UNIVERZITA J. E. PURKYNĚ V ÚSTÍ NAD LABEM

Fakulta sociálně ekonomická

Pravděpodobnost

Náhodné jevy, náhodná veličina

Studijní opora pro FSE UJEP

Kapitola 1

Náhodné jevy a jejich pravděpodobnost

Po prostudování této kapitoly byste měli být schopni:

- rozumět pojmu *náhodný jev* a umět jej vysvětlit na konkétních příkladech,
- rozumět vztahům mezi náhodnými jevy ($A = B$, $A \subset B$, neslučitelné jevy) a vysvětlit je na konkrétních příkladech,
- rozumět operacím s náhodnými jevy ($A \cup B$, $A \cap B$, $A \backslash B$ a použít je v konkrétních příkladech,
- rozumět pojmům *jistý jev*, *nemožný jev*, *opačný jev* a umět je použít v konkrétních příkladech,
- rozumět definici pojmu *pravděpodobnost náhodného jevu* a umět ji vysvětlit na konkétních příkladech,
- rozumět pojmu *nezávislé náhodné jevy* a umět jej vysvětlit na vhodných příkladech,
- rozumět vzorcům pro *pravděpodobnost sjednocení náhodných jevů* a umět je použít v odpovídajících situacích,
- rozumět vzorci pro *pravděpodobnost průniku dvou nezávislých náhodných jevů* a umět jej použít v odpovídajících situacích.

1.1 Náhodné jevy

Náhodný pokus

Náhodným pokusem rozumíme každé pozorování nějakého náhodného děje, každé zaznamenání, změření či konstatování jeho výsledku. Stručně - před proběhnutím náhodného děje nevíme, jaký bude výsledek tohoto pokusu, ale po ukončení tohoto náhodného děje jsme schopni výsledek zaznamenat, změřit či konstatovat.

Náhodným pokusem je proto jeden hod mincí, výběr a spočtení zmetků v krabici výrobků, měření teploty vzduchu v danou hodinu atd.

Náhodný jev

Náhodný jev je každý možný výsledek náhodného pokusu.

Těchto výsledků může být větší množství. K jejich odlišení je často nazýváme jmény ve formě písmen abecedy - řekneme, že nastal jev A, jev B atd.

1.1. Uvažujme náhodný pokus *hod kostkou* a sledujme, které číslo padne. Náhodné jevy mohou být například tyto výsledky hodu:

- Jev A ... nastane, když padne číslo 3 ———— $A = \{3\}$
- Jev B ... nastane, když padne sudé číslo ———— $B = \{2, 4, 6\}$
- Jev C ... nastane, když padne číslo dělitelné dvěma nebo třemi ———— $C = \{2, 3, 4, 6\}$
- Jev D ... nastane, když padnou čísla $2, 3, 4$ ———— $D = \{2, 3, 4\}$

Elementární jev

Elementární jevy jsou nejjednodušší výsledky náhodného pokusu, které se nedají složit z žádných jiných jevů. Množinu všech elementárních jevů budeme nazývat *prostorem elementárních jevů* a značit jej symbolem Ω.

V případě, kdy je náhodným pokusem *hod kostkou*, jsou elementární tyto jevy: $A = \{1\}$, $B = \{2\}$, $C = \{3\}$, $D = \{4\}$, $E = \{5\}$, $F = \{6\}$ a prostorem všech elementárních jevů je $\Omega = \{1, 2, 3, 4, 5, 6\}$.

1.1.1 Úlohy k řešení

1.2. Uveďte alespoň tři různé příklady náhodných pokusů.

1.3. Je dán náhodný pokus spočívající v náhodném výběru jedné osoby z množiny všech studentů FSE UJEP. Uveďte alespoň tři různé náhodné jevy k tomuto náhodnému pokusu.

1.4. Uvažujme náhodný pokus spočívající v náhodném výběru jedné karty z balíčku 32 karet. Uveďte alespoň tři různé náhodné jevy k tomuto náhodnému pokusu.

1.5. Uvažujme náhodný pokus v hodu 2 mincemi. Popište prostor elementárních jevů.

1.2 Vlastnosti náhodných jevů a operace s náhodnými jevy

- Symbolem $A \subset B$ značíme, že nastane-li jev A, nastane nutně i jev B. Říkáme, že jev A má za následek jev B.
  - V Příkladu 1.1 platí $B \subset C$. Jev B: *Padlo sudé číslo* má za následek jev C: *Padlo číslo dělitelné dvěma nebo třemi.* Množinově $\{2, 4, 6\} \subset \{2, 3, 4, 6\}$.
- Symbolem $A = B$ značíme, že **jevy A a B jsou si rovny**, tj. jev A nastane vždy, když nastane jev B a nikdy jindy.

- Řekneme, že nastalo **sjednocení jevů** A a B (značíme $A \cup B$) právě tehdy, nastal-li jev A nebo jev B, nebo oba současně.
- Řekneme, že nastal **průnik jevů** A a B (značíme $A \cap B$) právě tehdy, nastal-li jev A a současně nastal jev B.
- Jev **rozdíl jev**ů A a B (značíme $A \backslash B$) nastane právě tehdy, nastane-li jev A a nenastane jev B.
- Jev $\overline{A}$, tzv. **opačný** (doplňkový, komplementární) **jev** k jevu A, nastane právě tehdy, nenastane-li jev A.
- Náhodný jev, který nastane vždy, se nazývá **jev jistý**. Označuje se také Ω a je vlastně sjednocením všech jevů, které mohou nastat.
- Jev, který nenastane nikdy, nazýváme **jev nemožný** a značíme ho $\emptyset$.
- Jsou-li A a B takové jevy, že $A \cap B = \emptyset$, nazýváme je **neslučitelné** (disjunktní) **jevy**. Neslučitelné jevy tedy nemohou nastat současně.

1.6. Nově vyvinutý výrobek je podroben třem různým zkouškám. Jev A nastane tehdy, když náhodně vybraný výrobek ze zkušební série obstojí v první zkoušce, jev B nastane tehdy, když náhodně vybraný výrobek obstojí ve druhé zkoušce a jev C nastane, pokud náhodně vybraný výrobek vyhoví ve třetí zkoušce. Jak v množinové symbolice vyjádříme to, že náhodně vybraný výrobek obstojí:

a) jen v první zkoušce,

b) v první a ve druhé zkoušce, ale neobstojí ve třetí zkoušce,

c) ve všech třech zkouškách,

d) alespoň v jedné zkoušce,

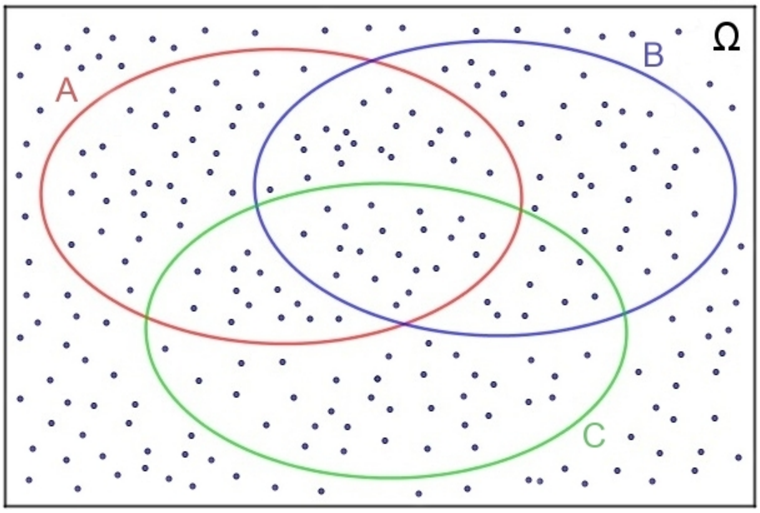
e) alespoň ve dvou zkouškách,

f) právě v jedné zkoušce,

g) právě ve dvou zkouškách,

h) maximálně dvakrát?

Řešení: Uvedenou situaci lze graficky znázornit na Obrázku 1.1. Uvnitř černého obdélníku (označení Ω) jsou černé tečky, které symbolicky představují testované výrobky. Každý z nich má svou kvalitu, podle které vyhoví, resp. nevyhoví jednotlivým zkouškám. Náhodným pokusem je zde výběr výrobku. Podle toho, o jak kvalitní výrobek se jedná, budou tyto výrobky vyhovovat jednotlivým testům a podle toho budou nastávat jednotlivé jevy A, B a C. Výrobky, které vyhoví v první zkoušce, jsou na obrázku zahrnuty v červeném oválu (množina A - prvky, při jejichž výběru nastane jev A). Analogicky - výrobky, které vyhoví ve druhé zkoušce, představují tečky, které se nacházejí v modrém oválu (množina B - prvky, při jejichž výběru nastane jev B)



Obrázek 1.1: Testování výrobků

a výrobky, které vyhoví ve třetí zkoušce, představují tečky, které se nacházejí v zeleném oválu (množina C - prvky, při jejichž výběru nastane jev C).

