při druhém podání, nebo udělá dvojchybu. Vypočtěte pravděpodobnosti jednotlivých možností.
b) Jaká je pravděpodobnost, že se mu podání podaří?
c) Podání se podařilo. Jaká je pravděpodobnost, že to bylo až na druhý pokus?
Výsledek: a) 1. podání: 0,6; 2. podání: 0,36; dvojchyba: 0,04 b) 0,96 c) $0{,}36/0{,}96 = 0{,}375$

Příklad 3.40. Studenti jsou rozděleni do tří skupin, přičemž první a druhá skupina jsou stejně početné, zatímco třetí skupina je dvojnásobné velikosti než první. Jistého testu se zúčastnilo 40 % studentů z první skupiny, 25% studentů ze druhé skupiny a 30 % studentů ze třetí skupiny. Víme, že studenti z první skupiny mívají na testech úspěšnost 60 %, studenti ze druhé skupiny mívají úspěšnost 35 % a ze třetí skupiny 70 %. Učitel si náhodně vybere jeden test a začne jej opravovat.
a) Jaká je pravděpodobnost, že opravuje úspěšný test?
b) Opravuje úspěšný test. Jaká je pravděpodobnost, že ho psal student ze druhé skupiny?
Výsledek: a) 0,598 b) přibližně 0,117

Příklad 3.41. V jedné dílně pracují Ruda Rychlý, Pepa Pečlivý a Lojza Lajdák. Ruda za den vyrobí dvakrát víc výrobků než Pepa a Pepa třikrát víc výrobků než Lojza. Z toho, co vyrobí Ruda, je 10 % zmetků, mezi Pepovými výrobky je 5 % zmetků a z Lojzových výrobků je zmetků polovina.
Jestliže náhodně vybraný výrobek z této dílny je zmetek, kdo je jeho nejpravděpodobnějším autorem? A kdo nejméně pravděpodobným?

Výsledek: Pravděpodobnosti autorství zmetku pro Rudu, Pepu a Lojzu jsou po řadě 0,48, 0,12 a 0,4. Nejpravděpodobnějším výrobcem zmetku je Ruda, nejméně pravděpodobným Pepa.

Příklad 3.42. Lovec vystřelí na medvěda. Pravděpodobnost zásahu je 0,4. Jestliže medvěda netrefí, rozzuřené zvíře na lovce zaútočí. Lovec vystřelí znovu, tentokrát má pravděpodobnost zásahu p. Jestliže medvěda nezasáhne ani tentokrát, medvěd jej sežere.
Určete hodnotu p tak, aby šance lovce a medvěda byly vyrovnané, tj. aby pravděpodobnosti, že medvěd bude zastřelen a že lovec bude sežrán, byly obě rovny 0,5.

Výsledek: $p = 1/6$ – vyjde jako řešení rovnice $0{,}4 + 0{,}6 \cdot p = 0{,}5$ nebo $0{,}6 \cdot (1 - p) = 0{,}5$.

Příklad 3.43. Učitel vybral náhodně test ze skupiny 5 těžkých a 2 lehkých testů. Poradí se s kolegou–expertem, který na první pohled rozezná těžký test s pravděpodobnosti $\frac{3}{4}$ a s pravděpodobnosti $\frac{1}{4}$ označí i lehký test za těžký. Určete pravděpodobnost, že kolega označí vybraný test za těžký.
Výsledek: $\frac{17}{28}$.

Příklad 3.44. Zuzanka má panenku Chou Chou. K této panence patří lahvička, na kterou ona reaguje. Baruška má panenku Baby Born a ta má stejnou lahvičku, na kterou ale Zuzančina panenka nereaguje. Holčičky si lahvičky pomíchaly a každá tu svou dokáže určit s pravděpodobností $\frac{1}{2}$. Zuzanka má v ruce lahvičku, o které tvrdí, že je její. S jakou pravděpodobností bude Chou Chou na ni reagovat?
Výsledek: 0,5.

Příklad 3.45. Zákazník si vybírá obraz ze skupiny 10 originálů a 2 kopií. Radí se s expertem, který rozezná originál od kopie s pravděpodobností $\frac{5}{6}$. Odborník právě tvrdí, že obraz je originál. Jaká je pravěpodobnost, že obraz je opravdu originál?
Výsledek: přibližně 0,96.

Příklad 3.46. V tramvaji si povídají dva studenti o tom, kdo byl na zkoušce. „Dnes bylo 8 z kombinace M-F a 4 z kombinace M-CH.“ „Kolik jich to udělalo?“ „Pět z M-F a tři z M-CH.“ „A co Eva?“ „Eva udělala.“ Student matematiky si to vyslechl, ale Evu nezná. Začal řešit, s jakou pravděpodobností je Eva z kombinace M-F. Zkuste si to taky spočítat.
Výsledek: přibližně 0,625.

Příklad 3.47. Na stavbu přivezli cihly ze čtyř cihelen v poměru 1 : 2 : 3 : 4. Jednotlivé cihelny vyrábějí kvalitní cihly s pravděpodobností v pořadí: 0,7; 0,6; 0,8; 0,9.
a) Jaká je pravděpodobnost, že náhodně vybraná cihla bude kvalitní?
b) V ruce mám kvalitní cihlu. Jaká je pravděpodobnost, že je ze třetí cihelny?
Výsledek: a) 0,79, b) $\frac{24}{79}$

Příklad 3.48. Zuzanka má tři balíčky bonbónů: karamelové, jahodové a čokoládové. Neposlušný Kamílek bonbónky promíchá a některé i sní, a tak nakonec v prvním balíčku je pět karamelových, jeden jahodový a dva čokoládové bonbónky, ve druhém balíčku je jeden karamelový, šest jahodových a jeden čokoládový a ve třetím balíčku jsou dva karamelové, jeden jahodový a šest čokoládových bonbónků. Pravděpodobnost, že bude Zuzanka hledat bonbón v prvním balíčku je 0,8, a pravděpodobnost hledání bonbónků ve druhém a třetím balíčku je stejná, tedy 0,1. Jaká je pravděpodobnost toho, že Zuzanka vybere karamelový bonbón?
Výsledek: $\frac{77}{144}$

Příklad 3.49. Zuzanka má tři balíčky barevných míčků: červené, modré a žluté. Neposlušný Kamílek míčky promíchá, takže nakonec v prvním balíčku je 20 červených, 8 modrých a 8 žlutých, v druhém balíčku 5 červených, 15 modrých a 16 žlutých a ve třetím balíčku je 11 červených, 13 modrých a 12 žlutých míčků. Pravděpodobnost, že Zuzanka vybere míček z prvního balíčku je 0,6, a pravděpodobnost vytažení míčků ze druhého a třetího balíčku je stejná, tedy 0,2. Jaká je pravděpodobnost toho, že Zuzanka vybere červený míček?
Výsledek: $\frac{19}{45}$

Příklad 3.50. Do textilní galanterie přivezli vlnu ze čtyř pletařských závodů v poměru 1 : 2 : 3 : 4. Pravděpodobnost toho, že vlna z jednotlivých závodů je kvalitně zabarvena, je po řadě 0,7, 0,6, 0,8 a 0,9. Určete pravděpodobnost toho, že náhodně vybrané klubíčko je kvalitně zabarveno.
Výsledek: 0,79.

Příklad 3.51. Na stavbu přivezli cihly ze čtyř cihelen v poměru 2 : 5 : 4 : 6. Jednotlivé cihelny vyrábějí kvalitní cihly s pravděpodobností v pořadí: 0,7; 0,6; 0,8; 0,9.
a) Jaká je pravděpodobnost, že náhodně vybraná cihla bude kvalitní?
b) V ruce mám kvalitní cihlu. Jaká je pravděpodobnost, že je ze třetí cihelny?

Výsledek: a) $\frac{13}{17}$, b) $\frac{32}{130}$.

Příklad 3.52. Do Tesca přivezli cukr ze tří balíren v poměru 2 : 3 : 5. Pravděpodobnost toho, že cukr z první balírny je kvalitně zabalen (nesype se z balíku), je 0,7, pravděpodonost, že cukr ze druhé a třetí balírny je kvalitně zabalen, je po řadě 0,8 a 0,9. Určete pravděpodobnost toho, že náhodně vybraný balíček cukru je kvalitně zabalen.
Výsledek: 0,83.

Příklad 3.53. * Pan Šťastný vyhrál ve Sportce. Vypočtěte pravděpodobnost, že to byla 1. cena. (Pozor, toto není totéž jako příklad 3.7 a).)

Výsledek: $P(1.\ \text{cena}/\text{vyhrál}) = \binom{6}{6} / \left(\binom{6}{6} + \binom{6}{5} \cdot \binom{43}{1} + \binom{6}{4} \cdot \binom{43}{2} + \binom{6}{3} \cdot \binom{43}{3} \right) \doteq 3{,}84 \cdot 10^{-6}$

Příklad 3.54. * Princ se vydal vysvobodit princeznu Zlatovlásku, kterou drží v zajetí zlý čaroděj. Čaroděj prince vlídně přivítá a řekne mu: „Dám ti Zlatovlásku, pokud ji dokážeš rozeznat.“ A ukáže mu tři zahalené postavy. Princi bohužel nepomůže žádná zlatá muška, ani „hlas srdce“ mu tentokrát nic nenapovídá. Proto zcela náhodně vybere jednu z postav: „Je to tahle!“
Verze A:
Princ má s sebou svého věrného mluvícího koně Zlatohříváka, a ten praví: „Pane, vyberte si radši tuhle.“ A ukáže na jinou dívku. Zlatohřívák ovšem kromě daru řeči žádné další magické schopnosti nemá a vybral také zcela náhodně. Čaroděj vše zlomyslně sleduje a pak odhalí zbývající, nevybranou dívku. Všichni vidí, že to Zlatovláska není. Princ se nyní musí definitivně rozhodnout mezi svou původní volbou a volbou koně Zlatohříváka.
Otázka: Co je nadějnější? Jeho původní volba, volba koně, nebo je to jedno?
Verze B:
Čaroděj, který samozřejmě velice dobře ví, která z dívek je Zlatovláska, chce prince pokoušet a řekne: „Vida, tuhle sis vybral. Možná je to ta pravá a možná také ne. Podívejme se…“ Přistoupí k jedné z nevybraných dívek a odhalí ji. Zlatovláska to není. Kromě dívky, kterou si princ vybral, teď zbývá ještě jedna. „Nechceš si to ještě rozmyslet? Není to třeba tahle?“ naléhá čaroděj na prince.
Otázka: Co má princ udělat? Zůstat u své původní volby, nebo dát na radu čaroděje? Nebo je to jedno? Je nějaký rozdíl mezi verzemi A a B?

Výsledek: Verze A: Je to jedno, pro obě dívky je teď pravděpodobnost, že je to Zlatovláska, 1/2.
Verze B: Předpokládejme, že čaroděj se absolutně nevyzná v pravděpodobnosti, o matematických aspektech problému vůbec nepřemýšlí a že chce prince prostě trápit. V tom případě ovšem čaroděj dělá chybu, protože princovy šance na úspěch zvyšuje. Jak? Když si princ zvolil jednu z dívek, pravděpodobnost, že je to Zlatovláska, byla 1/3. To znamená, že pravděpodobnost, že Zlatovláska je jedna ze zbývajících dívek, byla 2/3. Čaroděj ví, která je Zlatovláska, takže když jednu z dívek odhaluje, jde najisto (na rozdíl od koně Zlatohříváka ve verzi A). Vybere tu, která zaručeně Zlatovláska není. To ale znamená, že

pravděpodobnost 2/3 zůstává pro dívku, která je ještě zahalená a kterou princ nevybral. Jestliže se tedy princ vyzná v pravděpodobnosti, změní svou volbu a vybere si dívku, kterou mu nabízí čaroděj – bude mít dvakrát větší šanci na úspěch.
Jiná situace by ovšem nastala, kdybychom předpokládali, že čaroděj se v pravděpodobnosti vyzná, a začali celou věc rozebírat z psychologického hlediska. Do toho se však raději pouštět nebudeme. Obáváme se, že už předvedené čistě matematické řešení dokázalo mnoha lidem zamotat hlavu.

4 Diskrétní náhodná veličina

4.1 Základní pojmy a řešené příklady

Zatím jsme zkoumali výsledky náhodných procesů, které jsme popisovali slovně, např. „na kostce padla šestka“. Nyní se zaměříme na náhodné procesy, jejichž výsledek se dá popsat pomocí čísla, obvykle přirozeného.

Náhodná veličina (náhodná proměnná) Náhodná veličina (nebo též náhodná proměnná) je veličina X, jejíž hodnota je jednoznačně určena výsledkem náhodného pokusu.

Například počet kvalitních výrobků ve vyrobené denní dávce 100 ks je náhodná veličina X, která může nabýt hodnoty $x = 0, 1, 2, ..., 100$.

Diskrétní náhodná veličina
Náhodná veličina (proměnná) X se nazývá diskrétní, jestliže její obor hodnot (množina všech čísel, kterým se X může rovnat) je nanejvýš spočetná množina. To znamená, že X se buď může rovnat jen konečně mnoha hodnotám, nebo sice nekonečně mnoha hodnotám, ale tyto hodnoty lze seřadit do posloupnosti.
Hodnoty, kterých diskrétní náhodná veličina může nabývat, označíme $x_1, x_2, \dots$ a jejich počet označíme n, přičemž n může být i ∞.

Příkladem diskrétní náhodné veličiny s konečným oborem hodnot je např. součet ok při hodu dvěma kostkami. Příkladem diskrétní náhodné veličiny s nekonečným oborem hodnot je např. počet hodů mincí, než poprvé padne líc. Příkladem náhodné veličiny, která není diskrétní, je doba, po které přijde do obchodu další zákazník – takováto náhodná veličina (kdybychom mohli měřit s dokonalou přesností a nezaokrouhlovali např. na sekundy) by mohla nabýt jakékoli hodnoty z intervalu $(0, \infty)$, což je množina nespočetná – její prvky nelze uspořádat do posloupnosti.

Pravděpodobnostní funkce
Pravděpodobnostní funkce diskrétní náhodné veličiny X je funkce p, která je definována jako

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

(Čteme: „Hodnota funkce malé p v bodě malé x je rovna pravděpodobnosti, že náhodná veličina velké X se bude rovnat malému x.“)
Pro hodnoty pravděpodobnostní funkce platí

$$\sum_{i=1}^{n} p(x_i) = 1.$$

Dále se k popisu náhodných veličin používá distribuční funkce.

Distribuční funkce
Distribuční funkce náhodné veličiny X je funkce F, která je definována jako

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

(Čteme: „Hodnota funkce F v bodě malé x je rovna pravděpodobnosti, že náhodná veličina velké X nabude hodnoty menší nebo rovné malému x, tj. hodnoty z intervalu $(-\infty, x\rangle$.“) Pro distribuční funkci diskrétní náhodné veličiny platí

$$F(x) = \sum_{x_i \leq x} p(x_i).$$

Distribuční funkce má tyto základní vlastnosti: Je neklesající, zprava spojitá a

$$\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to \infty} F(x) = 1.$$

U diskrétní veličiny je distribuční funkce schodového tvaru – jedná se o funkci, která je po částech konstantní (na intervalech (x_i, x_{i+1})), pouze v bodech $x_1, x_2, x_3, \ldots$ dochází ke změně (ke schodu), kde velikost změny (= výška schodu) v bodě x_k je rovna právě hodnotě $p(x_k)$. Body vyznačené na pravém konci každého ze schodů prázdným kolečkem naznačují, že funkční hodnota distribuční funkce v bodě skoku je definována ne v bodě prázdného kolečka, ale až výše, na úrovni dalšího schodu – viz obrázek 4.1.

Příklad 4.1. Zuzanka má pytlík s kuličkami. Jsou v něm 3 bílé a 2 červené kuličky. Zuzanka vytahuje náhodně kuličky, po vytažení je do pytlíku nevrací. Své výběry ukončí, když vytáhne červenou kuličku. Náhodná veličina X udává počet tažení. Určete její rozdělení a distribuční funkci.

Řešení. Zřejmě $X \in \{1, 2, 3, 4\}$, Zuzanka může červenou kuličku vytáhnout napoprvé (1 tažení), napodruhé (vytáhne bílou kuličku a až pak červenou), napotřetí (dvakrát vytáhne bílou kuličku a pak červenou) nebo napočtvrté (postupně vytáhne všechny tři bílé kuličky a pak vytáhne jednu červenou). Jaká je pravděpodobnost, že napoprvé vytáhne červenou kuličku? Všech možností (tedy počet všech kuliček v pytlíku) je pět, zatímco příznivé možnosti (červené kuličky) jsou dvě. Proto

$$p(1) = \frac{2}{5} = 0{,}4.$$

Pro výpočet $p(2)$ uvažujeme následovně: Zuzanka nejdřív vytáhne bílou kuličku (všech kuliček je pět a bílé jsou tři, tedy pravděpodobnost je $\frac{3}{5}$) a pak vytáhne ze zbylých čtyř kuliček (tu vytaženou bílou už nevrací) jednu červenou kuličku (pravděpodobnost je $\frac{2}{4}$). Potom

$$p(2) = \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10} = 0{,}3.$$

Podobně zjistíme, že

$$p(3) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = \frac{1}{5} = 0{,}2.$$

Nejdřív z pěti kuliček (tří bílých a dvou červených) taháme bílou, pak ze zbylých čtyř (dvou bílých a dvou červených) taháme opět bílou a nakonec ze tří kuliček (jedné bílé a dvou červených) vytáhneme červenou. A na závěr

$$p(4) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot \frac{2}{2} = \frac{1}{10} = 0{,}1.$$

Pro kontrolu ještě spočítáme součet

$$p(1) + p(2) + p(3) + p(4) = 0{,}4 + 0{,}3 + 0{,}2 + 0{,}1 = 1{,}0.$$

Teď budeme zkoumat distribuční funkci. Nejprve vypočteme její hodnoty v „zajímavých“ bodech – v těch, kterým se náhodná veličina X může rovnat.

Podle definice je $F(1) = P(X \le 1)$. Určujeme tedy pravděpodobnost, že Zuzanka vytáhne nanejvýš jednu kuličku. To v našem případě znamená, že vytáhne právě jednu, takže

$$F(1) = P(X \le 1) = P(X = 1) = p(1) = 0{,}4.$$

Dále, $F(2) = P(X \le 2)$, tj. jedná se o pravděpodobnost, že Zuzanka vytáhne nanejvýš 2 kuličky neboli že vytáhne jednu nebo dvě. To dává

$$F(2) = P(X \le 2) = P(X = 1) + P(X = 2) = p(1) + p(2) = 0{,}4 + 0{,}3 = 0{,}7.$$

$F(3)$ bude pravděpodobnost, že Zuzanka vytáhne nanejvýš 3 kuličky:

$$F(3) = P(X \le 3) = p(1) + p(2) + p(3) = 0{,}4 + 0{,}3 + 0{,}2 = 0{,}9.$$

Podobně pak

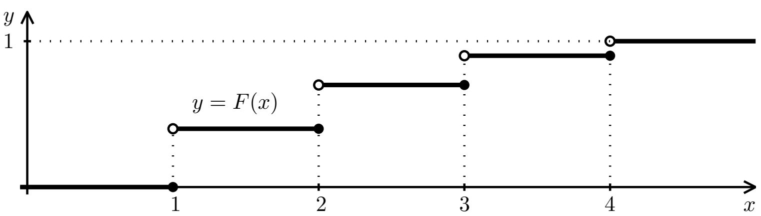
$$F(4) = P(X \le 4) = p(1) + p(2) + p(3) + p(4) = 1.$$

Nyní se podíváme, jak bude vypadat distribuční funkce v jiných bodech. Je zřejmé, že pro $x < 1$ bude $F(x) = 0$: Zuzanka např. nemůže vytáhnout nanejvýš -1 kuličku. Pro $x > 4$

bude naopak $F(x) = 1$: např. $F(5)$ je z definice $P(X \leq 5)$ a Zuzanka vytáhne určitě méně než 5 kuliček.
Je-li $x \in (1, 2)$, např. $x = 1{,}5$, počítáme pravděpodobnost, že Zuzanka vytáhne nanejvýš „jedna a něco" kuličky. To je možné jedině tak, že vytáhne právě jednu, a tedy $F(x) = p(1) = 0{,}4$. To znamená, že distribuční funkce je na tomto intervalu stejná jako pro $x = 1$.
Podobně pro $x \in (2, 3)$ dostaneme stejnou hodnotu jako $F(2)$, atd. Celkem pro $F(x)$ dostáváme

$$F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 1), \\ 0{,}4 & \text{pro } x \in \langle 1, 2), \\ 0{,}7 & \text{pro } x \in \langle 2, 3), \\ 0{,}9 & \text{pro } x \in \langle 3, 4), \\ 1 & \text{pro } x \in \langle 4, \infty). \end{cases}$$

Graf distribuční funkce je na obrázku 4.1. Vidíme, že distribuční funkce je schodovitá.



Obr. 4.1: K příkladu 4.1: Distribuční funkce náhodné veličiny X

První schůdek je v bodě $x_1 = 1$ a jeho výška je $p(1) = 0{,}4$. Další schůdek je v bodě $x_2 = 2$ a jeho výška je $p(2) = 0{,}3$. A tak dále, až poslední schůdek je v $x_n = x_4 = 4$ a jeho výška je $p(4) = 0{,}1$. Nalevo od prvního schůdku (tedy pro $x \to -\infty$) $F(x)$ nabývá hodnoty 0 a napravo od posledního schůdku (tedy pro $x \to \infty$) $F(x)$ nabývá hodnoty 1. Body na začátku schůdku jsou vyznačeny plným kolečkem a konci schůdku prázdným kolečkem, tedy tak, že funkce $F(x)$ je zprava spojitá.
Zkuste si tento příklad spočítat s jednou malou úpravou. Zuzanka bude kuličky do pytlíku vracet. Můžete se inspirovat následujícím příkladem.

Příklad 4.2. Pravděpodobnost, že student Martin přijde včas na vyučování, je $\frac{1}{4}$. Tato pravděpodobnost je stejná každý den, není závislá na předchozích dnech. Náhodná veličina X udává počet dnů, po kterých dojde k prvnímu pozdnímu příchodu. Určete její rozdělení.

Řešení. Zřejmě $X \in \{0, 1, 2, 3, ..., n, ...\}$, Martin může přijít hned první den pozdě (počet dnů do prvního pozdního příchodu je 0), nebo přijde pozdě druhý den (počet dnů do prvního pozdního příchodu je 1) a tak dále, tedy naše náhodná proměnná může nabýt nekonečně mnoho hodnot. Určeme pravděpodobnosti jednotlivých jevů.

- **0 dnů do prvního pozdního příchodu**
  Jev, že student přijde pozdě, je opačný k jevu, že přijde včas, tedy jeho pravděpodobnost je $1 - \frac{1}{4} = \frac{3}{4}$. Proto pravděpodobnost, že Martin přijde už první den pozdě, je $p(0) = \frac{3}{4}$.
- **1 den do prvního pozdního příchodu**
  Jestliže má být jeden den do prvního pozdního příchodu, tak Martin přišel první den včas a druhý den pozdě. Proto pravděpodobnost tohoto jevu bude součin pravděpodobností včasného a pozdního příchodu, tedy $p(1) = \frac{1}{4} \cdot \frac{3}{4}$.
- **2 dny do prvního pozdního příchodu**
  Martin přišel první dva dny včas, což může nastat s pravděpodobností $\frac{1}{4} \cdot \frac{1}{4} = \left(\frac{1}{4}\right)^2$. Třetí den přišel pozdě. Proto $p(2) = \left(\frac{1}{4}\right)^2 \cdot \frac{3}{4}$.
- **3 dny do prvního pozdního příchodu**
  První tři dny přišel Martin včas, což nastane s pravděpodobností $\frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} = \left(\frac{1}{4}\right)^3$. Čtvrtý den přišel pozdě a proto $p(3) = \left(\frac{1}{4}\right)^3 \cdot \frac{3}{4}$.
  .
  .
  .
- **n dnů do prvního pozdního příchodu**
  Po předchozích úvahách už umíme naše výsledky zobecnit: $p(n) = \left(\frac{1}{4}\right)^n \cdot \frac{3}{4}$.

Pro kontrolu ještě zjistíme součet hodnot pravděpodobnostní funkce, o kterém víme, že musí být 1. (Pozor, může se stát, že i při špatně určené pravděpodobnostní funkci tento součet vyjde 1, proto ani tato kontrola nemusí zaručit správné řešení!) Při výpočtu použijeme vztah pro součet nekonečné geometrické řady.

$$\frac{3}{4} + \frac{1}{4} \cdot \frac{3}{4} + \left(\frac{1}{4}\right)^2 \cdot \frac{3}{4} + ... + \left(\frac{1}{4}\right)^n \cdot \frac{3}{4} + ... =$$
$$= \frac{3}{4}\left(1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + ... + \left(\frac{1}{4}\right)^n + ...\right) =$$
$$= \frac{3}{4} \cdot \frac{1}{1 - \frac{1}{4}} = 1.$$

Příklad 4.3. Diskrétní náhodná veličina X udává počet opravených písemek za hodinu. Její distribuční funkce je dána takto:

$$F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 8), \\ 0{,}1 & \text{pro } x \in \langle 8, 9), \\ 0{,}3 & \text{pro } x \in \langle 9, 10), \\ 0{,}6 & \text{pro } x \in \langle 10, 11), \\ 0{,}9 & \text{pro } x \in \langle 11, 12), \\ 1 & \text{pro } x \in \langle 12, \infty). \end{cases}$$

Zjistěte pravděpodobnost toho, že za hodinu bude opravených
a) právě 10 písemek,
b) nanejvýš 10 písemek,
c) aspoň 10 písemek.

Řešení. a) Pravděpodobnost toho, že bude za hodinu opravených **právě** 10 písemek zjistíme z distribuční funkce snadno. Už víme, že $P(X = 10) = p(10)$ je výška schůdku v $x = 10$. A to je rozdíl hodnot, které nabývá funkce $F(x)$ na intervalu $\langle 10, 11)$ a na intervalu $\langle 9, 10)$, tedy $p(10) = 0{,}6 - 0{,}3 = 0{,}3$.
b) Pravděpodobnost toho, že bude za hodinu opravených **nanejvýš** 10 písemek se dá zapsat jako $P(X \leq 10)$, což je přesně $F(10)$. Protože pro interval $\langle 10, 11)$ nabývá $F(x)$ hodnoty 0,6, je $F(10) = 0{,}6$, což je odpověď na část b). Úlohu bychom mohli vyřešit i trochu jinak. Co znamená, že bude opraveno nanejvýš 10 písemek? Že jich bude opraveno 10 nebo 9 nebo 8 a tak dále až 0 (méně se už zřejmě nedá). Tedy hledaná pravděpodobnost bude součtem pravděpodobností $p(10), p(9), p(8), \dots, p(0)$. Na základě řešení části a) víme, že tyto pravděpodobnosti jsou výšky schůdků distribuční funkce v bodech $10, 9, 8, \dots, 0$. Z předpisu $F(x)$ vidíme, že v bodech 0 až 7 nejsou schůdky (funkce tam je konstantní), tedy jejich výška je nulová, výška schůdku v bodě 8 je 0,1, v bodě 9 je 0,2 a v bodě 10 je 0,3. Proto naše hledaná pravděpodobnost bude součet $0 + \dots + 0 + 0{,}1 + 0{,}2 + 0{,}3 = 0{,}6$, což je samozřejmě stejný výsledek, který jsme dostali užitím vědomostí o distribuční funkci.
c) Pravděpodobnost toho, že bude za hodinu opravených **aspoň** 10 písemek, je $P(X \geq 10)$. K výpočtu můžeme využít výsledky částí a) a b) – pozor, jev z části b) není doplňkový k jevu z části c)! Jedna z možností výpočtu je

$$\begin{aligned} P(X \geq 10) = P(X > 10) + P(X = 10) &= 1 - P(X \leq 10) + P(X = 10) = \\ &= 1 - F(10) + p(10) = 1 - 0{,}6 + 0{,}3 = 0{,}7. \end{aligned}$$

Výsledek můžeme zjistit i tak, že sečteme pravděpodobnosti $p(10), p(11), p(12)$, atd. Opět budeme sečítat výšky schůdků v jednotlivých bodech. Proto naše hledaná pravděpodobnost bude součet $0{,}3 + 0{,}3 + 0{,}1 + 0 + 0 + 0 + \dots = 0{,}7$. Odkud se vzaly nulové hodnoty? V bodě 10 je schůdek vysoký 0,3 a stejný schůdek je v bodě 11, v bodě 12 má schůdek výšku 0,1 a pak už funkce $F(x)$ svou hodnotu nemění, je konstantní a má hodnotu 1. Proto výšky dalších schůdků jsou nulové.

Příklad 4.4. Pravděpodobnost, že výrobek bude vyhovovat všem požadavkům normy, je 0,8. Popište rozdělení počtu nevyhovujících mezi 3 výrobky.

Pokus o řešení-bohužel nesprávný: Pro tři výrobky mohou nastat právě tyto čtyři (neslučitelné) možnosti: ani jeden nevyhovující, jeden nevyhovující, dva nevyhovující nebo tři nevyhovující. Postupně si je rozebereme:

- **Ani jeden nevyhovující výrobek**
  To znamená, že všechny tři výrobky budou vyhovovat všem požadavkům normy. Nezávadnost či závadnost nějakého výrobku nezávisí od nezávadnosti nebo závadnosti jiného výrobku, jsou to nezávislé jevy. Proto když máme například dva výrobky

v_1, v_2, a víme, že $P(v_1 \text{ je vyhovující}) = P(v_2 \text{ je vyhovující}) = 0,8$, tak pravděpodobnost, že tyto dva výrobky jsou vyhovující oba, je

$$P(v_1 \text{ je vyhovující}) \cdot P(v_2 \text{ je vyhovující}) = 0,8 \cdot 0,8.$$

My máme tři vyhovující výrobky, proto pravděpodobnost naší první možnosti je

$$P(v_1 \text{ je vyhovující}) \cdot P(v_2 \text{ je vyhovující}) \cdot P(v_3 \text{ je vyhovující}) = 0,8 \cdot 0,8 \cdot 0,8 = 0,512.$$

- **Jeden nevyhovující výrobek**
  Jestliže máme tři výrobky a jeden je nevyhovující, zbylé dva musí být vyhovující. Když pravděpodobnost vyhovujícího výrobku je $0,8$, tak pravděpodobnost nevyhovujícího výrobku je $1 - 0,8 = 0,2$. Stejnou úvahou jako v předchozí části zjistíme, že hledaná pravděpodobnost je součinem
  $$0,2 \cdot 0,8 \cdot 0,8 = 0,128.$$
- **Dva nevyhovující výrobky**
  Ze tří výrobků máme dva nevyhovující, tedy jeden je vyhovující. Proto naše pravděpodobnost je
  $$0,2 \cdot 0,2 \cdot 0,8 = 0,032.$$
- **Tři nevyhovující výrobky**
  To znamená, že všechny tři výrobky jsou nevyhovující, a podle předchozích úvah pro pravděpodobnost našeho jevu dostáváme
  $$0,2 \cdot 0,2 \cdot 0,2 = 0,008.$$

Kde se stala chyba? Nebo si někdo myslí, že jsme rozdělení našli správně? Naše čtyři možnosti jsou neslučitelné a žádná další možnost už nemůže nastat. Proto by součet uvedených pravděpodobností měl být 1, ale náš součet je

$$0,512 + 0,128 + 0,032 + 0,008 = 0,68 < 1.$$

Protože víme, že součet pravděpodobností všech možností je 1, mnohokrát při hledání rozdělení postupujeme tak, že vypočítáme pravděpodobnosti všech možností kromě té poslední, pak je sečteme (součet označme $p_{1\ldots n-1}$), a pravděpodobnost poslední možnosti vypočítame jako rozdíl $1 - p_{1\ldots n-1}$. Naštěstí jsme pravděpodobnost poslední možnosti spočítali zvlášť, jinak bychom chybu neodhalili.
Správné řešení: Rozdělení na čtyři možnosti, které jsme provedli v našem nesprávném řešení, je správné. Výpočty pravděpodobností první a poslední možnosti jsou také správné, chyby jsme udělali jenom při určení pravděpodobností u druhé a třetí možnosti. Všimněme si druhé možnosti, tedy jeden výrobek je nevyhovující a dva jsou vyhovující. Kde jsme udělali chybu? Podle zápisu našeho výpočtu $0,2 \cdot 0,8 \cdot 0,8 = 0,128$, jsme uvažovali jenom tu možnost, že nevyhovující výrobek byl první a zbylé dva byly vyhovující. Ale on mohl

být i druhý výrobek nevyhovující (k tomu první a třetí výrobek vyhovující), nebo třetí nevyhovující (a k tomu první dva vyhovující). Proto pravděpodobnost této možnosti je

$$0,2 \cdot 0,8 \cdot 0,8 + 0,8 \cdot 0,2 \cdot 0,8 + 0,8 \cdot 0,8 \cdot 0,2 = 0,384.$$

Podobně ve třetí možnosti (dva nevyhovující a jeden vyhovující výrobek) musíme uvažovat, že vyhovující výrobek může být první, druhý, nebo třetí, a proto pravděpodobnost této možnosti je

$$0,8 \cdot 0,2 \cdot 0,2 + 0,2 \cdot 0,8 \cdot 0,2 + 0,2 \cdot 0,2 \cdot 0,8 = 0,096.$$

Když si teď sečteme všechny pravděpodobnosti, tak konečně dostaneme

$$0,512 + 0,384 + 0,096 + 0,008 = 1.$$

A teď se budeme věnovat důležitým číselným charakteristikám náhodních veličin:

Střední hodnota
Střední hodnota diskrétní náhodné veličiny X se vypočítá jako

$$\mathrm{E}X = \sum_{i=1}^{n} x_i \cdot p(x_i) = x_1 \cdot p(x_1) + x_2 \cdot p(x_2) + \cdots .$$

Zhruba řečeno, střední hodnota náhodné proměnné udává, jaký asi bude průměr získaných hodnot náhodné proměnné při mnoha opakováních náhodného procesu. Určit střední hodnotu má velký význam při statistických analýzách různých náhodných procesů, třeba známý algoritmus třídění `quicksort` může počítat i velmi dlouho, ale střední hodnota času jeho běhu je nejlepší ze všech běžných třídících algoritmů.
Praktický výpočet středních hodnot v situacích složených výběrů lze výrazně ulehčit pomocí následujících dvou tvrzení.
Pro libovolné dvě náhodné proměnné X, Y platí

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

Pro libovolné dvě **nezávislé** náhodné proměnné X, Y platí

$$\mathrm{E}(X \cdot Y) = \mathrm{E}X \cdot \mathrm{E}Y.$$

Pro spolehlivý popis náhodné veličiny potřebujeme znát nejenom střed, kolem kterého se jednotlivé hodnoty soustřeďují, ale také jak daleko se od tohoto středu rozptylují. Rozptyl je definován jako střední hodnota kvadrátů odchylek od střední hodnoty. Odchylku od střední hodnoty, která má rozměr stejný jako náhodná veličina, zachycuje směrodatná odchylka.

Rozptyl
Pro diskrétní náhodnou veličinu můžeme **rozptyl** definovat vztahem

$$\mathrm{D}X = \sum_{i=1}^{n} [x_i - \mathrm{E}X]^2 p(x_i) = \sum_{i=1}^{n} x_i^2 p(x_i) - (\mathrm{E}X)^2,$$

kde x_i jsou hodnoty, kterých může náhodná veličina X nabývat s pravděpodobnostmi $p(x_i)$, a $\mathrm{E}X$ je střední hodnota veličiny X.

Příklad 4.5. Diskrétní náhodná veličina udává počet dnů, které student využije na přípravu na zkoušku. Hodnoty její pravděpodobnostní funkce jsou v následující tabulce:

x_i	0	1	2	3	4	5	6
$p(x_i)$	0,4	0,35	0,15	0,05	0,03	0,01	0,01

a) Určete očekávaný počet dnů, který student využije na přípravu.
b) Určete rozptyl dané náhodné veličiny.

Řešení. a) Z teorie víme, že součet hodnot pravděpodobnostní funkce musí být 1. Proto je třeba nejdřív to zkontrolovat. Jelikož

$$0{,}4 + 0{,}35 + 0{,}15 + 0{,}05 + 0{,}03 + 0{,}01 + 0{,}01 = 1{,}00,$$

víme, že pravděpodobnostní funkce je zadána korektně, a proto můžeme spočítat její střední hodnotu podle známého vzorce:

$$\mathrm{E}X = 0 \cdot 0{,}4 + 1 \cdot 0{,}35 + 2 \cdot 0{,}15 + 3 \cdot 0{,}05 + 4 \cdot 0{,}03 + 5 \cdot 0{,}01 + 6 \cdot 0{,}01 = 1{,}03.$$

Tedy očekávaný počet dnů, který student využije na přípravu, je 1,03. (Trochu málo, nebo ne?)
b) Zřejmě

$$\mathrm{D}X = \sum_{i=1}^{n} x_i^2 p(x_i) - (\mathrm{E}X)^2,$$

a

$$(\mathrm{E}X)^2 = 1,03^2 = 1,0609.$$

Dále

$$\sum_{i=1}^{n} x_i^2 p(x_i) = 0^2 \cdot 0{,}4 + 1^2 \cdot 0{,}35 + 2^2 \cdot 0{,}15 + 3^2 \cdot 0{,}05 + 4^2 \cdot 0{,}03 + 5^2 \cdot 0{,}01 + 6^2 \cdot 0{,}01 = 2{,}49,$$

proto hledaný rozptyl je

$$\mathrm{D}X = 2{,}49 - 1{,}0609 = 1{,}4291.$$

Příklad 4.6. Diskrétní náhodná veličina udává počet dnů bezporuchového provozu měřicího přístroje. Hodnoty její pravděpodobnostní funkce jsou v následující tabulce, ale dvě hodnoty neznáme.

x_i	0	1	2	3	4	5	6
$p(x_i)$	*	0,35	0,15	*	0,03	0,01	0,01

Určete je, když víte, že očekávaný počet dnů bezporuchového provozu je 1,5.

Řešení. Víme, že součet hodnot pravděpodobnostní funkce má být 1, proto

$$p(0) + p(1) + p(2) + p(3) + p(4) + p(5) + p(6) = 1,$$

z toho dostaneme

$$p(0) + p(3) = 0,45,$$

tedy

$$p(0) = 0,45 - p(3)$$

a po dosazení do vzorce na výpočet střední hodnoty, ($\mathrm{E}X = 1{,}5$), máme

$$1{,}5 = 0 \cdot (0{,}45 - p(3)) + 1 \cdot 0{,}35 + 2 \cdot 0{,}15 + 3 \cdot p(3) + 4 \cdot 0{,}03 + 5 \cdot 0{,}01 + 6 \cdot 0{,}01,$$

$$1{,}5 = 1 \cdot 0{,}35 + 2 \cdot 0{,}15 + 3 \cdot p(3) + 4 \cdot 0{,}03 + 5 \cdot 0{,}01 + 6 \cdot 0{,}01,$$

$$1{,}5 = 0{,}35 + 0{,}3 + 3 \cdot p(3) + 0{,}12 + 0{,}05 + 0{,}06,$$

$$1{,}5 = 0{,}88 + 3 \cdot p(3),$$

$$p(3) \doteq 0{,}2067,$$

$$p(0) \doteq 0{,}45 - 0{,}2067 \doteq 0{,}2433.$$

Zkuste si tuto úlohu spočítat pro stejné tabulkové hodnoty s tím, že střední hodnota bude 2,5. Pokud budete správně počítat, dostanete pro $p(0)$ zápornou hodnotu, což je nesmysl. Podobně nevhodný výsledek by bylo číslo větší než 1. Proto dřív než svůj výsledek podtrhnete, posuďte jeho smysluplnost!

Pro další příklady na střední hodnotu se podívejme na hody kostkami.

- Jaká je střední hodnota (průměr) čísel padlých na šestistěnné kostce?
  Jednoduchým výpočtem vyjde $\mathrm{E}K = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3{,}5$.
- Jaká je střední hodnota součtu čísel padlých na dvou šestistěnných kostkách?
  S využitím předchozího výsledku $\mathrm{E}(K_1 + K_2) = \mathrm{E}K_1 + \mathrm{E}K_2 = 3{,}5 + 3{,}5 = 7$.
- Jaká je střední hodnota součinu čísel padlých na dvou šestistěnných kostkách?
  S využitím předchozího výsledku $\mathrm{E}(K_1 \cdot K_2) = \mathrm{E}K_1 \cdot \mathrm{E}K_2 = 3{,}5 \cdot 3{,}5 = 12{,}25$.

Vraťme se nyní k našemu intuitivnímu pojetí pravděpodobnosti jako očekávání četnosti opakování jevu. Pokud například hodíme n-krát mincí, pak pravděpodobnost $\frac{1}{2}$ jedné strany mince znamená, že zhruba $n/2$-krát padne každá ze stran mince. Pokud je naše matematická formalizace pravděpodobnosti správná, měla by nám tentýž výsledek říci střední hodnota počtu hlav při n hodech mincí. Jak tomu skutečně je, ukazuje následující příklad.

Příklad 4.7. Jaký je průměrný počet hlav padlých při n hodech mincí?

Řešení. Pokud hlavě mince přiřadíme hodnotu 1, máme tak n náhodných proměnných $X_i \in \{0,1\}$, $i = 1,\dots,n$ příslušejících jednotlivým hodům mince. Celkový počet hlav je dán náhodnou proměnnou $X = X_1 + \dots + X_n$. Takže průměrný počet hlav je

$$\mathrm{E}X = \mathrm{E}(X_1 + \dots + X_n) = \mathrm{E}X_1 + \dots + \mathrm{E}X_n = \frac{1}{2} + \dots + \frac{1}{2} = \frac{n}{2}\,.$$

To přesně odpovídá našemu vnímání pravděpodobnosti jako relativní četnosti jevu.

Příklad 4.8. Kolik je třeba průměrně hodů mincí, aby vyšly tři stejné výsledky?

Řešení. Je snadno vidět, že nejdříve tři stejné výsledky mohou nastat po třech hodech a nejpozději po pěti hodech (z dvou hlav a dvou orlů pět hodů nesložíme). S jakou pravděpodobností získáme stejné výsledky při třech hodech? Jsou možnosti buď tří hlav, nebo tří orlů, takže

$$p(3) = 2 \cdot \frac{1}{8} = \frac{1}{4}\,.$$

Až po pěti hodech získáme 3 stejné výsledky, pokud první čtyři hody budou rozděleny dva na dva (na posledním hodu pak již vlastně nezáleží). To se může stát v $\binom{4}{2} = 6$ možnostech pro 4 hody, takže pravděpodobnost je

$$p(5) = 6 \cdot \frac{1}{16} = \frac{3}{8}\,.$$

Možonst, že 3 stejné výsledky získáme po čtyřech hodech, je doplňková k předchozím dvěma a v součtu musí mít pravděpodobnost 1, proto

$$p(4) = 1 - p(3) - p(5) = 1 - \frac{1}{4} - \frac{3}{8} = \frac{3}{8}\,.$$

Průměrný počet potřebných hodů je dle definice střední hodnoty

$$N = p(3)\cdot 3 + p(4)\cdot 4 + p(5)\cdot 5 = \frac{3}{4} + \frac{3}{2} + \frac{15}{8} = \frac{33}{8} = 4{,}125.$$

(To je o trochu více než 4, což by jeden mohl intuitivně očekávat.)

Příklad 4.9. (varovný) Jaký je průměrný součin čísel horní a spodní stěny stejné kostky při hodech?

Řešení. Jak už víme, střední hodnota čísla na horní stěně je 3,5 a dolní stěně samozřejmě taky 3,5. Střední hodnota jejich součinu však není $3{,}5 \cdot 3{,}5 = 12{,}25$, protože tyto dva jevy **nejsou nezávislé**.
Místo toho střední hodnotu součinu spočítáme podle definice

$$\frac{1}{6}(1\cdot 6 + 2\cdot 5 + 3\cdot 4 + 4\cdot 3 + 5\cdot 2 + 6\cdot 1) = \frac{1}{6}\cdot 56 = 9 + \frac{1}{3}.$$

Proto si dávejme dobrý pozor na nezávislost jevů při násobení středních hodnot!

4.2 Příklady pro samostatnou práci

Příklad 4.10. Diskrétní náhodná veličina X udává počet gramatických chyb v textu. Její distribuční funkce $F(x)$ je následovná:

$$F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0), \\ 0{,}5 & \text{pro } x \in \langle 0, 1), \\ 0{,}8 & \text{pro } x \in \langle 1, 2), \\ 0{,}9 & \text{pro } x \in \langle 2, 3), \\ 0{,}95 & \text{pro } x \in \langle 3, 4), \\ 1 & \text{pro } x \in \langle 4, \infty). \end{cases}$$

a) Určete střední hodnotu náhodné veličiny X.
b) Určete pravděpodobnost toho, že v textu budou právě 3 chyby.
c) Určete pravděpodobnost toho, že v textu budou méně než 2 chyby.

Výsledek: a) $\mathrm{E}X = 0{,}85$, b) $P(X = 3) = 0{,}05$, c) $P(X < 2) = 0{,}8$

Příklad 4.11. Diskrétní náhodná veličina X udává počet svetrů, které uplete studentka Petra za zimní semestr. Její distribuční funkce $F(x)$ je následovná:

$$F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0), \\ 0{,}05 & \text{pro } x \in \langle 0, 1), \\ 0{,}15 & \text{pro } x \in \langle 1, 2), \\ 0{,}65 & \text{pro } x \in \langle 2, 3), \\ 0{,}85 & \text{pro } x \in \langle 3, 4), \\ 1 & \text{pro } x \in \langle 4, \infty). \end{cases}$$

a) Určete její střední hodnotu.
b) Určete pravděpodobnost toho, že za semestr Petra uplete právě 2 svetry.
c) Určete pravděpodobnost toho, že za semestr Petra uplete méně než 1 svetr.

Výsledek: a) $\mathrm{E}X = 2{,}3$, b) $P(X = 2) = 0{,}5$, c) $P(X < 1) = 0{,}05$

Příklad 4.12. Diskrétní náhodná veličina X udává počet pozdních příchodů studenta Martina H. na vyučování za semestr. Její distribuční funkce $F(x)$ je následovná:

$$F(x) = \begin{cases} 0 & \text{pro } x \in \langle -\infty, 1), \\ 0{,}5 & \text{pro } x \in \langle 1, 2), \\ 0{,}7 & \text{pro } x \in \langle 2, 3), \\ 0{,}8 & \text{pro } x \in \langle 3, 4), \\ 0{,}9 & \text{pro } x \in \langle 4, 5), \\ 1 & \text{pro } x \in \langle 5, \infty). \end{cases}$$

a) Určete její střední hodnotu.
b) Určete pravděpodobnost toho, že za semestr bude mít Martin právě 4 pozdní příchody.
c) Určete pravděpodobnost toho, že za semestr bude mít více než 3 pozdní příchody.
Výsledek: a) $\mathrm{E}X = 2{,}1$, b) $P(X = 4) = 0{,}1$, c) $P(X > 3) = 0{,}2$

Příklad 4.13. Je známa distribuční funkce denního počtu obsazených pokojů v hotelu:

$$F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 7), \\ 0{,}5 & \text{pro } x \in \langle 7, 8), \\ 0{,}8 & \text{pro } x \in \langle 8, 9), \\ 0{,}9 & \text{pro } x \in \langle 9, 10), \\ 1 & \text{pro } x \in \langle 10, \infty). \end{cases}$$

a) Určete pravděpodobnost, že v náhodně zvolený den bude obsazených právě 9 pokojů.
b) Určete pravděpodobnost, že v náhodně zvolený den bude obsazených alespoň 8 pokojů.

Výsledek: a) $P(X = 9) = 0{,}1$, b) $P(X \geq 8) = 0{,}5$

Příklad 4.14. Pravděpodobnostní funkce náhodné proměnné X je daná tabulkou

x_i	10	11	12	13	14	15	16
$p(x_i)$	0,45	0,3	0,1	0,05	0,08	0,01	0,01

Určete její střední hodnotu.

Výsledek: $\mathrm{E}X = 11{,}08$

Příklad 4.15. Náhodná proměnná X je daná tabulkou

x_i	-1	0	1	2	3	4	5
$p(x_i)$	0,04	0,35	0,15	0,05	0,03	0,01	0,01

Určete její střední hodnotu a rozptyl.

Výsledek: Součet pravděpodobností není 1, úloha nemá smysl.

Příklad 4.16. Diskrétní náhodná veličina X udává počet padnutých ok na šestistěnné kostce.
a) Určete její pravděpodobnostní a distribuční funkci.
b) Určete její střední hodnotu.
c) Určete pravděpodobnost toho, že počet ok bude větší než 3.

Výsledek:

a) $P(X = i) = \frac{1}{6};\ i \in \{1, 2, 3, 4, 5, 6\}$, $F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 1), \\ \frac{i}{6} & \text{pro } x \in \langle i, i+1), i \in \{1, 2, ...5\}, \\ 1 & \text{pro } x \in \langle 6, \infty). \end{cases}$

b) $\mathrm{E}X = 3{,}5$, c) $P(X > 3) = 0{,}5$

Příklad 4.17. Diskrétní náhodná veličina X udává součet padnutých ok na dvou šestistěnných kostkách.
a) Určete její pravděpodobnostní funkci.
b) Určete její střední hodnotu.
c) Určete pravděpodobnost toho, že součet bude menší než 5.

Výsledek:
a) $P(X=2)=P(X=12)=\frac{1}{36}$, $P(X=3)=P(X=11)=\frac{1}{18}$, $P(X=4)=P(X==10)=\frac{1}{12}$, $P(X=5)=P(X=9)=\frac{1}{9}$, $P(X=6)=P(X=8)=\frac{5}{36}$, $P(X=7)=\frac{1}{6}$,
b) $\mathrm{E}X=\frac{20}{3}$, c) $P(X<5)=\frac{1}{6}$

Příklad 4.18. Náhodná veličina X udává absolutní hodnotu rozdílu čísel padlých na dvou kostkách.
a) Určete pravděpodobnostní rozdělení náhodné veličiny X.
b) Určete střední hodnotu $\mathrm{E}X$.

Výsledek:

a) $p(i)=\begin{cases}\frac{1}{6} & \text{pro } i=0, i=3,\\ \frac{5}{18} & \text{pro } i=1,\\ \frac{2}{9} & \text{pro } i=2,\\ \frac{1}{9} & \text{pro } i=4,\\ \frac{1}{18} & \text{pro } i=5.\end{cases}$ b) $\mathrm{E}X=\frac{35}{18}$

Příklad 4.19. Vytáčení telefonického připojení k internetu je maximálně šestkrát opakováno (tj. po úspěšném připojení, respektive po šesti neúspěšných vytáčeních se v pokusu o spojení nepokračuje). Jednotlivá vytáčení jsou navzájem nezávislá. Pravděpodobnost správného připojení při každém vytáčení je rovna 0,6. Veličina X udává počet vytáčení při daném pokusu o spojení.
a) Určete rozdělení pravděpodobnosti veličiny X.
b) Určete střední hodnotu veličiny X.

Výsledek:

a) $p(i)=\begin{cases}0{,}4^{i-1}\cdot 0{,}6 & \text{pro } i\in\{1,...,5\},\\ 0{,}4^5\cdot 0{,}6+0{,}4^6 & \text{pro } i=6,\end{cases}$ b) $\mathrm{E}X=1{,}65984$

Příklad 4.20. Balík je linkou nedostatečně zabalen s pstí 0,3. Jednotlivá balení jsou nezávislá na předchozích baleních. Balicí linka se zastaví poté, co balík je nedostatečně zabalen, v každém případě se však zastaví po zabalení pátého balíku. (X = počet zabalených balíků (ať už dobře, nebo špatně – počítají se všechny) před prvním zastavením linky).
a) Určete pravděpodobnostní funkci veličiny X.
b) Určete $\mathrm{E}X$, $\mathrm{D}X$.

Výsledek:

a) $p(i) = \begin{cases} 0{,}3 & \text{pro } i = 1, \\ 0{,}21 & \text{pro } i = 2, \\ 0{,}147 & \text{pro } i = 3, \\ 0{,}1029 & \text{pro } i = 4, \\ 0{,}2401 & \text{pro } i = 5, \end{cases}$ b) $\mathrm{E}X = 2{,}7731$, $\mathrm{D}X \doteq 2{,}4218$

Příklad 4.21. Honza chodí na pravidelné odběry krve. Pravděpodobnost, že výsledek odběru bude dobrý, je vždy 0,4 a nezávisí na předchozích odběrech. Jestliže budou dva po sobě jdoucí odběry dobré, nebude muset přijít na další odběry. Dále také v každém případě přestane chodit na odběry po čtyřech absolvovaných odběrech. Náhodná veličina X udává počet odběrů, které Honza absolvuje v daném období.
a) Určete pravděpodobnostní rozdělení náhodné veličiny X.
b) Určete střední hodnotu $\mathrm{E}X$.

Výsledek:
a) $p(i) = \begin{cases} 0{,}16 & \text{pro } i = 2, \\ 0{,}096 & \text{pro } i = 3, \\ 0{,}744 & \text{pro } i = 4, \end{cases}$ b) $\mathrm{E}X = 3{,}584$

Příklad 4.22. Honza má v zásobníku pět nábojů. Pravděpodobnost, že zasáhne terč, je vždy $0,4$ a nezávisí na předchozích pokusech. Honza bude střílet tak dlouho, dokud nezasáhne terč, pak střílet přestane. V každém případě přestane střílet, až mu dojde daných pět nábojů. Náhodná veličina X udává počet zbylých nábojů po vykonání popsaného experimentu.
a) Určete pravděpodobnostní rozdělení náhodné veličiny X.
b) Určete střední hodnotu $\mathrm{E}X$.

Výsledek:
a) $p(i) = \begin{cases} 0{,}4 & \text{pro } i = 4, \\ 0{,}24 & \text{pro } i = 3, \\ 0{,}144 & \text{pro } i = 2, \\ 0{,}0864 & \text{pro } i = 1, \\ 0{,}1296 & \text{pro } i = 0, \end{cases}$ b) $\mathrm{E}X = 2{,}6944$

Příklad 4.23. Střelec střílí do terče, v zásobě má 4 náboje. Pravděpodobnost zásahu je při každém výstřelu 0,6. Diskrétní náhodná veličina X udává počet zásahů do terče.
a) Určete její pravděpodobnostní funkci.
b) Určete její střední hodnotu.

Výsledek:

a) $p(i) = \begin{cases} 0{,}0256 & \text{pro } i = 0, \\ 0{,}1536 & \text{pro } i = 1, \\ 0{,}3456 & \text{pro } i = 2, \\ 0{,}3456 & \text{pro } i = 3, \\ 0{,}1296 & \text{pro } i = 4, \end{cases}$ b) $\mathrm{E}X = 2{,}4$

Příklad 4.24. Basketbalista hází trestné koše – s házením přestává po dvou úspěšných pokusech v těsném sledu za sebou, v každém případě ovšem přestává házet i po pátém pokusu. Pravděpodobnost úspěchu při každém hodu je 0,8. Jednotlivé hody jsou považovány za nezávislé. Veličina X udává počet hodů.
a) Určete pravděpodobnostní funkci veličiny X.
b) střední hodnotu veličiny X.

Výsledek:

a) $p(i) = \begin{cases} 0{,}64 & \text{pro } i = 2, \\ 0{,}128 & \text{pro } i = 3, \\ 0{,}128 & \text{pro } i = 4, \\ 0{,}104 & \text{pro } i = 5, \end{cases}$ b) $\mathrm{E}X = 2{,}5936$

Příklad 4.25. Hráči A, B hrají následující hru: Losují za regulérních pravidel losy s čísly 1 až 20 (vytažení libovolného čísla je stejně pravděpodobné). Pokud je vytaženo některé z čísel 1 až 4, dává hráč A hráči B 40 Kč (zisk hráče A je záporný). Pokud je taženo některé z čísel 5 až 12, dává hráč A hráči B 60 Kč (zisk hráče A je záporný). Pokud je taženo číslo 13 nebo 14, dává hráč B hráči A k Kč. Pokud je taženo číslo 15 až 20, dává hráč B hráči A 30 Kč.
a) Určete rozdělení veličiny X = zisk hráče A v jednom kole hry. Hodnotu k přitom zatím určovat nemusíte.
b) Určete hodnotu k tak, aby se jednalo o spravedlivou hru, tj. střední hodnota zisku hráče A byla rovna nule.

Výsledek:

a) $p(i) = \begin{cases} \frac{1}{5} & \text{pro } i = -40, \\ \frac{2}{5} & \text{pro } i = -60, \\ \frac{1}{10} & \text{pro } i = k, \\ \frac{3}{10} & \text{pro } i = 30, \end{cases}$ b) $k = 230$

Příklad 4.26. Budeme házet mincí, dokud nepadne znak. Diskrétní náhodná proměnná X udává počet hodů do prvního padnutí znaku. Určete pravděpodobnostní a distribuční funkci proměnné X.

Výsledek: $p(i) = \frac{1}{2^i}, i \in \mathbb{N}$, $F(x) = \begin{cases} 0 & \text{pro } x < 1, \\ \frac{2^n-1}{2^n} & \text{pro } x \in \langle n, n+1), \end{cases}$

Příklad 4.27. Diskrétní náhodná veličina X udává počet ok při hodu kostkou. Vypočítejte její rozptyl.

Výsledek: $35/12 \doteq 2{,}917$

Příklad 4.28. Jaká je střední hodnota počtu šestek padlých při hodu 10 šestistěnných kostek?

Výsledek: $5/3 \doteq 1{,}667$

Příklad 4.29. Jaký je průměrný počet hlav padlých při pěti hodech mincí?

Výsledek: 2,5

Příklad 4.30. Jaký je průměrný počet hlav padlých při n hodech mincí?

Výsledek: $\frac{n}{2}$

Příklad 4.31. Kolik je třeba průměrně hodů mincí, aby vyšly dva stejné výsledky?

Výsledek: 2,5

Příklad 4.32. (varovný) Jaký je průměrný součet čísel horní a spodní stěny stejné kostky při hodech?

Výsledek: 7

Příklad 4.33. * Kolik je třeba průměrně hodů mincí, aby padla první hlava?
Návod: Uvědomte si, že teoreticky hlava nemusí padnout nikdy, ale pravděpodobnost samých orlů jde k nule.

Výsledek: 2

5 Spojité náhodné veličiny

5.1 Základní pojmy a řešené příklady

V předchozí kapitole jsme se věnovali diskrétním náhodným veličinám, které mohou nabývat pouze konečného počtu hodnot, nebo je počet možných hodnot sice nekonečný, ale lze je uspořádat do posloupnosti, jednotlivě si je ukazovat. Spojité náhodné veličiny jsou jiného charakteru. Zjednodušeně řečeno, mohou nabývat všech hodnot z určitého intervalu.

Hustota pravděpodobnosti
U spojité náhodné veličiny se pravděpodobnost, že náhodná veličina X padne do určitého intervalu (a, b), počítá jako

$$P(X \in (a, b)) = \int_a^b f(x)\mathrm{d}x,$$

kde funkce f je tzv. hustota pravděpodobnosti náhodné veličiny X.

Příklad 5.1. Hustota pravděpodobnosti náhodné veličiny X má tvar:

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

Vypočtěte pravděpodobnosti $P(1{,}5 \leq X \leq 1{,}8)$, $P(X < 1{,}2)$, $P(X > 1{,}7)$, $P(X = 1{,}5)$, $P(X < 3)$. Dále najděte hodnotu a, pro kterou by platilo $P(X < a) = 0{,}8$.

Řešení. První tři pravděpodobnosti vypočítáme jako integrál z hustoty f přes příslušný interval. Integrály z nulové funkce bychom mohli rovnou vynechávat, ale zde, v našem prvním příkladu, je vypíšeme.

$$\begin{aligned} P(1{,}5 \leq X \leq 1{,}8) &= \int_{1{,}5}^{1{,}8} f(x)\,\mathrm{d}x = \int_{1{,}5}^{1{,}8} \left(x - \frac{1}{2}\right)\,\mathrm{d}x = \left[\frac{x^2}{2} - \frac{1}{2}\,x\right]_{1{,}5}^{1{,}8} = \\ &= \left(\frac{1{,}8^2}{2} - \frac{1}{2}\,1{,}8\right) - \left(\frac{1{,}5^2}{2} - \frac{1}{2}\,1{,}5\right) = 0{,}345, \end{aligned}$$

$$P(X<1{,}2)=\int_{-\infty}^{1,2} f(x)\,\mathrm{d}x=\int_{-\infty}^{1} 0\,\mathrm{d}x+\int_{1}^{1,2}\left(x-\frac{1}{2}\right)\,\mathrm{d}x=\left[\frac{x^2}{2}-\frac{1}{2}\,x\right]_1^{1,2}=0{,}12,$$

$$P(X>1{,}7)=\int_{1,7}^{\infty} f(x)\,\mathrm{d}x=\int_{1,7}^{2}\left(x-\frac{1}{2}\right)\,\mathrm{d}x+\int_{2}^{\infty} 0\,\mathrm{d}x=\left[\frac{x^2}{2}-\frac{1}{2}\,x\right]_{1,7}^{2}=0{,}405.$$

Pokud jde o $P(X=1{,}5)$, zde můžeme rovnou říct, že výsledek je nula, ale mohli bychom použít i integrál:

$$P(X=1{,}5)=\int_{1,5}^{1,5} f(x)\,\mathrm{d}x=0.$$

Poslední pravděpodobnost, $P(X<3)$, můžeme též určit bez jakéhokoli počítání. Výsledek musí být 1, protože náhodná veličina X menší než 3 určitě je. Pomocí integrálu bychom k výsledku došli takto:

$$\begin{aligned}P(X<3)&=\int_{-\infty}^{3} f(x)\,\mathrm{d}x=\int_{-\infty}^{1} 0\,\mathrm{d}x+\int_{1}^{2}\left(x-\frac{1}{2}\right)\,\mathrm{d}x+\int_{2}^{3} 0\,\mathrm{d}x=\\&=\left[\frac{x^2}{2}-\frac{1}{2}\,x\right]_1^{2}=1.\end{aligned}$$

Nakonec určíme konstantu a, pro kterou je $P(X<a)=0{,}8$ – hledáme vlastně 0,8-kvantil náhodné veličiny X. Je zřejmé, že $a\in(1,2)$. Musí platit

$$P(X<a)=\int_{1}^{a}\left(x-\frac{1}{2}\right)\,\mathrm{d}x=0{,}8.$$

Odtud dostáváme

$$\left[\frac{x^2}{2}-\frac{1}{2}\,x\right]_1^{a}=0{,}8\quad\Rightarrow\quad\frac{a^2}{2}-\frac{1}{2}\,a-\left(\frac{1^2}{2}-\frac{1}{2}\cdot 1\right)=0{,}8\quad\Rightarrow\quad a^2-a-1{,}6=0.$$

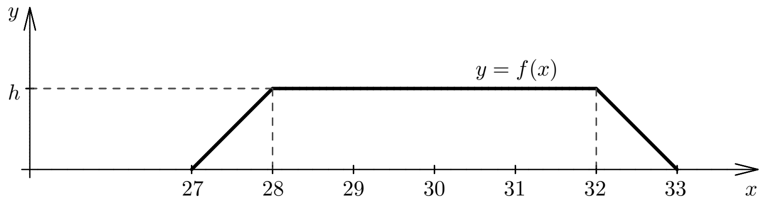
Řešením této kvadratické rovnice s neznámou a dostáváme

$$a_{1,2}=\frac{1\pm\sqrt{1+4\cdot 1{,}6}}{2}=\frac{1\pm\sqrt{7{,}4}}{2}\,.$$

Protože kořen $\frac{1-\sqrt{7,4}}{2}$ nenáleží do intervalu $(1,2)$, zůstává nám jediná možnost, a to

$$a=\frac{1+\sqrt{7{,}4}}{2}\doteq 1{,}86.$$

Příklad 5.2. Teplota ve skleníku je náhodná veličina X s „lichoběžníkovým" rozdělením pravděpodobnosti, graf její hustoty je na obrázku 5.1.
a) Určete hodnotu h vyznačenou v obrázku.
b) Vypočtěte pravděpodobnost, že teplota překročí 31,5°C.
c) Pod jakou mez se teplota dostane jen s pravděpodobností 0,05?

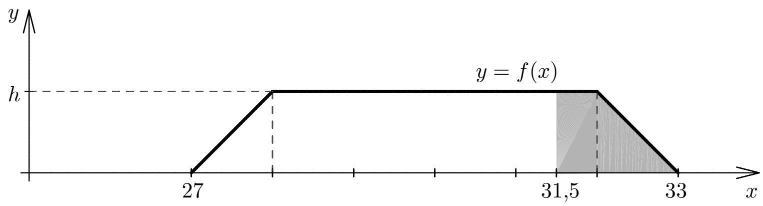


Obr. 5.1: K příkladu 5.2: Hustota zadané náhodné veličiny X

Řešení. a) Pro určení zatím neznámé hodnoty h (výšky lichoběžníka) využijeme faktu, že $P(X \in (-\infty, \infty)) = \int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$. To znamená, že obsah celého lichoběžníka musí být roven 1. Pro výpočet obsahu lichoběžníka platí vztah $S = \frac{h}{2}(a+c)$, kde h je výška a a, c jsou délky základen. V našem případě je $a = 6$, $c = 4$, S má být rovno jedné, a tedy

$$1 = \frac{h}{2}(6+4) \quad \Rightarrow \quad h = \frac{2}{10} = \frac{1}{5} = 0{,}2.$$

b) Máme za úkol vypočítat $P(X > 31{,}5)$. Tato pravděpodobnost je dána integrálem $\int_{31{,}5}^{\infty} f(x)\mathrm{d}x$ neboli obsahem oblasti vyznačené na obrázku 5.2.



Obr. 5.2: K příkladu 5.2, část b)

Můžeme postupovat dvěma způsoby: buď vypočteme přímo obsah oblasti, nebo najdeme funkční předpis pro hustotu, a tu pak zintegrujeme. Předvedeme obě možnosti.
Nejprve pomocí přímého výpočtu obsahu: Oblast se skládá z obdélníka a trojúhelníka. Výška je v obou případech $h = 0{,}2$, šířka obdélníka je 0,5 a základna trojúhelníka má délku 1. Celkem tedy

$$P(X > 31{,}5) = 0{,}5 \cdot 0{,}2 + \frac{1}{2} \cdot 1 \cdot 0{,}2 = 0{,}2.$$

(Tento výsledek jsme mohli určit i „od oka“, bez znalosti hodnoty h. Stačí si uvědomit, že vybarvená část tvoří jednu pětinu celkové plochy a že obsah celého lichoběžníka je 1.)
Nyní vyřešíme stejný problém pomocí integrálu z hustoty: Nejprve musíme najít funkční předpis pro hustotu. Z obrázku vidíme, že graf hustoty se skládá z několika částí. Vně intervalu $\langle 27, 33\rangle$ je $f(x) = 0$, na intervalu $\langle 28, 32\rangle$ je hustota konstantní, $f(x) = 0{,}2$. Na intervalu $\langle 27, 28)$ je grafem hustoty část přímky se směrnicí $k = 0{,}2$ (připomeňme, že směrnice přímky je tangens úhlu, který přímka svírá s kladným směrem osy x, a že tangens se vypočítá jako poměr protilehlé a přilehlé odvěsny pravoúhlého trojúhelníka).

To znamená, že funkční předpis na tomto intervalu bude ve tvaru $y = 0{,}2x + q$. Přímka prochází bodem $[27, 0]$, a proto

$$0 = 0{,}2 \cdot 27 + q \quad \Rightarrow \quad q = -0{,}2 \cdot 27 \quad \Rightarrow \quad y = 0{,}2(x - 27).$$

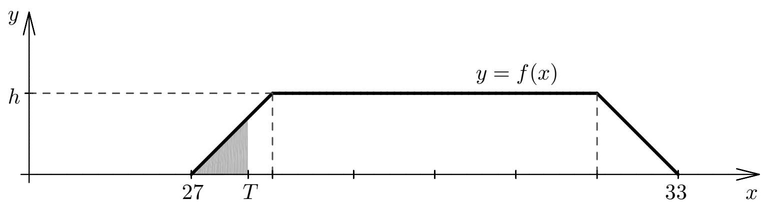
Podobným způsobem bychom zjistili, že pro interval $(32, 33\rangle$ je $f(x) = -0{,}2(x - 33)$. Celkem tedy máme

$$f(x) = \begin{cases} 0 & \text{pro } x < 27, \\ 0{,}2(x-27) & \text{pro } 27 \leq x < 28, \\ 0{,}2 & \text{pro } 28 \leq x \leq 32, \\ -0{,}2(x-33) & \text{pro } 32 < x \leq 33, \\ 0 & \text{pro } x > 33 \end{cases}$$

Požadovanou pravděpodobnost teď vypočteme příslušným integrálem:

$$P(X > 31{,}5) = \int_{31{,}5}^{32} 0{,}2\mathrm{d}x - 0{,}2 \int_{32}^{33} (x - 33)\mathrm{d}x = 0{,}2\,[x]_{31{,}5}^{32} - 0{,}2 \left[\frac{(x-33)^2}{2}\right]_{32}^{33} = 0{,}2.$$

c) Potřebujeme najít mezní hodnotu teploty, označme ji T, pro kterou by platilo $P(X < T) = 0{,}05$. To znamená, že hledáme T, pro které by obsah oblasti vyznačené na obrázku 5.3 byl roven 0,05.



Obr. 5.3: K příkladu 5.2, část c)

Je evidentní, že T bude někde mezi 27 a 28 (protože $P(X < 28) = 0{,}1$, což už je víc než 0,05). Pro výpočet T použijeme hustotu náhodné veličiny X, ale kdo chce, může zkusit najít T pouze pomocí obsahu vyznačeného trojúhelníka.

$$P(X < T) = \int_{27}^{T} 0{,}2(x - 27)\mathrm{d}x = 0{,}2 \left[\frac{(x-27)^2}{2}\right]_{27}^{T} = 0{,}1(T - 27)^2$$

$$0{,}1(T - 27)^2 = 0{,}05 \quad \Rightarrow \quad (T - 27)^2 = 0{,}5 \quad \Rightarrow \quad T - 27 = \pm\sqrt{0{,}5}$$

Protože T je určitě větší než 27, přichází v úvahu pouze $+\sqrt{0{,}5}$. Mezní hodnota, pod kterou teplota klesne jen s pravděpodobností 0,05, je proto $T = 27 + \sqrt{0{,}5} \doteq 27{,}7$.