Z obrázku lze vyčíst, které výrobky například vyhoví pouze v první zkoušce, které vyhoví ve všech třech zkouškách, které vyhoví pouze první a druhé zkoušce atd. Pokuste se na obrázku určit, které body vyhovují uvedeným případům a) až h).

a) Jestliže výrobek obstojí pouze v první zkoušce, znamená to: obstojí v první zkoušce, tj. nastane jev A. Zároveň neobstojí ve druhé zkoušce, tj. nenastane jev B, tj. nastane opačný jev k jevu B, tj. nastane jev $\overline{B}$. Zároveň výrobek nevyhoví ve třetí zkoušce, tedy nenastane jev C, tj. nastane jev $\overline{C}$. Jevy A, $\overline{B}$ a $\overline{C}$ nastávají současně, proto jev, kdy výrobek obstojí pouze v první zkoušce množinově vyjádříme zápisem $A \cap \overline{B} \cap \overline{C}$.

b) Víme, že vybraný výrobek obstojí v první a ve druhé zkoušce, tj. nastane jev A a B, ale neobstojí ve třetí zkoušce, tj. nastane jev $\overline{C}$. Všechny tyto jevy přitom nastanou současně, jedná se tedy o průnik těchto jevů $A \cap B \cap \overline{C}$.

c) Pokud vybraný výrobek vyhoví ve všech třech zkouškách, znamená to, že vyhověl v první zkoušce - nastal je A, zároveň vyhoví ve druhé zkoušce - nastal jev B a zároveň vyhověl ve třetí zkoušce - nastal jev C. Protože jevy nastanou současně, jedná se o jejich průnik $A \cap B \cap C$.

d) Výrok *výrobek obstojí alespoň v jedné zkoušce* znamená totéž, jako že výrobek obstojí v první zkoušce (bez ohledu na to jak (ne)dopadly ostatní zkoušky) - nastane jev A, nebo ve druhé zkoušce (bez ohledu na to jak (ne)dopadly ostatní zkoušky) - nastane jev B, nebo ve třetí zkoušce (bez ohledu na to jak (ne)dopadly ostatní zkoušky) - nastane jev C. Stačí, aby nastala alespoň jedna z uvedených situací a bude pravda, že výrobek vyhověl v alespoň jedné zkoušce. Tedy výrobek obstál alespoň v jedné zkoušce, pokud nastane alespoň jeden z jevů A, B, resp. C. Jedná se tedy o sjednocení jevů - množinově $A \cup B \cup C$.

e) Výrok *výrobek obstojí alespoň ve dvou zkouškách* znamená totéž, jako že výrobek obstojí v první a druhé zkoušce (bez ohledu na to jak (ne)dopadla třetí zkouška) - nastane jev

$A \cap B$, nebo v první a třetí zkoušce (bez ohledu na to jak (ne)dopadla druhá zkouška) - nastane jev $A \cap C$, nebo ve druhé a třetí zkoušce (bez ohledu na to jak (ne)dopadla první zkouška) - nastane jev $B \cap C$. Stačí, aby nastala alespoň jedna z uvedených situací a bude pravda, že výrobek vyhověl alespoň ve dvou zkouškách. Tedy výrobek obstál alespoň ve dvou zkouškách, pokud nastane alespoň jeden z jevů $A \cap B$, $A \cap C$, resp. $B \cap C$. Jedná se o sjednocení uvedených jevů - množinově $(A \cap B) \cup (A \cap C) \cup (B \cap C)$.

f) Výrobek obstojí právě v jedné zkoušce, když vyhoví v první zkoušce a zároveň nevyhoví ve druhé a třetí zkoušce - tj. nastane jev $A \cap \overline{B} \cap \overline{C}$, nebo vyhoví ve druhé zkoušce a zároveň nevyhoví v první a třetí zkoušce - tj. nastane jev $\overline{A} \cap B \cap \overline{C}$, nebo vyhoví ve třetí zkoušce a zároveň nevyhoví v první a druhé zkoušce - tj. nastane jev $\overline{A} \cap \overline{B} \cap C$. Uvažovaný jev nastane při výskytu kterékoliv z popsaných situací - jedná se tedy o jejich sjednocení. Množinový zápis zní $(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$.

g) Výrobek obstojí právě ve dvou zkouškách, když vyhoví v první a druhé zkoušce a zároveň nevyhoví ve třetí zkoušce - tj. nastane jev $A \cap B \cap \overline{C}$, nebo vyhoví v první a třetí zkoušce a zároveň nevyhoví ve druhé zkoušce - tj. nastane jev $A \cap \overline{B} \cap C$, nebo vyhoví ve druhé a třetí zkoušce a zároveň nevyhoví v první zkoušce - tj. nastane jev $\overline{A} \cap B \cap C$. Uvažovaný jev nastane při výskytu kterékoliv z popsaných situací - jedná se tedy o jejich sjednocení. Množinový zápis zní $(A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C)$.

h) Výrobek v testech obstojí nejvýše dvakrát, pokud neobstojí v žádném testu, nebo obstojí právě v jednom testu, nebo obstojí právě ve dvou testech. Sloučením (tj. sjednocením) všech těchto případů bychom dostali žádané vyjádření. Jednodušší však bude následující pohled. Pokud výrobek obstojí nejvýše ve dvou testech - jak vypadá příslušný opačný jev? Tomu odpovídá situace, že výrobek obstál ve všech třech testech, tj jevu $A \cap B \cap C$. Opačný jev má tedy vyjádření $\overline{A \cap B \cap C}$, což je výraz, který jsme měli určit.

1.3 Definice pravděpodobnosti náhodného jevu

1.3.1 Pravděpodobnost náhodného pokusu

Pro náhodný jev A definujeme jeho pravděpodobnost vztahem

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

kde $m(A)$, resp. $m(\Omega)$ znamená mohutnost množiny A, resp. Ω.

Výše uvedená definice potřebuje několik poznámek. Jedná se o tzv. klasickou definici pravděpodobnosti. Lze ji použít v případě, že každý z prvků množiny Ω má stejnou šanci, že nastane - což nemusí být vždy splněno. V případě množin s konečným počtem prvků většinou mohutnost množiny odpovídá počtu prvků této množiny, viz Příklad 1.7. Nicméně, lze se setkat i s úlohami, ve kterých jak množina A, tak množina Ω mají nekonečně mnoho prvků, a pak mohutností množiny rozumíme její jinou míru, viz např. Příklad 1.8 .

1.7. Náhodným pokusem je náhodný výběr jedné osoby při zasedání poslanecké sněmovny ČR (předpokládejme, že nikdo z poslanců a poslankyň nechybí). Náhodný jev A nastane tehdy, je-li touto osobou žena. Jaká je pravděpodobnost jevu A?

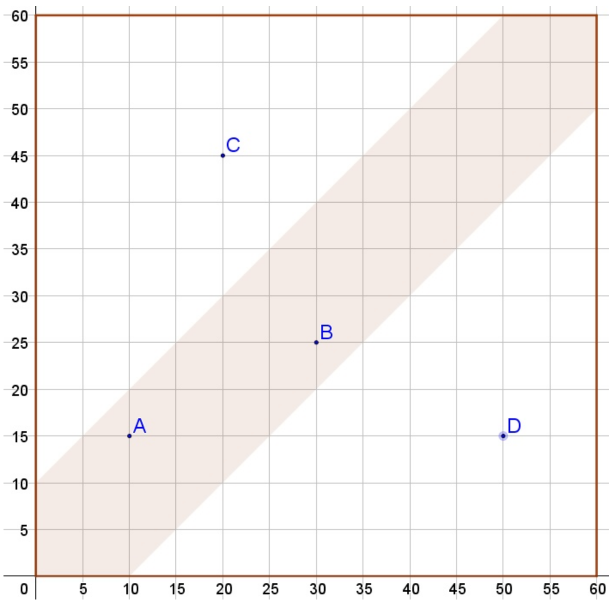
Řešení: Víme, že poslanecká sněmovna ČR má celkem 200 členů, tj. $m(\Omega) = 200$, z nichž je 45 žen (stav v březnu 2020), tj. $m(A) = 45$. Dle definice je

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{45}{200} = 0,225.$$

Pravděpodobnost výběru ženy při výběru z řad členů poslanecké sněmovny je rovna $0,225$.

1.8. Adam s Evou se domluvili, že se spolu setkají zítra u kina Hraničář, a to v době mezi 12. a 13. hodinou. Každý z nich bude na druhého čekat 10 minut, pokud ten druhý do té doby nepřijde, první odchází a nesetkají se. Jaká je pravděpodobnost, že se zítra za uvedených podmínek setkají?

Řešení: Při popisu situace si můžeme pomoci Obrázkem 1.2. Dobu mezi 12. a 13. hodinou jsme si rozdělili na 60 minut. V zobrazeném grafu si pomocí bodů (viz bod A, bod B, bod C, bod D) znázorníme okamžiky příchodu obou partnerů tak, že x-ová, resp. y-ová souřadnice každého bodu představuje čas příchodu Adama, resp. Evy. Z obrázku je zřejmé, že bod A představuje



Obrázek 1.2: Ilustrační obrázek k Příkladu 1.8

situaci, ve které Adam přišel ve 12:10 a Eva ve 12:15. Eva přišla 5 minut po Adamovi, ten na ni v rámci desetiminutové čekací doby počkal, a proto se oba setkali. Oproti tomu bod D znázorňuje situaci, ve které Eva přišla ve 12:15, čekala 10 minut a pak odešla aniž by potkala Adama, který se dostavil ve 12:50. Bod D tedy představuje situaci, kdy se oba partneři nesetkali. Stejně tak můžeme říci, že bod B je situace, ve které setkání proběhlo a v bodě C k setkání nedošlo (proč?).

Každý bod ze zobrazeného čtverce představuje jednu možnost příchodu každého z partnerů. Některé body přitom odpovídají situaci, kdy se Adam s Evou setkali, některé (zbývající) body představují stav, kdy se oba partněři minuli. „Úhlopříčný pruh“ na obrázku obsahuje ty body, které znázorňují situaci, ve které se Adam s Evou potkají. Body mimo tento pruh symbolizují stav, kdy se nesetkají.

Zadání úlohy tak můžeme převést do abstraktnější roviny, kdy náhodným pokusem je výběr jednoho bodu ze zobrazeného čtverce (tento bod představuje náhodné příchody obou partnerů v průběhu jedné hodiny). Jev A nastane, leží-li tento bod ve zvýrazněném úhlopříčném pruhu. Prostorem všech elementárních jevů Ω je zde množina všech bodů čtverce, jev A symbolizují body uvnitř zvýrazněného pruhu. Jak Ω, tak A má nekonečně mnoho prvků, proto není vhodné použít jako mohutnost (velikost) obou množin počet jejich prvků. Namísto toho lze jako jejich velikost uvažovat obsah plochy, kterou obě množiny vyplňují - v tomto případě je to vhodná míra velikosti obou množin.