Distribuční funkce a její vztah s hustotou
Univerzální definice distribuční funkce náhodné veličiny X je

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

U spojité náhodné veličiny se hodnoty distribuční funkce počítají jako

$$F(x) = P(X \in (-\infty, x\rangle) = \int\limits_{-\infty}^{x} f(t)\mathrm{d}t.$$

Hustota f se proto z distribuční funkce F spočítá jako

$$f(x) = F'(x).$$

V bodech, kde $F'(x)$ není definována, můžeme $f(x)$ zvolit libovolně.

Příklad 5.3. Hustota pravděpodobnosti náhodné veličiny X má tvar:

$$f(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ \sin x & \text{pro } 0 < x \leq \frac{\pi}{2}, \\ 0 & \text{pro } x > \frac{\pi}{2}. \end{cases}$$

a) Vypočtěte $P(X \leq -1)$, $P(X \leq \frac{\pi}{4})$, $P(X \leq \frac{\pi}{3})$ a $P(X \leq 2)$.
b) Určete předpis pro distribuční funkci náhodné veličiny X.

Řešení. Hledání distribuční funkce studentům často působí problémy. Proto zde budeme postupovat pomalu a opatrně. Všem, kdo by distribuční funkci uměli najít hned, bez zbytečného zdržování, se omlouváme.
a) Budeme postupovat obdobně jako v příkladu 5.1. Zdá se, že tato část příkladu nepřináší nic nového, je však míněna jako příprava na část b). V rámci této přípravy teď jako integrační proměnnou místo x použijeme t:

$$P(X \leq -1) = \int\limits_{-\infty}^{-1} f(t)\mathrm{d}t = \int\limits_{-\infty}^{-1} 0\mathrm{d}t = 0$$

$$P(X \leq \frac{\pi}{4}) = \int_{-\infty}^{\pi/4} f(t)\mathrm{d}t = \int_{0}^{\pi/4} \sin t\mathrm{d}t = [-\cos t]_0^{\pi/4} = -\frac{\sqrt{2}}{2} - (-1) \doteq 0{,}293$$

$$P(X \leq \frac{\pi}{3}) = \int_{-\infty}^{\pi/3} f(t)\mathrm{d}t = \int_{0}^{\pi/3} \sin t\mathrm{d}t = [-\cos t]_0^{\pi/3} = -\frac{1}{2} - (-1) = 0{,}5$$

$$P(X \leq 2) = \int_{-\infty}^{2} f(t)\mathrm{d}t = \int_{0}^{\pi/2} \sin t\mathrm{d}t + \int_{\pi/2}^{2} 0\mathrm{d}t = [-\cos t]_0^{\pi/2} = -0 - (-1) = 1$$

b) Distribuční funkce je definována jako $F(x) = P(X \leq x)$. To znamená, že v části a) už jsme vypočítali hodnoty $F(-1)$, $F(\pi/4)$, $F(\pi/3)$ a $F(2)$. Zde máme najít obecný předpis pro $F(x)$. Platí

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

To už zde sice bylo uvedeno ve vzorcích v rámečku, při pohledu na řešení části a) ale možná bude jasnější, co se tímto vzorcem myslí. Též už je asi jasné, že distribuční funkce bude vypadat jinak, je-li x (neboli horní mez integrálu) menší než 0, je-li v intervalu $\langle 0, \pi/2\rangle$ a je-li větší než $\pi/2$. Proto výpočet rozdělíme na tři části:

Pro $x < 0$:

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \int_{-\infty}^{x} 0\mathrm{d}t = 0.$$

Pro $x \in \langle 0, \pi/2\rangle$ (reprezentanty tohoto případu byly výpočty pro $x = \pi/4$ a $x = \pi/3$) :

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \int_{-\infty}^{0} 0\mathrm{d}t + \int_{0}^{x} \sin t\mathrm{d}t = [-\cos t]_0^x = -\cos x + 1 = 1 - \cos x.$$

Pro $x > \pi/2$:

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \int_{-\infty}^{0} 0\mathrm{d}t + \int_{0}^{\pi/2} \sin t\mathrm{d}t + \int_{\pi/2}^{x} 0\mathrm{d}t = 1.$$

Celkem jsme dostali předpis pro distribuční funkci

$$F(x) = \begin{cases} 0 & \text{pro } x < 0, \\ 1 - \cos x & \text{pro } 0 \leq x \leq \pi/2, \\ 1 & \text{pro } x > \pi/2. \end{cases}$$

Kdybychom nyní do této funkce dosadili za x např. $\pi/4$, dostali bychom stejnou hodnotu jako v části a). Můžete si též všimnout, že funkce F je spojitá, její jednotlivé části na sebe navazují.

Upozornění na častou chybu: V prostřední fázi výpočtu studenti občas napíšou:
Pro $x \in \langle 0, \pi/2\rangle$ je $F(x) = \int_0^{\pi/2} \sin x\mathrm{d}x = \cdots$
To však není správně. Právě uvedený integrál udává pravděpodobnost, že náhodná veličina X patří do intervalu $\langle 0, \pi/2\rangle$. To ale vůbec není to, co chceme spočítat. Při výpočtu $F(x)$ počítáme pravděpodobnost, že náhodná veličina X je menší nebo rovná x, a v tomto případě víme, že tohle x – horní mez integrálu – je z intervalu $\langle 0, \pi/2\rangle$, jako tomu bylo například pro $x = \pi/3$.

Jiná častá chyba: Některým studentům se vzorec $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$ patrně nelíbí a použít jej nechtějí. Místo toho si řeknou: „Když hustotu f dostanu jako F', tak je F integrál z hustoty a hotovo!“ A napíšou:

$$F(x) = \int f(x)\mathrm{d}x = \int \sin x\mathrm{d}x = -\cos x.$$

To je špatně, což ukážeme na jednoduchém příkladu. Vypočtěme pomocí takto získané distribuční funkce $P(X \leq \pi/3)$:

$$P(X \leq \pi/3) = F(\pi/3) = -\cos(\pi/3) = -0{,}5$$

Pravděpodobnost nám vyšla záporně!! (Pokud někoho tento fakt nezarazil, nechť se vrátí k první kapitole o pravděpodobnosti.)
Oprava této chyby – jiný způsob nalezení $F(x)$: Právě popsaný způsob (nalezení distribuční funkce F pomocí neurčitého integrálu z hustoty f) se ve skutečnosti použít dá, musíme být ale opatrní. Před chvílí jsme totiž zapomněli na integrační konstantu, ono „$+c$“ na závěr. Máme

$$F(x) = \int f(x)\mathrm{d}x = \int \sin x \mathrm{d}x = -\cos x + c.$$

Konstantu c nyní určíme tak, aby hodnoty funkce F vycházely správně. Například víme, že $F(\pi/2)$ musí být 1 (protože $P(X < \pi/2) = 1$). Odtud

$$F(\pi/2) = 1 \quad \Rightarrow \quad -\cos(\pi/2) + c = 1 \quad \Rightarrow \quad 0 + c = 1 \quad \Rightarrow \quad c = 1.$$

Distribuční funkce pro $x \in \langle 0, \pi/2 \rangle$ je proto

$$F(x) = -\cos x + 1.$$

Stejně dobře jsme mohli pro určení c využít faktu, že $F(0)$ musí být 0 (pro tento konkrétní příklad; obecně to být pravda nemusí). Opět bychom dostali, že $c = 1$.

Nyní předvedeme ještě jeden příklad na hledání distribuční funkce. Hustota tentokrát bude rozdělena na více částí.

Příklad 5.4. Najděte distribuční funkci náhodné veličiny X z příkladu 5.2 (příklad se skleníkem).

Řešení. Už jsme zjistili, že hustota zkoumané náhodné veličiny X je

$$f(x) = \begin{cases} 0 & \text{pro } x < 27, \\ 0{,}2(x-27) & \text{pro } 27 \leq x < 28, \\ 0{,}2 & \text{pro } 28 \leq x \leq 32, \\ -0{,}2(x-33) & \text{pro } 32 < x \leq 33, \\ 0 & \text{pro } x > 33 \end{cases}$$

Budeme hledat distribuční funkci pro jednotlivé intervaly:
Pro $x < 27$ je zřejmě $F(x) = 0$.
Pro $x \in \langle 27, 28)$:

$$F(x) = \int_{27}^{x} 0{,}2(t-27)\mathrm{d}t = 0{,}2\left[\frac{(t-27)^2}{2}\right]_{27}^{x} = 0{,}1(x-27)^2.$$

(Poznamenejme, že primitivní funkce se samozřejmě mohla vyjádřit i jako $0{,}2(\frac{t^2}{2} - 27t)$. Dosazení mezí by pak vedlo k „ošklivějšímu“ tvaru výsledku, do kterého by se pracněji dosazovaly konkrétní hodnoty x.)
Pro $x \in \langle 28, 32\rangle$:

$$\begin{aligned} F(x) &= \int_{27}^{28} 0{,}2(t-27)\mathrm{d}t + \int_{28}^{x} 0{,}2\mathrm{d}t = 0{,}2\left[\frac{(t-27)^2}{2}\right]_{27}^{28} + 0{,}2\,[t]_{28}^{x} = \\ &= 0{,}1 + 0{,}2(x-28). \end{aligned}$$

Pro $x \in \langle 32, 33\rangle$:

$$\begin{aligned} F(x) &= \int_{27}^{28} 0{,}2(t-27)\mathrm{d}t + \int_{28}^{32} 0{,}2\mathrm{d}t + \int_{32}^{x} (-0{,}2)(t-33)\mathrm{d}t = \\ &= 0{,}2\left[\frac{(t-27)^2}{2}\right]_{27}^{28} + 0{,}2\,[t]_{28}^{32} - 0{,}2\left[\frac{(t-33)^2}{2}\right]_{32}^{x} = 0{,}1 + 0{,}2\cdot 4 - \\ &\quad -0{,}2\left[\frac{(t-33)^2}{2}\right]_{32}^{x} = 0{,}9 - 0{,}1((x-33)^2 - 1) = 1 - 0{,}1(x-33)^2. \end{aligned}$$

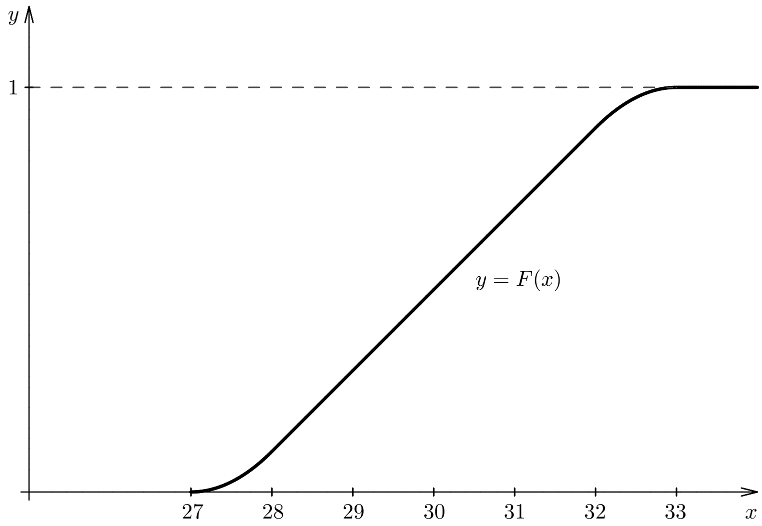
Pro $x > 33$ můžeme říci rovnou, že bude $F(x) = 1$. Kdo by však chtěl vidět výpočet rozepsaný, má příležitost:

$$\begin{aligned} F(x) &= \int_{27}^{28} 0{,}2(t-27)\mathrm{d}t + \int_{28}^{32} 0{,}2\mathrm{d}t + \int_{32}^{33} (-0{,}2)(t-33)\mathrm{d}t + \int_{33}^{x} 0\mathrm{d}t = \\ &= 0{,}2\left[\frac{(t-27)^2}{2}\right]_{27}^{28} + 0{,}2\,[t]_{28}^{32} - 0{,}2\left[\frac{(t-33)^2}{2}\right]_{32}^{33} = \\ &= 0{,}1 + 0{,}2\cdot 4 - 0{,}1(0-1) = 1. \end{aligned}$$

Celkem jsme dostali pro distribuční funkci předpis

$$F(x) = \begin{cases} 0 & \text{pro } x < 27, \\ 0{,}1(x-27)^2 & \text{pro } 27 \leq x < 28, \\ 0{,}1 + 0{,}2(x-28) & \text{pro } 28 \leq x \leq 32, \\ 1 - 0{,}1(x-33)^2 & \text{pro } 32 < x \leq 33, \\ 1 & \text{pro } x > 33. \end{cases}$$

Graf distribuční funkce vidíme na obrázku 5.4
Můžete se pokusit právě nalezenou distribuční funkci spočítat znovu, ale bez integrování, pomocí obrázku 5.1. Stačí si uvědomit, že při výpočtu $F(x)$ neboli $P(X \leq x)$ počítáme vlastně obsah plochy pod grafem hustoty f od $-\infty$ po příslušné x. Např. je-li $x \in \langle 28, 32\rangle$, bereme obsah levého trojúhelníka (což je 0,1) a přidáváme k němu obsah obdélníka s výškou 0,2 a šířkou $x - 28$. Výsledek je pak $F(x) = 0{,}1 + 0{,}2(x-28)$ – porovnejte s výsledkem získaným integrací.



Obr. 5.4: Distribuční funkce náhodné veličiny z příkladu 5.4

Příklad 5.5. Náhodná veličina X má distribuční funkci

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} x$$

a) Vypočtěte následující pravděpodobnosti: $P(X < 1)$, $P(X > 1{,}5)$, $P(0{,}5 < X \leq 1{,}5)$ a $P(-0{,}5 \leq X \leq 1)$.
b) Najděte hodnotu x, kterou náhodná veličina X překročí (směrem nahoru) jen s pravděpodobností 0,01.
c) Najděte interval souměrný podle počátku, do kterého náhodná veličina X padne s pravděpodobností 0,9.
d) Určete hustotu pravděpodobnosti náhodné veličiny X.

Řešení. a) Ještě jednou připomeňme, že $F(x) = P(X \leq x)$. V zadání je ale nerovnost ostrá, $X < 1$. Protože však u spojitých náhodných veličin platí $P(X = x) = \int_x^x f(t)\mathrm{d}t = 0$, nemusíme na to brát ohled. Proto

$$P(X < 1) = P(X \leq 1) - P(X = 1) = F(1) - 0 = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} 1 = \frac{1}{2} + \frac{1}{\pi}\cdot\frac{\pi}{4} \doteq 0{,}75.$$

Pro výpočet $P(X > 1{,}5)$ použijeme pravděpodobnost jevu opačného. Opačný jev k jevu $[X > 1{,}5]$ je $[X \leq 1{,}5]$, a tedy

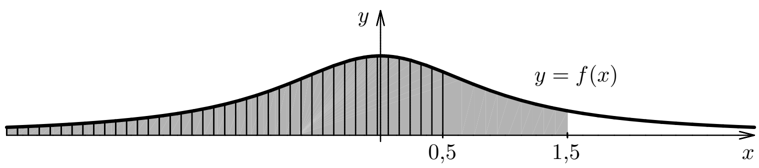
$$P(X > 1{,}5) = 1 - P(X \leq 1{,}5) = 1 - F(1{,}5) = 1 - \left(\frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} 1{,}5\right) \doteq 0{,}187.$$

Má-li být $0{,}5 < X \leq 1{,}5$, znamená to, že X musí být menší nebo rovné 1,5, a přitom nesmí být menší nebo rovné 0,5. Proto

$$P(0{,}5 < X \leq 1{,}5) \quad = \quad P(X \leq 1{,}5) - P(X \leq 0{,}5) = F(1{,}5) - F(0{,}5) =$$

$$= \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} 1{,}5 - \left(\frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} 0{,}5\right) \doteq 0{,}165.$$

Protože právě předvedená úvaha některým studentům činí potíže, vysvětlíme vše ještě pomocí obrázku 5.5. Na tomto obrázku je znázorněna hustota náhodné veličiny X (funkční předpis pro ni zatím neznáme, ale to teď nijak nevadí). Jak víme, pravděpodobnost, že $0{,}5 < X \leq 1{,}5$, je rovna obsahu plochy pod grafem hustoty na intervalu $(0{,}5; 1{,}5\rangle$. Dále, hodnota distribuční funkce v bodě 0,5, tj. pravděpodobnost, že $X \leq 0{,}5$, je rovna obsahu plochy pod grafem hustoty na intervalu $(-\infty; 0{,}5\rangle$. V našem obrázku je tato plocha vyznačena svislým šrafováním. Podobně, hodnota distribuční funkce v bodě 1,5 je rovna obsahu plochy, která je v obrázku 5.5 šedě vybarvena. Nás zajímá obsah plochy, která je šedá, ale nikoli šrafovaná. Opět se dostáváme k tomu, že od sebe musíme odečíst $F(1{,}5)$ a $F(0{,}5)$.



Obr. 5.5: K příkladu 5.5 – hustota náhodné veličiny X

Ještě zbývá vypočítat $P(-0{,}5 \leq X \leq 1)$. Protože pravděpodobnosti, že by se X rovnalo nějaké jedné konkrétní hodnotě, jsou nulové, můžeme tento příklad řešit stejně jako ten předchozí:

$$P(-0{,}5 \leq X \leq 1) = F(1) - F(-0{,}5) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} 1 - \left(\frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}(-0{,}5)\right) \doteq 0{,}398.$$

b) Hledáme x, pro které by platilo $P(X > x) = 0{,}01$. Už jsme ukázali, že $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$. Proto musíme najít x, pro které bude $1 - F(x) = 0{,}01$. Dosazením do funkce F dostáváme:

$$1 - \left(\frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} x\right) = 0{,}01 \quad \Rightarrow \quad \operatorname{arctg} x = \pi\left(\frac{1}{2} - 0{,}01\right)$$

Tedy

$$x = \operatorname{tg}(0{,}49\pi) \doteq 31{,}82.$$

c) Hledáme interval, označme jej $(-a, a)$, pro který by platilo $P(-a < X < a) = 0{,}9$. Při výpočtu využijeme faktu, že funkce arctg je lichá.

$$\begin{aligned}
P(-a < X < a) &= F(a) - F(-a) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} a - \left(\frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}(-a)\right) = \\
&= \frac{1}{\pi}(\operatorname{arctg} a - \operatorname{arctg}(-a)) = \frac{2}{\pi}\operatorname{arctg} a
\end{aligned}$$

$$\frac{2}{\pi}\operatorname{arctg} a = 0{,}9 \quad \Rightarrow \quad a = \operatorname{tg}\frac{0{,}9\pi}{2} \doteq 6{,}31$$

Hledaný interval je tedy $(-6{,}31; 6{,}31)$.
d) Platí, že $f(x) = F'(x)$, a tedy v našem případě

$$f(x) = \left(\frac{1}{2} + \frac{1}{\pi} \operatorname{arctg} x\right)' = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

Kdo by chtěl, může teď části příkladu a), b), c) vyřešit pomocí hustoty.

V předchozím příkladu jsme ukázali výpočty pravděpodobnosti různých typů nerovností. Vše shrneme do rámečku:

Výpočty různých pravděpodobností pomocí distribuční funkce
Pro jakoukoli náhodnou veličinu platí

$$\begin{aligned} P(X \le x) &= F(x) \\ P(X > x) &= 1 - F(x) \\ P(a < X \le b) &= F(b) - F(a). \end{aligned}$$

Protože pro spojité náhodné veličiny je $P(X = x) = 0$, můžeme pro spojité náhodné veličiny všude nahradit ostré nerovnosti neostrými a naopak.

Střední hodnota, rozptyl a směrodatná odchylka
Střední hodnota spojité náhodné veličiny X se vypočítá jako

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

rozptyl jako

$$\mathrm{d}x = \int_{-\infty}^{\infty} x^2 \cdot f(x)\mathrm{d}x - (\mathrm{E}X)^2$$

a směrodatná odchylka je $\sqrt{\mathrm{d}x}$.

Příklad 5.6. Hustota pravděpodobnosti náhodné veličiny X má tvar:

$$f(x) = \begin{cases} 0 & \text{pro } x \le 0, \\ \frac{1}{2}\sin x & \text{pro } 0 < x \le \pi, \\ 0 & \text{pro } x > \pi. \end{cases}$$

a) Vypočtěte střední hodnotu, rozptyl a směrodatnou odchylku náhodné veličiny X.
b) Vypočtěte pravděpodobnost, že náhodná veličina X překročí (směrem nahoru) svou střední hodnotu více než o $\pi/4$.
c) Vypočtěte pravděpodobnost, že se náhodná veličina X bude od své střední hodnoty lišit nanejvýš o dvojnásobek směrodatné odchylky.

Řešení. a) Střední hodnota:

$$\begin{aligned}\mathrm{E}X &= \int_{-\infty}^{\infty} x\cdot f(x)\mathrm{d}x = \frac{1}{2}\int_0^{\pi} x\sin x\mathrm{d}x \quad \begin{vmatrix} u = x & u' = 1 \\ v' = \sin x & v = -\cos x\end{vmatrix} \\ &= \frac{1}{2}\left([-x\cos x]_0^{\pi} + \int_0^{\pi}\cos x\mathrm{d}x\right) = \frac{1}{2}\left(\pi + [\sin x]_0^{\pi}\right) = \frac{\pi}{2}\,.\end{aligned}$$

Tento výsledek jsme mohli i uhodnout, protože graf hustoty je souměrný podle přímky $x = \pi/2$, takže se dá čekat, že „průměrně“ bude náhodná veličina X nabývat hodnoty $\pi/2$.

Rozptyl:

$$\mathrm{d}x = \int_{-\infty}^{\infty} x^2\cdot f(x)\mathrm{d}x - (\mathrm{E}X)^2 = \frac{1}{2}\int_0^{\pi} x^2\sin x\mathrm{d}x - \left(\frac{\pi}{2}\right)^2.$$

Nejprve zvlášť vypočteme integrál z $x^2\sin x$, a pak se vrátíme k výpočtu DX:

$$\int_0^{\pi} x^2\sin x\mathrm{d}x = \begin{vmatrix} u = x^2 & u' = 2x \\ v' = \sin x & v = -\cos x\end{vmatrix} = \left[-x^2\cos x\right]_0^{\pi} + 2\int_0^{\pi} x\cos x\mathrm{d}x =$$

$$= \begin{vmatrix} u = x & u' = 1 \\ v' = \cos x & v = \sin x\end{vmatrix} = \pi^2 + 2\left([x\sin x]_0^{\pi} - \int_0^{\pi}\sin x\mathrm{d}x\right) = \pi^2 - 4$$

Rozptyl je pak

$$\mathrm{d}x = \frac{1}{2}\left(\pi^2 - 4\right) - \left(\frac{\pi}{2}\right)^2 = \frac{\pi^2}{4} - 2 \doteq 0{,}467.$$

Směrodatná odchylka:

$$\sqrt{\mathrm{d}x} = \sqrt{\frac{\pi^2}{4} - 2} \doteq 0{,}684.$$

Upozornění na častou chybu: Při výpočtu rozptylu často člověk správně zapíše začátek výpočtu: $\mathrm{d}x = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x - (\mathrm{E}X)^2$, ale pak se soustředí na výpočet integrálu a na odečtení $(\mathrm{E}X)^2$ zapomene. Nevíme, jak této chybě zabránit. Snad jen doporučíme čtenáři, ať si poctivě počítá příklady. Jestliže se této chyby párkrát dopustí, dokud je to „nanečisto“, při písemce se mu to snad už nestane.

b) Budeme počítat pravděpodobnost, že X bude větší než $\mathrm{E}X + \pi/4$:

$$\begin{aligned}P(X > \mathrm{E}X + \tfrac{\pi}{4}) &= P(X > \tfrac{3\pi}{4}) = \tfrac{1}{2}\int_{3\pi/4}^{\pi}\sin x\mathrm{d}x = \tfrac{1}{2}\left[-\cos x\right]_{3\pi/4}^{\pi} = \\ &= -\frac{1}{2}\left(-1 - \left(-\frac{\sqrt{2}}{2}\right)\right) = \frac{2-\sqrt{2}}{4} \doteq 0{,}146.\end{aligned}$$

c) Vypočteme pravděpodobnost, že X bude v intervalu $\left\langle \mathrm{E}X - 2\sqrt{\mathrm{d}x}, \mathrm{E}X + 2\sqrt{\mathrm{d}x}\right\rangle$:

$$P(\mathrm{E}X - 2\sqrt{\mathrm{d}x} \le X \le \mathrm{E}X + 2\sqrt{\mathrm{d}x}) = P(\tfrac{\pi}{2} - 2\sqrt{\tfrac{\pi^2}{4} - 2} \le X \le \tfrac{\pi}{2} + 2\sqrt{\tfrac{\pi^2}{4} - 2}) \doteq$$

$$P(0{,}636 \le X \le 2{,}506) = \tfrac{1}{2}\int_{0{,}636}^{2{,}506}\sin x\mathrm{d}x \doteq 0{,}804.$$

Na výpočet střední hodnoty a rozptylu předvedeme ještě jeden příklad, tentokrát s hustotou definovanou po částech.

Příklad 5.7. Hustota pravděpodobnosti náhodné veličiny X má tvar:

$$f(x) = \begin{cases} \frac{1}{6}\,x & \text{pro } x \in \langle 0,3\rangle \\ -\frac{1}{2}(x-4) & \text{pro } x \in (3,4\rangle \\ 0 & \text{jinak.} \end{cases}$$

a) Vypočtěte střední hodnotu a rozptyl náhodné veličiny X.
b) Vypočtěte $P(X < \mathrm{E}X)$, $P(X > \mathrm{E}X + 1)$ a $P(X > 2\mathrm{E}X)$.

Řešení. a) Střední hodnota:

$$\begin{aligned} \mathrm{E}X &= \int_{-\infty}^{\infty} x \cdot f(x)\mathrm{d}x = \int_0^3 \tfrac{1}{6}\,x^2 \mathrm{d}x + \int_3^4 \left(-\tfrac{1}{2}(x^2-4x)\right)\mathrm{d}x = \\ &= \frac{1}{6}\left[\frac{x^3}{3}\right]_0^3 - \frac{1}{2}\left[\frac{x^3}{3}\right]_3^4 + 2\left[\frac{x^2}{2}\right]_3^4 = \tfrac{27}{18} - \tfrac{1}{6}(64-27) + 16 - 9 = \tfrac{7}{3} \doteq 2{,}333 \end{aligned}$$

Rozptyl:

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

Integrál z $x^2 f(x)$ vypočteme zvlášť pro jednotlivé intervaly a pak vše dáme dohromady:

$$\int_0^3 \tfrac{1}{6}\,x^3 \mathrm{d}x = \frac{1}{6}\left[\frac{x^4}{4}\right]_0^3 = \tfrac{1}{24}\,(81-0) = \tfrac{27}{8}$$

$$\int_3^4 \left(-\tfrac{1}{2}(x^3-4x^2)\right)\mathrm{d}x = -\frac{1}{2}\left[\frac{x^4}{4}\right]_3^4 + 2\left[\frac{x^3}{3}\right]_3^4 = -\tfrac{1}{8}\,(256-81) + \tfrac{2}{3}\,(64-27) = \tfrac{67}{24}$$

Celkem tedy

$$\int_{-\infty}^{\infty} x^2 \cdot f(x)\mathrm{d}x = \tfrac{27}{8} + \tfrac{67}{24} = \tfrac{148}{24} = \tfrac{37}{6}\,,$$

$$\mathrm{d}x = \tfrac{37}{6} - \left(\tfrac{7}{3}\right)^2 = \tfrac{13}{18} \doteq 0{,}722.$$

b) Při výpočtu zadaných pravděpodobností musíme dbát na to, ve kterém intervalu se pohybujeme. Uvědomme si, že $10/3 \doteq 3{,}333$ a $14/3 \doteq 4{,}667$.

$$P(X < \mathrm{E}X) = P\left(X < \tfrac{7}{3}\right) = \int_0^{7/3} \tfrac{1}{6}\,x\mathrm{d}x = \frac{1}{6}\left[\frac{x^2}{2}\right]_0^{7/3} = \tfrac{1}{12}\cdot\left(\tfrac{7}{3}\right)^2 = \tfrac{49}{108} \doteq 0{,}454$$

$$\begin{aligned} P(X > \mathrm{E}X + 1) &= P\left(X > \tfrac{10}{3}\right) = \int_{10/3}^{4} \left(-\tfrac{1}{2}(x-4)\right)\mathrm{d}x = -\tfrac{1}{4}\left[x^2\right]_{10/3}^4 + 2\left[x\right]_{10/3}^4 = \\ &= -\tfrac{1}{4}\left(16 - \tfrac{100}{9}\right) + 2\left(4 - \tfrac{10}{3}\right) = \tfrac{1}{9} \doteq 0{,}111 \end{aligned}$$

$$P(X > 2\mathrm{E}X) = P\left(X > \tfrac{14}{3}\right) = 0.$$

5.2 Příklady pro samostatnou práci

Příklad 5.8. Náhodná veličina X má hustotu

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

Vypočtěte pravděpodobnosti: a) $P(X \in (1/2, 3/4))$; b) $P(X < 0{,}3)$; c) $P(X > 4/5)$; d) $P(X < 2)$; e) $P(X > 3)$.

Výsledek: a) $35/128 \doteq 0{,}273$; b) 0,1635; c) $43/125 \doteq 0{,}344$; d) 1; e) 0

Příklad 5.9. Náhodná veličina X má distribuční funkci

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 3, \\ \frac{1}{3}x - 1 & \text{pro } 3 < x \leq 6, \\ 1 & \text{pro } x > 6. \end{cases}$$

Vypočtěte pravděpodobnosti: a) $P(X < 4)$; b) $P(X > 5{,}5)$; c) $P(3{,}5 < X < 5)$; d) $P(X > 2)$; e) $P(X > 7)$.

Výsledek: a) 1/3; b) 1/6; c) 1/2; d) 1; e) 0

Příklad 5.10. Hustota pravděpodobnosti náhodné veličiny X má tvar:

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

Najděte distribuční funkci náhodné veličiny X a pak pomocí ní vypočítejte tytéž pravděpodobnosti, jaké se počítaly v příkladu 5.1 (všimněte si, že jde o náhodnou veličinu se stejnou hustotou).

Výsledek: $F(x) = 0$ pro $x \leq 1$; $F(x) = (x^2 - x)/2$ pro $1 < x \leq 2$; $F(x) = 1$ pro $x > 2$. Pravděpodobnosti viz příklad 5.1.

Příklad 5.11. Je dána funkce

$$f(x) = \begin{cases} a - x^2 & \text{pro } x \in \langle 0, 1\rangle, \\ 0 & \text{jinak.} \end{cases}$$

a) Určete konstantu a tak, aby funkce $f(x)$ byla hustotou pravděpodobnosti nějaké náhodné veličiny.
b) Určete střední hodnotu a rozptyl příslušné náhodné veličiny.

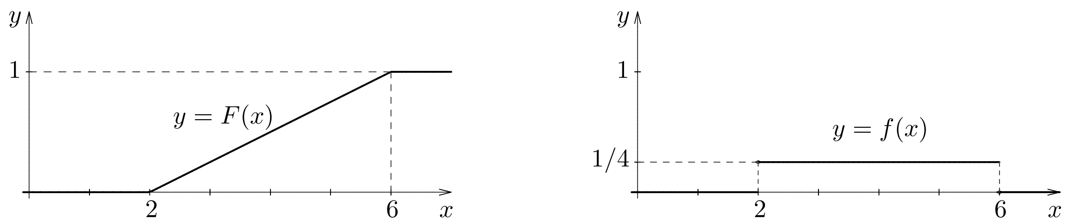
Výsledek: a) $a = 4/3$ (najde se na základě podmínky $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$); b)$\mathrm{E}X = 5/12$, $\mathrm{d}x = 17/240$.

Příklad 5.12. Náhodná veličina X má distribuční funkci

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 2, \\ \frac{1}{4}(x-2) & \text{pro } 2 < x \leq 6, \\ 1 & \text{pro } x > 6. \end{cases}$$

a)Určete hustotu pravděpodobnosti náhodné veličiny X.
b)Znázorněte graficky hustotu a distribuční funkci.
c)Určete střední hodnotu, rozptyl a směrodatnou odchylku.

Výsledek: a) $f(x) = 1/4$ pro $2 < x < 6$; $f(x) = 0$ jinak; b) viz obrázek 5.6; c) $\mathrm{E}X = 4$, $\mathrm{d}x = 4/3$, $\sqrt{\mathrm{d}x} = 2\sqrt{3}/3 \doteq 1{,}155$



Obr. 5.6: K příkladu 5.12 – distribuční funkce a hustota náhodné veličiny X

Příklad 5.13. Náhodná veličina X má distribuční funkci

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ x^2 & \text{pro } 0 < x \leq 1, \\ 1 & \text{pro } x > 1. \end{cases}$$

a) Určete hodnotu a, kterou X překročí směrem dolů jen s pravděpodobností 0,25.
b) Určete hodnotu b, kterou X překročí směrem nahoru jen s pravděpodobností 0,1.
c) Určete střední hodnotu a rozptyl náhodné veličiny X.

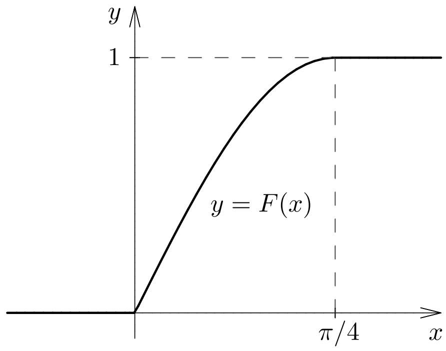
Výsledek: a) $a = 0{,}5$; b) $b = 3\sqrt{10}/10 \doteq 0{,}949$; c) $\mathrm{E}X = 2/3$, $\mathrm{d}x = 1/18$

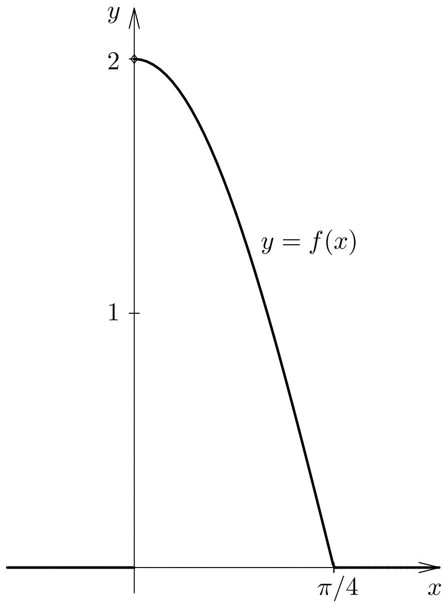
Příklad 5.14. Náhodná veličina X má distribuční funkci

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ \sin 2x & \text{pro } 0 < x \leq \frac{\pi}{4}, \\ 1 & \text{pro } x > \frac{\pi}{4}. \end{cases}$$

a)Určete hustotu pravděpodobnosti náhodné veličiny X.
b)Znázorněte graficky hustotu a distribuční funkci.
c)Určete střední hodnotu a rozptyl.