Množina Ω vyplňuje čtverec, jehož obsah snadno určíme - je $m(\Omega) = 60 \cdot 60 = 3\,600$. Množina A vyplňuje šestiúhelník, jehož obsah lze určit několika způsoby. V tomto případě zřejmě nejsnažší (tj. hodně elementární) způsob spočívá v určení počtu vybarvených čtverců o rozměrech 5×5 jednotek. Je zde 34 plně vybarvených čtverců a 20 z poloviny vybarvených čtverců, které odpovídají 10 plně vybarveným čtvercům, tj. dohromady 44 čtverců, každý z nich o obsahu $5 \cdot 5 = 25$. Pro velikost množiny A tak platí $m(A) = 44 \cdot 25 = 1\,100$. Jiný (obecnější) způsob určení obsahu vyznačené plochy může spočívat např. v rozdělení šestiúhelníku na dva shodné lichoběžníky a určení obsahu těchto lichoběžníků odečtením ploch dvou trojúhelníků.

Pravděpodobnost jevu A je rovna

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{1100}{3600} = 0,30\overline{5}.$$

Pravděpodobnost, že se Adam s Evou setkají za podmínek uvedených v zadání úlohy činí $0,30\overline{5}$.

1.3.2 Operace s náhodnými jevy

V následující části si ukážeme postup výpočtu pravděpodobnosti jevů pomocí klasické definice a rozšíříme jej na jevy, které vzniknou pomocí operací s náhodnými jevy. Mohutností množiny zde budeme rozumět počet prvků dané množiny.

1.9. Náhodný pokus spočívá v hodu kostkou. Dále jsou definovány následující náhodné jevy.

a) Jev A nastane, jestliže na kostce padne číslo 3.

b) Jev B nastane, jestliže na kostce padne sudé číslo.

c) Jev C nastane, jestliže na kostce padne číslo větší než 3.

Vypočtěte pravděpodobnosti jevů A, B a C a určete a interpretujte pravděpodobnosti jevů $\overline{C}$, $B \cap C$, $B \cup C$.

Výpočet $P(A)$

$$\begin{aligned} A &= \{3\} \\ \Omega &= \{1,\,2,\,3,\,4,\,5,\,6\} \\ m(A) &= 1 \\ m(\Omega) &= 6 \\ P(A) &= \frac{m(A)}{m(\Omega)} = \frac{1}{6} \end{aligned}$$

Pravděpodobnost jevu A je rovna $\frac{1}{6}$.

Výpočet $P(B)$

$$\begin{aligned} B &= \{2,\ 4,\ 6\} \\ \Omega &= \{1,\ 2,\ 3,\ 4,\ 5,\ 6\} \\ m(B) &= 3 \\ m(\Omega) &= 6 \\ P(B) &= \frac{m(B)}{m(\Omega)} = \frac{3}{6} = \frac{1}{2} \end{aligned}$$

Pravděpodobnost jevu B je rovna $\frac{1}{2}$.

Výpočet $P(C)$

$$\begin{aligned} C &= \{4,\ 5,\ 6\} \\ \Omega &= \{1,\ 2,\ 3,\ 4,\ 5,\ 6\} \\ m(C) &= 3 \\ m(\Omega) &= 6 \\ P(C) &= \frac{m(C)}{m(\Omega)} = \frac{3}{6} = \frac{1}{2} \end{aligned}$$

Pravděpodobnost jevu C je rovna $\frac{1}{2}$.

Výpočet $P(\overline{C})$

$$\begin{aligned} C &= \{4,\ 5,\ 6\} \\ \Omega &= \{1,\ 2,\ 3,\ 4,\ 5,\ 6\} \\ \overline{C} &= \Omega \backslash C = \{1,\ 2,\ 3\} \\ m(\overline{C}) &= 3,\ m(\Omega) = 6 \\ P(\overline{C}) &= \frac{m(\overline{C})}{m(\Omega)} = \frac{3}{6} = \frac{1}{2} \end{aligned}$$

(Výsledek a jeho interpretace:) Pravděpodobnost jevu, že nepadne číslo větší než tři, je rovna $\frac{1}{2}$.

Výpočet $P(B \cap C)$

$$\begin{aligned} B &= \{2,\ 4,\ 6\} \\ C &= \{4,\ 5,\ 6\} \\ B \cap C &= \{4,\ 6\} \\ m(B \cap C) &= 2,\ m(\Omega) = 6 \\ P(B \cap C) &= \frac{m(B \cap C)}{m(\Omega)} = \frac{2}{6} = \frac{1}{3} \end{aligned}$$

Výsledek a jeho interpretace

Pravděpodobnost padnutí sudého čísla většího než tři (tj. případu kdy nastane jev B a současně jev C) je rovna $\frac{1}{3}$.

Výpočet $P(B \cup C)$

$$
\begin{aligned}
B &= \{2,\, 4,\, 6\} \\
C &= \{4,\, 5,\, 6\} \\
B \cup C &= \{2,\, 4,\, 5,\, 6\} \\
m(B \cup C) &= 4,\ m(\Omega) = 6 \\
P(B \cup C) &= \frac{m(B \cup C)}{m(\Omega)} = \frac{4}{6} = \frac{2}{3}
\end{aligned}
$$

Výsledek a jeho interpretace
Pravděpodobnost padnutí buď sudého čísla nebo čísla většího než tři (tj. případu, kdy nastane alespoň jeden z jevů B, resp. C) je rovna $\frac{2}{3}$.

1.4 Pravděpodobnost průniku náhodných jevů

V této části se budeme věnovat obecnějším úvahám o pravděpodobnostech. Opustíme způsob výpočtu spočívající ve výčtu všech elementárních jevů, které vedou k nastání stanovených jevů a zaměříme se na obecná pravidla, která platí pro počítání s pravděpodobnostmi jevů a jejich průniku, sjednocení, resp. rozdílu.

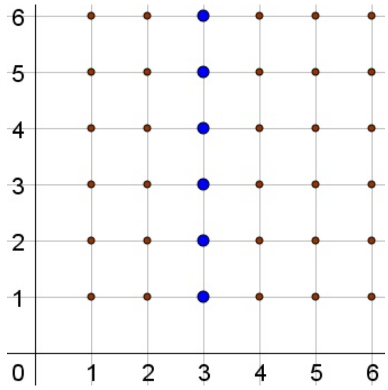
Věta **??** se někdy používá k definici nezávislosti jevů, někdy se používá k výpočtu pravděpodobnosti jevů, o kterých víme, že jsou nezávislé. Zde si vysvětlíme, co rozumíme pojmem nezávislé jevy a pak si uvedeme, kterak počítat pravděpodobnost jejich průniku, tj. pravděpodobnost toho, že oba jevy nastanou současně.

Předpokládejme, že jsou dány jevy A a B, přičemž známe pravděpodobnosti obou těchto jevů. Jevy A a B lze považovat za nezávislé, jestliže znalost o tom, zda nastal jev A, nevede k změně pravděpodobnosti jevu B a naopak, znalost o tom, zda nastal jev B, nevede ke změně pravděpodobnosti jevu A. Pro ukázku uveďme dva ilustrační příklady.

1.10. Budeme vycházet ze zadání Příkladu 1.8, kde náhodným pokusem bylo hození kostkou a byly definovány jevy A, B a C. Připomeňme, že $B = \{2, 4, 6\}$, $P(B) = 1/2$ a $C = \{4, 5, 6\}$, $P(C) = 1/2$. Nyní si představte situaci, že někdo hodí kostkou. Vy nevíte, jaké číslo na kostce padlo. Pravděpodobnost hodu sudého čísla je $1/2$, pravděpodobnost hodu čísla většího než 3 je $1/2$. Nyní Vám osoba házející kostkou řekla, že při hodu kostkou nastal jev B, tj. že padlo sudé číslo. Jaká je nyní pravděpodobnost, že nastal jev C, tj. že padlo číslo větší než 3? Nyní již víme, že mohlo padnout pouze jedno z čísel 2, 4, 6. Proto je $\Omega = \{2, 4, 6\}$. Z těchto případů jsou větší než tři pouze čísla 4 a 6, tedy $C = \{4, 6\}$. Pravděpodobnost toho, že padlo číslo větší než tři, při znalosti toho, že padlo sudé číslo, je rovna $2/3$ a tedy znalost, že nastal jev B změnila pravděpodobnost jevu C. Jevy proto nejsou nezávislé, jsou tzv. závislé.

1.11. Náhodný pokus spočívá v hodu dvěma kostkami - pro jejich rozlišení budeme uvažovat hod modrou a červenou kostkou. Jev A nastane, jestliže na modré kostce padne číslo 3. Jev

B nastane, jestliže na červené kostce padne číslo 4. Situaci budeme ilustrovat na Obrázku 1.3. Na něm můžeme vidět 36 bodů (no, spíše puntíků), z nichž x-ová souřadnice každého bodu odpovídá číslu, které padlo na modré kostce, y-ová souřadnice každého bodu odpovídá číslu, které padlo na červené kostce. Každý bod tak odpovídá jednomu konkrétnímu výsledku hodu dvěma kostkami. Pokud například na modré kostce padne číslo 2 a na červené kostce padne číslo 5, dostáváme bod o souřadnicích (2,5). Jev A nastane v těchto případech



Obrázek 1.3: Ilustrace hodu dvěma kostkami

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

Je $m(A) = 6$, $m(\Omega) = 36$, proto

$$P(A) = \frac{6}{36} = \frac{1}{6}.$$

Na Obrázku 1.3 jsou případy, kdy nastal jev A, zobrazeny modrými body. Je zřejmé, že modré puntíky tvoří šestinu z celkového počtu, což potvrzuje výpočet $P(A)$.

Analogicky, pro jev B platí

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

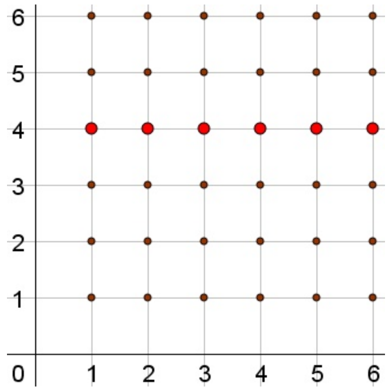
Je $m(B) = 6$, $m(\Omega) = 36$, proto

$$P(B) = \frac{6}{36} = \frac{1}{6}.$$

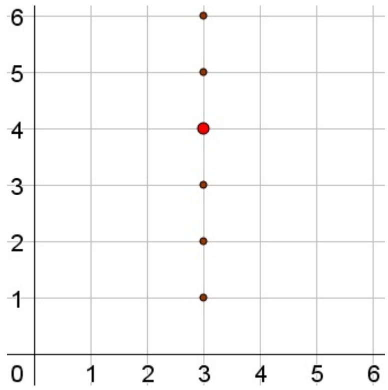
Nyní se někdo podíval na hozené kostky a řekl, že nastal jev A, tj. na modré kostce padlo číslo tři. Jaká je nyní pravděpodobnost, že nastal jev B, tj. že při tomtéž hodu padlo na červené kostce číslo čtyři? Nyní již množinu Ω tvoří pouze šest možností $\Omega = \{(3,1),\ (3,2),\ (3,3),\ (3,4),\ (3,5),\ (3,6)\}$, viz Obrázek 1.5. Z uvedených možností vede k jevu B jediná, a to $(3,4)$. Pokud nastal jev A, je $m(\Omega) = 6$, $m(B) = 1$, proto

$$P(B) = \frac{1}{6}.$$

Viz Obrázek 1.5. Vidíme, že pravděpodobnost jevu B je stejná bez ohledu na to, zda máme informaci o tom, že nastal či nenastal jev A. Jevy A a B jsou proto nezávislé. Nyní by již měla být zřejmá idea pojmu závislé, resp. nezávislé jevy a můžeme přikročit k následující větě.



Obrázek 1.4: Vyznačení jevu B



Obrázek 1.5: Vyznačení jevu B pokud nastal jev A

Věta 1.4.1. *Jsou-li A a B nezávislé jevy, potom platí rovnost*

$$P(A \cap B) = P(A) \cdot P(B). \tag{1.1}$$

Uvedená věta říká, jak určit pravděpodobnost, že dva jevy nastanou současně za předpokladu, že jsou tyto jevy vzájemně nezávislé. Poznamenejme ještě, že v praxi často není hned zřejmé, zda jevy jsou, či nejsou nezávislé a je nutné toto ověřit měřením. Při tomto měření v podstatě ověřujeme, zda platí vztah (1.1).

Ukažme si platnost věty 1.4.1 na již probraných příkladech. Uvažujme jevy B a C z Příkladu 1.9. Víme (viz Příklad 1.10), že jevy B a C jsou závislé. Dále jsme již vypočítali (viz Příklad 1.9), že platí:

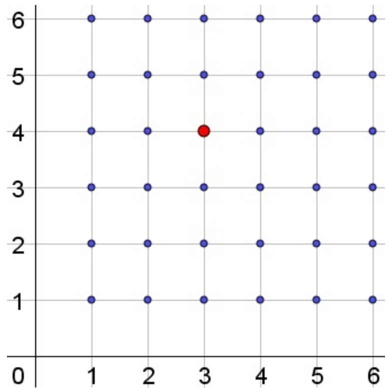
$$P(B) = \frac{1}{2}, \quad P(C) = \frac{1}{2} \quad P(B \cap C) = \frac{1}{3}.$$

Všimněte si, že

$$P(B \cap C) = \frac{1}{3} \neq \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B) \cdot P(C).$$

Pravděpodobnost průniku jevů v tomto případě **není** rovna součinu pravděpodobností těchto jevů, neboť jevy nejsou vzájemně nezávislé. Pokud byste tedy použili vztah (1.1) v případě závislých jevů, dostanete špatný výsledek.

Nyní uvažujme jevy A a B z Příkladu 1.11. Jev $A \cap B$ představuje jev, při kterém nastane jev A a současně nastane i jev B, tj. hodíme dvěma kostkami - na modré kostce padne číslo tři a současně na červené kostce padne číslo čtyři, tj. průnikem obou jevů je jediný případ $(3, 4)$ z celkového počtu všech 36 možných výsledků, viz Obrázek 1.6. Proto je $m(A \cap B) = 1$ a



Obrázek 1.6: Grafické vyjádření $P(A \cap B)$

$m(\Omega) = 36$ a platí:

$$P(A \cap B) = \frac{m(A \cap B)}{m(\Omega)} = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(A) \cdot P(B).$$

Pravděpodobnost průniku jevů v tomto případě **je** rovna součinu pravděpodobností těchto jevů, neboť z Příkladu 1.11 víme, že oba jevy jsou vzájemně nezávislé.

Můžeme se ptát, jak se vypočítá pravděpodobnost průniku jevů v případě, kdy jsou oba jevy závislé. Takový vzorec existuje, nicméně k jeho vyslovení budeme potřebovat pojem podmíněné pravděpodobnosti, který dosud nemáme zavedený.

1.5 Pravděpodobnost sjednocení náhodných jevů

Níže vyslovíme tři věty, které ukazují jak počítat pravděpodobnost, že nastal alespoň jeden z uvažovaných jevů, tj. pravděpodobnost sjednocení náhodných jevů.

Věta 1.5.1. *Pro libovolné jevy A a B platí vztah*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{1.2}$$

Speciálně, jsou-li A a B nezávislé jevy, potom je

$$P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B). \tag{1.3}$$

Speciálně, jsou-li A a B neslučitelné jevy, potom je

$$P(A \cup B) = P(A) + P(B). \tag{1.4}$$

Dále následuje věta, která je přímým důsledkem rovnice (1.4) a faktu, že jevy A a $\overline{A}$ jsou neslučitelné a přitom jejich sjednocení tvoří množinu Ω.

Věta 1.5.2. *Pro pravděpodobnosti jevů A a $\overline{A}$ platí následující vzorec:*

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

1.12. Vypočtěte pravděpodobnosti, že při hodu dvěma hracími kostkami padne

a) na obou kostkách šestka,

b) alespoň jedna šestka,

c) právě jedna šestka,

d) ani jedna šestka.

Řešení: Nejprve obecný komentář. Náhodným pokusem je hod dvěma kostkami, např. modrou a červenou.

- Náhodný jev A nastane, pokud na první (modré) kostce padne šestka.
- Náhodný jev B nastane, pokud na druhé (červené) kostce padne šestka.

Již víme, viz Příklad 1.11, že jevy A a B jsou navzájem nezávislé a platí:

$$P(A) = \frac{1}{6}, \quad P(B) = \frac{1}{6}.$$

Nyní následuje řešení konkrétních bodů.

a) Na obou kostkách padne šestka právě tehdy, když nastane jev A a současně nastane jev B. Jedná se o průnik jevů A a B a vzhledem k nezávislosti obou jevů lze psát

$$P(A \cap B) = P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

b) K uvedenému případu dojde, když nastane alespoň jeden z jevů A a B. Jedná se tedy o sjednocení jevů A a B a vzhledem k nezávislosti obou jevů lze psát

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A) \cdot P(B) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}.$$

c) Tvrzení znamená, že buď padne šestka na první kostce a na druhé nepadne, nebo na první kostce nepadne šestka a na druhé ano.

$$\begin{aligned} P[(A \cap \overline{B}) \cup (\overline{A} \cap B)] &= P(A \cap \overline{B}) + P(\overline{A} \cap B) - P(\overline{A} \cap B \cap A \cap \overline{B}) = \\ &= P(A) \cdot P(\overline{B}) + P(\overline{A}) \cdot P(B) + P(\emptyset) = \\ &= \frac{1}{6} \cdot \frac{5}{6} + \frac{5}{6} \cdot \frac{1}{6} - 0 = \frac{10}{36} \end{aligned}$$

Komentář k výpočtu: Jev $\overline{A} \cap B \cap A \cap \overline{B}$ je nemožný s nulovou pravděpodobností. Pokud např. počítáme průnik jevů A a $\overline{A}$, tak chceme, aby nastal jev A a současně jev opačný k jevu A, což není možné.

d) Na první kostce nastane jev $\overline{A}$ a současně na druhé kostce nastane jev $\overline{B}$. Proto je

$$P(\overline{A} \cap \overline{B}) = P(\overline{A}) \cdot P(\overline{B}) = \frac{5}{6} \cdot \frac{5}{6} = \frac{25}{36}.$$

Lze řešit i následujícím způsobem. Jev „žádná šestka“ je doplněk k jevu „alespoň jedna šestka“. Proto platí

$$P(\overline{A} \cap \overline{B}) = 1 - P(A \cup B) = 1 - \frac{11}{36} = \frac{25}{36}.$$

1.13. Přítomnost nežádoucích příměsí v mléčném výrobku se testuje pomocí dvou nezávislých zkoušek. První zkouška odhalí nežádoucí příměs s pravděpodobností $p_1 = 0,96$, druhá zkouška s pravděpodobností $p_2 = 0,98$. Jaká je pravděpodobnost, že nežádoucí příměs

a) odhalí alespoň jedna zkouška,

b) odhalí právě jedna z obou zkoušek,

c) neodhalí žádná zkouška?

Obecný komentář: Náhodným pokusem je testování výrobku pomocí dvou nezávislých zkoušek.

- Náhodný jev A nastane, pokud výrobek má nežádoucí příměs a tu odhalí první zkouška.
- Náhodný jev B nastane, pokud výrobek má nežádoucí příměs a tu odhalí druhá zkouška.

Ze zadání příkladu (zkoušky jsou na sobě nezávislé) víme, že jevy A a B jsou navzájem nezávislé a platí:

$$P(A) = 0,96, \quad P(B) = 0,98.$$

Nyní následuje řešení konkrétních bodů. *Řešení:*

a) Jev „alespoň jedna zkouška“ je sjednocením jevů A a B. Proto je

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) = \\ &= P(A) + P(B) - P(A) \cdot P(B) = \\ &= 0,96 + 0,98 - 0,96 \cdot 0,98 = 0,9992. \end{aligned}$$

b) Jev „právě jedna zkouška“ znamená, že nastane A a nenastane B, nebo nenastane A a nastane B. Proto je

$$\begin{aligned} P[(A \cap \overline{B}) \cup (\overline{A} \cap B)] &= P(A \cap \overline{B}) + P(\overline{A} \cap B) - P(\overline{A} \cap B \cap A \cap \overline{B}) = \\ &= P(A) \cdot P(\overline{B}) + P(\overline{A}) \cdot P(B) + P(\emptyset) = \\ &= 0,96 \cdot 0,02 + 0,98 \cdot 0,04 = 0,0584. \end{aligned}$$

c) Jev „žádná zkouška“ nastane tehdy, nenastane-li A a nenastane B, resp. nastane $\overline{A}$ a současně nastane $\overline{B}$. Je

$$P(\overline{A} \cap \overline{B}) = P(\overline{A}) \cdot P(\overline{B}) = 0,02 \cdot 0,04 = 0,0008.$$

Lze počítat i jako opačný jev k jevu „příměs odhalí alespoň jedna zkouška“, potom je

$$P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - 0,9992 = 0,0008.$$

1.5.1 Úlohy k samostatné práci

1.14. Jaká je pravděpodobnost, že při hodu dvěma kostkami bude součet padlých čísel roven číslu osm? *Nápověda*: Použijte schéma analogické Obrázkům 1.3 až 1.6.