Výsledek: a) $f(x) = 2\cos 2x$ pro $0 < x < \pi/4$, $f(x) = 0$ jinak; b) viz obrázek 5.6;
c) $\mathrm{E}X = -1/2 + \pi/4$, $\mathrm{d}x = -3/4 + \pi/4$





Obr. 5.7: K příkladu 5.14 – distribuční funkce a hustota náhodné veličiny X

Příklad 5.15. Je dána funkce

$$f(x) = \begin{cases} \frac{a}{x} & \text{pro } x \in \langle 1, \mathrm{e} \rangle, \\ 0 & \text{jinak.} \end{cases}$$

a) Určete konstantu a tak, aby funkce $f(x)$ byla hustotou pravděpodobnosti.
b) Určete předpis pro distribuční funkci příslušné náhodné veličiny.
c) Určete střední hodnotu a rozptyl.

Výsledek: a) $a = 1$; b) $F(x) = 0$ pro $x < 1$, $F(x) = \ln x$ pro $1 < x < \mathrm{e}$, $F(x) = 1$ pro $x > \mathrm{e}$; c) $\mathrm{E}X = \mathrm{e} - 1$, $\mathrm{d}x = -\mathrm{e}^2/2 + 2\mathrm{e} - 3/2 \doteq 0{,}242$

Příklad 5.16. Hustota pravděpodobnosti náhodné proměnné X má tvar

$$f(x) = \begin{cases} c \cos x & \text{pro } -\frac{\pi}{2} < x \leq \frac{\pi}{2}, \\ 0 & \text{jinak.} \end{cases}$$

a) Určete konstantu c.
b) Určete předpis pro distribuční funkci $F(x)$.
c) Najděte interval souměrný kolem nuly, ve kterém náhodná veličina X bude ležet s pravděpodobností 0,95

Výsledek: a) $c = 1/2$; b) $F(x) = 0$ pro $x < -\pi/2$, $F(x) = (1 + \sin x)/2$ pro $-\pi/2 \leq x < \pi/2$, $F(x) = 1$ pro $x \geq \pi/2$; c) $(-\arcsin(19/20), \arcsin(19/20)) \doteq (-1{,}253;\ 1{,}253)$

Příklad 5.17. Hustota pravděpodobnosti náhodné proměnné X má tvar

$$f(x) = \begin{cases} 6x(1-x) & \text{pro } 0 < x \leq 1, \\ 0 & \text{jinak.} \end{cases}$$

a) Ověřte, že funkce f opravdu může být hustotou nějaké náhodné veličiny.
b) Určete předpis pro distribuční funkci $F(x)$.
c) Určete střední hodnotu a rozptyl.
d) Určete pravděpodobnost, že se náhodná veličina od své střední hodnoty liší více než o $1/3$.

Výsledek: a) ano, f je nezáporná funkce a $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$; b) $F(x) = 0$ pro $x < 0$, $F(x) = -2x^3 + 3x^2$ pro $0 \leq x < 1$, $F(x) = 1$ pro $x \geq 1$; c) $\mathrm{E}X = 1/2$, $\mathrm{d}x = 1/20$; d) $4/27 \doteq 0{,}148$

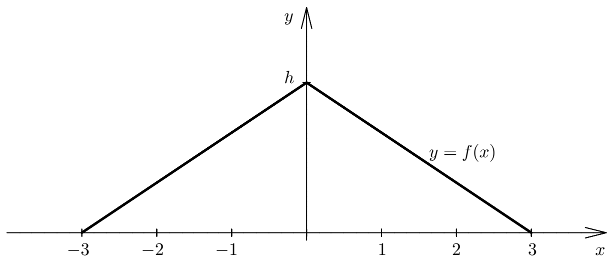
Příklad 5.18. Hustota pravděpodobnosti náhodné proměnné X má tvar

$$f(x) = \begin{cases} a(3x-4) & \text{pro } 1 < x \leq 2, \\ 0 & \text{jinak.} \end{cases}$$

a) Určete konstantu a a pak načrtněte graf funkce f.
b) Určete předpis pro její distribuční funkci $F(x)$.
c) Určete $P(0 < X < \frac{1}{2})$.
d) Určete střední hodnotu a rozptyl.

Výsledek: Příklad nemá řešení. Z podmínky $\int_{\infty}^{\infty} f(x)\mathrm{d}x = 1$ by vyšlo $a = 2$, jenže funkce f nabývá na části intervalu $\langle 1, 2\rangle$ záporných hodnot, což se u hustoty nesmí stát. Části b), c), d) proto nemá význam počítat.

Příklad 5.19. Chyba určitého měření je náhodná veličina X s „trojúhelníkovým" rozdělením pravděpodobnosti. Graf její hustoty je na obrázku 5.8.
a) Určete hodnotu h.
b) Najděte funkční předpis pro hustotu f.
c) Najděte funkční předpis pro distribuční funkci F.
d) Vypočtěte pravděpodobnost, že chyba bude v intervalu $(-1, 1)$.
e) Najděte interval souměrný kolem nuly, v němž bude chyba s pravděpodobností 0,99.



Obr. 5.8: K příkladu 5.19 – hustota náhodné veličiny X

Výsledek: a) $h = 1/3$; b) $f(x) = (x+3)/9$ pro $-3 < x < 0$, $f(x) = -(x-3)/9$ pro $0 \leq x < 3$, $f(x) = 0$ jinak; c) $F(x) = 0$ pro $x < -3$, $F(x) = (x+3)^2/18$ pro $-3 \leq x < 0$, $F(x) = 1 - (x-3)^2/18$ pro $0 \leq x < 3$, $F(x) = 1$ pro $x \geq 3$; d) $5/9$; e) $(-2{,}7; 2{,}7)$

Příklad 5.20. Náhodná proměnná X má distribuční funkci:

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 1, \\ \ln x & \text{pro } 1 < x \leq a, \\ 1 & \text{pro } x > a. \end{cases}$$

a) Určete konstantu a.
b) Určete hustotu pravděpodobnosti náhodné proměnné.
c) Určete střední hodnotu a rozptyl.

Výsledek:
a) $a = \mathrm{e}$,
b) $f(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 1\rangle \cup (\mathrm{e}, \infty), \\ \frac{1}{x} & \text{pro } 1 < x \leq \mathrm{e}, \end{cases}$ c) $\mathrm{E}X = \mathrm{e} - 1, \quad \mathrm{d}x = \frac{(\mathrm{e}-1)(3-\mathrm{e})}{2}$

Příklad 5.21. Náhodná proměnná X má distribuční funkci:

$$F(x) = \begin{cases} 0 & \text{pro } x \leq -2, \\ \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{2} & \text{pro } -2 < x \leq 2, \\ 1 & \text{pro } x > 2. \end{cases}$$

a)Určete hustotu pravděpodobnosti náhodné proměnné.
b)Pravděpodobnost toho, že náhodná proměnná nabýva hodnoty z intervalu $(-1, 1)$.

Výsledek: a) $f(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, -2\rangle \cup (2, \infty) \\ \frac{1}{\pi\sqrt{4-x^2}} & \text{pro } -2 < x \leq 2, \end{cases}$ b) $\frac{1}{3}$.

Příklad 5.22. Náhodná proměnná X má distribuční funkci:

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ a + b \sin x & \text{pro } 0 < x \leq \frac{\pi}{2}, \\ 1 & \text{pro } x > \frac{\pi}{2}. \end{cases}$$

a) Určete konstanty a, b.
b) Určete hustotu pravděpodobnosti náhodné proměnné.
c) Určete $P(0 < X < \frac{\pi}{4})$.

Výsledek:
a) $a = 0, b = 1$, b) $f(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0\rangle \cup \left(\frac{\pi}{2}, \infty\right), \\ \cos x & \text{pro } 0 < x \leq \frac{\pi}{2}, \end{cases}$ c) $P(0 < X < \frac{\pi}{4}) = \frac{1}{\sqrt{2}}$.

Příklad 5.23. Náhodná proměnná X má distribuční funkci:

$$F(x) = \begin{cases} a + b.\mathrm{e}^{-x} & \text{pro } x > 0, \\ 0 & \text{pro } x \leq 0, \end{cases}$$

a) Určete konstanty a, b.
b) Určete hustotu pravděpodobnosti náhodné proměnné.
c) Určete $P(0 < X < 3)$.

Výsledek: a) $a = 1, b = -1$, b) $f(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ \mathrm{e}^{-x} & \text{pro } x > 0, \end{cases}$ c) $P(0 < X < 3) = 1 - \frac{1}{\mathrm{e}^3}$.

Příklad 5.24. Náhodná proměnná X má distribuční funkci:

$$F(x) = \begin{cases} a + \frac{b}{1+x^2} & \text{pro } x > 0, \\ 0 & \text{pro } x \leq 0, \end{cases}$$

a) Určete konstanty a, b.
b) Určete hustotu pravděpodobnosti náhodné proměnné.

Výsledek: a) $a = 1, b = -1$, b) $f(x) = \frac{2}{(1+x^2)^2}$, $\forall x > 0$.

Příklad 5.25. Náhodná proměnná X má distribuční funkci:

$$F(x) = a + b \operatorname{arctg} ax$$

a) Určete konstanty a, b.
b) Určete hustotu pravděpodobnosti náhodné proměnné.

Výsledek: a) $a = \frac{1}{2}$, $b = \frac{1}{\pi}$, b) $f(x) = \frac{2}{\pi(x^2+4)}$; $\forall x \in \mathbb{R}$.

Příklad 5.26. Hustota pravděpodobnosti náhodné proměnné X má tvar:

$$f(x) = \begin{cases} 0 & \text{pro } x \leq 1, \\ a(x-1) & \text{pro } 1 < x \leq 2, \\ a(x-2) & \text{pro } 2 < x \leq 3, \\ 0 & \text{pro } x > 3. \end{cases}$$

a) Určete konstantu a.
b) Určete předpis pro její distribuční funkci $F(x)$.
c) Určete $P(1 < X < \frac{3}{2})$.

Výsledek:

a) $a = 1$, b) $F(x) = \begin{cases} 0 & \text{pro } x \leq 1, \\ \frac{x^2}{2} - x + \frac{1}{2} & \text{pro } 1 < x \leq 2, \\ \frac{x^2}{2} - 2x + \frac{5}{2} & \text{pro } 2 < x \leq 3, \\ 1 & \text{pro } x > 3, \end{cases}$ c) $P(1 < X < \frac{3}{2}) = \frac{1}{8}$.

Příklad 5.27. Hustota pravděpodobnosti náhodné proměnné X má tvar:

$$f(x)=\begin{cases}0 & \text{pro } x\le -1,\\ a(x+1) & \text{pro } -1<x\le 0,\\ a(1-x) & \text{pro } 0<x\le 1,\\ 0 & \text{pro } x>1.\end{cases}$$

a) Určete konstantu a.
b) Určete předpis pro její distribuční funkci $F(x)$.
c) Určete $P(0\le X<\frac{1}{2})$.

Výsledek: a) $a=1$, b) $F(x)=\begin{cases}0 & \text{pro } x\le -1,\\ \frac{x^2}{2}+x+\frac{1}{2} & \text{pro } -1<x\le 0,\\ x-\frac{x^2}{2}+\frac{1}{2} & \text{pro } 0<x\le 1,\\ 1 & \text{pro } x>1,\end{cases}$ c) $P(0\le X<\frac{1}{2})=\frac{3}{8}$.

Příklad 5.28. Hustota pravděpodobnosti náhodné proměnné X má tvar:

$$f(x)=\begin{cases}0 & \text{pro } x\le 0,\\ ax & \text{pro } 0<x\le 1,\\ a & \text{pro } 1<x\le 2,\\ a(3-x) & \text{pro } 2<x\le 3,\\ 0 & \text{pro } x>3.\end{cases}$$

a) Určete konstantu a.
b) Určete předpis pro její distribuční funkci $F(x)$.
c) Určete $P(0<X\le\frac{1}{2})$.

Výsledek: a) $a=\frac{1}{2}$, b) $F(x)=\begin{cases}0 & \text{pro } x\le 0,\\ \frac{x^2}{4} & \text{pro } 0<x\le 1,\\ \frac{1}{2}x-\frac{1}{4} & \text{pro } 1<x\le 2,\\ \frac{3}{2}x-\frac{x^2}{4}-\frac{5}{4} & \text{pro } 2<x\le 3,\\ 1 & \text{pro } x>3.\end{cases}$ c) $P(0<X\le\frac{1}{2})=\frac{1}{16}$.

Příklad 5.29. Hustota pravděpodobnosti náhodné proměnné X má tvar:

$$f(x)=\begin{cases}0 & \text{pro } x\le 0,\\ ax^2 & \text{pro } 0<x\le 1,\\ \frac{1}{x} & \text{pro } 1<x\le \mathrm{e},\\ 0 & \text{pro } x>\mathrm{e}.\end{cases}$$

a) Určete konstantu a.
b) Určete předpis pro její distribuční funkci $F(x)$.

Výsledek: a) $a = 0$, b) $F(x) = \begin{cases} 0 & \text{pro } x \leq 1, \\ \ln x & \text{pro } 1 < x \leq \mathrm{e}, \\ 1 & \text{pro } x > \mathrm{e}. \end{cases}$

Příklad 5.30. Hustota pravděpodobnosti náhodné proměnné X má tvar:

$$f(x) = \frac{a}{1 + x^2}.$$

a) Určete konstantu a.
b) Určete předpis pro její distribuční funkci $F(x)$.
c) Určete $P(-1 < X < 1)$.

Výsledek: a) $a = \frac{1}{\pi}$, b) $F(x) = \frac{1}{\pi}\mathrm{arctg}\, x + \frac{1}{2}$; $\forall x \in \mathbb{R}$, c) $P(-1 < X < 1) = \frac{1}{2}$.

Příklad 5.31. Hustota pravděpodobnosti náhodné proměnné X má tvar:

$$f(x) = \frac{4a}{\mathrm{e}^x + \mathrm{e}^{-x}}.$$

a) Určete konstantu a.
b) Určete předpis pro její distribuční funkci $F(x)$.

Výsledek: a) $a = \frac{1}{2\pi}$, b) $F(x) = \frac{2}{\pi}\,\mathrm{arctg}\, \mathrm{e}^x$; $\forall x \in \mathbb{R}$.

6 Významná diskrétní rozdělení pravděpodobnosti

V této části probereme některá z nejznámějších diskrétních rozdělení pravděpodobností – **binomické, geometrické a hypergeometrické**. Podívejme se nejdřív na dva následující jednoduché příklady.

Příklad 6.1. Jaká je pravděpodobnost události A, že při 5-násobném nezávislém opakování hodu pravidelnou kostkou padne šestka při druhém a čtvrtém hodu, zatímco při prvním, třetím a pátém ne?

Řešení. Počet všech možností při pěti nezávislých hodech je $6 \cdot 6 \cdot 6 \cdot 6 \cdot 6 = 6^5$. Můžeme si je představit jako uspořádané pětice z čísel $1, 2, 3, 4, 5, 6$. Událost A reprezentují pětice, které mají na druhém a čtvrtém místě šestku a na zbylých místech jsou čísla z množiny $\{1, 2, 3, 4, 5\}$. Počet takových pětic je $5 \cdot 1 \cdot 5 \cdot 1 \cdot 5 = 5^3$. Proto

$$P(A) = \frac{5^3}{6^5} = \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3.$$

Příklad 6.2. Jaká je pravděpodobnost události B, že při 5-násobném nezávislém opakování hodu pravidelnou kostkou padne šestka právě dvakrát?

Řešení. Událost B reprezentují pětice, které mají na dvou místech šestku a na zbylých třech místech jsou čísla z množiny $\{1, 2, 3, 4, 5\}$. Šestky můžou padnout první dvě, nebo první a třetí, první a čtvrtá, a tak dále až poslední dvě, schematicky můžeme zapsat následovně:

$$(66---), (6-6--), (6--6-), (6---6), (-66--), \dots, (---66).$$

Pravděpodobnost všech uvedených možností je stejná a její hodnotu jsme vypočítali v předchozím příkladu. Kolik je vlastně uvedených možností? Z pěti míst vybíráme vždy dvě místa pro šestky, což můžeme udělat $\binom{5}{2}$ způsoby. Proto

$$P(B) = \binom{5}{2} \cdot \frac{5^3}{6^5} = \binom{5}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3 = 0{,}16.$$

Není problém výsledek zobecnit. Nechť událost U může nastat s pravděpodobností p. Jaká je pravděpodobnost, že při n stejných a nezávislých pokusech se bude událost U

opakovat právě k-krát? Poučeni řešením předchozích příkladů už bez problémů umíme tuto pravděpodobnost určit. Stačí si jen uvědomit, že pravděpodobnost toho, že nenastane událost U, je $1-p$. Potom

$$P(U \text{ nastane k-krát}) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}.$$

Tento výsledek je dost důležitý a má i svůj název - **Bernoulliho schéma**. My jej hned aplikujeme, budeme se totiž zabývat nejznámějším diskrétním rozdělením pravděpodobnosti - **binomickým rozdělením**.
Začněme hned definicí binomického rozdělení, kterou pak osvětlíme na několika příkladech.

Binomické rozdělení
Uvažujme experiment takové povahy, že mohou nastat jen dva různé výsledky, které se navzájem vylučují (nemůže k nim dojít současně): „úspěch“ a „neúspěch“ („úspěch“ nemusí znamenat nic světoborného; označuje se tímto termínem proto, že se jedná o ten ze dvou možných výsledků, na který se ve svých úvahách chceme zaměřit).
Pravděpodobnost úspěchu je p, pravděpodobnost neúspěchu $1-p$.
Experiment n-krát nezávisle zopakujeme. „Nezávisle“ znamená, že výskyt úspěchu při předchozím opakování experimentu nemá vliv na to, zda při dalších opakováních nastane úspěch nebo neúspěch.
Náhodná veličina X, která udává počet výskytů úspěchu při n nezávislých opakováních experimentu, má tzv. binomické rozdělení pravděpodobnosti s parametry n a p a nabývá hodnot z množiny $\{0, 1, 2, \ldots, n\}$ s pravděpodobností

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

Skutečnost, že veličina X má binomické rozdělení s parametry n, p, budeme označovat

$$X \sim \mathrm{Bi}(n, p).$$

Střední hodnota a rozptyl binomického rozdělení
Má-li náhodná veličina X binomické rozdělení s parametry n, p, pak

$$\begin{aligned} \mathrm{E}X &= np, \\ \mathrm{D}X &= np(1-p). \end{aligned}$$

Někdy se hodnoty veličiny s binomickým rozdělením uvádějí nikoliv v četnostech i (např. 12 úspěchů ze 20 pokusů), ale v podílech úspěšnosti $\frac{i}{n}$ (např. $\frac{12}{20}$). Toto **binomické rozdělení podílů úspěšnosti** má stejné parametry n, p, ale díky jiným hodnotám, kterých nabývá, je zde jiná střední hodnota a rozptyl:

$$\begin{aligned} \mathrm{E}X &= \sum_0^n \frac{i}{n} \cdot p(i) = \frac{1}{n} \cdot (\text{střední hodnota veličiny četností}) = \frac{1}{n} \cdot np = p. \\ \mathrm{D}X &= \left(\sum_0^n \frac{i^2}{n^2} \cdot p(i) \right) - (EX)^2 = \frac{1}{n^2} \left(\sum_0^n i^2 p(i) \right) - p^2 = \frac{p(1-p)}{n}. \end{aligned}$$

Poznámka: Náhodná veličina X má *alternativní rozdělení* s parametrem $p \in (0,1)$ jestliže nabývá pouze hodnoty $\{0,1\}$ a

$$P(X = 1) = p, \; P(X = 0) = 1 - p.$$

Alternativní rozdělení je speciálním případem binomického rozdělení.
Podívejme se nyní na konkrétní příklady.

Příklad 6.3. Pravděpodobnost, že náhodně vybraný výrobek bude kvalitní, je 0,75. Vybereme náhodně pět výrobků. Náhodná veličina X udává počet kvalitních výrobků v uvedené vzorce pěti výrobků. Určete její rozdělení.

Řešení. Nejdřív si je třeba uvědomit, že mezi vybranými výrobky nemusí být ani jeden kvalitní, nebo může být jeden kvalitní, dva kvalitní a tak dále až pět kvalitních výrobků, tedy $X \in \{0,1,2,3,4,5\}$. Určeme pravděpodobnosti jednotlivých situací.

- **0 kvalitních výrobků**
  Všechny výrobky jsou nekvalitní a pravděpodobnost, že jeden je nekvalitní je $1 - - 0{,}75$. Proto
  $$p(0) = (1 - 0{,}75)^5,$$
  což můžeme taky zapsat
  $$p(0) = \binom{5}{0} \cdot 0{,}75^0 \cdot (1 - 0{,}75)^5 = \binom{5}{0} \cdot 0{,}75^0 \cdot 0{,}25^5.$$
  To nám dává návod na určení dalších pravděpodobností.

- **1 kvalitní výrobek**
  V tomto případě jsou čtyři výrobky nekvalitní, a proto
  $$p(1) = \binom{5}{1} \cdot 0{,}75^1 \cdot (1 - 0{,}75)^4 = \binom{5}{1} \cdot 0{,}75^1 \cdot 0{,}25^4.$$

- **2 kvalitní výrobky**
  Z pěti výrobků máme dva kvalitní, tři nekvalitní, a proto
  $$p(2) = \binom{5}{2} \cdot 0{,}75^2 \cdot (1 - 0{,}75)^3 = \binom{5}{2} \cdot 0{,}75^2 \cdot 0{,}25^3.$$
  Podobně už určíme další pravděpodobnosti.

- **3 kvalitní výrobky**

$$p(3) = \binom{5}{3} \cdot 0{,}75^3 \cdot (1 - 0{,}75)^2 = \binom{5}{3} \cdot 0{,}75^3 \cdot 0{,}25^2.$$

- **4 kvalitní výrobky**

$$p(4) = \binom{5}{4} \cdot 0{,}75^4 \cdot (1 - 0{,}75)^1 = \binom{5}{4} \cdot 0{,}75^4 \cdot 0{,}25^1.$$

- **5 kvalitních výrobků**

$$p(5) = \binom{5}{5} \cdot 0{,}75^5 \cdot (1 - 0{,}75)^0 = \binom{5}{5} \cdot 0{,}75^5 \cdot 0{,}25^0$$

Součet jednotlivých pravděpodobností musí být 1, protože se jedná o jevy neslučitelné, které zahrnují všechny situace, které můžou nastat. Náš součet je opravdu roven 1, vyzkoušejte si to! Pokud si ovšem vzpomenete ze střední školy na *binomickou větu*, nemusíte nic počítat.

$$(a+b)^n = \binom{n}{0} \cdot a^0 \cdot b^n + \binom{n}{1} \cdot a^1 \cdot b^{n-1} + \binom{n}{2} \cdot a^2 \cdot b^{n-2} + \ldots \binom{n}{n} \cdot a^n \cdot b^0.$$

Stačí si uvědomit, že v našem případě je $a = 0{,}75$, $b = 0{,}25$ a $n = 5$. Proto náš součet je $(0{,}75 + 0{,}25)^5 = 1^5 = 1$.

Příklad 6.4. V předmětu IMA psali studenti 6 testů. Naše náhodná veličina X udává počet úspěšně napsaných testů studenta M.H. Pravděpodobnost úspěšnosti na testu studenta M.H. je vždy 0,85. Zjistěte rozdělení veličiny X a určete pravděpodobnost toho, že student M.H.
a) bude úspěšný právě 3-krát,
b) bude úspěšný alespoň 2-krát,
c) nebude úspěšný víc než 2-krát,
d) bude úspěšný jenom na prvním testu,
e) bude úspěšný na prvním testu.

Řešení. Nejdřív určíme rozdělení veličiny X. Zřejmě student M.H. může být úspěšný 0 až 6-krát, proto $X \in \{0, 1, 2, 3, 4, 5, 6\}$.

- **0 úspěšných testů**
  Všechny testy jsou neúspěšné a pravděpodobnost, že jeden je neúspěšný, je $1 - - 0{,}85 = 0{,}15$. Proto

  $$p(0) = (0{,}15)^6,$$

  což můžeme taky zapsat jako

  $$p(0) = \binom{6}{0} \cdot 0{,}85^0 \cdot 0{,}15^6,$$

  a to je návod na určení dalších pravděpodobností:

- **1 úspěšný test**

$$p(1) = \binom{6}{1} \cdot 0{,}85^1 \cdot 0{,}15^5.$$

- **2 úspěšné testy**

$$p(2) = \binom{6}{2} \cdot 0{,}85^2 \cdot 0{,}15^4.$$

- **3 úspěšné testy**

$$p(3) = \binom{6}{3} \cdot 0{,}85^3 \cdot 0{,}15^3.$$

- **4 úspěšné testy**

$$p(4) = \binom{6}{4} \cdot 0{,}85^4 \cdot 0{,}15^2.$$

- **5 úspěšných testů**

$$p(5) = \binom{6}{5} \cdot 0{,}85^5 \cdot 0{,}15^1.$$

- **6 úspěšných testů**

$$p(6) = \binom{6}{6} \cdot 0{,}85^6 \cdot 0{,}15^0.$$

Hned můžeme odpovědět na část a). Pravděpodobnost, že student M.H. bude úspěšný právě 3-krát, je $P(A) = p(3) = \binom{6}{3} \cdot 0{,}85^3 \cdot 0{,}15^3 \doteq 0{,}04145$.

Jev, že student M.H. bude úspěšný alespoň 2-krát, znamená, že M.H. bude úspěšný 2, 3, 4, 5 nebo 6-krát. Proto pro výsledek části b) by bylo třeba sečíst $P(B) = p(2) + p(3) + p(4) + p(5) + p(6)$, což není právě nejpříjemnější. Jednodušší je ale podívat se na doplňkový jev, tedy M.H. by byl úspěšný nejvíc 1-krát (nejvíc 1-krát = ani jednou nebo právě 1-krát) a pravděpodobnost tohoto jevu je součet $p(0) + p(1)$. Pravděpodobnost našeho jevu je pak $P(B) = 1 - (p(0) + p(1)) \doteq 0{,}9996$,

Část c) vlastně znamená, že student M.H. bude úspěšný maximálně 2-krát, proto pravděpodobnost tohoto jevu je součet $P(C) = p(0) + p(1) + p(2) \doteq 0{,}005885$.

V části d) se jedná o situaci, kdy první test student M.H. zvládně úspěšně ($p = 0{,}85$), ale další testy budou určitě neúspěšné ($p = 0{,}15 \cdot 0{,}15 \cdot 0{,}15 \cdot 0{,}15 \cdot 0{,}15$), proto

$$P(D) = 0{,}85 \cdot 0{,}15 \cdot 0{,}15 \cdot 0{,}15 \cdot 0{,}15 \cdot 0{,}15 = 0{,}85^1 \cdot 0{,}15^5 = 0{,}000064547.$$

Čím se liší části d) a e)? Není to stejné? V části d) je jasné, že kromě prvního testu jsou všechny zbylé neúspěšné. V části e) víme, že první test je úspěšný, a o úspěšnosti zbylých testů nevíme nic. Zřejmě rozdíl mezi pravděpodobnostmi v částech d) a e) bude. V části e) nás tedy zajímá jenom první test, a proto $P(E) = 0{,}85$.

Příklad 6.5. Spočítejte očekávanou střední hodnotu úspěšně napsaných testů a rozptyl z předchozího příkladu na binomické rozdělení (příklad 6.4 o písemkách z IMA).

Řešení. Protože se jedná o binomické rozdělení, počet všech pokusů (počet všech napsaných písemek) je 6 ($n = 6$) a pravděpodobnost úspěšně napsaného testu je $p = 0{,}85$. Proto je

$$\mathrm{E}X = n \cdot p = 6 \cdot 0{,}85 = 5{,}1$$

a

$$\mathrm{D}X = n \cdot p \cdot (1 - p) = 6 \cdot 0{,}85 \cdot 0{,}15 = 0{,}765.$$

Kdybychom měli zjistit očekávanou hodnotu neúspěšně napsaných testů, počítali bychom s doplňkovou pravděpodobností

$$\mathrm{E}X = n \cdot (1 - p) = 6 \cdot 0{,}15 = 0{,}9.$$

V případě rozptylu si stačí uvědomit, že v obou případech je výpočet stejný. Všimněte si součtu uvedených středních hodnot.

A teď zkusíme úlohu 6.4 trochu pozměnit.

Příklad 6.6. V předmětu IMA psali studenti 6 testů. Jestliže student nenapíše úspěšně nějaký test, nemůže psát už další testy (naštěstí je to jen příklad). Naše náhodná veličina X udává počet napsaných testů studenta M.H. Pravděpodobnost úspěšnosti na testu studenta M.H. je vždy 0,85. Zjistěte rozdělení veličiny X.

Řešení. Naše náhodná veličina X nabývá hodnoty z množiny $\{1, 2, 3, 4, 5, 6\}$. Čím se příklad odlišuje od předchozího? Jestliže student nenapíše např. první test, další už psát nebude. To znamená, že přesně známe pořadí úspěšných a neúspěšných testů pro jednotlivé hodnoty veličiny X. Proto

- **1 test**
  to znamená, že první test byl neúspěšný a další už student M.H. nepsal. Pravděpodobnost, že test napíše neúspěšně, je $1 - 0{,}85 = 0{,}15$. Tedy

  $$p(1) = 0{,}15$$

- **2 testy**
  první test napsal úspěšně a druhý neúspěšně, proto

  $$p(2) = 0{,}85 \cdot 0{,}15.$$

  Podobně určíme pravděpodobnost pro 3, 4 a 5 testů

- **3 testy**

$$p(3) = 0{,}85^2 \cdot 0{,}15.$$

- **4 testy**

$$p(4) = 0{,}85^3 \cdot 0{,}15.$$

- **5 testů**

$$p(5) = 0{,}85^4 \cdot 0{,}15.$$

- **6 testů**
  v tomto případě bude trochu změna. Je třeba si uvědomit, že šest napsaných testů nemusí znamenat, že napsal všech šest testů úspěšně. Také se mohlo stát, že prvních pět napsal úspěšně a poslední test neúspěšně. Proto

  $$p(6) = 0{,}85^5 \cdot 0{,}15 + 0{,}85^6.$$

  Tento výsledek ještě můžeme upravit jako

  $$p(6) = 0{,}85^5 \cdot (0{,}15 + 0{,}85) = 0{,}85^5 \cdot 1 = 0{,}85^5.$$

  Poslední výsledek může být interpretován takto: Student se propracuje k 6. testu, jestliže prvních 5 napsal úspěšně, což má pravděpodobnost právě $0{,}85^5$. Na výsledku šestého testu už pak nezáleží.

Zřejmě naše náhodná veličina nemá binomické rozdělení, ale toho si už všímavý čtenář pravděpodobně všiml. Součet jednotlivých pravděpodobností je 1, ověřte si to!
Vyřešili jsme dvě na první pohled stejné úlohy a každá z nich nás přivedla k jinému rozdělení pravděpodobností, proto doporučujeme si vždy pečlivě přečíst zadání úloh.

V předchozí úloze náhodná veličina představovala počet pokusů při nezávislém opakování pokusů po první výskyt neúspěchu včetně toho neúspěšného. Kdyby celkový počet pokusů nebyl omezen (v tomto příkladu omezen byl – studenti psali 6 testů) a nepočítali bychom závěrečný neúspěšný pokus, dostali bychom tzv. geometrické rozdělení pravděpodobnosti.

Geometrické rozdělení
Jestliže náhodná veličina X udává počet úspěchů při nezávislém opakování pokusu po první výskyt neúspěchu a pravděpodobnost úspěchu je p, pro její pravděpodobnosti platí

$$P(X = k) = p^k \cdot (1 - p) \quad \text{pro } k = 0, 1, 2, \ldots$$

Taková náhodná veličina má tzv. geometrické rozdělení pravděpodobnosti. Samozřejmě, tato ukončovací podmínka může být chápána i opačně jako čekání na úspěch (pak p je pravděpodobnost neúspěchu). Pro její střední hodnotu a rozptyl platí

$$\mathrm{E}X = \frac{p}{1-p}, \qquad \mathrm{D}X = \frac{p}{(1-p)^2}.$$

Zkusíme si příklad tohoto typu.

Příklad 6.7. Jeden zlý učitel předmětu IMA dozkušuje ty studenty, kteří mají celkem 49 bodů, tím způsobem, že jim klade otázky tak dlouho, dokud nedostane nesprávnou odpověď. Pak nešťastníka vyhodí (naštěstí je to jen příklad). Naše náhodná veličina X udává počet úspěšně zodpovězených otázek studenta M. H. Dosti zjednodušeně budeme předpokládat, že pravděpodobnost úspěšné odpovědi je pro M. H. vždy 0,1. Zjistěte rozdělení veličiny X.

Řešení. Naše náhodná veličina X nabývá hodnoty z množiny $\{0, 1, 2, \dots\}$ (můžou ho, chudáka, zkoušet až do nekonečna).

- **0 úspěšných odpovědí**
  to znamená, že hned na první otázku odpověděl špatně a letěl. Pravděpodobnost, že odpoví špatně, je $1 - 0{,}1 = 0{,}9$. Tedy
  $$p(0) = 0{,}9 = 0{,}1^0 \cdot 0{,}9$$
- **1 úspěšná odpověď**
  to znamená, že na první otázku uměl odpovědět a na druhou už ne. Tedy
  $$p(1) = 0{,}1^1 \cdot 0{,}9$$
- **2 úspěšné odpovědi**
  první dvě otázky zodpověděl úspěšně a třetí neúspěšně, proto
  $$p(2) = 0{,}1^2 \cdot 0{,}9.$$
- **Obecně dostáváme pro k úspěšných odpovědí**
  $$p(k) = 0{,}1^k \cdot 0{,}9, \quad k = 0, 1, 2, \dots$$

Jednalo se o geometrické rozdělení.

Jak již bylo napsáno v posledním rámečku, u geometrického rozdělení někdy počítáme úspěšné pokusy před prvním neúspěchem a někdy neúspěšné pokusy před prvním úspěchem. Jestliže bereme v potaz, na kolikátý pokus se nám něco podařilo, pak příslušná náhodná veličina nemá sice přesně geometrické rozdělení, ale pravděpodobnosti se počítají velmi podobně. Ukážeme si to na následujícím příkladu.

Příklad 6.8. Opakovaně házíme kostkou, dokud nepadne šestka. Náhodná veličina X udává, na kolikátý pokus šestka padla. Najděte pravděpodobnostní funkci náhodné veličiny X a vypočtěte její střední hodnotu.