1.15. Během dne mají být doplněny zásoby ve skladu, přičemž zásobování probíhá pomocí dvou kamionů, jejichž vyložení trvá přesně 1 hodinu. Sklad přijímá kamiony k vyskladnění v době od 8:00 do 16:00. Kamiony přijíždí nezávisle na sobě náhodně kdykoliv během přijímací doby. Pokud kamion přijede a vykládací rampa je již obsazena jiným kamionem, musí tento počkat na vyložení nákladu předchozího kamionu. Jaká je pravděpodobnost, že jeden kamion bude muset čekat na vyložení předchozího kamionu? *Nápověda*: Vidíte analogii s Adamem a Evou v Příkladu 1.8?

1.16. V ruletě mohou padnout čísla $0, 1, 2, \ldots, 36$. Hráč při jednom kole hry učinil celkem tři sázky, přičemž vsadil na sudá čísla, na červenou barvu a na první tucet. S jakou pravděpodobností vyhraje alespoň jednu z těchto 3 sázek? *Poznámka*: Při řešení sledujte hrací plán uvedený na Obrázku 1.7.

Obrázek 1.7: Hrací plán pro ruletu

1.17. Víme, že platí následující rovnosti: $P(A) = 0.2$, $P(B) = 0.6$, $P(A \cap B) = 0.2$. Rozhodněte, zda jsou jevy A a B:

- nezávislé,
- neslučitelné,

a svou odpověď zdůvodněte.

1.18. Náhodný pokus spočívá ve dvou hodech kostkou. Jev A nastane, jestliže při prvním hodu padne číslo 3. Jev B nastane tehdy, padne-li při druhém hodu číslo větší než při prvním hodu. Rozhodněte, zda jsou jevy A a B:

- nezávislé,
- neslučitelné,

a svou odpověď zdůvodněte.

1.19. Řidič na cestě do svého zaměstnání projíždí přes tři křižovatky řízené semafory. Na každém z nich mu svítí zelená s pravděpodobností $1/4$. Vypočtěte pravděpodobnost, že

a) „chytne zelenou vlnu“, tj. že na všech semaforech bude svítit zelená,

b) „chytne červenou vlnu“, tj. že na všech semaforech bude svítit červená,

c) bude čekat pouze na prvním semaforu.

Poznámka: Předpokládáme, že semafory jsou řízené nezávisle na sobě a řidič křižovatkou projíždí pouze při zeleném světle na semaforu.

1.20. Dva na sobě nezávislí kontroloři kontrolují účetnictví jedné firmy. V účtech je chyba. První kontrolor ji nalezne s pravděpodobností 0,85, druhý kontrolor ji nalezne s pravděpodobností 0,93. Jaká je pravděpodobnost, že

a) chybu odhalí alespoň jeden kontrolor,

b) chybu odhalí oba kontroloři,

c) chybu neodhalí ani jeden kontrolor.

1.21. Studenti mají v daném dni cvičení ze tří různých předmětů. Na každý předmět mají jiného kantora, z nichž každý přijde do hodiny pozdě s pravděpodobností 0.1, resp. 0.2, resp. 0.3. Vypočtěte pravděpodobnost, že:

a) alespoň jeden kantor přijde včas,

b) alespoň jeden kantor přijde pozdě.

1.22. Zařízení se skládá z 5 stejných, na sobě nezávislých součástek. Pravděpodobnost, že každá z nich bude bezchybně pracovat alespoň 40 hodin, je rovna 0.95. Jaká je pravděpodobnost, že zařízení bude fungovat alespoň 40 hodin, stačí-li, aby fungovaly alespoň 4 součástky?

1.23. Kolik hodů mincí je třeba provést, aby pravděpodobnost, že padne alespoň jednou líc, byla větší než $0,9$?

Kapitola 2

Náhodná veličina, rozdělení pravděpodobnosti náhodné veličiny

V této kapitole se budeme zabývat pojmem *náhodná veličina* a vlastnostmi náhodné veličiny. Budeme přitom vycházet z videopřednášek v Moodlu, které připravila přednášející dr. Šimsová - tzn. že byste tato videa měli mít shlédnuta, než začnete číst tento text. Po prostudování kapitoly byste měli být schopni:

- rozumět pojmu *náhodná veličina*,
- při zadání konkrétního příkladu náhodné veličiny určit její *výběrový prostor*,
- rozlišovat *diskrétní* náhodnou veličinu a *spojitou* náhodnou veličinu,
- rozumět pojmu *pravděpodobnostní funkce* diskrétní náhodné veličiny,
- rozumět pojmu *distribuční funkce* náhodné veličiny,
- rozumět pojmu *rozdělení pravděpodobnosti* diskrétní náhodné veličiny,
- pracovat s *charakteristikami* diskrétní náhodné veličiny, mezi které patří zejména:
  - první obecný moment, tedy *střední hodnota* náhodné veličiny,
  - druhý obecný moment náhodné veličiny,
  - druhý centrovaný moment, neboli *rozptyl* náhodné veličiny,
  - *směrodatná odchylka* náhodné veličiny,
  - modální hodnota, neboli *modus* náhodné veličiny,
  - *p% kvantil* náhodné veličiny,
- pracovat s pojmem náhodná veličina s *alternativním rozdělením* pravděpodobnosti,
- pracovat s pojmem náhodná veličina s *binomickým rozdělením* pravděpodobnosti,
- pracovat s pojmem náhodná veličina s *Poissonovým rozdělením* pravděpodobnosti,
- pracovat s pojmem náhodná veličina s *hypergeometrickým rozdělením* pravděpodobnosti.

2.1 Náhodná veličina, pravděpodobnostní funkce

V níže uvedeném příkladu představíme pojmy náhodná veličina, výběrový prostor náhodné veličiny a pravděpodobnostní funkce (též funkce pravděpodobnosti) náhodné veličiny.

2.1. Studenti skládají písemku sestávající ze tří úloh. Za každou z nich mohou získat až dva body. Celkem tedy mohou při písemce získat nejvýše šest bodů. Víme, že jeden student získal 0 bodů, dva získali 1 bod, 4 studenti získali 2 body, šest studentů 3 body, čtyři studenti 4 body, dva studenti 5 bodů a jeden student 6 bodů. Náhodným pokusem je náhodný výběr jedné písemky. Označme symbolem $\mathbb{X}$ počet bodů, které má tato vybraná písemka. Jaké hodnoty $\mathbb{X}$ můžeme očekávat a jaké jsou jejich pravděpodobnosti?

Řešení: Ze zadání plyne, že veličina $\mathbb{X}$ nabývá hodnoty pouze z množiny $M = \{0, 1, 2, 3, 4, 5, 6\}$. Může tedy být $\mathbb{X} = 0$, resp. $\mathbb{X} = 1, \ldots, \mathbb{X} = 6$.

Vypočteme pravděpodobnost, že náhodně vybraná písemka je ohodnocena nula body. Matematicky tuto situaci vyjádříme zápisem $P(\mathbb{X} = 0)$. Tedy symbol $P(\mathbb{X} = 0)$ znamená pravděpodobnost jevu, že $\mathbb{X}$ bude rovno nule. Podobně, symbol $P(\mathbb{X} = 1)$ znamená pravděpodobnost, že $\mathbb{X}$ bude rovno jedné; tedy pravděpodobnost, že při náhodném výběru jedné písemky vytáhneme tu s jedním bodem.

Pro výpočet pravděpodobnosti uvedeme hodnoty ze zadání do přehledné tabulky, viz Tabulka 2.1. Z tabulky snadno vyčteme, že studentů je celkem 20. Písemku s nulou bodů napsal

počet bodů z písemky	0	1	2	3	4	5	6	celkem
počet žáků s daným počtem bodů	1	2	4	6	4	2	1	20

Tabulka 2.1: Počty studentů s daným počtem bodů z písemné práce

jeden z nich, proto pravděpodobnost výběru písemky s nulou bodů je rovna

$$P(\mathbb{X} = 0) = \frac{1}{20} = 0,05.$$

Podobným způsobem určíme, že (viz ostatní hodnoty v Tabulce 2.1)

$$P(\mathbb{X} = 1) = \frac{2}{20} = 0.10, \quad P(\mathbb{X} = 2) = \frac{4}{20} = 0.20, \quad P(\mathbb{X} = 3) = \frac{6}{20} = 0.30,$$
$$P(\mathbb{X} = 4) = \frac{4}{20} = 0.20, \quad P(\mathbb{X} = 5) = \frac{2}{20} = 0.10, \quad , P(\mathbb{X} = 6) = \frac{1}{20} = 0.05.$$

Dále tedy v této úloze budeme používat symbol $\mathbb{X}$ pro označení veličiny znamenající počet bodů z písemné práce, symbol x pro konkrétní hodnotu této veličiny a označení $P(\mathbb{X} = x) = p$ bude znamenat, že pravděpodobnost že veličina $\mathbb{X}$ je rovna konkrétnímu číslu x, má hodnotu p. Nalezené výsledky jsou shrnuty v Tabulce 2.2.

x	0	1	2	3	4	5	6	Σ
$P(\mathbb{X} = x)$	0.05	0.10	0.20	0.30	0.20	0.10	0.05	1

Tabulka 2.2: Pravděpodobnosti konkrétních hodnot veličiny $\mathbb{X}$

Přestože jsme v příkladu počítali pravděpodobnosti výskytu jednotlivých situací, podstata úlohy spočívá v něčem jiném - v zavedení a osvojení si názvosloví a symboliky používané v souvislosti s náhodnou veličinou. Zde symbol:

- $\mathbb{X}$ ... vyjadřuje označení konkrétní náhodné veličiny (v dalším textu budeme k označení náhodné veličiny používat právě tento font písma),
- x ... znamená konkrétní hodnotu veličiny $\mathbb{X}$, kterou jsme obdrželi při *realizaci* konkrétního náhodného pokusu,
- $P(\mathbb{X} = x)$ ... uvádí hodnotu pravděpodobnosti, že náhodná veličina $\mathbb{X}$ při konkrétní realizaci náhodného pokusu nabude hodnotu x.

Úlohy k samostatnému procvičení

V následujících úlohách určete množinu M možných hodnot náhodné veličiny $\mathbb{X}$ a pokud je to ze zadání možné, vypočtěte hodnoty pravděpodobnostní funkce pro jednotlivé hodnoty x.

2.2. Náhodným pokusem je hod dvěma hracími kostkami. Náhodná veličina $\mathbb{X}$ má hodnotu jedna, pokud na obou kostkách padla stejná čísla, a hodnotu nula, pokud padla různá čísla.

2.3. Náhodným pokusem je hod dvěma hracími kostkami. Náhodnou veličinou $\mathbb{X}$ je počet sudých čísel, která při jednom hodu padnou.

2.4. Náhodným pokusem je hod dvěma hracími kostkami. Náhodnou veličinou $\mathbb{X}$ je součet čísel, která při jednom hodu padnou.

Všimněte si, že ve všech třech předchozích úlohách byla náhodným pokusem stále stejná situace a pokaždé jsme dokázali náhodnou veličinu definovat různým způsobem. Ve všech těchto příkladech byste měli být schopni určit pravděpodobnostní funkci.

2.5. Náhodným pokusem je příchod na autobusovou zastávku a čekání na příjezd autobusu, který jezdí v pravidelných intervalech po 20 minutách . Náhodnou veličinou $\mathbb{X}$ je doba, po kterou musíme čekat, než přijede autobus, na který čekáme.

2.6. Náhodným pokusem je příchod na autobusovou zastávku a čekání na příjezd autobusu. Náhodnou veličinou $\mathbb{X}$ je počet autobusů jiných linek, které zastavily na zastávce do příjezdu „našeho“ busu.

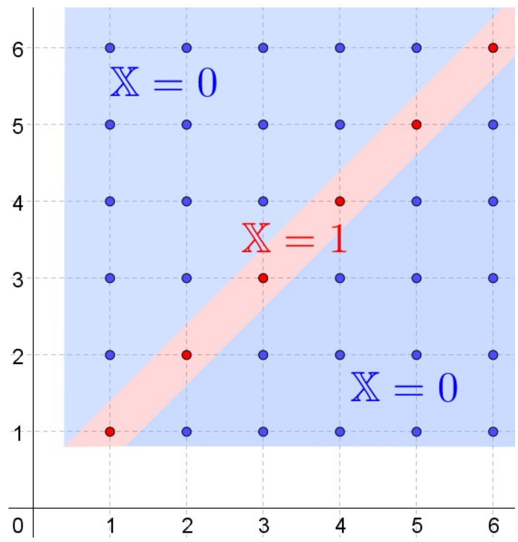
2.7. Náhodným pokusem je přejímka zboží. Náhodnou veličinou $\mathbb{X}$ je počet vadných kusů v zásilce.

2.8. Náhodným pokusem je přejímka zboží. Náhodnou veličinou $\mathbb{X}$ je celková hmotnost zásilky.

V předchozích úlohách se opět některé náhodné pokusy opakovaly a byli jsme schopni k nim stanovit různé náhodné veličiny. V posledních čtyřech příkladech máme málo informací k tomu, abychom určili pravděpodobnostní funkci. Všimněte si, že v Příkladech 2.5 a 2.8 množina M všech možných hodnot x veličiny $\mathbb{X}$ odpovídala intervalu reálných čísel, ve zbylých případech se množina M dala určit výčtem hodnot.

Definice 2.1.1. Mějme dán nějaký náhodný pokus a s ním spojený prostor všech elementárních jevů Ω. Vytvořme rozklad množiny Ω, tj. úplný systém náhodných jevů $A_1, A_2, A_3, \ldots, A_n$ tak, aby každému náhodnému jevu $A_1, A_2, A_3, \ldots, A_n$ bylo možné přiřadit jistou číselnou hodnotu. Tímto přiřazením čísla jednotlivým jevům jsme vytvořili tzv. *náhodnou veličinu.* Náhodná veličina je zobrazení podmnožin z daného rozkladu množiny Ω na množinu reálných čísel.

K ilustraci Definice 2.1.1 si ukažme toto zobrazení v případě náhodné veličiny z Příkladu 2.2. Množinu Ω zde tvoří všechny dvojice možných výsledků hodu dvěma kostkami. Bod na pozici $[2, 3]$ představuje ten výsledek náhodného pokusu, při kterém na první kostce padla dvojka a



Obrázek 2.1: Rozklad množiny Ω na úplný systém náhodných jevů v Příkladu 2.2

na druhé kostce trojka. Ze zadání je zřejmé, že prostor všech elementárních jevů tvoří celkem 36 prvků: $\Omega = \{[1, 1], [1, 2], [1, 3], \ldots [6, 5], [6, 6]\}$, kde ve výrazu $[x, y]$ symbol x znamená číslo padlé na první kostce, y znamená číslo padlé na druhé kostce. Množina Ω se nám rozpadne na dvě podmnožiny.

V první (na Obrázku 2.1 jsou body (spíše puntíky, ale to nechme stranou) reprezentující tuto podmnožinu vyznačeny červeným podbarvením a samotné body mají červenou barvu) se nacházejí prvky symbolizující případy, kdy na obou kostkách padla stejná čísla:

$$[1, 1], \ [2, 2], \ [3, 3], \ [4, 4], \ [5, 5], \ [6, 6],$$

pro které je $\mathbb{X} = 1$. Zbývající body představují případy, kdy na první kostce padlo jiné číslo než na druhé kostce. Ty mají modré podbarvení a body jsou na obrázku vyznačeny modrou barvou. Pro tyto modré body je $\mathbb{X} = 0$.

Volně řečeno – množina Ω se nám rozpadla na dvě části – modrou a červenou – kde modré přiřazujeme hodnotu $\mathbb{X} = 0$ a červené části přiřazujeme hodnotu $\mathbb{X} = 1$. Tolik tedy ke smyslu Definice 2.1.1.

2.1.1 Dělení na diskrétní a spojitou náhodnou veličinu

V následující části občas použijeme slovní spojení *spočetná množina*. O množině můžeme říci, že je spočetná, pokud lze prvkům této množiny přiřadit přirozená čísla tak, aby každý prvek množiny měl přiřazeno přirozené číslo a žádné přirozené číslo se přitom nepoužilo více než jednou. V podstatě jde o to, že spočetná množina je taková množina, ve které můžeme stanovit pořadí jejích prvků - toto je první prvek množiny, druhý prvek množiny, třetí prvek množiny, ... Taková množina může mít nekonečně mnoho prvků - stejně jako množina všech přirozených čísel. Příkladem spočetných množin jsou např. množina všech sudých čísel, množina všech prvočísel atd. Na druhou stranu – ne každá množina s nekonečně mnoha prvky je spočetná. Například jakýkoli neprázdný interval reálných čísel je nespočetná množina.

Náhodné veličiny můžeme dle charakteru jejich hodnot dělit dva základní typy:

- *diskrétní* náhodná veličina (množina M je dána výčtem hodnot),
- *spojitá* náhodná veličina (množina M odpovídá intervalu reálných čísel).

Definice 2.1.2. Diskrétní náhodná veličina nabývá hodnoty pouze z konečné nebo spočetné množiny.

Příklady diskrétní náhodné veličiny:

- číslo padlé při hodu kostkou ... $M = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$,
- počet zmetků vyrobených na výrobní lince za jednu hodinu ... $M = \{0,\ 1,\ 2,\ 3,\ \ldots\}$,
- počet prodaných novin ve stánku za jeden den ... $M = \{0,\ 1,\ 2,\ 3,\ 4,\ \ldots\}$,
- počet bodů z písemky ze statistiky u náhodně vybraného studenta ... $M = \{0,\ 1,\ \ldots,\ 15\}$.

Definice 2.1.3. Spojitá náhodná veličina nabývá libovolnou hodnotu z konečného nebo nekonečného intervalu reálných čísel.

Příklady spojité náhodné veličiny:

- *doba* čekání na obsluhu ... $M = \langle 0, T \rangle$ (kde T je konstanta závislá na rychlosti obsluhy, resp. míře naší trpělivosti),
- *rozměr* náhodně vybrané součástky,
- *výdrž* baterie od posledního nabití,
- *hmotnost* náhodně vybraného člověka.

Úlohy k samostatnému procvičení

2.9. Vraťte se k Příkladu 2.2 až Příkladu 2.8 a rozhodněte, zda náhodné veličiny uvedené v těchto příkladech jsou diskrétní či spojité náhodné veličiny.

2.10. Náhodným pokusem je prodej jahod balených v košíku. Náhodnou veličinou $\mathbb{X}$ je hmotnost jahod v jednom prodaném košíku. Rozhodněte o typu (diskrétní $\times$ spojitá) náhodné veličiny.

2.11. Náhodným pokusem je prodej jahod balených v košíku. Náhodnou veličinou $\mathbb{X}$ je množství jahod v jednom prodaném košíku. Rozhodněte o typu náhodné veličiny.

2.12. Náhodným pokusem je prodej a nákup akcií na burze. Náhodnou veličinou $\mathbb{X}$ je cena jedné akcie konkrétní firmy na konci obchodního dne. Rozhodněte o typu náhodné veličiny.