Řešení. Kdyby X udávala počet hodů **před** prvním výskytem šestky, mohla by nabývat hodnot $0, 1, 2, \ldots$ a měla by geometrické rozdělení s parametrem $p = 5/6$. Pro náš příklad však X nemůže být rovno 0 (nemůžeme hodit šestku na nultý pokus). V nejlepším případě nám šestka padne prvním pokusem a pravděpodobnost toho je

$$p(1) = \frac{1}{6}\,.$$

Pravděpodobnost, že poprvé šestka padne na druhý pokus, je

$$p(2) = \frac{5}{6} \cdot \frac{1}{6}$$

a obecně

$$p(k) = \left(\frac{5}{6}\right)^{k-1} \frac{1}{6}\,, \quad k = 1, 2, \ldots$$

Pro výpočet střední hodnoty je vhodné si uvědomit, že kdybychom měli náhodnou veličinu Y udávající počet „nešestek“ před první šestkou, tak střední hodnota této náhodné veličiny by byla

$$\mathrm{E}Y = \frac{p}{1-p} = \frac{\frac{5}{6}}{1 - \frac{5}{6}} = 5.$$

Pro naši náhodnou veličinu X platí $X = Y + 1$ – jestliže šestka padla poprvé např. na třetí pokus, tj. $X = 3$, pak před ní padly dvě „nešestky“, tj. $Y = 2$, apod. Proto střední hodnota X je

$$\mathrm{E}X = \mathrm{E}(Y + 1) = \mathrm{E}Y + 1 = 5 + 1 = 6.$$

Na závěr uvedeme ještě jeden příklad:

Příklad 6.9. Mezi stovkou výrobků je 20 zmetků. Náhodně vybereme 10 výrobků. Náhodná veličina X udává počet zmetků mezi vybranými výrobky. Určete její rozdělení.

Řešení. Zřejmě X nabývá hodnoty $\{0, 1, 2, ..., 10\}$ a nejedná se ani o binomické, ani geometrické rozdělení. Výpočet jednotlivých pravděpodobností by nám neměl způsobit problémy, v předchozích kapitolách jsme se již podobnou myšlenkou zabývali.

- **0 zmetků**
  to znamená, že všech 10 výrobků, které vybereme, bude kvalitních. Všech kvalitních je $100 - 20 = 80$. Proto

$$p(0) = \frac{\binom{20}{0}\binom{100-20}{10-0}}{\binom{100}{10}} = \frac{\binom{80}{10}}{\binom{100}{10}}.$$

- **1 zmetek**
  jeden výrobek je některý z 20 zmetků a zbylých 9 výrobků budeme vybírat z 80 kvalitních výrobků. Tedy

$$p(1) = \frac{\binom{20}{1}\binom{80}{9}}{\binom{100}{10}}$$

- **2 zmetky**
  dva výrobky jsou zmetky a zbylých 8 jsou kvalitní, proto
$$p(2) = \frac{\binom{20}{2}\binom{80}{8}}{\binom{100}{10}}.$$
  A tak dále, až nakonec
- **10 zmetků**
  analogicky dostaneme
$$p(10) = \frac{\binom{20}{10}\binom{80}{0}}{\binom{100}{10}}.$$

Tato jednoduchá úloha nás přivedla k dalšímu diskrétnímu rozdělení pravděpodobností. Jedná se o hypergeometrické rozdělení náhodné veličiny, kdy při opakování náhodného pokusu je výskyt sledovaného jevu závislý na výsledcích předcházejících pokusů. Jde tedy o pokusy, které jsou na sobě závislé. Typickým představitelem je výběr prvků bez vracení.

Hypergeometrické rozdělení
Uvažujeme situaci, kdy máme určitý soubor prvků. Celkový počet prvků je N. Z toho M prvků má sledovanou vlastnost. Ze souboru vybereme bez vracení n prvků. Náhodná veličina X udává počet prvků se sledovanou vlastností ve vybrané n-tici. Takováto náhodná veličina má hypergeometrické rozdělení s parametry N, M, n, kde N a M jsou přirozená čísla a $n < N$. Pravděpodobnostní funkce je určena předpisem

$$P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad k = \max\{0, n-(N-M)\}, \ldots, \min\{n, M\}.$$

Pro střední hodnotu a rozptyl hypergeometrického rozdělení platí

$$\mathrm{E}X = n\frac{M}{N}, \qquad \mathrm{D}X = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}.$$

V dalším příkladu si ukážeme ještě jinou možnost výpočtu pravděpodobnostní funkce u hypergeometrického rozdělení. Na základě toho pak uvidíme jistou souvislost mezi rozdělením hypergeometrickým a binomickým.

Příklad 6.10. a) Máme košík s 10 jablky, z toho 30 % je červivých. Náhodně vybereme 4 jablka. Náhodná veličina X udává, kolik z nich je červivých. Najděte pravděpodobnostní funkci náhodné veličiny X, její střední hodnotu a rozptyl.
b) Ze sadu bylo sklizeno 10 000 jablek, z toho 30 % je červivých. Náhodná veličina X opět udává počet červivých jablek ve vybrané čtveřici. Najděte pravděpodobnostní funkci, střední hodnotu a rozptyl náhodné veličiny X.

Řešení. a) Pro 10 jablek celkem: Máme 3 jablka červivá a zbylých 7 je zdravých.

- **0 červivých**,
  tj. všechna zdravá. Pravděpodobnost bychom mohli počítat pomocí kombinačních čísel:
  $$p(0) = \frac{\binom{7}{4}}{\binom{10}{4}} = \frac{35}{210} = \frac{1}{6}\,.$$
  Můžeme však k problému přistoupit i jinak. Budeme jablka brát postupně. Všechna 4 musí být zdravá. Pravděpodobnost, že první jablko, které jsme vybrali, je zdravé, je 7/10. Když vybíráme druhé jablko, celkem zbývá jablek už jen 9, z toho zdravých je 6. Proto pravděpodobnost, že druhé jablko je zdravé, je 6/9. Podobně pak pro třetí a čtvrté jablko. Pravděpodobnost, že všechna jablka jsou zdravá, je
  $$p(0) = \frac{7}{10}\cdot\frac{6}{9}\cdot\frac{5}{8}\cdot\frac{4}{7} = \frac{1}{6}\,.$$
  Dospěli jsme ke stejnému výsledku postupem, který se na první pohled může zdát jednodušší než ten s kombinačními čísly. Pro nenulové počty červivých jablek se ale objeví komplikace.

- **1 červivé**
  Pomocí kombinačních čísel dostaneme $p(1) = \binom{3}{1}\cdot\binom{7}{3}/\binom{10}{4} = 1/2$. Teď budeme jablka brát postupně. Musí být jedno červivé a všechna ostatní zdravá. Pravděpodobnost, že první jablko je červivé, je 3/10. Druhé červivé má pravděpodobnost 7/9 atd. Zdálo by se, že výsledek bude
  $$\frac{3}{10}\cdot\frac{7}{9}\cdot\frac{6}{8}\cdot\frac{5}{7} = \frac{1}{8}\,.$$
  Vyšlo to jinak, něco je špatně! Problém je podobný jako u binomického rozdělení: právě vypočtená hodnota není pravděpodobnost, že **jedno** jablko ze čtyř je červivé, ale že **první** jablko je červivé a ostatní zdravá. Musíme ještě probrat možnosti, že červivé je druhé, třetí a čtvrté jablko. Celkem dostaneme
  $$p(1) = \frac{3}{10}\cdot\frac{7}{9}\cdot\frac{6}{8}\cdot\frac{5}{7} + \frac{7}{10}\cdot\frac{3}{9}\cdot\frac{6}{8}\cdot\frac{5}{7} + \frac{7}{10}\cdot\frac{6}{9}\cdot\frac{3}{8}\cdot\frac{5}{7} + \frac{7}{10}\cdot\frac{6}{9}\cdot\frac{5}{8}\cdot\frac{3}{7}$$
  Když si jednotlivé sčítance prohlédneme, zjistíme, že ve jmenovateli je vždy $10\cdot9\cdot8\cdot7$. V čitateli vždy násobíme čísla 3, 7, 6 a 5, mění se jen jejich pořadí. Celkem tedy máme
  $$p(1) = 4\cdot\frac{3}{10}\cdot\frac{7}{9}\cdot\frac{6}{8}\cdot\frac{5}{7} = 4\cdot\frac{1}{8} = \frac{1}{2}\,.$$

- **2 červivá**
  Pomocí kombinačních čísel: $p(2) = \binom{3}{2}\cdot\binom{7}{2}/\binom{10}{4} = 3/10$.

  Po jednotlivých kusech: Pravděpodobnost, že první dvě jablka jsou červivá a druhá dvě zdravá, je
  $$\frac{3}{10}\cdot\frac{2}{9}\cdot\frac{7}{8}\cdot\frac{6}{7} = \frac{1}{20}$$

Toto číslo musíme vynásobit počtem všech možností, jak mohla být dvě červivá jablka ve vybrané čtveřici rozmístěna, což je $\binom{4}{2}$. Celkem

$$p(2) = \binom{4}{2} \cdot \frac{3}{10} \cdot \frac{2}{9} \cdot \frac{7}{8} \cdot \frac{6}{7} = 6 \cdot \frac{1}{20} = \frac{3}{10}.$$

- **3 červivá**

$$p(3) = \binom{4}{3} \cdot \frac{3}{10} \cdot \frac{2}{9} \cdot \frac{1}{8} \cdot \frac{7}{7} = \frac{1}{30}.$$

Můžete se přesvědčit, že pomocí kombinačních čísel vyjde totéž.

Pravděpodobnostní funkce je tímto popsána, protože více než 3 červivá jablka ve vybrané čtveřici být nemohou.

- **Střední hodnota a rozptyl**
Dosadíme do vzorců s $N = 10$, $M = 3$ a $n = 4$ (případně bychom mohli použít univerzální vzorce $EX = \sum x_i p(x_i)$, podobně pro DX):

$$\mathrm{E}X = 4 \cdot \frac{3}{10} = \frac{6}{5} = 1{,}2, \quad \mathrm{D}X = 4 \cdot \frac{3}{10} \cdot \left(1 - \frac{3}{10}\right) \cdot \frac{10-4}{10-1} = \frac{14}{25} = 0{,}56.$$

b) Pro 10 000 jablek celkem, z toho 3 000 červivých a 7 000 zdravých:
Nebudeme probírat celou pravděpodobnostní funkci, na ukázku vezmeme např. $p(2)$. Provedeme-li výpočet způsobem předvedeným v části a), dostaneme

$$p(2) = \binom{4}{2} \frac{3\,000}{10\,000} \cdot \frac{2\,999}{9\,999} \cdot \frac{7\,000}{9\,998} \cdot \frac{6\,999}{9\,997}.$$

Proč tento příklad vlastně počítáme? Vždyť je to stejné jako a), jen s většími čísly! Uvidíme hned – vyčíslíme jednotlivé zlomky:

$$\frac{3\,000}{10\,000} = 0{,}3, \quad \frac{2\,999}{9\,999} \doteq 0{,}3, \quad \frac{7\,000}{9\,998} \doteq 0{,}7, \quad \frac{6\,999}{9\,997} \doteq 0{,}7$$

Je vidět, že to, že už jsme jedno jablko vybrali, téměř neovlivní pravděpodobnost červivosti pro druhé jablko. Je to dáno velkým celkovým počtem jablek – když jedno ubereme, nestane se skoro nic. U předchozího příkladu s 10 jablky pravděpodobnost červivosti druhého jablka závisela na tom, jestli první jablko červivé bylo nebo nebylo.
Hodnotu $p(2)$ můžeme přibližně vyjádřit jako

$$p(2) \doteq \binom{4}{2} \cdot 0{,}3^2 \cdot 0{,}7^2,$$

což je pravděpodobnost pro binomické rozdělení. Vyčíslit tento výraz dá mnohem méně práce než počítat s původním vzorcem pro hypergeometrické rozdělení a nepřesnost, které se tím dopustíme, je velmi malá.

Vypočteme ještě střední hodnotu a rozptyl. Nejprve přesně, tj. pomocí vzorců pro hypergeometrické rozdělení. Za M/N dosazujeme $3\,000/10\,000 = 0{,}3$:

$$\mathrm{E}X = 4 \cdot 0{,}3 = 1{,}2, \quad \mathrm{D}X = 4 \cdot 0{,}3 \cdot (1 - 0{,}3) \cdot \frac{10\,000 - 4}{10\,000 - 1} \doteq 0{,}8397.$$

Pro srovnání, kdybychom použili vzorce pro binomické rozdělení s $n = 4$ a $p = 0{,}3$, dostali bychom $\mathrm{E}X$ stejně, zatímco rozptyl by byl

$$\mathrm{D}X = 4 \cdot 0{,}3 \cdot 0{,}7 = 0{,}84$$

čili vyšel by skoro stejně jako pro rozdělení hypergeometrické.
Právě předvedený příklad můžeme zobecnit:

Vztah hypergeometrického a binomického rozdělení
Máme-li náhodnou veličinu X s hypergeometrickým rozdělení, kde čísla N a M jsou velká, zatímco n je vzhledem k nim malé, můžeme s náhodnou veličinou X počítat, jako kdyby měla binomické rozdělení s parametry n a $p = M/N$.

Tento fakt se často používá. Předvedeme si to na příkladu.

Příklad 6.11. Senátor Kadrnožka před volbami tvrdí, že jej bude volit 70 % voličů. Předpokládejme, že má pravdu. Provedeme průzkum mezi 10 lidmi odhodlanými jít volit. Jaká je pravděpodobnost, že pro Kadrnožku jich bude právě 7?

Řešení. Kdybychom chtěli tento příklad řešit přesně, museli bychom vědět, kolik lidí v Kadrnožkově volebním obvodu půjde k volbám – to by bylo dosazeno za N. Z toho 70 % by bylo M. Pak bychom použili vzorec pro hypergeometrické rozdělení. My počet voličů neznáme, avšak lze předpokládat, že bude velký. Nebo aspoň velký vzhledem k 10. Můžeme tedy říci, že u každého jednotlivého člověka je pravděpodobnost 70 %, že bude volit Kadrnožku. A proto

$$p(7) = \binom{10}{7} \cdot 0{,}7^7 \cdot 0{,}3^3 \doteq 0{,}267.$$

6.1 Příklady na samostatnou práci

Zde uvádíme neřešené úlohy na různé typy rozdělění náhodné proměnné. Zřejmě to bude náročnější než v předchozích kapitolách, kde jste měli dopředu dané, o jaký typ rozdělení se jedná, ale taková je realita života. Tak mnoho zdaru!

Příklad 6.12. Pravděpodobnost toho, že student David přijde pozdě do školy, je každý den 0,2. Náhodná proměnná X udává počet jeho pozdních příchodů v průběhu 20 pracovních dnů.
a) Určete její pravděpodobnostní funkci.
b) Určete její střední hodnotu a rozptyl.
c) Určete pravděpodobnost toho, že David bude mít právě 7 pozdních příchodů.

Výsledek: a) $p(k) = \binom{20}{k} \cdot 0{,}2^k \cdot 0{,}8^{20-k}$; b) $\mathrm{E}X = 4$, $\mathrm{D}X = 3{,}2$; c) $\binom{20}{7} \cdot 0{,}2^7 \cdot 0{,}8^{13} \doteq 0{,}055$

Příklad 6.13. Házíme 6-krát šestistěnnou kostkou. Náhodná veličina X udává počet padnutých jedniček.
a) Určete její pravděpodobnostní funkci.
b) Určete její střední hodnotu a rozptyl.
c) Určete pravděpodobnost toho, že jednička padne právě třikrát.
d) Určete pravděpodobnost toho, že jednička padne minimálně třikrát.
e) Určete pravděpodobnost toho, že jednička padne nejvíc pětkrát.

Výsledek: a) $p(k) = \binom{6}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{6-k}$; b) $\mathrm{E}X = 1$, $\mathrm{D}X = 5/6$; c) přibl. 0,0536; d) přibl. 0,0623; e) $1 - 1/6^6 \doteq 0{,}99998$

Příklad 6.14. Házíme čtyřikrát mincí. Náhodná veličina X udává, kolikrát padl znak (rub mince).
a) Určete její pravděpodobnostní a distribuční funkci.
b) Určete její střední hodnotu a rozptyl.
c) Určete pravděpodobnost toho, že znak padne právě dvakrát.
d) Určete pravděpodobnost toho, že znak padne minimálně dvakrát.
e) Určete pravděpodobnost toho, že znak padne nejvíc dvakrát.

Výsledek: a) $p(k) = \binom{4}{k} \cdot \left(\frac{1}{2}\right)^k \cdot \left(\frac{1}{2}\right)^{4-k}$, $F(x) = 0$ pro $x < 0$, $F(x) = 1/16$ pro $x \in \langle 0, 1)$, $F(x) = 5/16$ pro $x \in \langle 1, 2)$, $F(x) = 11/16$ pro $x \in \langle 2, 3)$, $F(x) = 15/16$ pro $x \in \langle 3, 4)$, $F(x) = 1$ pro $x \geq 4$; b) $\mathrm{E}X = 2$, $\mathrm{D}X = 1$; c) 3/8; d) 11/16; e) 11/16

Příklad 6.15. Třikrát vystřelíme na cíl. Pravděpodobnost zásahu při každém výstřelu je $p = 0{,}7$. Náhodná proměnná X udává počet zásahů cíle.
a) Určete její pravděpodobnostní a distribuční funkci.
b) Určete její střední hodnotu a rozptyl.
c) Určete pravděpodobnost toho, že budeme mít právě 2 zásahy cíle.

Výsledek: a) $p(k) = \binom{3}{k} \cdot 0{,}7^k \cdot 0{,}3^{3-k}$, $F(x) = 0$ pro $x < 0$, $F(x) = 0{,}027$ pro $x \in \langle 0, 1)$, $F(x) = 0{,}216$ pro $x \in \langle 1, 2)$, $F(x) = 0{,}657$ pro $x \in \langle 2, 3)$, $F(x) = 1$ pro $x \geq 3$; b) $\mathrm{E}X = 2{,}1$, $\mathrm{D}X = 0{,}63$; c) 0,441

Příklad 6.16. (Příklad statistického řízení procesu:) Každou hodinu je vybrán vzorek 20 součástí procesu perforace (= prorážení) kovu. Průměrně 4% výsledků vyžadují dodatečné úpravy. Náhodná veličina X označuje počet součástí z vybraných dvaceti, které vyžadují dodatečné úpravy. Pokud veličina X překročí svou střední hodnotu o více než trojnásobek své směrodatné odchylky, pracovní linka se musí zastavit a opravit.
a) Jaká je pst, že X překročí svou střední hodnotu o více než trojnásobek své směrodatné odchylky?
b) Jaká je pst, že X překročí hodnotu 1 aspoň v jedné z následujících pěti hodin provozu?

Výsledek: a) $P(X > \mathrm{E}X + 3\sqrt{\mathrm{D}X}) = P(X > 3{,}43) \doteq 0{,}0074$;
b) Mezivýsledek: $P(X > 1) \doteq 0{,}1897$, výsledek: $1 - (1 - 0{,}1897)^5 \doteq 0{,}6506$

Příklad 6.17. Zuzanka hází míčkem na cíl, pravděpodobnost zásahu je 0,6. Dostala 7 míčků a všechny chce vyzkoušet. Náhodná proměnná X udává počet zásahů cíle.
a) Určete její pravděpodobnostní funkci.
b) Určete její střední hodnotu a rozptyl.
c) Určete pravděpodobnost toho, že Zuzanka zasáhne cíl alespoň dvakrát.

Výsledek:
a) $p(i) = \binom{7}{i} \cdot 0{,}6^i \cdot 0{,}4^{7-i},\ i \in \{0, 1, \dots, 7\}$,
b) $\mathrm{E}X = 4,2;\ \mathrm{D}X = 1,68$ c) $P(X \geq 2) = 0{,}9811584$.

Příklad 6.18. Pravděpodobnost výskytu jistého slova v jazyku je 0,05. Kolik slov musíme mít v textu, aby se v něm s pravděpodobností 0,99 tohle slovo vyskytlo alespoň jednou?

Výsledek: 90

Příklad 6.19. Výrobní podnik expedoval zásilku, která obsahuje 20 výrobků. Pravděpodobnost toho, že se jeden výrobek během přepravy poškodí, je 0,1. Diskrétní náhodná veličina X udává počet poškozených výrobků.
a) Určete její pravděpodobnostní funkci.
b) Určete její střední hodnotu.
c) Určete pravděpodobnost toho, že se během přepravy poškodí více než 3 výrobky.

Výsledek:
a)$p(i) = \binom{20}{i} \cdot 0{,}1^i \cdot 0{,}9^{20-i},\ i \in \{0, 1, \dots, 20\}$
b) $\mathrm{E}X = 0{,}5$.

Příklad 6.20. Automatická linka produkuje 95% výrobků první kvality a 5% zmetků. Ze 100 kusů vybereme náhodně 10 kusů. Diskrétní náhodná veličina X udává počet zmetků z vybraných deseti kusů.
a) Určete její pravděpodobnostní funkci.
b) Určete její střední hodnotu.

Výsledek:
a)$p(i) = \frac{\binom{95}{10-i} \cdot \binom{5}{i}}{\binom{100}{10}},\ i \in \{0, 1, \dots, 5\}$
b) $\mathrm{E}X = 0{,}5$.

Příklad 6.21. Elektronická váha v automatizovaném provozu zastaví balicí linku poté, co zaznamená dva balíky v těsném sledu za sebou s jinou hmotností, než je požadováno. V každém případě je balicí linka zastavena po zabalení pátého balíku. Při každém balení má balík požadovanou hmotnost s pravděpodobností 0,9. Jednotlivá balení jsou považována za nezávislá na ostatních baleních. Veličina X udává počet balíků s požadovanou hmotností zabalených před prvním zastavením linky.
a) Určete pravděpodobnostní funkci veličiny X.
b) Určete střední hodnotu veličiny X.

Výsledek:
a) $p(0) = 0{,}01$; $p(1) = 0{,}0099$; $p(2) = 0{,}01053$; $p(3) = 0{,}05103$; $p(4) = 0{,}32805$; $p(5) =$ $= 0{,}59049$
b) $\mathrm{E}X = 4,4487$.

Příklad 6.22. * Pravděpodobnost zásahu terče 0,5. Kolik pokusů o zásah musíme udělat, aby s pravděpodobností 0,75 byl terč zasažen právě dvakrát?

Výsledek: $n = 3$.

7 Exponenciální a Poissonovo rozdělení pravděpodobnosti

V této kapitole se seznámíme s dalšími dvěma typy rozdělení pravděpodobnosti, které jsou využívány v úlohách technické praxe. I když Poissonovo rozdělení je diskrétní a exponenciální rozdělení spojité, existuje mezi nimi blízký vztah - každé z nich sice používáme k popisu jiné veličiny, ale hodnoty těchto veličin měříme v jedné a téže situaci.

Poissonovo rozdělení
Náhodná veličina X má Poissonovo rozdělení pravděpodobnosti s parametrem $\lambda > 0$ (píšeme $X \sim \mathrm{Po}(\lambda)$), jestliže nabývá hodnot $0, 1, 2, 3, \ldots$ s pravděpodobností

$$p(k) = P(X = k) = \frac{\lambda^k}{k!} \cdot \mathrm{e}^{-\lambda} \quad \text{pro } k = 0, 1, 2, 3, \ldots .$$

Tento vzorec používáme, když se ptáme na pravděpodobnost, že nastane právě k událostí za jednu časovou jednotku, přičemž víme, že průměrně nastane za tuto časovou jednotku λ událostí. Nemusí ale jít jen o jednotky časové, může jít i o jednotky délky, obsahu a pod.
Střední hodnota a rozptyl Poissonova rozdělení jsou

$$\mathrm{E}X = \lambda, \quad \mathrm{D}X = \lambda.$$

Občas je třeba zjistit pravděpodobnost, že nastane právě k událostí za t časových jednotek, když víme, že průměrně nastane λ událostí za jednu časovou jednotku. Snadno určíme, že za t časových jednotek nastane průměrně $\lambda \cdot t$ událostí, a proto pro náhodnou veličinu Y udávající počet událostí za t časových jednotek platí

$$p(k) = P(Y = k) = \frac{(\lambda \cdot t)^k}{k!} \cdot \mathrm{e}^{-\lambda \cdot t} \quad \text{pro } k = 0, 1, 2, 3, \ldots .$$

Samozřejmě, problém můžeme vyřešit i tak, že si pro novou časovou jednotku přepočítáme λ a použijeme předchozí vzoreček.
Na následujícím příkladu si podrobně vysvětlíme význam parametrů λ a k ve vzorcích.

Příklad 7.1. Do restaurace přijde průměrně 20 lidí za hodinu. Určete pravděpodobnost toho, že

- v průběhu 5 minut přijdou alespoň 2 lidé,
- v průběhu 15 minut nepřijde nikdo,
- v průběhu 5 minut nepřijde nikdo,
- v průběhu hodiny přijde právě 20 lidí,
- v průběhu hodiny přijde právě 15 lidí.

Řešení. Víme, že za hodinu přijde průměrně 20 lidí, proto pro časovou jednotku 1 hodina je $\lambda = 20$. Postupně vyřešíme jednotlivé úkoly:

- v průběhu 5 minut přijdou alespoň 2 lidé:

  V tomto případě vidíme, že časové jednotky nejsou stejné, proto si musíme uvědomit, že 5 minut je $\frac{1}{12}$ z hodiny a že bude vhodné použít druhý vzoreček. Zřejmě bude výhodnější použít vědomosti o doplňkovém jevu:

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1),$$

  a proto

$$P(X \geq 2) = 1 - \frac{\left(20 \cdot \frac{1}{12}\right)^0}{0!} \cdot \mathrm{e}^{-20 \cdot \frac{1}{12}} - \frac{\left(20 \cdot \frac{1}{12}\right)^1}{1!} \cdot \mathrm{e}^{-20 \cdot \frac{1}{12}} = 1 - \mathrm{e}^{-\frac{5}{3}} \cdot \left(1 + \frac{5}{3}\right).$$

  Postupovat jsme mohli i tak, že bychom určili λ pro novou časovou jednotku, tedy pro časovou jednotku 5 minut. Když za hodinu přijde průměrně 20 lidí, tak za 5 minut to bude průměrně $\frac{5}{3}$ lidí. A už můžeme použít první vzoreček, samozřejmě opět využijeme doplňkový jev:

$$P(X \geq 2) = 1 - \frac{\left(\frac{5}{3}\right)^0}{0!} \cdot \mathrm{e}^{-\frac{5}{3}} - \frac{\left(\frac{5}{3}\right)^1}{1!} \cdot \mathrm{e}^{-\frac{5}{3}} = 1 - \mathrm{e}^{-\frac{5}{3}} \cdot \left(1 + \frac{5}{3}\right).$$

- v průběhu 15 minut nepřijde nikdo:

  Máme zjistit $P(X = 0)$ a naše časové jednotky jsou opět různé, první výpočet uděláme pro $\lambda = 20$ a $t = \frac{1}{4}$ (15 minut je $\frac{1}{4}$ z hodiny).

$$P(X = 0) = \frac{\left(20 \cdot \frac{1}{4}\right)^0}{0!} \cdot \mathrm{e}^{-20 \cdot \frac{1}{4}} = \mathrm{e}^{(-5)}.$$

  Úloha se dala řešit i tak, že si zjistíme, kolik lidí přijde průměrně za 15 minut. To je čtvrtina z 20, tedy 5 lidí. Proto

$$P(X = 0) = \frac{(5)^0}{0!} \cdot \mathrm{e}^{-5} = \mathrm{e}^{(-5)}.$$

- v průběhu 5 minut nepřijde nikdo:

  Využijeme zkušenosti z předchozích příkladů:

$$P(X=0)=\frac{\left(20\cdot\frac{1}{12}\right)^0}{0!}\cdot\mathrm{e}^{-20\cdot\frac{1}{12}}=\mathrm{e}^{\left(-\frac{5}{3}\right)}$$

  nebo

$$P(X=0)=\frac{\left(\frac{5}{3}\right)^0}{0!}\cdot\mathrm{e}^{-\frac{5}{3}}=\mathrm{e}^{\left(-\frac{5}{3}\right)}.$$

- v průběhu hodiny přijde právě 20 lidí:

  V tomto případě máme stejné časové jednotky, proto

$$P(X=20)=\frac{20^{20}}{20!}\cdot\mathrm{e}^{-20}.$$

- v průběhu hodiny přijde právě 15 lidí:

  Zřejmě

$$P(X=15)=\frac{20^{15}}{15!}\cdot\mathrm{e}^{-20}.$$

Nyní předvedeme příklad, kdy nepůjde zrovna o události v čase.

Příklad 7.2. Ve sbírce příkladů se vyskytuje průměrně 1 chyba na pět stránek. Určete pravděpodobnost, že

- na stránce, kterou si zrovna prohlížíme, je právě jedna chyba,
- na stránce, kterou si zrovna prohlížíme, je aspoň jedna chyba.
- na stránce, kterou si zrovna prohlížíme, je více než jedna chyba.
- Dále určete, jaký je očekávaný (střední) počet chyb na deseti stránkách a jaká je pravděpodobnost, že na deseti stránkách se vyskytne právě tento počet chyb.

Řešení. Pro první tři úkoly označíme X počet chyb na jedné stránce. Za jednotku (i když ne zrovna časovou) tedy považujeme jednu stranu textu. Parametr λ je pak roven $1/5 = 0{,}2$. Další výpočet probíhá obvyklým způsobem.

- Právě jedna chyba na stránce:

$$P(X=1)=\frac{0{,}2^1}{1!}\cdot\mathrm{e}^{-0{,}2}\doteq 0{,}164.$$

- Aspoň jedna chyba na stránce:

$$P(X>0)=1-P(X=0)=1-\frac{0{,}2^0}{0!}\cdot\mathrm{e}^{-0{,}2}\doteq 0{,}181.$$

- Více než jedna chyba:
  Můžeme použít předchozí výsledky:
  $$P(X > 1) = P(X > 0) - P(X = 1) \doteq 0{,}017.$$
- Na deseti stránkách očekáváme dvě chyby. Označíme-li Y počet chyb na deseti stranách, pak $Y \sim Po(2)$ a
  $$P(Y = 2) = \frac{2^2}{2!} \cdot \mathrm{e}^{-2} \doteq 0{,}271.$$

Exponenciální rozdělení
Spojitá náhodná veličina X má exponenciální rozdělení s parametrem $\lambda > 0$ (píšeme $X \sim \mathrm{Exp}(\lambda)$), jestliže má hustotu f danou předpisem

$$f(t) = \begin{cases} 0 & \text{pro } t < 0, \\ \lambda \cdot \mathrm{e}^{-\lambda t} & \text{pro } t \geq 0. \end{cases}$$

Její distribuční funkce je

$$F(t) = \begin{cases} 0 & \text{pro } t < 0, \\ 1 - \mathrm{e}^{-\lambda t} & \text{pro } t \geq 0. \end{cases}$$

Náhodná veličina s exponenciálním rozdělením popisuje dobu čekání na další událost (resp. délku časového intervalu mezi dvěma událostmi), jestliže průměrně nastává λ událostí za jednotku času.
Střední hodnota a rozptyl jsou

$$\mathrm{E}X = \frac{1}{\lambda}, \quad \mathrm{D}X = \frac{1}{\lambda^2}.$$

Užití exponenciálního rozdělení objasníme na následujícím příkladu.

Příklad 7.3. Do restaurace přijde průměrně 20 lidí za hodinu. Určete pravděpodobnost toho, že

- v průběhu 15 minut nepřijde nikdo,
- v průběhu 5 minut nepřijde nikdo,

Řešení.
- v průběhu 15 minut nepřijde nikdo:

  Nejdřív si uvědomíme, že za 15 minut přijde průměrně 5 lidí, proto – pro časovou jednotku 15 minut – je $\lambda = 5$. Úlohu jsme již vyřešili užitím Poissonova rozdělení, ale snadno se dá aplikovat i exponenciální rozdělení, stačí si uvědomit, že když v průběhu 15 minut nepřijde nikdo, znamená to, že mezi dvěma příchody bude doba delší nebo rovna 15 minut. Proto
  $$P(X \geq 1) = 1 - P(X < 1) = \mathrm{e}^{(-5)}.$$

- v průběhu 5 minut nepřijde nikdo:

  Využijeme zkušenosti z předchozí části, stačí určit parametr λ. Ale to už umíme, $\lambda = \frac{5}{3}$. Proto

$$P(X \geq 1) = 1 - P(X < 1) = \mathrm{e}^{(-\frac{5}{3})}.$$

Zkusíme trochu jiný příklad:

Příklad 7.4. Roční dítě, když se nechá bez dozoru, dovede rozbít nebo pokazit průměrně 4 hračky za hodinu. Jak dlouho ho může maminka nechat bez dozoru, aby s pravděpodobností 0,9 nenastal problém?

Řešení. Zřejmě potřebujeme zjistit čas, za který nedojde k rozbití hračky s pravděpodobností 0,9, a víme, že jestliže za časovou jednotku zvolíme 1 hodinu, pak $\lambda = 4$. Potom

$$P(X > t) = 1 - P(X \leq t) = 1 - P(X < t) = 1 - F(t) = 0{,}9$$

Proto

$$F(t) = 0{,}1,$$

teda

$$1 - e^{-4t} = 0{,}1 \Rightarrow t = \frac{\ln 0{,}9}{-4} \doteq 0{,}0263$$

Čas $t \doteq 0{,}0263$, co je přibližně 1,58 minut.

Poissonovo rozdělení se občas hodí např. na aproximaci binomického rozdělení, a to v případě, že n je „velké“ a p „malé“. Za parametr λ se bere součin $n \cdot p$. Vyvětlíme to na následujícím příkladu.

Příklad 7.5. Dělnice obsluhuje 800 vřeten, na které navíjí přízi. Pravděpodobnost roztržení příze za čas t je na každém z nich 0,005. Jaká je pravděpodobnost toho, že se za čas t příze roztrhne právě na čtyrech vřetenech?

Řešení. Jedná se o binomické rozdělení, proto

$$P(X = 4) = \binom{800}{4} \cdot 0{,}005^4 \cdot (1 - 0{,}005)^{800-4} \doteq 0{,}1959.$$

Aproximujme binomické rozdělení Poissonovým, $\lambda = n \cdot p = 800 \cdot 0{,}005 = 4$, proto

$$P(X = 4) = \mathrm{e}^{-4} \cdot \frac{4^4}{4!} \doteq 0{,}1954.$$

Vidíme, že výsledky se liší jen nepatrně a výpočet pomocí Poissonova rozdělení byl přitom mnohem méně pracný.
Pokusíme se vysvětlit, proč lze tuto aproximaci použít. Na problém se podíváme z jiné strany: Jak se asi mohlo přijít na to, že pro každé vřeteno je pravděpodobnost přetržení příze zrovna 0,005? Patrně se za čas t průměrně trhá příze na 4 vřetenech (pak je pro jedno vřeteno pravděpodobnost přetržení příze $4/800 = 0{,}005$. Počítáme tedy události,

přičemž víme, že průměrně jich nastane $\lambda = 4$ za danou jednotku času, a to je přesně situace pro použití Poissonova rozdělení.

V případě Poissonova rozdělení je $EX = DX = \lambda$. Když aproximujeme binomické rozdělení Poissonovým, mělo by pro binomické rozdělení platit něco podobného. Proto tuto aproximaci použijeme, když není velký rozdíl mezi $\mathrm{E}X$ a $\mathrm{D}X$. Protože u binomického rozdělení s parametry n a p je $\mathrm{E}X = n \cdot p$ a $\mathrm{D}X = n \cdot p \cdot (1-p)$, jsou si tyto hodnoty blízké v případě, že $1-p$ je blízké jedné, tj. p je blízké nule.

7.1 Příklady pro samostatnou práci

Příklad 7.6. Do restaurace přijde průměrně 20 zákazníků za půlhodinu. Určete pravděpodobnost toho, že
a) v průběhu 5 minut přijdou alespoň dva zákazníci,
b) v průběhu 15 minut nepřijde ani jeden zákazník,
c) v průběhu 5 minut nepřijde ani jeden zákazník,
d) v průběhu jedné hodiny přijde právě 20 zákazníků,
e) v průběhu jedné hodiny přijde právě 15 zákazníků.

Výsledek: a) $1 - \frac{13}{3} \cdot \mathrm{e}^{\frac{-10}{3}}$; b) e^{-10}; c) $\mathrm{e}^{\frac{-10}{3}}$; d) $\frac{40^{20}}{20!} \cdot \mathrm{e}^{-40}$; e) $\frac{40^{15}}{15!} \cdot \mathrm{e}^{-40}$

Příklad 7.7. Do kanceláře přijdou průměrně 2 studenti za hodinu. Určete pravděpodobnost toho, že
a) v průběhu 5 minut přijdou alespoň dva studenti,
b) v průběhu 15 minut nepřijde ani jeden student,
c) v průběhu jedné hodiny přijdou právě 2 studenti,
d) doba mezi dvěma po sebe jdoucími příchody studentů je v intervalu $\langle 10min, 50min\rangle$.

Výsledek: a) $1 - \frac{7}{6} \cdot \mathrm{e}^{\frac{-1}{6}}$; b) $\mathrm{e}^{-0{,}5}$; c) $2 \cdot \mathrm{e}^{-2}$; d) přibližně 0,5278

Příklad 7.8. Do třídy přijde průměrně 5 opozdilých studentů za vyučovací hodinu (50 minut) . Určete pravděpodobnost toho, že
a) v průběhu 5 minut přijdou alespoň dva opozdilci,
b) v průběhu 5 minut nepřijde ani jeden student,
c) v průběhu 15 minut přijde alespoň jeden student,
d) v průběhu jedné vyučovací hodiny přijde právě 20 studentů,
e) v průběhu jedné vyučovací hodiny nepřijde ani jeden opozdilý student.

Výsledek: a) $1 - \frac{3}{2} \cdot \mathrm{e}^{\frac{-1}{2}}$; b) $\mathrm{e}^{-\frac{1}{2}}$; c) $1 - \mathrm{e}^{\frac{-3}{2}}$; d) $\frac{5^{20}}{20!} \cdot \mathrm{e}^{-5}$; e) e^{-5}

Příklad 7.9. Roční dítě je nutno průměrně 7-krát za den přebalit. Když se nechá bez dozoru, dovede rozbít nebo pokazit průměrně 4 hračky za hodinu. Jak dlouho ho může maminka nechat bez dozoru, aby s pravděpodobností 0,9 nenastal ani jeden z uvedených problémů?

Výsledek: Přibližně 1,473 min

Příklad 7.10. Náhodná veličina X má Poissonovo rozdělení se střední hodnotou $EX = 3$. Určete
a) pravděpodobnost toho, že náhodná veličina nabývá hodnoty menší než její střední hodnota,
b) pravděpodobnost toho, že náhodná veličina nabývá kladné hodnoty.

Výsledek: a) 0,423; b) 0,950

Příklad 7.11. Do telefonní ústředny přijde za hodinu průměrně 120 hovorů. Určete jaká je pravděpodobnost toho, že za dvě minuty přijdou do ústředny právě dva hovory.

Výsledek: $8 \cdot e^{-4}$

Příklad 7.12. Telefonní ústředna obsluhuje 3000 účastníků. Pravděpodobnost, že libovolný účastník bude telefonovat v průběhu hodiny je $p = 0,002$. Náhodná veličina X udává počet účastníků, kteří budou v průběhu hodiny telefonovat. Vypočítejte pravděpodobnost toho, že v průběhu hodiny budou telefonovat právě čtyři účastníci.

Výsledek: $P(X = 4) = 54 \cdot e^{-6}$

Příklad 7.13. Výrobní zařízení má poruchu v průměru jednou za 2000 hodin. Předpokládejme, že „doba čekání“ na poruchu je náhodná veličina s exponenciálním rozdělením. Určete hodnotu t tak, aby pravděpodobnost, že přístroj bude pracovat delší dobu než t, byla 0,99.

Výsledek: Přibližně 20,5 hodin

Příklad 7.14. Na jednom metru vlny se průměrně vyskytují dva uzlíky. Vypočtěte pravděpodobnost, že na dvou metrech vlny bude tři až pět uzlíků.

Výsledek: Přibližně 0,547

Příklad 7.15. Počet povrchových kazů na plastových deskách použitých v interiéru auta má Poissonovo rozdělení se střední hodnotou 0,005 kazu na decimetr čtvereční. Předpokládejme, že celková plocha těchto desek v jednom autě je 1 metr čtvereční.
a) Jaká je pravděpodobnost, že v autě není žádný kaz?
b) Jestliže si podnik objedná pět aut, jaká je pravděpodobnost, že se v žádném z nich povrchové kazy neobjeví?

Výsledek: a) $P(X = 0) = \mathrm{e}^{-0,5} \doteq 0{,}607$; b) $(\mathrm{e}^{-0,5})^5 \doteq 0{,}082$

Příklad 7.16. Učitel zapisuje body 200 studentům do informačního systému. Ze zkušenosti ví, že u každého záznamu je pravděpodobnost chyby 0,01. Vypočtěte pravděpodobnost, že učitel udělá dvě až tři chyby
a) pomocí binomického rozdělení,
b) pomocí Poissonova rozdělení.

Výsledek: Přibližně: a) 0,453; b) 0,451

8 Normální rozdělení

8.1 Základní pojmy a řešené příklady

Normální rozdělení
Spojitá náhodná veličina X s hustotou

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\,\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in (-\infty, \infty),$$

kde $\mu \in (-\infty, \infty)$ a $\sigma > 0$ jsou konstanty, se nazývá náhodná veličina s normálním rozdělením. Platí

$$\mathrm{E}X = \mu, \qquad \mathrm{D}X = \sigma^2.$$

To, že náhodná veličina X má normální rozdělení se střední hodnotou μ a rozptylem σ^2, zapisujeme jako

$$X \sim \mathrm{N}(\mu, \sigma^2).$$

Standardizované normální rozdělení
Jesliže má náhodná veličina normální rozdělení se střední hodnotou $\mu = 0$ a rozptylem $\sigma^2 = 1$, říkáme, že tato náhodná veličina má standardizované normální rozdělení a značíme ji U. Platí tedy

$$U \sim \mathrm{N}(0, 1).$$

Distribuční funkce náhodné veličiny U se značí Φ, tj. (podle obecné definice distribuční funkce) $\Phi(u) = P(U \leq u)$. Hodnoty funkce Φ jsou tabelovány (viz tabulka 9.1).
Jestliže X je náhodná veličina s normálním rozdělením se střední hodnotou μ a rozptylem σ^2, $X \sim \mathrm{N}(\mu, \sigma^2)$, pak

$$\frac{X - \mu}{\sigma} = U.$$

Vzorce pro výpočty pravděpodobností nerovností různého typu jsme už uváděli pro obecnou náhodnou veličinu se spojitým rozdělením. Nyní vše zopakujeme speciálně pro náhodnou veličinu U a její distribuční funkci Φ.

Výpočty různých pravděpodobností pro U pomocí Φ

$$P(U \leq u) = \Phi(u)$$
$$P(U > u) = 1 - \Phi(u)$$
$$P(u_1 < U \leq u_2) = \Phi(u_2) - \Phi(u_1).$$

Ve všech uvedených vztazích lze zaměnit ostré a neostré nerovnosti.
Pro funkci Φ navíc platí vztah

$$\Phi(-u) = 1 - \Phi(u).$$

Výpočty různých pravděpodobností pro obecné normální rozdělení
Jestliže $X \sim \mathrm{N}(\mu, \sigma^2)$, pak pro výpočty různých pravděpodobností využijeme transformační vztah mezi X a U. Konkrétně,

$$P(X \leq x) = P(U \leq \tfrac{x-\mu}{\sigma})$$
$$P(X > x) = P(U > \tfrac{x-\mu}{\sigma})$$
$$P(x_1 < X \leq x_2) = P(\tfrac{x_1-\mu}{\sigma} < U \leq \tfrac{x_2-\mu}{\sigma}).$$

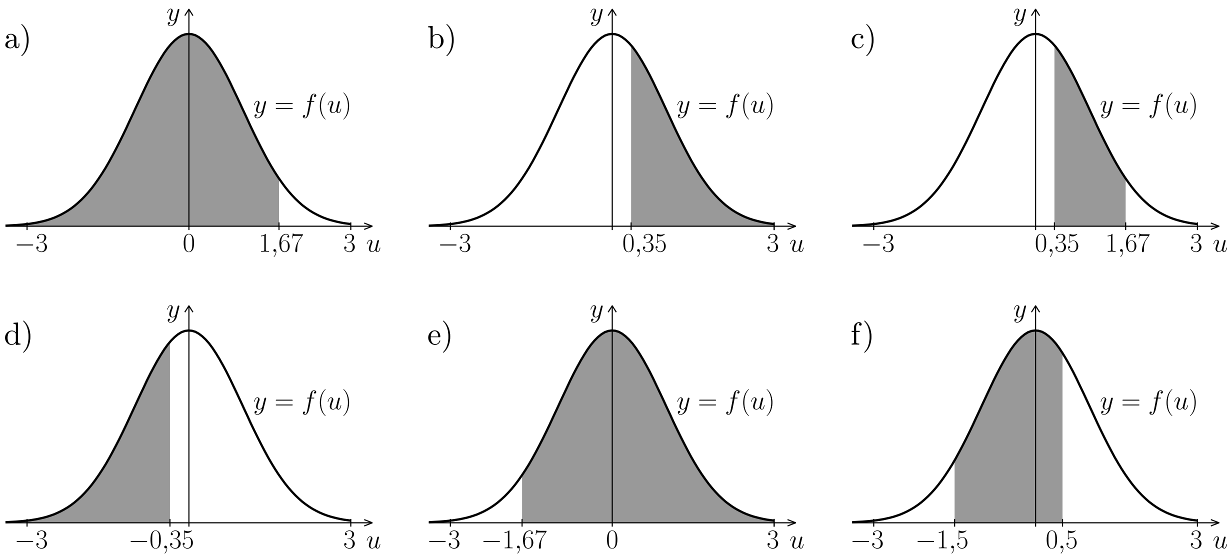
Dál pak počítáme s náhodnou veličinou U, viz předchozí rámeček.

Příklad 8.1. Vypočtěte pravděpodobnosti:

a) $P(U < 1{,}67)$ b) $P(U > 0{,}35)$ c) $P(0{,}35 < U < 1{,}67)$
d) $P(U < -0{,}35)$ e) $P(U > -1{,}67)$ f) $P(-1{,}5 < U < 0{,}5)$

Řešení. Budeme postupovat podle výše uvedených vzorců, příslušné hodnoty funkce Φ najdeme vždy v tabulce.
a) $P(U < 1{,}67) = \Phi(1{,}67) \doteq 0{,}953$
b) $P(U > 0{,}35) = 1 - \Phi(0{,}35) \doteq 1 - 0{,}637 = 0{,}363$
c) $P(0{,}35 < U < 1{,}67) = \Phi(1{,}67) - \Phi(0{,}35) \doteq 0{,}953 - 0{,}637 = 0{,}316$
d) $P(U < -0{,}35) = \Phi(-0{,}35) = 1 - \Phi(0{,}35) \doteq 1 - 0{,}637 = 0{,}363$
e) $P(U > -1{,}67) = 1 - \Phi(-1{,}67) = 1 - (1 - \Phi(1{,}67)) = \Phi(1{,}67) \doteq 0{,}953$
f) $P(-1{,}5 < U < 0{,}5) = \Phi(0{,}5) - \Phi(-1{,}5) = \Phi(0{,}5) - (1 - \Phi(1{,}5)) = \Phi(0{,}5) - 1 + \Phi(1{,}5) \doteq$
$\doteq 0{,}691 - 1 + 0{,}933 = 0{,}624$

Jako ilustrace k právě uvedeným výpočtům slouží obrázky 8.1. Zdůrazněme, že je na nich vždy graf hustoty normálního rozdělení s $\mu = 0$ a $\sigma^2 = 1$. Příslušnou pravděpodobnost počítáme jako obsah plochy pod grafem hustoty na odpovídajícím intervalu. Na obrázku a) šedě vybarvená plocha přímo udává hodnotu distribuční funkce v bodě 1,67. Na obrázku b) by hodnota $\Phi(0{,}35)$ byl obsah bílé plochy. Obsah vybarvené plochy na obrázku c)



Obr. 8.1: K příkladu 8.1

dostaneme tak, že odečteme obsah šedé plochy z obrázku a) a bílé plochy z obrázku b). Porovnáním dvojic obrázků a) a e), resp. b) a d) vidíme, že výsledky musely vyjít stejně.
Upozornění na častou chybu: Někteří studenti mívají problémy s pravděpodobnostmi nerovností typu $a < U < b$ (viz část c)). Občas má někdo tendenci takovouto pravděpodobnost začít počítat např. jako

$$P(0{,}35 < U < 1{,}67) = P(U < 1{,}67) + P(U > 0{,}35) = \cdots \qquad \text{(toto je špatně)}$$

nebo dělá jiné nešťastné pokusy. Máte-li k tomuto sklony, zkuste si nakreslit příslušný obrázek a uvědomit si, jaké obsahy by se zde sčítaly – výsledek by vyšel větší než 1, což u pravděpodobnosti nelze.
Jiná častá chyba: Při opravování písemek se nezřídka setkáváme se zcela nesmyslnými zápisy (i když konečný výsledek je pak třeba numericky dobře), jako např.

$$P(1 < U < 2) = \Phi(1) < P(U) < \Phi(2) = \cdots \qquad \text{Toto je úplně špatně!!!}$$

Je-li někdo schopen napsat takovouto věc, znamená to, že se naprosto nezorientoval v základních pojmech a snaží se jen zoufale napodobit jakýsi postup, kterému vůbec nerozumí. Marně se pak nešťastný student brání, že „to přece měl dobře, jenom to špatně zapsal“…

Příklad 8.2. Najděte hodnotu u, pro kterou platí
a) $P(U < u) = 0{,}9$ b) $P(U > u) = 0{,}05$ c) $P(U > u) = 0{,}8$ d) $P(-u < U < u) = 0{,}9$

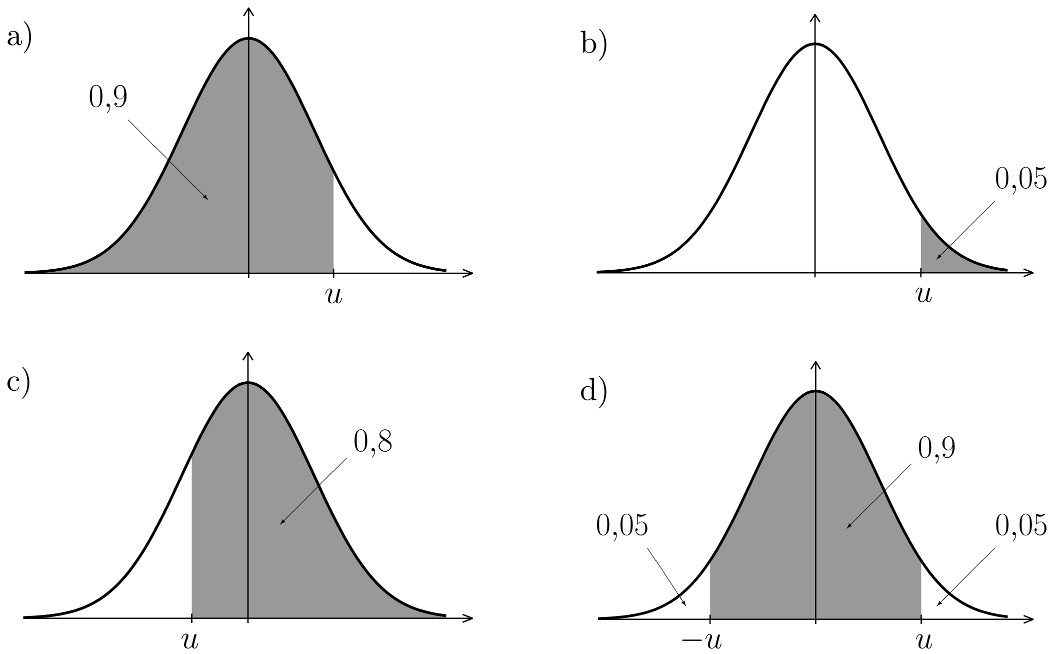
Řešení. Pro názornost si můžeme nejprve prohlédnout obrázky 8.2.
a) Máme najít u, pro které je $P(U < u) = 0{,}9$. Protože $P(U < u)$ je přímo $\Phi(u)$, znamená to, že hledáme u, pro které platí

$$\Phi(u) = 0{,}9.$$

Pohledem do tabulky zjistíme (příslušný výřez viz vpravo), že

$$u \doteq 1{,}29.$$

u	$\Phi(u)$
$\vdots$	$\vdots$
1,29	0,9014747
$\vdots$	$\vdots$



Obr. 8.2: K příkladu 8.2

Upozornění na častou chybu: Občas někteří studenti vypočítají omylem $\Phi(0{,}9)$ a vyhlásí, že $u \doteq 0{,}816$. Pozor, $\Phi(0{,}9)$ nás teď vůbec nezajímá! Musíme najít to u, pro které je $\Phi(u)$ rovno 0,9 - viz výše uvedená část tabulky.

b) Zadanou pravděpodobnost vyjádříme pomocí funkce Φ:

$$P(U > u) = 1 - P(U < u) = 1 - \Phi(u)$$

Odtud máme

$$1 - \Phi(u) = 0{,}05 \quad \Rightarrow \quad \Phi(u) = 0{,}95 \quad \Rightarrow \quad u \doteq 1{,}65.$$

c) Budeme postupovat obdobně jako v části b). Opět je zadaná pravděpodobnost, že U je větší než u. Musí proto platit

$$1 - \Phi(u) = 0{,}8 \quad \Rightarrow \quad \Phi(u) = 0{,}2.$$

Nyní může nastat mírný problém – hodnotu u, pro kterou by bylo $\Phi(u) = 0{,}2$, v tabulce přímo nenajdeme. Je zřejmé, že u bude záporné, využijeme proto vztahu mezi $\Phi(u)$ a $\Phi(-u)$ (s tím, že $-u$ je teď kladné):

$$1 - \Phi(-u) = 0{,}2 \quad \Rightarrow \quad \Phi(-u) = 0{,}8 \quad \Rightarrow \quad -u \doteq 0{,}85 \quad \Rightarrow \quad u \doteq -0{,}85.$$

Příklad jsme mohli řešit i tak, že bychom si představili, že by se obrázek 8.2 c) otočil kolem osy y – vybarvená plocha by pak udávala $P(U < -u)$ ($-u$ je opačná hodnota k hledanému zápornému u) a přímo bychom tak dospěli k rovnici $\Phi(-u) = 0{,}8$.

d) Zde máme dvě možnosti:

Můžeme vyjádřit zadanou pravděpodobnost pomocí Φ,

$$P(-u < U < u) = \Phi(u) - \Phi(-u) = \Phi(u) - (1 - \Phi(u)) = 2\Phi(u) - 1,$$

a dosadit tento výraz do zadané rovnice $P(-u < U < u) = 0{,}9$:

$$2\Phi(u) - 1 = 0{,}9 \quad \Rightarrow \quad 2\Phi(u) = 1{,}9 \quad \Rightarrow \quad \Phi(u) = 0{,}95 \quad \Rightarrow \quad u \doteq 1{,}65.$$

Jiná možnost je uvědomit si, že má-li obsah šedě vybarvené plochy na obrázku 8.2 d) být roven 0,9, znamená to, že oba dva přečnívající kousky (vlevo a vpravo) mají dohromady obsah 0,1, a tím pádem každý z nich musí mít obsah 0,05. Hledáme proto u, pro které by platilo $P(U > u) = 0{,}05$, což je ovšem úloha, kterou jsme už vyřešili v části b). Srovnejte též obrázky 8.2 b) a d), z nich je přímo vidět, že výsledky těchto dvou příkladů musí být stejné.

Příklad 8.3. Výsledek měření proudu je náhodná veličina X s normálním rozdělením se střední hodnotou $\mu = 10\,\mathrm{mA}$ a rozptylem $\sigma^2 = 4\,(\mathrm{mA})^2$. Vypočtěte pravděpodobnost, že bude naměřeno

a) méně než 13 mA b) více než 9 mA c) méně než 12 mA, ale více než 5 mA

Řešení. Podle zadání je $X \sim \mathrm{N}(10, 4)$. Abychom mohli vypočítat požadované pravděpodobnosti, musíme náhodnou veličinu X převést na U. Obecně platí

$$U = \frac{X - \mu}{\sigma}.$$

Ze zadání známe nikoli přímo σ, ale σ^2. Snadno vypočteme, že $\sigma = \sqrt{4} = 2\,\mathrm{mA}$. Vztah mezi X a U je tedy

$$U = \frac{X - 10}{2}.$$

Funkční závislost náhodné veličiny U na náhodné veličině X je znázorněna na obrázku 8.3. Nyní přistupme k výpočtu jednotlivých pravděpodobností.

a) Máme vypočítat $P(X < 13)$. Z obrázku 8.3 vidíme, že hodnotám X menším než 13 odpovídají hodnoty U menší než $\frac{13-10}{2} = 1{,}5$. Proto platí

$P(X < 13) = P(U < \frac{13-10}{2}) = P(U < 1{,}5) = \Phi(1{,}5) \doteq 0{,}933.$

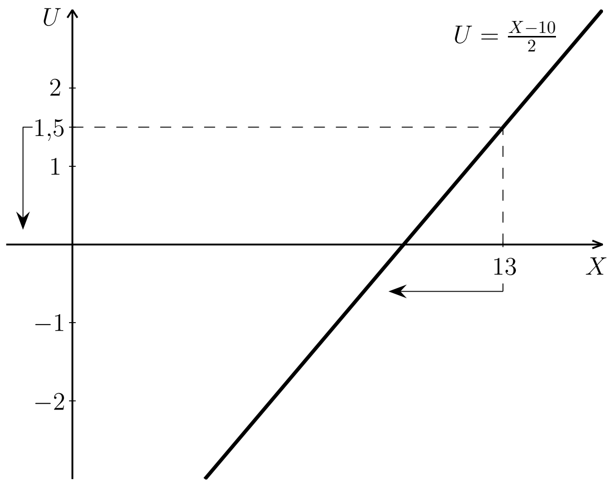
b) Počítáme $P(X > 9)$. Podobně jako v části a) si můžeme uvědomit, že hodnotám X větším než 9 odpovídají hodnoty U větší než $\frac{9-10}{2} = -0{,}5$. Proto

$$\begin{aligned}P(X > 9) &= P(U > \tfrac{9-10}{2}) = P(U > -0{,}5) = 1 - P(U < -0{,}5) = 1 - \Phi(-0{,}5) = \\ &= 1 - (1 - \Phi(0{,}5)) = \Phi(0{,}5) \doteq 0{,}691.\end{aligned}$$

c) Meze pro U určíme podobně jako v předchozích částech a pak už můžeme počítat pomocí Φ:

$$\begin{aligned}P(5 < X < 12) &= P(\tfrac{5-10}{2} < U < \tfrac{12-10}{2}) = P(-2{,}5 < U < 1) = \Phi(1) - \Phi(-2{,}5) = \\ &= \Phi(1) - (1 - \Phi(2{,}5)) = \Phi(1) + \Phi(2{,}5) - 1 \doteq 0{,}841 + 0{,}994 - 1 = 0{,}835.\end{aligned}$$

Upozornění na častou chybu: Ve vztahu mezi náhodnými veličinami X a U se ve jmenovateli zlomku vyskytuje σ, tj. směrodatná odchylka, zatímco v zadáních příkladů bývá často uveden rozptyl neboli σ^2. Musíme si proto vždy pečlivě přečíst zadání, a v případě, že je zadán rozptyl, jej odmocnit (a nic neodmocňovat, je-li zadána přímo směrodatná odchylka).



Obr. 8.3: K příkladu 8.3 – převod mezi hranicemi pro X a U

Příklad 8.4. Měřicí přístroj je zatížen jednak systematickou chybou 1 a dále chybou náhodnou. Víme, že náhodné chyby mají normální rozdělení se směrodatnou odchylkou $\sigma = 0{,}3$.
a) Nad jakou hodnotu se chyba dostane jen s pravděpodobností 0,01?
b) Pod jakou hodnotou bude chyba s pravděpodobností 0,02?
c) V jakých mezích (souměrných kolem 1) se dá očekávat chyba při 75 % měření?

Řešení. Celková chyba měření – označme ji X – má normální rozdělení se střední hodnotou $\mu = 1$ a směrodatnou odchylkou $\sigma = 0{,}3$. Tedy $X \sim \mathrm{N}(1;\ 0{,}3^2)$, a proto

$$U = \frac{X-1}{0{,}3}\,.$$

a) Hledáme x, pro které by platilo $P(X > x) = 0{,}01$. Stejně jako v předchozím příkladu náhodnou veličinu X převedeme na náhodnou veličinu U:

$$P(X > x) = 0{,}01 \quad\Rightarrow\quad P(U > \tfrac{x-1}{0{,}3}) = 0{,}01.$$

Označme $u = \frac{x-1}{0{,}3}$. Pak

$$P(U > u) = 0{,}01 \quad\Rightarrow\quad 1 - \Phi(u) = 0{,}01 \quad\Rightarrow\quad \Phi(u) = 0{,}99 \quad\Rightarrow\quad u \doteq 2{,}33.$$

Nyní dopočítáme hledanou hodnotu x:

$$\frac{x-1}{0{,}3} = 2{,}33 \quad\Rightarrow\quad x = 2{,}33 \cdot 0{,}3 + 1 = 1{,}699.$$

Dá se tedy čekat, že hodnotu 1,699 přesáhne chyba jen při 1 % měření.
b) Budeme postupovat obdobně jako v části a). Najdeme příslušnou hraniční hodnotu u a pak pomocí ní ze vztahu $u = \frac{x-1}{0{,}3}$ vypočítáme mezní hodnotu x, pro kterou platí $P(X < x) = 0{,}02$:

$$P(U < u) = 0{,}02 \quad\Rightarrow\quad \Phi(u) = 0{,}02.$$

Odtud vidíme (protože $0{,}02 < 0{,}5$), že u bude záporné. Použijeme proto příslušný vztah:

$$1 - \Phi(-u) = 0{,}02 \quad \Rightarrow \quad \Phi(-u) = 0{,}98 \quad \Rightarrow \quad -u \doteq 2{,}06 \quad \Rightarrow \quad u \doteq -2{,}06.$$

Zbývá vypočítat x:

$$\frac{x-1}{0{,}3} = -2{,}06 \quad \Rightarrow \quad x = -2{,}06 \cdot 0{,}3 + 1 = 0{,}382.$$

c) Zde hledáme dvě hraniční hodnoty, označme je x_1 a x_2, pro které platí

$$P(x_1 < X < x_2) = 0{,}75.$$

Navíc víme, že tyto hodnoty mají být souměrné kolem 1, tzn. že $x_1 = 1 - d$ a $x_2 = 1 + d$, kde $d > 0$ je nějaká konstanta, kterou zatím neznáme. Mohli bychom také říct, že hledáme právě toto d, a předchozí rovnici přepsat jako

$$P(-d < X - 1 < d) = 0{,}75.$$

Lze tušit (viz obrázek 8.4), že při přepočítávání mezních hodnot x_1 a x_2 na mezní hodnoty pro U dostaneme čísla navzájem opačná. Pro x_1 to bude

$$u_1 = \frac{x_1 - 1}{0{,}3} = \frac{1 - d - 1}{0{,}3} = -\frac{d}{0{,}3},$$

zatímco pro x_2 je to

$$u_2 = \frac{x_2 - 1}{0{,}3} = \frac{1 + d - 1}{0{,}3} = \frac{d}{0{,}3}\,.$$

Opravdu je tedy $u_1 = -u_2$. Pro zjednodušení přeznačíme u_2 na u a u_1 je pak rovno $-u$. Má platit, že

$$P(-u < U < u) = 0{,}75,$$

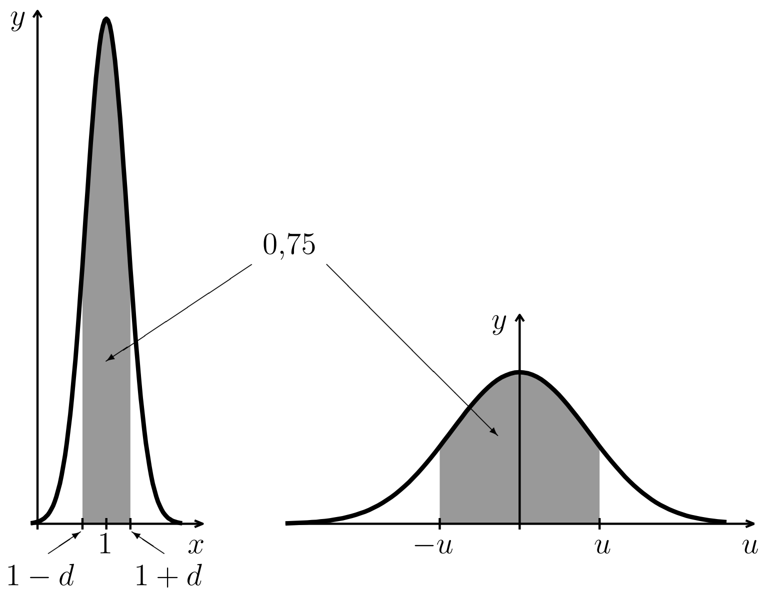
viz obrázek 8.4. Podobný úkol jsme řešili v příkladu 8.2, části d), kde jsme ukázali, že $P(-u < U < u) = 2\Phi(u) - 1$. Využijeme-li tohoto faktu, dostáváme:

$$2\Phi(u) - 1 = 0{,}75 \quad \Rightarrow \quad \Phi(u) = 0{,}875 \quad \Rightarrow \quad u \doteq 1{,}16.$$

Nyní můžeme dopočítat x_1 a x_2 (případně bychom mohli napřed spočítat d, a x_1 a x_2 pak pomocí něj):

$$\begin{aligned} \tfrac{x_1-1}{0{,}3} &= -1{,}16 &\Rightarrow x_1 &= -1{,}16 \cdot 0{,}3 + 1 = 0{,}652 \\ \tfrac{x_2-1}{0{,}3} &= 1{,}16 &\Rightarrow x_2 &= 1{,}16 \cdot 0{,}3 + 1 = 1{,}348. \end{aligned}$$

Dá se tedy čekat, že v 75 % měření bude chyba mezi 0,652 a 1,348.
Pro kontrolu bychom nyní mohli u každé z částí a), b), c) vypočítat zpětně příslušnou pravděpodobnost. Např. u části c) bychom počítali $P(0{,}652 < X < 1{,}348)$ a mělo by vyjít přibližně 0,75 (zcela přesně by to nebylo, protože při předchozím výpočtu jsme používali zaokrouhlené hodnoty).



Obr. 8.4: K příkladu 8.4, část c)

8.2 Příklady pro samostatnou práci

Příklad 8.5. Vypočtěte pravděpodobnosti:

a) $P(U > 1{,}23)$	c) $P(U > -6/7)$	e) $P(1 < U < 2)$	g) $P(\|U\| < 1{,}5)$
b) $P(U < 0{,}45)$	d) $P(U < -9/8)$	f) $P(-3 < U < -2)$	h) $P(\|U\| > 2)$

Výsledek: (Vše zaokrouhleno na 3 místa) a) 0,109; b) 0,674; c) 0,804; d) 0,130; e) 0,136; f) 0,021; g) 0,866; h) 0,046

Příklad 8.6. Aniž byste uvedené pravděpodobnosti počítali, doplňte místo znaku $\lesseqgtr$ jeden ze znaků $<$, $>$ nebo $=$, podle toho, v jakém vztahu tyto pravděpodobnosti jsou.

a) $P(U < 0{,}8) \lesseqgtr P(U < 0{,}9)$	e) $P(0{,}9 < U < 1{,}1) \lesseqgtr P(1{,}9 < U < 2{,}1)$
b) $P(U < 0{,}7) \lesseqgtr P(U > 0{,}7)$	f) $P(1 < U < 2) \lesseqgtr P(U < 2) + P(U > 1)$
c) $P(U > -2) \lesseqgtr P(U < 2)$	g) $P(1 < U < 2) \lesseqgtr 1 - P(U < 1) - P(U > 2)$
d) $P(U < -3) \lesseqgtr P(U < 3)$	h) $2P(U > 1) \lesseqgtr 1 - P(-1 < U < 1)$

Výsledek: a) $<$; b) $>$; c) $=$; d) $<$; e) $>$; f) $<$; g) $=$; h) $=$

Příklad 8.7. Najděte hodnotu u, pro kterou platí

a) $P(U > u) = 0{,}25$	c) $P(U < u) = 0{,}99$	e) $P(0 < U < u) = 0{,}35$
b) $P(U < u) = 0{,}1$	d) $P(-u < U < u) = 0{,}99$	f) $P(1 < U < u) = 0{,}2$

Výsledek: a) 0,68; b) -1,29; c) 2,33; d) 2,58; e) 1,04; f) nemá řešení

Příklad 8.8. Náhodná veličina X má normální rozdělení se střední hodnotou $\mu = 5$ a směrodatnou odchylkou $\sigma = 4$. Vypočtěte následující pravděpodobnosti:

a) $P(X < 11)$ c) $P(3 < X < 7)$ e) $P(\mu - \sigma < X < \mu + 2\sigma)$
b) $P(X > 0)$ d) $P(0 < X < \mu)$ f) $P(|X - \mu| < 0{,}4)$

Výsledek: a) 0,933; b) 0,894; c) 0,383; d) 0,394; e) 0,819; f) 0,080

Příklad 8.9. Náhodná veličina X má normální rozdělení, $X \sim \mathrm{N}(7, 9)$. Určete hodnotu x, pro kterou platí:
a) $P(X > x) = 0{,}5$ c) $P(X > x) = 0{,}95$ (pokuste se využít výsledku b))
b) $P(X < x) = 0{,}95$ d) $P(0 < X < x) = 0{,}95$ e) $P(-x < X - 7 < x) = 0{,}95$

Výsledek: a) 7; b) 11,95; c) 2,05; d) 12,28; e) 5,88

Příklad 8.10. Náhodná veličina X má normální rozdělení se střední hodnotou μ a rozptylem σ^2. Pro určité x platí, že $P(X < x) = 0{,}75$. Rozhodněte, jestli by se tato pravděpodobnost zvětšila nebo zmenšila, kdyby
a) μ bylo větší,
b) σ bylo větší,
c) x bylo větší.

Výsledek: Pravděpodobnost se a) zmenší; b) zmenší; c) zvětší.