2.1.2 Zákon rozdělení pravděpodobnosti diskrétní náhodné veličiny

V této kapitole se vrátíme k příkladu, který použila dr. Šimsová ve videopřednášce a budeme demonstrovat zákon rozdělení pravděpodobnosti diskrétní náhodné veličiny.

Již víme, že jevy, které tvoří rozklad množiny Ω, jsou navzájem neslučitelné. Pro pravděpodobnost jejich sjednocení platí

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Sjednocení všech jevů, které tvoří rozklad množiny Ω, je rovno přímo množině Ω a víme, že $P(\Omega) = 1$ (jedná se o jistý jev).

Věta 2.1.4. Zákon rozdělení pravděpodobnosti *diskrétní náhodné veličiny* $\mathbb{X}$ *říká*

$$\sum_{i \in I} P(A_i) = \sum_{x_i \in M} P(\mathbb{X} = x_i) = 1. \tag{2.1}$$

2.13. Náhodným pokusem je hod dvěma kostkami. Náhodnou veličinou $\mathbb{X}$ je počet šestek, které při hodu padly. Určeme možné hodnoty náhodné veličiny $\mathbb{X}$ a pravděpodobnostní funkci náhodné veličiny $\mathbb{X}$. Ověřme zákon rozdělení pravděpodobnosti diskrétní náhodné veličiny.

Řešení: Množinu Ω nyní znázorníme dalším možným způsobem – jako uspořádané dvojice čísel v Tabulce 2.3, kde první hodnota říká jaké číslo padlo na první kostce a druhá hodnota uvádí jaké číslo padlo na druhé kostce. Množina Ω se nyní rozpadla na tři části, které jsme barevně

1, 1	1, 2	1, 3	1, 4	1, 5	1, 6
2, 1	2, 2	2, 3	2, 4	2, 5	2, 6
3, 1	3, 2	3, 3	3, 4	3, 5	3, 6
4, 1	4, 2	4, 3	4, 4	4, 5	4, 6
5, 1	5, 2	5, 3	5, 4	5, 5	5, 6
6, 1	6, 2	6, 3	6, 4	6, 5	6, 6

Tabulka 2.3: Zobrazení výsledků hodu dvěma kostkami

zvýraznili. Červeným písmem jsou vyznačeny ty případy, kdy je $\mathbb{X} = 0$, zeleným písmem jsou vyznačeny ty případy, kdy je $\mathbb{X} = 1$ a modrým písmem je vyznačen případ $\mathbb{X} = 2$. Množina M možných hodnot náhodné veličiny má tři prvky a platí $M = \{0,\ 1,\ 2\}$. Vzhledem k počtu jednotlivých případů v Tabulce 2.3 snadno odvodíme následující pravděpodobnosti.

$A_1 \ldots$ nehozena ani jedna šestka $\ldots$ $\mathbb{X} = 0$ $\ldots$ $P(\mathbb{X} = 0) = 25/36$

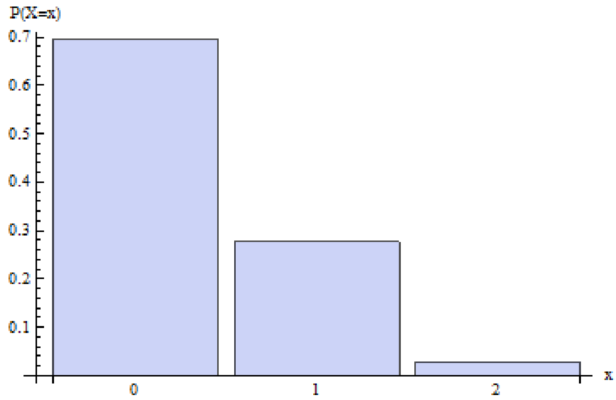
$A_2 \ldots$ hozena jedna šestka $\ldots$ $\mathbb{X} = 1$ $\ldots$ $P(\mathbb{X} = 1) = 10/36$

$A_3 \ldots$ hozeny dvě šestky $\ldots$ $\mathbb{X} = 2$ $\ldots$ $P(\mathbb{X} = 2) = 1/36$

Nyní ověříme, zda platí zákon rozdělení pravděpodobnosti (což je nadsázka – víme, že platí). Je

$$\sum_{x_i \in M} P(\mathbb{X} = x_i) = P(\mathbb{X} = 0) + P(\mathbb{X} = 1) + P(\mathbb{X} = 2) = \frac{25}{36} + \frac{10}{36} + \frac{1}{36} = \frac{36}{36} = 1.$$

Rozdělení pravděpodobnosti náhodné veličiny můžeme znázornit graficky, viz Obrázek 2.2. Na vodorovné ose leží hodnoty x, na svislé ose nanášíme pravděpodobnosti jednotlivých hodnot x. Zde je $\frac{25}{36} \doteq 0.69$, $\frac{10}{36} \doteq 0.28$ a $\frac{1}{36} \doteq 0.03$. Vzhledem k tomu, že šířka jednotlivých sloupců je rovna jedné, je součet obsahů vyznačených obdélníků roven jedné.



Obrázek 2.2: Rozdělení pravděpodobnosti mezi jednotlivé hodnoty náhodné veličiny $\mathbb{X}$

2.2 Charakteristiky náhodné veličiny

2.2.1 Modus (modální hodnota) náhodné veličiny

Definice 2.2.1. Nejpravděpodobnější hodnotu náhodné veličiny $\mathbb{X}$ nazýváme *modus* náhodné veličiny $\mathbb{X}$. Je to hodnota x náhodné veličiny $\mathbb{X}$, která má nejvyšší pravděpodobnost, že nastane.

2.14. Určíme modus náhodné veličiny uvedené v Příkladu 2.13.

Řešení: K určení modální hodnoty potřebujeme znát pravděpodobnosti jednotlivých hodnot náhodné veličiny. Tyto určíme z pravděpodobnostní funkce $P(\mathbb{X} = x)$, krátce $P(x)$. Víme, že platí

$$P(\mathbb{X} = x) = P(x) = \begin{cases} \dfrac{25}{36} & \dots \quad x = 0, \\ \dfrac{10}{36} & \dots \quad x = 1, \\ \dfrac{1}{36} & \dots \quad x = 2, \\ 0 & \dots \quad \text{pro zbývající hodnoty } x. \end{cases}$$

Porovnáním hodnot pravděpodobností vidíme, že nejvyšší pravděpodobnost nastává pro $x = 0$, proto je modální hodnotou náhodné veličiny číslo nula, tj. $\widehat{x} = 0$.

Úlohy k samostatnému řešení

2.15. Vypočtěte modální hodnotu pro náhodnou veličinu uvedenou v Příkladu 2.2 na straně 19.

2.16. Vypočtěte modální hodnotu pro náhodnou veličinu uvedenou v Příkladu 2.3 na straně 19.

2.17. Vypočtěte modální hodnotu pro náhodnou veličinu uvedenou v Příkladu 2.4 na straně 19.

2.2.2 Distribuční funkce náhodné veličiny

Definice 2.2.2. *Distribuční funkce* $F(x)$ náhodné veličiny $\mathbb{X}$ uvádí, jaká je pravděpodobnost, že náhodná veličina $\mathbb{X}$ nabude hodnotu nejvýše x, tj.

$$F(x) = P(\mathbb{X} \leq x).$$

2.18. Nalezneme předpis distribuční funkce pro náhodnou veličinu z Příkladu 2.13.

Řešení: Použijeme výsledky dané úlohy. Víme, že pravděpodobnost je nenulová pouze v hodnotách $x = 0$, $x = 1$ a $x = 2$. To znamená, že například v případu $F(1,4) = P(\mathbb{X} \leq 1,4)$ můžeme započítat pouze ty hodnoty $x \leq 1,4$, které mají nenulovou pravděpodobnost. To jsou hodnoty $x = 0$ a $x = 1$. Proto

$$F(1,4) = P(\mathbb{X} \leq 1,4) = P(\mathbb{X} = 0) + P(\mathbb{X} = 1) = \frac{25}{36} + \frac{10}{36} = \frac{35}{36}.$$

Mezi významné hodnoty distribuční funkce patří následující příklady.

$$\begin{aligned}
&F(x) = P(\mathbb{X} \leq x)\\
&F(0) = P(\mathbb{X} \leq 0) = P(0) = \frac{25}{36}\\
&F(1) = P(\mathbb{X} \leq 1) = P(0) + P(1) = \frac{25}{36} + \frac{10}{36} = \frac{35}{36}\\
&F(2) = P(\mathbb{X} \leq 2) = P(0) + P(1) + P(2) = \frac{25}{36} + \frac{10}{36} + \frac{1}{36} = \frac{36}{36} = 1
\end{aligned}$$

S uvážením výše uvedených poznámek dostáváme pro zadanou náhodnou veličinu $\mathbb{X}$ následující předpis distribuční funkce $F(x)$.

$$F(x) = \begin{cases} 0 & \ldots \quad x < 0 \\ \frac{25}{36} & \ldots \quad 0 \leq x < 1 \\ \frac{35}{36} & \ldots \quad 1 \leq x < 2 \\ 1 & \ldots \quad 2 \leq x \end{cases}$$

Úlohy k samostatnému řešení

2.19. Určete předpis distribuční funkce pro náhodnou veličinu uvedenou v Příkladu 2.2 na straně 19.

2.20. Určete předpis distribuční funkce pro náhodnou veličinu uvedenou v Příkladu 2.3 na straně 19.

2.21. Určete předpis distribuční funkce pro náhodnou veličinu uvedenou v Příkladu 2.4 na straně 19.

Následující úlohy jsou věnovány použití distribuční funkce diskrétní náhodné veličiny při řešení slovních úloh.

x	0	1	2	3	4	Σ
$P(\mathbb{X} = x)$	0.34	0.46	0.10	0.07	0.03	1

Tabulka 2.4: Rozdělení pravděpodobností k Příkladu 2.22

2.22. Náhodným pokusem je výběr osoby v určité oblasti. Náhodná veličina $\mathbb{X}$ představuje počet sourozenců této vybrané osoby. Pravděpodobnostní rozdělení $\mathbb{X}$ je uvedeno v Tabulce 2.4. Určete předpis distribuční funkce této náhodné veličiny a vypočtěte pravděpodobnost, že počet sourozenců náhodně vybrané osoby je:

a) menší než 3,

b) nejvýše roven 3,

c) alespoň 3,

d) větší než 1 a menší než 4.