Příklad 8.11. Náhodná veličina X má normální rozdělení se střední hodnotou μ a rozptylem σ^2. Pro $\alpha = 0{,}05$ je mezní hodnota x, pro kterou platí $P(X > x) = \alpha$, rovna 1. Které z následujících výroků platí?
a) Pro $\alpha = 0{,}01$ by mezní hodnota x vyšla menší než 1.
b) Pro $\alpha = 0{,}01$ by mezní hodnota x vyšla větší než 1.
c) Střední hodnota μ je určitě menší než 1.
d) Rozptyl σ^2 je určitě menší než 1.

Výsledek: Platí b) a c).

Příklad 8.12. Při kontrole jakosti přebíráme součástku tehdy, jestliže se její rozměr pohybuje v mezích 26–27 mm. Rozměry součástek mají normální rozdělení se střední hodnotou $\mu = 26{,}4$ mm a směrodatnou odchylkou $\sigma = 0{,}2$ mm. Jaká je pravděpodobnost, že rozměr součástky náhodně vybrané ke kontrole bude v požadovaných mezích?

Výsledek: 0,976

Příklad 8.13. Doba, kterou potřebuje buňka jistého typu, aby se rozdělila na dvě buňky, je náhodná veličina s normálním rozdělením se střední hodnotou 1 hodina a směrodatnou odchylkou 5 minut.
a) Jaká je pravděpodobnost, že se buňka rozdělí dříve než za 45 minut?
b) Jaká je pravděpodobnost, že buňce bude dělení trvat déle než 65 minut?
c) Do jaké doby se rozdělí 99% buněk?

Výsledek: a) 0,001; b) 0,159; c) 71,65 min

Příklad 8.14. V konzervárně plní sklenice s džemem. Hmotnost obsahu naplněné sklenice má normální rozdělení se střední hodnotou 250 g a směrodatnou odchylkou 5 g. Produkt je v normě, je-li jeho hmotnost v rozmezí 250 ± 8 g.

a) Jaká je pravděpodobnost, že náhodně vybraná sklenice je v normě?
b) V jakém rozmezí (souměrném kolem střední hodnoty) je hmotnost 95% sklenic?
c) Jestliže náhodně vybereme čtyři sklenice, jaká je pravděpodobnost, že všechny jsou v normě?
d) Jaká je pravděpodobnost, že aspoň jedna z náhodně vybraných čtyř sklenic není v normě?

Výsledek: a) 0,990; b) 240,2–259,8 g; c) $0,990^4 \doteq$ 0,961; d) $1 - 0,990^4 \doteq$ 0,039

Příklad 8.15. Reakční doba řidiče po spatření překážky je náhodná veličina s normálním rozdělením se střední hodnotou $\mu = 0{,}4$ sekundy a směrodatnou odchylkou $\sigma = 0{,}05$ sekundy.
a) Jaká je pravděpodobnost, že řidič zareaguje později než za 0,5 s?
b) Jaká je pravděpodobnost, že reakční doba bude v intervalu (0,4; 0,5)?
c) Jakou hodnotu reakční doba překročí v 90 % případů?

Výsledek: a) 0,023; b) 0,477; c) 0,336 sekundy

Příklad 8.16. Životnost laseru má normální rozdělení se střední hodnotou $\mu = 7000$ hodin a směrodatnou odchylkou $\sigma = 600$ hodin.
a) Jaká je pravděpodobnost, že laser selže dříve než za 6000 hodin?
b) Do jaké doby selže 95 % laserů?
c) V určitém zařízení pracují nezávisle na sobě 3 lasery. Jaká je pravděpodobnost, že po 7000 hodinách bude zařízení ještě fungovat?

Výsledek: a) 0,048; b) do 7987 hodin; b) 0,125

Příklad 8.17. Hmotnost páru speciálních běžeckých bot v určité velikosti má normální rozdělení se střední hodnotou $\mu = 340$ gramů a směrodatnou odchylkou $\sigma = 15$ gramů.
a) Jaká je pravděpodobnost, že hmotnost bot překročí 370 g?
b) Jaká by musela být směrodatná odchylka σ, aby výrobce mohl tvrdit, že 99,9 % párů bot nepřekročí hmotnost 370 g?
c) Kdyby σ bylo 15 g, jaká by musela být střední hodnota μ, aby výrobce mohl tvrdit, že 99,9 % párů bot nepřekročí hmotnost 370 g?

Výsledek: a) 0,023; b) přibližně 9,375 g; c) přibližně 322 g

Příklad 8.18. * Výpočtem příslušných integrálů dokažte, že pro náhodnou veličinu s hustotou $f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\,\mathrm{e}^{-(x-\mu)^2/(2\sigma)}$ pro $x \in (-\infty, \infty)$ je střední hodnota rovna μ a rozptyl roven σ^2. (Návod: využijte toho, že $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$.)

8.3 Aproximace binomického rozdělení normálním rozdělením

Připomeňme, že náhodná veličina X má binomické rozdělení pravděpodobnosti s parametry n a p, píšeme $X \sim \mathrm{Bi}(n, p)$, jestliže X udává počet úspěchů v n nezávislých pokusech, přičemž v každém jednotlivém pokusu úspěch nastává s pravděpodobností p (viz kapitola 6). Zopakujme též, že pro takovouto náhodnou veličinu je střední hodnota $\mathrm{E}X = np$ a rozptyl $\mathrm{D}X = np(1-p)$. Protože pro velký počet pokusů n se výpočty pravděpodobností technicky komplikují, je výhodné použít místo binomického rozdělení pravděpodobnosti rozdělení normální.

Aproximace binomického rozdělení normálním

Jestliže náhodná veličina X má binomické rozdělení, $X \sim \mathrm{Bi}(n,p)$, n je velké a p není příliš blízké nule ani jedničce, pak se uvedené rozdělení dá přibližně nahradit normálním rozdělením se stejnou střední hodnotou a rozptylem, jako mělo původní binomické rozdělení. Neboli

$$X \sim \mathrm{Bi}(n,p) \quad \Rightarrow \quad \text{přibližně } X \sim \mathrm{N}(\mu,\sigma^2), \quad \mu = np, \quad \sigma^2 = np(1-p).$$

Pro výpočet pravděpodobnosti pak platí

$$P(x_1 < X \leq x_2) \doteq P(\frac{x_1-\mu}{\sigma} < U \leq \frac{x_2-\mu}{\sigma}) = \Phi(\frac{x_2-\mu}{\sigma}) - \Phi(\frac{x_1-\mu}{\sigma}).$$

Formulace „n je velké a p není příliš blízké nule ani jedničce“ je ovšem dosti vágní. V literatuře lze nalézt např. podmínku, že aproximace je dobrá, je-li

$$np > 5 \quad \text{a současně} \quad n(1-p) > 5.$$

Aproximace s korekcí

Aproximaci binomického rozdělení normálním lze vylepšit pomocí tzv. korekce (viz příklad 8.20 a obrázky 8.5).

Jestliže x_1 a x_2 jsou celá čísla, pravděpodobnost $P(x_1 \leq X \leq x_2)$ počítáme jako

$$\begin{aligned} P(x_1 \leq X \leq x_2) &= P(x_1 - 0{,}5 < X < x_2 + 0{,}5) \doteq \\ &\doteq P\left(\frac{(x_1-0{,}5)-\mu}{\sigma} < U < \frac{(x_2+0{,}5)-\mu}{\sigma}\right) = \\ &= \Phi\left(\frac{(x_2+0{,}5)-\mu}{\sigma}\right) - \Phi\left(\frac{(x_1-0{,}5)-\mu}{\sigma}\right). \end{aligned}$$

Pozor, zde na typu nerovností záleží, protože X je diskrétní náhodná veličina! V případě $P(x_1 < X < x_2)$ by byl výsledek jiný.

Příklad 8.19. Pravděpodobnost narození chlapce je 0,515[1]. Jaká je pravděpodobnost, že

[1] Chlapců se opravdu rodí více než děvčat, poměr je údajně zhruba 106:100, což dává právě onu pravděpodobnost 0,515.

a) mezi 10 novorozenci bude více děvčat než chlapců,
b) mezi 10 000 novorozenci bude více děvčat než chlapců,
c) mezi 100 000 novorozenci bude relativní četnost chlapců v mezích od 0,513 do 0,515?

Řešení. a) Tato část má sloužit hlavně pro osvěžení binomického rozdělení. Zde je počet „pokusů“ (tj. narozených dětí) vcelku malý, takže pravděpodobnost spočítáme přesně. Má-li být mezi 10 novorozenci více děvčat než chlapců, musí být děvčat alespoň 6, tj. může jich být 6, 7, 8, 9, nebo 10. Označíme-li X počet narozených děvčat, pak X má binomické rozdělení s parametry $n = 10$ a $p = 1 - 0{,}515 = 0{,}485$ (počítáme děvčata). Požadovaná pravděpodobnost je

$$P(X \geq 6) = p(6) + p(7) + p(8) + p(9) + p(10) = \binom{10}{6} \cdot 0{,}485^6 \cdot 0{,}515^4 +$$
$$+ \binom{10}{7} \cdot 0{,}485^7 \cdot 0{,}515^3 + \binom{10}{8} \cdot 0{,}485^8 \cdot 0{,}515^2 + 10 \cdot 0{,}485^9 \cdot 0{,}515 + 0{,}485^{10} \doteq 0{,}341.$$

b) Označme X počet děvčat mezi 10 000 novorozenci. V tomto případě $X \sim Bi(10\,000; 0{,}485)$. Děvčat má být nadpoloviční většina, a proto jich musí být alespoň 5 001. Kdybychom tuto část řešili stejně jako a), zadaná pravděpodobnost by se počítala jako

$$P(X \geq 5\,001) = \sum_{k=5\,001}^{10\,000} p(k) = \binom{10\,000}{5\,001} \cdot 0{,}485^{5\,001} \cdot 0{,}515^{4\,999} + \cdots + 0{,}485^{10\,000},$$

což by bylo velmi pracné. Příklad proto vyřešíme pomocí aproximace normálním rozdělením. Nejprve vypočteme střední hodnotu a rozptyl náhodné veličiny X:

$$\mathrm{E}X = 10\,000 \cdot 0{,}485 = 4\,850, \quad \mathrm{D}X = 10\,000 \cdot 0{,}485 \cdot (1 - 0{,}485) = 2\,497{,}75.$$

Můžeme říct, že přibližně

$$X \sim \mathrm{N}(\mu = 4\,850; \sigma^2 = 2\,497{,}75).$$

Další výpočet už bude podobný jako např. v příkladu 8.3. Náhodnou veličinu X ztransformujeme (pozor, rozptyl $\mathrm{D}X = \sigma^2$ musíme opět odmocnit, abychom dostali σ):

$$U = \frac{X - \mu}{\sigma} = \frac{X - 4\,850}{\sqrt{2\,497{,}75}}$$

a vypočteme pravděpodobnost způsobem obvyklým pro normální rozdělení:

$$P(X \geq 5\,001) = P\left(U \geq \frac{5\,001 - 4\,850}{\sqrt{2\,497{,}75}}\right) \doteq P(U \geq 3) = 1 - P(U < 3)$$
$$= 1 - \Phi(3) \doteq 1 - 0{,}99865 = 0{,}00135.$$

Pro srovnání, přesným výpočtem pomocí binomického rozdělení by vyšlo 0,00130 – kvůli tomuto porovnání jsme uvedli tolik desetinných míst.

c) Tentokrát označíme X počet chlapců mezi 100 000 narozenými dětmi. Pro náhodnou veličinu X platí, že $X \sim \mathrm{Bi}(100\,000; 0{,}515)$, a tedy

$$EX = 100\,000 \cdot 0{,}515 = 51\,500, \quad DX = 100\,000 \cdot 0{,}515 \cdot 0{,}485 = 24\,977{,}5$$

Je-li relativní četnost (tj. poměr počtu narozených chlapců ku počtu všech dětí) mezi 0,513 a 0,515, znamená to, že chlapců se narodilo 51 300 až 51 500. Pro výpočet pravděpodobnosti opět použijeme aproximaci normálním rozdělením. Přibližně je

$$X \sim \mathrm{N}(51\,500; 24\,977{,}5), \quad \text{a tedy} \quad U = \frac{X - 51\,500}{\sqrt{24\,977{,}5}}\,.$$

Zadaná pravděpodobnost je

$$P(51\,300 \le X \le 51\,500) = P\left(\frac{51\,300 - 51\,500}{\sqrt{24\,977{,}5}} \le U \le \frac{51\,500 - 51\,500}{\sqrt{24\,977{,}5}}\right) \doteq$$
$$\doteq P(-1{,}27 \le U \le 0) = \Phi(0) - \Phi(-1{,}27) = \Phi(0) - (1 - \Phi(1{,}27)) \doteq 0{,}5 - (1 - 0{,}898) = 0{,}398.$$

(Přesným výpočtem s binomickým rozdělením by vyšlo 0,399.)

Příklad 8.20. V učebně je 25 počítačů. Každý z nich může být mimo provoz s pravděpodobností 0,1. Jaká je pravděpodobnost, že mimo provoz budou 2 až 3 počítače? Úlohu řešte
a) přesně,
b) pomocí normálního rozdělení,
c) pomocí normálního rozdělení s korekcí.
Porovnejte výsledky příkladů b) a c) s výsledkem a).

Řešení. Označme X počet pokažených počítačů v učebně. Platí, že $X \sim \mathrm{Bi}(25; 0{,}1)$.
a) Výsledek vypočteme jako součet pravděpodobností, že pokažené počítače budou právě dva a právě tři:

$$P(2 \le X \le 3) = p(2) + p(3) = \binom{25}{2} 0{,}1^2 \cdot 0{,}9^{23} + \binom{25}{3} 0{,}1^3 \cdot 0{,}9^{22} \doteq 0{,}492.$$

b) Vypočteme střední hodnotu a rozptyl náhodné veličiny X:

$$\mathrm{E}X = \mu = 25 \cdot 0{,}1 = 2{,}5, \quad \mathrm{D}X = \sigma^2 = 25 \cdot 0{,}1 \cdot 0{,}9 = 2{,}25 \quad \Rightarrow \quad \sigma = \sqrt{2{,}25} = 1{,}5.$$

Nyní pravděpodobnost vypočteme pomocí normálního rozdělení. Transformační vztah mezi X a U je $U = \frac{X - 2{,}5}{1{,}5}$.

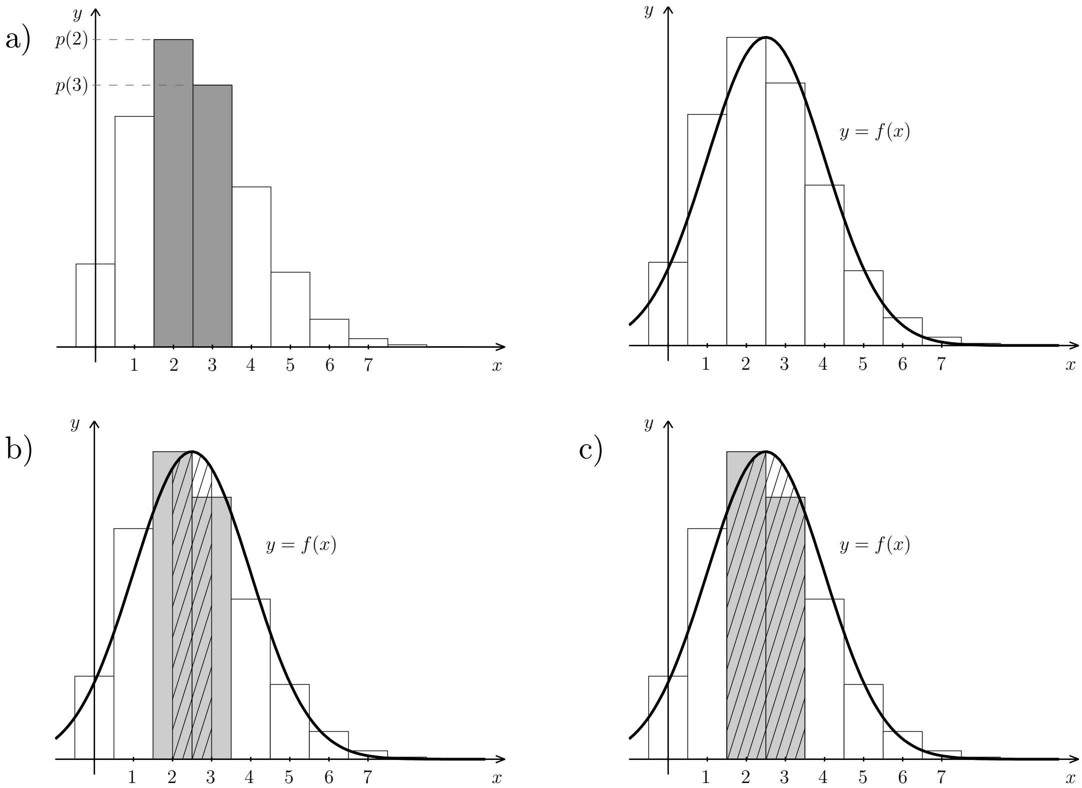
$$P(2 \le X \le 3) \doteq P\left(\frac{2 - 2{,}5}{1{,}5} \le U \le \frac{3 - 2{,}5}{1{,}5}\right) \doteq P(-0{,}33 \le U \le 0{,}33) =$$
$$= \Phi(0{,}33) - \Phi(-0{,}33) = \Phi(0{,}33) - (1 - \Phi(0{,}33)) = 2\Phi(0{,}33) - 1 \doteq 2 \cdot 0{,}629 - 1 = 0{,}258.$$

Vidíme, že výsledek se od přesné hodnoty, získané v části a), značně liší.

c) Použijeme korekci:

$$P(2 \leq X \leq 3) = P(1{,}5 < X < 3{,}5) \doteq P\left(\frac{1{,}5-2{,}5}{1{,}5} \leq U \leq \frac{3{,}5-2{,}5}{1{,}5}\right) \doteq$$
$$\doteq P(-0{,}67 \leq U \leq 0{,}67) = \Phi(0{,}67) - \Phi(-0{,}67) = 2\Phi(0{,}67) - 1 \doteq 2 \cdot 0{,}749 - 1 = 0{,}498.$$

Tentokrát je výsledek mnohem přesnější, liší se až na třetím desetinném místě.
Jako ilustrace k tomuto příkladu slouží obrázek 8.5. Na obrázku vlevo nahoře je histogram pravděpodobností binomického rozdělení. Pravděpodobnost vypočtená v části a) je reprezentována obsahy dvou vybarvených obdélníků. Oba tyto obdélníky mají šířku 1, výšky jsou rovny $p(2)$, resp. $p(3)$. Součet obsahů těchto obdélníku je proto právě $p(2) + p(3)$. Na obrázku vpravo nahoře je pak histogram spolu s grafem hustoty normálního rozdělení se stejnou střední hodnotou a rozptylem, jako má původní binomické rozdělení. Na obrázku vlevo dole vyšrafovaná plocha pod grafem hustoty určuje pravděpodobnost vypočtenou v části b), obrázek vpravo dole se vztahuje k části c). Odtud vidíme, proč je výsledek b) špatný, zatímco výsledek c) se hodně blíží přesné hodnotě. Kdybychom podobný obrázek nakreslili k příkladu 8.19 (příklad s narozenými dětmi) b) nebo c), jednotlivé sloupce histogramu by měly velmi malou výšku (např. pro část b) předchozího příkladu je $p(5001) \doteq 8 \cdot 10^{-5}$, takže korekce by se příliš neprojevila.



Obr. 8.5: K příkladu 8.20

Příklad 8.21. Pod jakou hodnotou zůstane počet šestek při 1000 hodech kostkou s pravděpodobností 0,9?

Řešení. Označme X počet šestek při 1000 hodech kostkou. Platí $X \sim \mathrm{Bi}(1000,\ 1/6)$.
Pro začátek poznamenejme, že zadání není formulováno úplně přesně. Najít x, pro které by přesně platilo $P(X < x) = 0{,}9$, se u diskrétní náhodné veličiny nemusí podařit. Ve skutečnosti se budeme snažit najít nejmenší celé číslo x, pro které bude platit $P(X < x) \geq 0{,}9$. (Tj. s hodnotou x, pro kterou by platilo $P(X < x) = 0{,}85$, bychom nebyli spokojeni, zatímco s x, pro které je $P(X < x) = 0{,}95$, už ano).
Kdybychom měli hraniční hodnotu hledat pomocí binomického rozdělení, bylo by to značně pracné. Hledali bychom první celé číslo x, pro které by platilo

$$P(X < x) = \sum_{k=0}^{x-1} \binom{1000}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{1000-k} \geq 0{,}9.$$

Využijeme proto aproximaci binomického rozdělení normálním:

$$\mathrm{E}X = \mu = 1000 \cdot \frac{1}{6} \doteq 166{,}67, \quad \mathrm{D}X = \sigma^2 = 1000 \cdot \frac{1}{6} \cdot \frac{5}{6} \doteq 138{,}89.$$

Přibližně $X \sim \mathrm{N}(166{,}67;\ 138{,}89)$, a tedy $U = \frac{X - 166{,}67}{\sqrt{138{,}89}}$. Hledáme x, pro které platí $P(X < x) = 0{,}9$. Budeme postupovat obdobně jako v příkladu 8.4. Nejprve najdeme hraniční hodnotu pro náhodnou veličinu U:

$$P(U < u) = 0{,}9 \quad \Rightarrow \quad \Phi(u) = 0{,}9 \quad \Rightarrow \quad \doteq 1{,}29.$$

Odtud vypočteme x:

$$x = 1{,}29 \cdot \sqrt{138{,}89} + 166{,}67 \doteq 182.$$

Protože x má být hranice pro počet šestek, hodnotu jsme zaokrouhlili na celé číslo. Platí tedy, že s pravděpodobností zhruba 90% bude při 1000 hodech kostkou počet šestek menší než 182.
Protože jsme si výpočet zjednodušili použitím normálního rozdělení, není zaručeno, že $P(X < 182)$ už opravdu přesáhne 0,9. Při přesném výpočtu pomocí binomického rozdělení bychom zjistili, že $P(X < 182)$ je ve skutečnosti jen 0,895 a že hodnota 0,9 je dosažena až pro 183. S takovouto chybou se však dokážeme smířit, uvážíme-li, jak pracný by byl přesný výpočet.

8.4 Příklady pro samostatnou práci

Příklad 8.22. Stokrát hodíme kostkou. Jaká je pravděpodobnost, že šestka padne alespoň 20-krát? Vypočtěte
a) pomocí normálního rozdělení,
b) pomocí normálního rozdělení s korekcí,
c) přesně (použijte počítač).

Výsledek: a) 0,187 ; b) 0,224 ; c) 0,220

Příklad 8.23. 250 soustruhů pracuje nezávisle na sobě. Každý z nich je v provozu 80 % z celkové pracovní doby. Jaká je pravděpodobnost, že v náhodně vybraném okamžiku je v provozu
a) 190 až 220 soustruhů,
b) 74 až 86 procent soustruhů?
Vypočtěte pomocí normálního rozdělení.

Výsledek: a) 0,942 (s korekcí 0,951); b) 0,982 (s korekcí 0,986)

Příklad 8.24. Pravděpodobnost, že se zasazený strom ujme, je 0,85. Jaká je pravděpodobnost, že z 500 zasazených stromů se jich ujme
a) aspoň 420,
b) nanejvýš 440?
Vypočtěte pomocí normálního rozdělení.

Výsledek: a) 0,736 (s korekcí 0,755) ; b) 0,970 (s korekcí 0,974)

Příklad 8.25. * Pravděpodobnost, že se zasazený strom ujme, je 0,85. Kolik stromů musíme vysadit, aby jich s pravděpodobností 90 % přežilo aspoň 100?
Vypočtěte pomocí normálního rozdělení.

Výsledek: 124

Příklad 8.26. 12 000-krát hodíme mincí. Jaká je pravděpodobnost, že relativní četnost líců bude 0,5015?
Vypočtěte pomocí normálního rozdělení.

Výsledek: 0,008 – počítáno s korekcí. Bez korekce vyjde 0

Příklad 8.27. Podle údajů ze sčítání lidu v roce 1970 bylo zhruba 75 % domácností vybaveno televizorem[1]. Náhodně bylo vybráno 400 domácností.
a) Jaká je pravděpodobnost, že z vybraných 400 domácností má televizor 290–305?
b) Určete, v jakých mezích (souměrných kolem střední hodnoty) bude počet domácností s televizorem (ze 400 vybraných) s pravděpodobností 95 %.

Výsledek: a) 0,594 (s korekcí 0,625); b) 283 až 317

Příklad 8.28. Lojza vyplňuje test IQ, který se skládá z 50 otázek. U každé otázky je 5 možností pro odpověď, správná je vždy právě jedna možnost. Lojza vše vyplní náhodně.
a) Jaká je pravděpodobnost, že správně odpoví právě na 10 otázek? (Vypočtěte pomocí binomického rozdělení a pomocí normálního rozdělení s korekcí.)
b) Nad jakou hranici počtu správných odpovědí se Lojza dostane jen s pravděpodobností 0,01?

Výsledek: a) přesně, tj. pomocí Bi: 0,140; pomocí N s korekcí: 0,143 (bez korekce by vyšlo 0); b) 17

[1] Opravdu je to tak, viz www stránky Českého statistického úřadu.

Příklad 8.29. Každý ze 300 přístrojů se skládá ze tří součástek, které se mohou nezávisle na sobě pokazit. Když se některá součástka pokazí, přístroj přestane pracovat. Pravděpodobnost, že se během dne pokazí první součástka, je 0,1. Pro druhou součástku je pravděpodobnost závady 0,05 a pro třetí 0,15.
a) Jaká je pravděpodobnost, že se během dne pokazí méně než 100 přístrojů?
b) Pod jakou hranicí zůstane počet pokažených přístrojů s pravděpodobností 0,9?

Výsledek: Mezivýsledek: pravděpodobnost poruchy je pro každý jednotlivý přístroj 0,27325. a) 0,986 (s korekcí 0,988); b) 92

9 Statistické testy

9.1 Základní pojmy a řešené příklady

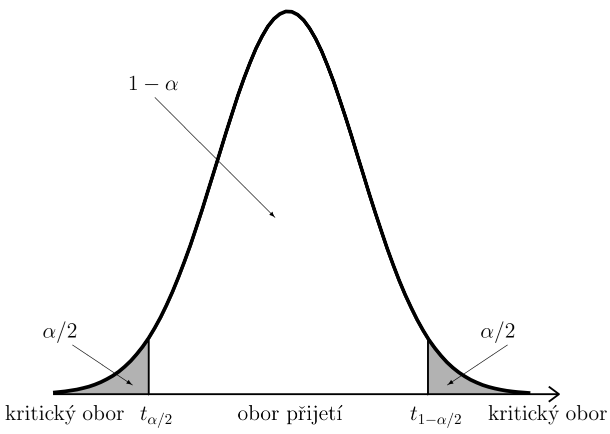
Statistické testy, které probíráme v našem kurzu, slouží k tomu, abychom rozhodli, jestli by určitý parametr mohl být roven určité hodnotě, $\theta = \theta_0$, nebo jestli okolnosti nasvědčují spíše tomu, že tento parametr bude větší nebo menší než hodnota θ_0. Budeme se věnovat dvěma případům: v prvním z nich budeme testovat parametr p binomického rozdělení a ve druhém střední hodnotu μ normálního rozdělení. Existuje ale i řada dalších testů. Obecně se testování provádí v těchto krocích:

1. Vyslovíme **nulovou hypotézu** H_0 a **alternativní hypotézu** H_1.
2. Stanovíme **testové kritérium** – náhodnou veličinu T, podle které chceme o platnosti nulové hypotézy H_0 rozhodnout.
3. Předpokládáme, že platí H_0, a najdeme **kritický obor** W, do kterého testové kritérium T padne jen se zvolenou malou pravděpodobností α. Hodnotu α nazýváme **hladinou významnosti testu**. Kritický obor W je tedy stanoven tak, aby

   $$P(T \in W|\text{platí } H_0) = \alpha,$$

   a jeho hranici (hranice) tvoří odpovídající kvantil (kvantily) náhodné veličiny T. Připomeňme, že α-kvantil náhodné veličiny T je to číslo t_α, pro které je $P(T \leq t_\alpha) = F(t_\alpha) = \alpha$.
4. Zjistíme hodnotu testového kritéria (zpracujeme výsledek konkrétního pokusu či měření).
5. Jestliže empirická (tj. pokusem získaná) hodnota kritéria leží v kritickém oboru, zamítáme hypotézu H_0 ve prospěch alternativní hypotézy H_1 – hypotéza H_1 byla prokázána. Pokud naměřená hodnota v kritickém oboru neleží, hypotézu H_0 nezamítáme a hypotéza H_1 se neprokázala.

Testy rozlišujeme jednostranné a oboustranné, podle toho, co se snažíme prokázat.



Obr. 9.1: Oboustranný test

Jestliže testujeme nulovou hypotézu $H_0\colon \theta = \theta_0$ proti alternativní hypotéze

$$H_1\colon \theta \neq \theta_0,$$

provádíme **oboustranný test**. Kritický obor je pak tvaru (viz obrázek 9.1)

$$W = \left(t_{\min}, t_{\alpha/2}\right) \cup \left(t_{1-\alpha/2}, t_{\max}\right).$$

Jestliže je alternativní hypotéza tvaru

$$H_1\colon \theta > \theta_0,$$

provádíme **jednostranný, a to pravostranný test**. Kritický obor je pak tvaru (viz obrázek 9.2)

$$W = (t_{1-\alpha}, t_{\max}).$$

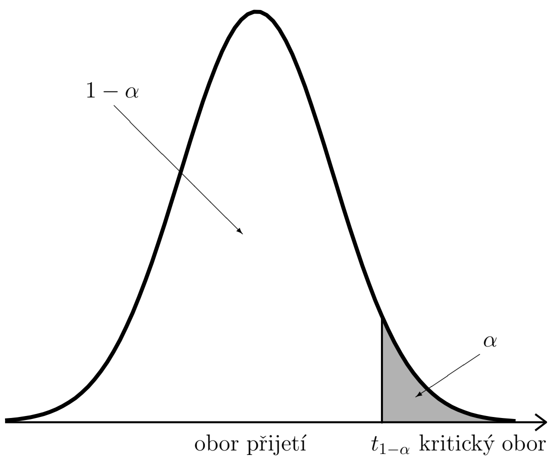
Jestliže je alternativní hypotéza tvaru

$$H_1\colon \theta < \theta_0,$$

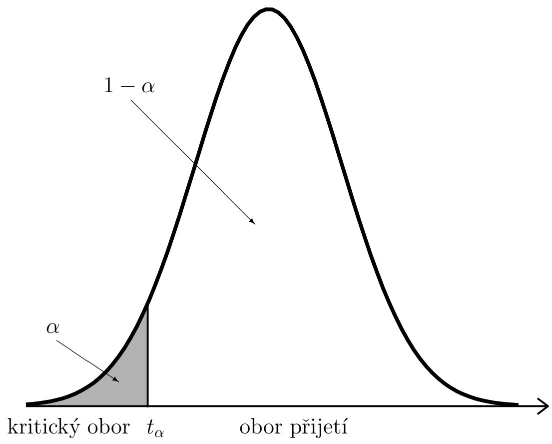
provádíme **jednostranný, a to levostranný test**. Kritický obor je pak tvaru (viz obrázek 9.3)

$$W = (t_{\min}, t_{\alpha}).$$

Výsledek testování může být správný nebo se můžeme dopustit špatného závěru.



Obr. 9.2: Pravostranný test



Obr. 9.3: Levostranný test

Chyba 1. druhu nastane, jestliže nulová hypotéza H_0 platí, ale my ji zamítneme. Pravděpodobnost chyby 1. druhu je rovna hladině významnosti testu α,

$$P(H_0 \text{ zamítneme}|H_0 \text{ platí}) = \alpha.$$

Chyba 2. druhu nastane, jestliže nulová hypotéza H_0 neplatí (čili platí H_1), a přitom H_0 není zamítnuta. Pravděpodobnost chyby 2. druhu označíme β. S chybou 2. druhu souvisí tzv. **síla testu**. Je to pravděpodobnost, že správně zamítneme H_0, když platí alternativní hypotéza H_1,

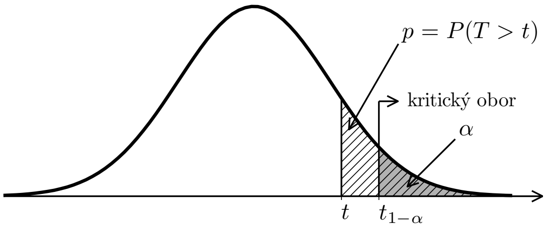
$$\text{síla jednostranného testu} = P(H_0 \text{ zamítneme}|H_0 \text{ neplatí}) = 1 - \beta.$$

Místo hledání hranic kritického oboru můžeme pro rozhodnutí, jestli výsledek pokusu v kritickém oboru leží nebo neleží, použít tzv. p-hodnotu.

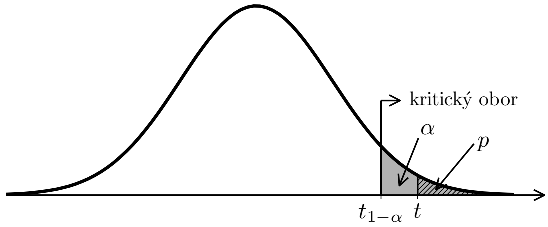
p-hodnota (nezaměňovat s parametrem p binomického rozdělení!) je přesná nebo též pozorovaná hladina významnosti. Udává nejmenší hodnotu hladiny významnosti, na které bychom nulovou hypotézu H_0 zamítli.
Vypočítáme ji jako pravděpodobnost, že testové kritérium nabude pozorované hodnoty nebo hodnoty extrémnější. Porovnáním s α pak můžeme rozhodnout, jestli pozorovaná hodnota leží nebo neleží v kritickém oboru.

Test hypotézy o parametru p binomického rozdělení

Již víme, že pro náhodnou veličinu $X \sim \mathrm{Bi}(n, p)$ parametr p udává pravděpodobnost úspěchu v jednom pokusu. Pomocí statistického testu se budeme snažit rozhodnout, jestli by tento parametr mohl být roven určitému p_0, nebo zda je prokazatelně vyšší či nižší.



Obr. 9.4: Testování pomocí p-hodnoty, pravostranný test, H_0 nezamítáme



Obr. 9.5: Testování pomocí p-hodnoty, pravostranný test, H_0 zamítáme

Často se tento test používá v případě, kdy se snažíme rozhodnout, jestli $p_0 \cdot 100\,\%$ populace má jistou vlastnost V. Náhodně nezávisle na sobě vybereme n objektů. Na základě toho, kolik z nich má vlastnost V, utvoříme závěr ohledně p.
Jako testové kritérium můžeme použít náhodnou veličinu X udávající, kolik objektů z vybraných n kusů má vlastnost V. Za předpokladu, že $p = p_0$, platí $X \sim \mathrm{Bi}(n, p_0)$. Prakticky se ale spíše využívá výpočet pomocí normované náhodné veličiny U nebo pomocí p-hodnoty.

Nulová hypotéza: $H_0\colon p = p_0$
Testové kritérium je

$$U = \frac{X - np_0}{\sqrt{np_0(1-p_0)}} = \frac{\frac{X}{n} - p_0}{\sqrt{p_0(1-p_0)}} \cdot \sqrt{n} \overset{\textit{přibl.}}{\sim} \mathrm{N}(0,1), \tag{9.1}$$

kde náhodná veličina X udává, kolik objektů z celkového počtu n má vlastnost V. Hodnotu získanou měřením (průzkumem, pokusem) označíme u.

- **Oboustranný test**
  Alternativní hypotéza $H_1\colon p \neq p_0$
  Kritický obor: $W = \big(-\infty, u_{\alpha/2}\big) \cup \big(u_{1-\alpha/2}, \infty\big) = \{u\colon |u| > u_{1-\alpha/2}\}$
  Obor přijetí H_0: $\big\langle u_{\alpha/2}, u_{1-\alpha/2}\big\rangle = \big\langle -u_{1-\alpha/2}, u_{1-\alpha/2}\big\rangle = \{u\colon |u| \leq u_{1-\alpha/2}\}$
  $-u_{1-\alpha/2} \leq u \leq u_{1-\alpha/2} \Rightarrow H_0$ nezamítáme; $|u| > u_{1-\alpha/2} \Rightarrow H_0$ zamítáme.
- **Pravostranný test**
  Alternativní hypotéza $H_1\colon p > p_0$
  Kritický obor: $W = (u_{1-\alpha}, \infty) = \{u\colon u > u_{1-\alpha}\}$
  Obor přijetí H_0: $(-\infty, u_{1-\alpha}\rangle = \{u\colon u \leq u_{1-\alpha}\}$
  $u \leq u_{1-\alpha} \Rightarrow H_0$ nezamítáme; $u > u_{1-\alpha} \Rightarrow H_0$ zamítáme.
- **Levostranný test**
  Alternativní hypotéza $H_1\colon p < p_0$
  Kritický obor: $W = (-\infty, u_\alpha) = \{u\colon u < u_\alpha\}$
  Obor přijetí H_0: $\langle u_\alpha, \infty) = \{u\colon u \geq u_\alpha\}$
  $u \geq u_\alpha \Rightarrow H_0$ nezamítáme; $u < u_\alpha \Rightarrow H_0$ zamítáme.

Příklad 9.1. Podle expertního předpokladu bude mít o nový výrobek zájem 20 % zákazníků. Ze 400 dotázaných zákazníků projevilo zájem 62 zákazníků. Na hladině významnosti $\alpha = 0{,}05$ testujte hypotézu o reálnosti předpokladu, tedy $H_0\colon p = 0{,}2$.

Řešení. Předvedeme řešení třemi metodami. Vyberte si, která vám vyhovuje nejlépe, a tu pak používejte. Bohužel se setkáváme s tím, že studenti z nepochopení počítají u zkoušky všemi třemi postupy (navíc často s rozdílnými výsledky) v domnění, že jde o tři postupné kroky výpočtu. Přitom se jedná o tři alternativy, jak lze výpočet provést, a všechny nás musí dovést ke stejnému závěru.
V každém případě budeme testovat nulovou hypotézu

$$H_0\colon p = 0{,}2$$

proti alternativní hypotéze

$$H_1\colon p \neq 0{,}2,$$

provádíme tedy oboustranný test.
I. metoda: Jako testové kritérium použijeme náhodnou veličinu X, která udává, kolik zákazníků ze 400 dotázaných projevilo zájem o nový výrobek. Při platnosti hypotézy H_0 platí

$$X \sim \mathrm{Bi}(400;\, 0{,}2), \quad \text{tedy přibližně} \quad X \sim \mathrm{N}(80, 64). \tag{9.2}$$

Najdeme hranice x_1 a x_2, mezi nimiž se X nachází s pravděpodobností $1 - \alpha = 0{,}95$:

$$P(x_1 \leq X \leq x_2) = 0{,}95$$

Podobné úkoly už jsme řešili v kapitolách 8.1 a 8.3, a proto výpočet nebudeme podrobně rozepisovat. Jen připomeňme, že při hledání souměrného intervalu musíme najít 0,975-kvantil, a nikoli 0,95-kvantil, viz příklad 8.2, část d), kde se totéž řešilo pro interval obsahující 90 % hodnot.

$$x_1 = 80 - 1{,}96 \cdot \sqrt{64} \doteq 64, \quad x_2 = 80 + 1{,}96 \cdot \sqrt{94} \doteq 96.$$

To znamená, že při provádění průzkumu mezi 400 lidmi by počet zájemců o nový výrobek měl být téměř vždy (s pravděpodobností 0,95) mezi 64 a 96. My jsme ale zjistili, že zájem má pouze 62 zákazníků, a to je podezřele málo. Odborněji formulovaný závěr je tento:
Pokusem získaná hodnota 62 leží v kritickém oboru, protože $62 < 64$. Hypotézu H_0 zamítáme.
Prokázalo se, že zájem zákazníků je menší než 20 %.
Tato metoda asi nejlépe vystihuje podstatu věci – hledání mezí přímo pro počet objektů se zkoumanou vlastností a rozhodnutí, jestli výsledek pokusu v těchto mezích leží. V praxi se ale spíše využívá následující metoda, kde se výsledek pokusu normuje, náhodná veličina X se převede na náhodnou veličinu U se standardizovaným normálním rozdělením.

II. metoda – pomocí normované hodnoty (postup z rámečku): Opět předpokládáme, že $p = 0{,}2$. Pak pro náhodnou veličina X udávající počet zájemců o nový výrobek mezi 400 dotázanými platí (9.2), a tedy

$$U = \frac{X - 80}{\sqrt{64}} = \frac{X - 80}{8} \sim \mathrm{N}(0, 1).$$

Interval, v němž je náhodná veličina U s pravděpodobností $1-\alpha = 0{,}95$, je

$$\langle u_{0,025}; u_{0,975}\rangle \doteq \langle -1{,}96;\ 1{,}96\rangle\,.$$

Tento interval je oborem přijetí hypotézy H_0.
Kritický obor je pak

$$W = (-\infty;\ -1{,}96) \cup (1{,}96; \infty)\,.$$

V průzkumu projevilo zájem $x = 62$ zákazníků. Odpovídající normovaná hodnota je

$$u = \frac{62-80}{8} = -2{,}25.$$

Tato hodnota náleží do kritického oboru, a proto hypotézu H_0 o předpokladu 20%zájmu zamítáme na hladině významnosti $\alpha = 0{,}05$. Skutečný zájem bude pravděpodobně menší.
III. metoda – pomocí p-hodnoty: Meze kritického oboru mají u oboustranného testu tu vlastnost, že

$$P(X < x_1) = \alpha/2, \qquad P(X > x_2) = \alpha/2.$$

Na chvíli zapomeňme, že už jsme x_1, x_2 určili v postupu I. O tom, jestli 62 do těchto mezí patří nebo ne, rozhodneme jiným způsobem, aniž bychom x_1, x_2 znali. Vypočítáme pravděpodobnost, že X nabude hodnoty 62 nebo extrémnější, tzn. v našem případě menší: Protože $62 < \mu = 80$, je 62 buď menší než x_1, nebo je v oboru přijetí. Je vyloučené, že by 62 leželo až za x_2.

$$P(X \leq 62) = P(U < -2{,}25) = \Phi(-2{,}25) = 1 - \Phi(2{,}25) \doteq 1 - 0{,}988 = 0{,}012.$$

Výsledek menší než $\alpha/2 = 0{,}025$ ukazuje, že 62 leží pod spodní hranicí x_1 oboru přijetí nulové hypotézy, tj. v kritickém oboru, a H_0 zamítáme. Z tohoto výsledku též vidíme, že na hladině významnosti $\alpha' = 0{,}01$ bychom H_0 nezamítli.
Pravděpodobnost, kterou jsme právě vypočítali, tj. 0,012, by ale nebyla p-hodnota. V souladu s definicí p-hodnoty coby nejmenší hladiny významnosti, na které bychom H_0 zamítli, musíme za p-hodnotu vzít u oboustranného testu dvojnásobek vypočítané pravděpodobnosti, 0,024 – kdybychom oboustranný test prováděli na této hladině významnosti, 62 by vyšlo přesně na dolní hranici oboru přijetí.
Jestliže se vám to zdá nějaké zamotané, můžete **u oboustranného testu** počítat p-hodnotu algoritimicky takto:

$$p\text{-hodnota} = 2P(U > |u|),\ \text{ kde } u \text{ je normovaná hodnota příslušná pozorované hodnotě.}$$

Je-li p-hodnota menší než hladina významnosti α, výsledek pokusu leží v kritickém oboru a H_0 zamítáme. Jestliže je větší nebo rovna α, je výsledek pokusu v oboru přijetí a H_0 nezamítáme.

Příklad 9.2. Výrobce integrovaných obvodů udává, že množství vadných výrobků nepřesahuje 2 %. Při testování 300 náhodně vybraných integrovaných obvodů bylo nalezeno 8 vadných kusů. Na hladině významnosti $\alpha = 0{,}05$ testujte, zda je tvrzení výrobce pravdivé.

Řešení. Tentokrát provedeme jednostranný test. Testujeme nulovou hypotézu

$$H_0\colon p = 0{,}02$$

proti alternativní hypotéze

$$H_1\colon p > 0{,}02.$$

Řešení tentokrát předvedeme pouze metodou z rámečku (v předchozím příkladu označenou II). Ostatní metody si může čtenář zkusit sám, závěr musí být vždy tentýž.
Předpokládáme, že platí H_0. Pak má náhodná veličina X udávající počet vadných výrobků mezi 300 vybranými binomické rozdělení,

$$X \sim \mathrm{Bi}(300;\ 0{,}02), \quad \mathrm{E}X = 300 \cdot 0{,}02 = 6, \quad \mathrm{D}X = 300 \cdot 0{,}02 \cdot 0{,}98 = 5{,}88.$$

Protože np i $n(1-p)$ jsou větší než 5, můžeme binomické rozdělení aproximovat normálním. Přibližně tedy

$$X \sim \mathrm{N}(6;\ 5{,}88), \quad U = \frac{X-6}{\sqrt{5{,}88}}.$$

Normovaná hodnota příslušná pokusem zjistěné hodnotě $x = 8$ je

$$u = \frac{8-6}{\sqrt{5{,}88}} \doteq 0{,}82.$$

Protože provádíme pravostranný test, kritický obor je tentokrát $(u_{0,95}; \infty)$, viz obrázek 9.2. Z tabulky zjistíme, že $u_{0,95} \doteq 1{,}65$. Tedy naměřená hodnota 0,82 v kritickém oboru neleží,

$$0{,}82 < 1{,}65, \quad 0{,}82 \notin (1{,}65; \infty)$$

a hypotézu H_0 nezamítáme. Na hladině významnosti $\alpha = 0{,}05$ se nepodařilo prokázat, že by vadných výrobků bylo více než 2 %, takže tvrzení výrobce může být pravdivé.

Příklad 9.3. Určete sílu testu z předchozího příkladu, jestliže ve skutečnosti jsou mezi integrovanými obvody 4 % vadných.

Řešení. Síla testu je pravděpodobnost, že zamítnu H_0, tj. v našem případě $p = 0{,}02$, jestliže ve skutečnosti platí H_1, zde $p = 0{,}04$.
Nejprve zapomeňme na H_1 a odpovězme si na otázku: Kdy v testu z příkladu 9.2 zamítnu H_0? Jestliže výsledek pokusu leží v kritickém oboru, který byl pro normovanou náhodnou veličinu U interval $(1{,}65; \infty)$. Původní náhodná veličinu X udávající počet vadných integrovaných obvodů mezi 300 náhodně vybranými má za předpokladu, že $p = 0{,}02$, přibližně $\mathrm{N}(6;\ 5{,}88)$, a kritický obor pro ni je

$$W = (1{,}65\sqrt{5{,}88} + 6;\ \infty) \doteq (10, \infty).$$

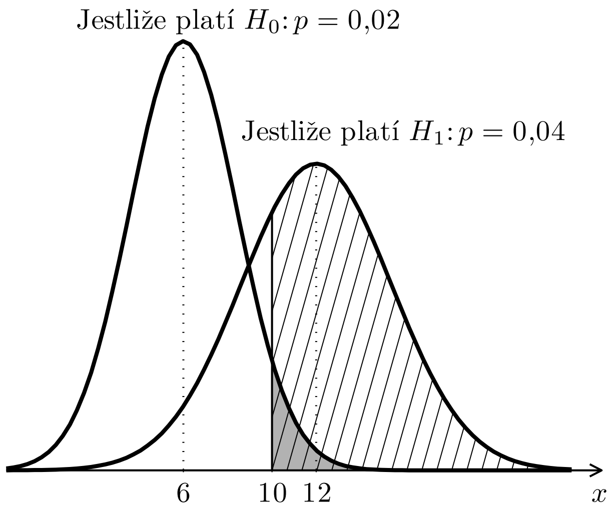
My nyní vypočítáme pravděpodobnost, že jestliže ve skutečnosti jsou vadných výrobků 4 %, pak počet vadných výrobků mezi 300 náhodně vybranými bude ležet v intervalu $(10, \infty)$. V tomto případě je

$$X \sim \mathrm{Bi}(300;\ 0{,}04) \quad \text{neboli přibližně} \quad X \sim \mathrm{N}(12; 11{,}52),$$

a tedy

$$P(X > 10) = P\left(U > \frac{10-12}{\sqrt{11{,}52}}\right) \doteq P(U > -0{,}59) = 1 - \Phi(-0{,}59) = \Phi(0{,}59) \doteq 0{,}722.$$

Síla testu je 0,722. Na obrázku 9.6 je síla testu obsah vyšrafované plochy.



Obr. 9.6: K příkladu 9.3, síla testu

9.2 Příklady pro samostatnou práci

Příklad 9.4. Stokrát jsme hodili kostkou, přitom 25-krát padla šestka. Testujte hypotézu H_0, že kostka je poctivá, tj. že pravděpodobnost, že na ní padne šestka, je $p = 1/6$, proti alternativní hypotéze, že $p \neq 1/6$ (oboustranný test), a to na hladině významnosti
a) $\alpha = 0{,}05$
b) $\alpha = 0{,}01$.
c) Určete nejmenší hodnotu významnosti, na které bychom H_0 zamítli (tj. p-hodnotu).

Výsledek: a) H_0 zamítáme, šestka padá s vyšší pravděpodobností; b) H_0 nezamítáme, neprokázalo se, že by šestka padala s jinou pravděpodobností než $1/6$; c) 0,026
Lepší, než počítat jen šestky, by bylo zkoumat, jestli všechna čísla na kostce padají se stejnou pravděpodobností. Takový test skutečně lze provést, bohužel však přesahuje rámec našeho kurzu.

Příklad 9.5. Ze 200 výrobků vyrobených novou technologií bylo k zmetků. Testujte na hladině významnosti $\alpha = 0{,}01$, zda nová technologie změnila zmetkovitost oproti dřívějším dlouhodobě zjištěným 10 % (proveďte oboustranný test), jestliže
a) $k = 6$, b) $k = 13$, c) $k = 25$, d) $k = 32$.

Výsledek: Nulovou hypotézu H_0, že $p = 0{,}1$, a) zamítáme (zmetkovitost se snížila), b), c) nezamítáme (neprokázalo se, že by zmetkovitost byla jiná než 10 %), d) zamítáme (zmetkovitost se zvýšila).

Příklad 9.6. Při průzkumu veřejného mínění v souboru 500 dotazovaných respondentů jich 180 vyjádřilo nespokojenost se svou poslední volbou v parlamentních volbách (tj. dnes by volili jinou stranu). Lze na základě tohoto průzkumu tvrdit, že 40 % voličů je nespokojených se svou poslední volbou? Testujte na hladině významnosti $\alpha = 0{,}05$.

Výsledek: Nulovou hypotézu, že nespokojených je 40 % voličů, nezamítáme.

Příklad 9.7. (Souvisí s příkladem 9.6.) Kolik lidí z 500 dotazovaných respondentů by muselo vyjádřit nespokojenost se svou poslední volbou, abychom na hladině významnosti $\alpha = 0{,}05$ byli oprávněni tvrdit, že se počet voličů nespokojených se svou poslední volbou významně liší od 40 %?

Výsledek: Méně než 179 nebo více než 221.

Příklad 9.8. Testovali jsme hypotézu $H_0\colon p = p_0$ proti $H_1\colon p \neq p_0$ na hladině významnosti $\alpha = 0{,}01$. Nulovou hypotézu jsme na této hladině významnosti
a) zamítli; b) nezamítli.
Jak by dopadl tentýž test na hladině významnosti $\alpha = 0{,}05$?

Výsledek: a) H_0 bychom zamítli; b) bez znalosti konkrétních hodnot nelze rozhodnout.

Příklad 9.9. Testovali jsme hypotézu $H_0\colon p = p_0$ proti $H_1\colon p \neq p_0$ na hladině významnosti $\alpha = 0{,}05$. Nulovou hypotézu jsme na této hladině významnosti
a) zamítli; b) nezamítli.
Jak by dopadl tentýž test na hladině významnosti $\alpha = 0{,}01$?

Výsledek: a) Bez znalosti konkrétních hodnot nelze rozhodnout; b) H_0 bychom nezamítli.

Příklad 9.10. $*$ Vraťme se k testu z příkladů 9.1 a 9.2. Předpokládejme, že čtyřprocentní chybovost obvodů je pro nás nepřijatelná. Kolik obvodů musíme otestovat, aby síla testu z příkladu 9.2 (neboli pravděpodobnost, že na hladině významnosti $\alpha = 0{,}05$ správně zamítneme hypotézu o 2 % vadných obvodů, jestliže ve skutečnosti jsou vadná 4 %) byla 90 %?

Výsledek: $n = 586$

9.3 Test hypotézy o střední hodnotě normálního rozdělení

Máme k dispozici n empiricky zjištěných hodnot $x_1, x_2, \ldots, x_n$ získaných nezávisle na sobě. Na základě těchto n měření se snažíme rozhodnout, jestli střední hodnota zkoumané náhodné veličiny může být rovna určitému μ_0, nebo jestli dosahuje prokazatelně vyšší či nižší hodnoty.

Předpokládejme, že zkoumaná náhodná veličina X má normální rozdělení, $X \sim \mathrm{N}(\mu, \sigma^2)$, nebo si nejsme jisti, jestli je má, ale máme k dispozici velký počet měření. Pak můžeme počítat s tím, že výběrový průměr $\overline{X}$ má opět normální rozdělení,

$$\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right), \text{ tzn. } U = \frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\overline{X} - \mu}{\sigma}\sqrt{n} \sim N(0, 1).$$

Jestliže rozptyl σ^2 neznáme, což je v praxi obvyklé, pak použijeme jeho odhad s^2. Náhodná veličina

$$T = \frac{\overline{X} - \mu}{\sqrt{S^2}}\sqrt{n}, \quad \text{kde} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X - \overline{X})^2,$$

už ale nemá normální rozdělení, nýbrž tzv. Studentovo t-rozdělení s $(n-1)$ stupni volnosti, $T \sim \mathrm{t}(n-1)$. Pro velké n se Studentovo rozdělení přibližuje ke standardizovanému normálnímu rozdělení.

Test hypotézy o střední hodnotě μ, jestliže rozptyl známe nebo je rozsah výběru velký

Nulová hypotéza: $H_0\colon \mu = \mu_0$

Testové kritérium je při známém rozptylu σ^2

$$U = \frac{\overline{X} - \mu_0}{\sqrt{\sigma^2}} \cdot \sqrt{n} \sim \mathrm{N}(0,1), \tag{9.3}$$

nebo při neznámém rozptylu a velkém rozsahu výběru

$$U = \frac{\overline{X} - \mu_0}{\sqrt{S^2}} \cdot \sqrt{n} \overset{\textit{přibl.}}{\sim} \mathrm{N}(0,1). \tag{9.4}$$

Hodnotu získanou pokusem označíme u.

- **Oboustranný test**
  Alternativní hypotéza $H_1\colon \mu \neq \mu_0$
  Kritický obor: $W = \big(-\infty, u_{\alpha/2}\big) \cup \big(u_{1-\alpha/2}, \infty\big) = \{u\colon |u| > u_{1-\alpha/2}\}$
  Obor přijetí H_0: $\big\langle u_{\alpha/2}, u_{1-\alpha/2}\big\rangle = \big\langle -u_{1-\alpha/2}, u_{1-\alpha/2}\big\rangle = \{u\colon |u| \leq u_{1-\alpha/2}\}$
  $-u_{1-\alpha/2} \leq u \leq u_{1-\alpha/2} \Rightarrow H_0$ nezamítáme; $|u| > u_{1-\alpha/2} \Rightarrow H_0$ zamítáme.
- **Pravostranný test**
  Alternativní hypotéza $H_1\colon \mu > \mu_0$
  Kritický obor: $W = (u_{1-\alpha}, \infty) = \{u\colon u > u_{1-\alpha}\}$
  Obor přijetí H_0: $(-\infty, u_{1-\alpha}\rangle = \{u\colon u \leq u_{1-\alpha}\}$
  $u \leq u_{1-\alpha} \Rightarrow H_0$ nezamítáme; $u > u_{1-\alpha} \Rightarrow H_0$ zamítáme.
- **Levostranný test**
  Alternativní hypotéza $H_1\colon \mu < \mu_0$
  Kritický obor: $W = (-\infty, u_\alpha) = \{u\colon u < u_\alpha\}$
  Obor přijetí H_0: $\langle u_\alpha, \infty) = \{u\colon u \geq u_\alpha\}$
  $u \geq u_\alpha \Rightarrow H_0$ nezamítáme; $u < u_\alpha \Rightarrow H_0$ zamítáme.

Test lze provádět taktéž pomocí p-hodnoty.

Příklad 9.11. Výrobce tvrdí, že jím vyrobené žárovky mají životnost v průměru 1 000 hodin. Potřebujeme se ujistit, že tomu tak opravdu je. Otestovali jsme 60 náhodně vybraných žárovek. Jejich průměrná životnost byla $\overline{x} = 998{,}23$ hodin a výběrový rozptyl byl $s^2 = 11\,083{,}45$. Na hladině významnosti $\alpha = 0{,}05$ testujte hypotézu $H_0\colon \mu = 1\,000$, a to proti alternativní hypotéze
a) $H_1\colon \mu \neq 1\,000$ (oboustranný test),
b) $H_1\colon \mu < 1\,000$ (jednostranný test).

Řešení. Určíme normovanou hodnotu odpovídající naměřené hodnotě $\overline{x} = 998{,}23$:

$$u = \frac{998{,}23 - 1\,000}{\sqrt{11\,083{,}45}}\sqrt{60} \doteq -0{,}13$$

a) U oboustranného testu je obor přijetí nulové hypotézy

$$\langle u_{0,025}; u_{0,975}\rangle = \langle -1{,}96;\ 1{,}96\rangle\,,$$

kritický obor je

$$W = (-\infty; -1{,}96) \cup (1{,}96;\ \infty).$$

Platí

$$-0{,}13 \in \langle -1{,}96;\ 1{,}96\rangle\,, \quad -0{,}13 \neq W,$$

a proto hypotézu H_0 nezamítáme na hladině významnosti $\alpha = 0{,}05$. Není tedy důvod tvrdit, že by průměrná životnost žárovek byla významně jiná než 1 000 hodin.

b) U jednostranného testu je obor přijetí

$$\langle u_{0,05}; \infty) = \langle -1{,}65;\ \infty)\,,$$

kritický obor je

$$W = (-\infty;\ -1{,}65).$$

Platí

$$-0{,}13 \in (-1{,}65;\ \infty\rangle\,, \quad -0{,}13 \neq W,$$

a proto hypotézu H_0 nezamítáme na hladině významnosti $\alpha = 0{,}05$. Není tedy důvod tvrdit, že by průměrná životnost žárovek byla významně nižší než 1 000 hodin.

9.4 Příklady pro samostatnou práci

Příklad 9.12. 40 vzorků nerostu bylo testováno na obsah aktivní látky. Průměrná hodnota byla $\overline{x} = 25{,}775$ promile, výběrový rozptyl je $s^2 = 10{,}987$. Lze na základě těchto měření tvrdit, že ve zkoumaném nerostu je 25 promile aktivní látky? Proveďte oboustranný test na hladině významnosti $\alpha = 0{,}05$.

Výsledek: a) Hypotézu $H_0\colon \mu = 25$ nezamítáme.

Příklad 9.13. Ředitel jedné velké kocourkovské továrny láká nové zaměstnance a tvrdí, že průměrný plat v jeho podniku je 1 100 Kk (korun kocourkovských). Náhodně jsme vybrali 50 zaměstnanců a zjistili, že jejich průměrný plat je 1 028,40 Kk a výběrová směrodatná odchylka s je 126,253 Kk. Na hladině významnosti $\alpha = 0{,}01$ testujte hypotézu H_0, že průměrný plat je skutečně 1 100 Kk. Proveďte oboustranný test.

Výsledek: Hypotézu $H_0\colon \mu = 1\,100$ zamítáme, průměrný plat je nižší

Příklad 9.14. Na balíčku kořenicí směsi je uvedeno, že obsah soli je průměrně 40 procent. Potravinářská inspekce náhodně vybrala 45 balíčků a otestovala na množství soli. Průměrné množství bylo $\overline{x} = 41{,}67$ procent a výběrová směrodatná odchylka byla $s = 3{,}342$. Na hladině významnosti $\alpha = 0{,}05$ testujte, zda průměrný obsah soli v kořenicí směsi nepřesahuje 40 procent, proveďte jednostranný test, tj. $H_1\colon \mu > 40$.

Výsledek: Hypotézu $H_0\colon \mu = 40$ zamítáme, obsah soli je vyšší než 40 %.

Příklad 9.15. Střední doba bezporuchové činnosti určitého typu přístrojů by měla být alespoň 1 000 hodin. Přitom z dřívějších měření víme, že směrodatná odchylka doby bezporuchové činnosti je $\sigma = 100$ hodin. Z velké skupiny těchto přístroju jsme náhodně vybrali 25 kusů. Průměrná doba jejich bezporuchové činnosti byla 970 hodin. Je tím prokázáno, že celá skupina nevyhovuje požadavku, aby přístroje pracovaly alespoň 1 000 hodin bez poruchy? Proveďte jednostranný test, a to na hladině významnosti
a) $\alpha = 0{,}1$,
b) $\alpha = 0{,}01$.

Výsledek: a) Hypotézu $H_0\colon \mu = 1\,000$ zamítáme, prokázala se alternativní hypotéza $H_1\colon \mu < 1\,000$.
b) Na hladině významnosti 0,01 H_0 nezamítáme; na této hladině významnosti se neprokázalo, že by stroje fungovaly významně kratší dobu.

Příklad 9.16. Určete nejmenší hladinu významnosti, pro kterou bychom v příkladu 9.15 hypotézu $H_0\colon \mu = 1\,000$ zamítli.

Výsledek: 0,067 (jedná se vlastně o p-hodnotu)

Příklad 9.17. Jaká by musela být průměrná doba bezporuchové činnosti 25 přístrojů z příkladu 9.15, abychom na hladině významnosti $\alpha = 0{,}01$ mohli tvrdit, že $\mu < 1\,000$ hodin?

Výsledek: Menší než 953,4 hodiny.

Příklad 9.18. Testovali jsme hypotézu $H_0\colon \mu = \mu_0$ proti $H_1\colon \mu \neq \mu_0$ na hladině významnosti $\alpha = 0{,}05$. Na základě $\overline{x}$ a s^2 vypočítaných z $n = 50$ naměřených hodnot jsme hypotézu H_0
a) zamítli; b) nezamítli.
Jak by test dopadl, kdyby tytéž hodnoty $\overline{x}$ a s^2 byly získány z většího výběru, např. $n = 100$?

Výsledek: a) H_0 bychom zamítli; b) bez znalosti konkrétních hodnot nelze rozhodnout.

Příklad 9.19. Testovali jsme hypotézu $H_0\colon \mu = \mu_0$ proti $H_1\colon \mu \neq \mu_0$ na hladině významnosti $\alpha = 0{,}05$. Na základě $\overline{x}$ a s^2 vypočítaných z $n = 100$ naměřených hodnot jsme hypotézu H_0
a) zamítli; b) nezamítli.
Jak by test dopadl, kdyby tytéž hodnoty $\overline{x}$ a s^2 byly získány z menšího výběru, např. $n = 50$?

Výsledek: a) Bez znalosti konkrétních hodnot nelze rozhodnout; b) H_0 bychom nezamítli.

Tab. 9.1: **Hodnoty distribuční funkce $\Phi(u)$ - 1.část.**

u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$
0,00	0,5000000	0,30	0,6179114	0,60	0,7257469	0,90	0,8159399	1,20	0,8849303
0,01	0,5039894	0,31	0,6217195	0,61	0,7290691	0,91	0,8185887	1,21	0,8868606
0,02	0,5079783	0,32	0,6255158	0,62	0,7323711	0,92	0,8212136	1,22	0,8887676
0,03	0,5119665	0,33	0,6293000	0,63	0,7356527	0,93	0,8238145	1,23	0,8906514
0,04	0,5159534	0,34	0,6330717	0,64	0,7389137	0,94	0,8263912	1,24	0,8925123
0,05	0,5199388	0,35	0,6368307	0,65	0,7421539	0,95	0,8289439	1,25	0,8943502
0,06	0,5239222	0,36	0,6405764	0,66	0,7453731	0,96	0,8314724	1,26	0,8961653
0,07	0,5279032	0,37	0,6443088	0,67	0,7485711	0,97	0,8339768	1,27	0,8979577
0,08	0,5318814	0,38	0,6480273	0,68	0,7517478	0,98	0,8364569	1,28	0,8997274
0,09	0,5358564	0,39	0,6517317	0,69	0,7549029	0,99	0,8389129	1,29	0,9014747
0,10	0,5398278	0,40	0,6554217	0,70	0,7580363	1,00	0,8413447	1,30	0,9031995
0,11	0,5437953	0,41	0,6590970	0,71	0,7611479	1,01	0,8437524	1,31	0,9049021
0,12	0,5477584	0,42	0,6627573	0,72	0,7642375	1,02	0,8461358	1,32	0,9065825
0,13	0,5517168	0,43	0,6664022	0,73	0,7673049	1,03	0,8484950	1,33	0,9082409
0,14	0,5556700	0,44	0,6700314	0,74	0,7703500	1,04	0,8508300	1,34	0,9098773
0,15	0,5596177	0,45	0,6736448	0,75	0,7733726	1,05	0,8531409	1,35	0,9114920
0,16	0,5635595	0,46	0,6772419	0,76	0,7763727	1,06	0,8554277	1,36	0,9130850
0,17	0,5674949	0,47	0,6808225	0,77	0,7793501	1,07	0,8576903	1,37	0,9146565
0,18	0,5714237	0,48	0,6843863	0,78	0,7823046	1,08	0,8599289	1,38	0,9162067
0,19	0,5753454	0,49	0,6879331	0,79	0,7852361	1,09	0,8621434	1,39	0,9177356
0,20	0,5792597	0,50	0,6914625	0,80	0,7881446	1,10	0,8643339	1,40	0,9192433
0,21	0,5831662	0,51	0,6949743	0,81	0,7910299	1,11	0,8665005	1,41	0,9207302
0,22	0,5870604	0,52	0,6984682	0,82	0,7938919	1,12	0,8686431	1,42	0,9221962
0,23	0,5909541	0,53	0,7019440	0,83	0,7967306	1,13	0,8707619	1,43	0,9236415
0,24	0,5948349	0,54	0,7054015	0,84	0,7995458	1,14	0,8728568	1,44	0,9250663
0,25	0,5987063	0,55	0,7088403	0,85	0,8023375	1,15	0,8749281	1,45	0,9264707
0,26	0,6025681	0,56	0,7122603	0,86	0,8051055	1,16	0,8769756	1,46	0,9278550
0,27	0,6064199	0,57	0,7156612	0,87	0,8078498	1,17	0,8789995	1,47	0,9292191
0,28	0,6102612	0,58	0,7190427	0,88	0,8105703	1,18	0,8809999	1,48	0,9305634
0,29	0,6140919	0,59	0,7224047	0,89	0,8132671	1,19	0,8829768	1,49	0,9318879

Tab. 9.2: **Hodnoty distribuční funkce $\Phi(u)$ - 2.část.**

u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$
1,50	0,9331928	1,80	0,9640697	2,10	0,9821356	2,40	0,9918025	4,50	0,9999966
1,51	0,9344783	1,81	0,9648521	2,11	0,9825708	2,41	0,9920237	5,00	0,9999997
1,52	0,9357445	1,82	0,9656205	2,12	0,9829970	2,42	0,9922397	5,50	0,9999999
1,53	0,9369916	1,83	0,9663750	2,13	0,9834142	2,43	0,9924506		
1,54	0,9382198	1,84	0,9671159	2,14	0,9838226	2,44	0,9926564		
1,55	0,9394392	1,85	0,9678432	2,15	0,9842224	2,45	0,9928572		
1,56	0,9406201	1,86	0,9685572	2,16	0,9846137	2,46	0,9930531		
1,57	0,9417924	1,87	0,9692581	2,17	0,9849966	2,47	0,9932443		
1,58	0,9429466	1,88	0,9699460	2,18	0,9853713	2,48	0,9934309		
1,59	0,9440826	1,89	0,9706210	2,19	0,9857379	2,49	0,9936128		
1,60	0,9452007	1,90	0,9712834	2,20	0,9860966	2,50	0,9937903		
1,61	0,9463011	1,91	0,9719334	2,21	0,9864474	2,51	0,9939634		
1,62	0,9473839	1,92	0,9725711	2,22	0,9867906	2,52	0,9941323		
1,63	0,9484493	1,93	0,9731966	2,23	0,9871263	2,53	0,9942969		
1,64	0,9494974	1,94	0,9738102	2,24	0,9874545	2,54	0,9944574		
1,65	0,9505285	1,95	0,9744119	2,25	0,9877755	2,55	0,9946139		
1,66	0,9515428	1,96	0,9750021	2,26	0,9880894	2,56	0,9947664		
1,67	0,9525403	1,97	0,9755808	2,27	0,9883962	2,57	0,9949151		
1,68	0,9535213	1,98	0,9761482	2,28	0,9886962	2,58	0,9950600		
1,69	0,9544860	1,99	0,9767045	2,29	0,9889893	2,59	0,9952012		
1,70	0,9554345	2,00	0,9772499	2,30	0,9892759	2,60	0,9953388		
1,71	0,9563671	2,01	0,9777844	2,31	0,9895559	2,70	0,9965330		
1,72	0,9572838	2,02	0,9783083	2,32	0,9898296	2,80	0,9974449		
1,73	0,9581849	2,03	0,9788217	2,33	0,9900969	2,90	0,9981342		
1,74	0,9590705	2,04	0,9793248	2,34	0,9903581	3,00	0,9986501		
1,75	0,9599408	2,05	0,9798178	2,35	0,9906133	3,20	0,9993129		
1,76	0,9607961	2,06	0,9803007	2,36	0,9908625	3,40	0,9996631		
1,77	0,9616364	2,07	0,9807738	2,37	0,9911060	3,60	0,9998409		
1,78	0,9624620	2,08	0,9812372	2,38	0,9913437	3,80	0,9999277		
1,79	0,9632730	2,09	0,9816911	2,39	0,9915758	4,00	0,9999683		