Řešení: Ze zadání příkladu plyne, že distribuční funkce bude své hodnoty měnit pouze pro $x = 0$, $x = 1$, $x = 2$, $x = 3$ a $x = 4$, v ostatních hodnotách x je pravděpodobnost rovna nule. Proto například

$$F(2,7) = P(\mathbb{X} \leq 2,7) = P(\mathbb{X} = 0) + P(\mathbb{X} = 1) + P(\mathbb{X} = 2) = 0,34 + 0,46 + 0,10 = 0,90.$$

Pro předpis $F(x)$ tak platí:

$$F(x) = \begin{cases} 0 & \dots & x < 0 \\ 0.34 & \dots & 0 \leq x < 1 \\ 0.80 & \dots & 1 \leq x < 2 \\ 0.90 & \dots & 2 \leq x < 3 \quad \dots \quad \text{a proto } F(2,7) = 0,90 \\ 0.97 & \dots & 3 \leq x < 4 \\ 1 & \dots & 4 \leq x \end{cases}$$

Dále se budeme věnovat podúlohám a) až d).

a) Pravděpodobnost, že počet sourozenců $\mathbb{X}$ náhodně vybrané osoby je menší než 3 znamená $P(\mathbb{X} < 3) = P(0) + P(1) + P(2) = 0,34 + 0,46 + 0,10 = 0,90$.

b) Pro pravděpodobnost, že počet sourozenců $\mathbb{X}$ je nejvýše roven 3, platí následující rovnosti $P(\mathbb{X} \leq 3) = P(0) + P(1) + P(2) + P(3) = 0,34 + 0,46 + 0,10 + 0,07 = 0,97$.

c) Pro pravděpodobnost, že počet sourozenců $\mathbb{X}$ je roven alespoň 3, platí následující rovnosti $P(\mathbb{X} \geq 3) = P(3) + P(4) = 0,07 + 0,03 = 0,10$.

d) Pro pravděpodobnost, že počet sourozenců $\mathbb{X}$ je větší než 1 a současně menší než 4, platí $P(1 < \mathbb{X} < 4) = P(2) + P(3) = 0,10 + 0,07 = 0,17$.

2.23. Zjistěte, zda funkce $P(x)$ může být pravděpodobnostní funkcí náhodné veličiny $\mathbb{X}$, kde

a) $P(x) = 1/4$ pro $x = 0, 1, 2, 3,$

b) $P(x) = 1/3$ pro $x = 0, 1, 2, 3,$

c) $P(x) = x/4$ pro $x = 0, 1, 3,$

d) $P(x) = (x - 5)/10$ pro $x = 0, 5, 10, 15,$

e) $P(x) = x^2/10$ pro $x = -1, 0, 3.$

Řešení: Při hledání odpovědí si musíme uvědomit, jaké podmínky musí splňovat funkce, aby mohla být pravděpodobnostní funkcí náhodné veličiny.

- Pravděpodobnosti všech hodnot veličiny $\mathbb{X}$ musí ležet v rozmezí od nuly (včetně) do jedné (včetně), tj. $0 \leq P(x) \leq 1$ – viz definice pojmu pravděpodobnost.
- Součet pravděpodobností všech hodnot musí být roven jedné – viz zákon rozdělení pravděpodobností náhodné veličiny. Zde stačí sečíst pouze pravděpodobnosti, které jsou nenulové – ty nulové do součtu ničím nepřispějí.

V dalších krocích proto budeme ověřovat tyto podmínky a pokud je předpis funkce takový, že s těmito podmínkami není v rozporu, budeme jej považovat za možnou pravděpodobnostní funkci.

a) Předpis funkce lze „přeložit do češtiny“ tímto způsobem. Možné hodnoty náhodné veličiny jsou pouze čísla $x = 0$, $x = 1$, $x = 2$ a $x = 3$ a pravděpodobnost každé této hodnoty je rovna $\frac{1}{4}$. Pravděpodobnosti pro ostatní hodnoty x jsou rovny nule. První podmínka je splněna. Pro ověření druhé podmínky sečteme všechny nenulové pravděpodobnosti. Je $P(0) + P(1) + P(2) + P(3) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$. Druhá podmínka je také splněna a uvedená funkce může být pravděpodobnostní funkcí.

b) Nyní jsou možné hodnoty náhodné veličiny opět pouze čísla $x = 0$, $x = 1$, $x = 2$ a $x = 3$ a pravděpodobnost každé této hodnoty je rovna $\frac{1}{3}$, v ostatních případech je pravděpodobnost rovna nule. První podmínka je splněna – hodnoty pravděpodobností jsou v rozmězí od nuly do jedné. Druhou podmínku ověříme sečtením nenulových pravděpodobností. Je

$$\sum_{x \in M} P(x) = P(0) + P(1) + P(2) + P(3) = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = \frac{4}{3} \neq 1.$$

Součet pravděpodobností vyšel vyšší než jedna, druhá podmínka je porušena, a funkce proto nemůže být pravděpodobnostní funkcí náhodné veličiny.

c) Nyní jsou možné hodnoty náhodné veličiny pouze čísla $x = 0$, $x = 1$ a $x = 3$. Pravděpodobnost každé této hodnoty závisí na hodnotě x podle vzorce $P(x) = \frac{x}{4}$. Proto je $P(0) = \frac{0}{4} = 0$, $P(1) = \frac{1}{4}$ a $P(3) = \frac{3}{4}$, v ostatních případech je pravděpodobnost rovna nule (a pro $x = 0$ vlastně také). První podmínka je splněna – pravděpodobnosti všech hodnot x leží v rozmezí od nuly do jedné. K ověření druhé podmínky sečteme nenulové pravděpodobnosti. Je

$$\sum_{x \in M} P(x) = P(1) + P(3) = \frac{1}{4} + \frac{3}{4} = 1.$$

Druhá podmínka je také splněna, funkce může být pravděpodobnostní funkcí náhodné veličiny.

d) Možné hodnoty náhodné veličiny jsou čísla $x = 0$, $x = 5$, $x = 10$ a $x = 15$. Pravděpodobnost každé této hodnoty závisí na hodnotě x podle vzorce $P(x) = \frac{x-5}{10}$. Proto je $P(0) = \frac{0-5}{10} = -\frac{1}{2}$, $P(5) = \frac{5-5}{10} = 0$, $P(10) = \frac{10-5}{10} = \frac{1}{2}$ a $P(15) = \frac{15-5}{10} = 1$, v ostatních případech (a pro $x = 5$) je pravděpodobnost rovna nule. Pravděpodobnost hodnoty $x = 0$ nám vyšla záporná, je $P(0) = -\frac{1}{2}$, proto první podmínka není splněna a funkce nemůže být funkcí pravděpodobnosti. Toto platí dokonce přesto, že součet pravděpodobností je roven jedné, viz
$$\sum_{x \in M} P(x) = P(0) + P(10) + P(15) = \frac{-5}{10} + \frac{5}{10} + \frac{10}{10} = \frac{10}{10} = 1.$$
Funkce proto není pravděpodobnostní funkcí.

e) Možné hodnoty náhodné veličiny jsou čísla $x = -1$, $x = 0$ a $x = 3$. To, že jedna z hodnot je záporná, nevadí – problémem by bylo, kdyby její pravděpodobnost vycházela záporná. Dosazením do vzorce pro pravděpodobnost dostaneme $P(-1) = \frac{(-1)^2}{10} = \frac{1}{10}$, $P(0) = \frac{0^2}{10} = 0$ a $P(3) = \frac{3^2}{10} = \frac{9}{10}$; pravděpodobnosti zbývajících hodnot jsou rovny nule. Pravděpodobnosti všech hodnot jsou v rozmezí od nuly do jedné, proto je první podmínka splněna. Druhou podmínku ověříme součtem nenulových pravděpodobností. Je
$$\sum_{x \in M} P(x) = P(-1) + P(3) = \frac{1}{10} + \frac{9}{10} = 1.$$
Součet pravděpodobností je roven jedné, druhá podmínka je splněna a funkce může být pravděpodobnostní funkcí náhodné veličiny.

2.2.3 Momentové charakteristiky náhodné veličiny

V této části si připomeneme a procvičíme pojmy obecný moment k-tého řádu a centrovaný moment k-tého řádu (viz videopřednáška dr. Šimsové Diskrétní náhodná veličina 2 od času cca 5:22).

Definice 2.2.3. Pro diskrétní náhodnou veličinu $\mathbb{X}$ s možnými hodnotami $x_1, x_2, \ldots, x_n$ a jejich pravděpodobnostmi $p_i = P(x_i) = P(\mathbb{X} = x_i)$ definujeme k-tý obecný moment vztahem
$$m_0^k = \sum_{x_i \in M} x_i^k \cdot P(x_i).$$

Velmi používaným obecným momentem je *první obecný moment*, kterému častěji říkáme *střední hodnota* a místo označení m_0^1 používáme k jeho označení symbol μ. Setkat se také můžeme s označením $E(\mathbb{X})$ a názvem *očekávaná hodnota* (z anglického výrazu *expected value*), viz následující definice.

Definice 2.2.4. Pro diskrétní náhodnou veličinu $\mathbb{X}$ s možnými hodnotami $x_1, x_2, \ldots, x_n$ a jejich pravděpodobnostmi $p_i = P(x_i) = P(\mathbb{X} = x_i)$ definujeme střední hodnotu $E(\mathbb{X})$ vztahem
$$E(\mathbb{X}) = \mu = \sum_{x_i \in M} x_i \cdot P(x_i). \tag{2.2}$$

Střední hodnotu náhodné veličiny $\mathbb{X}$ lze chápat jako teoretickou hodnotu, která předpovídá průměr z realizací náhodné veličiny při dostatečně vysokém počtu opakování příslušného náhodného pokusu – v tom smyslu, že čím více pokusů proběhlo, tím více se průměr z realizací náhodné veličiny blíží střední hodnotě této veličiny. Tuto poznámku upřesníme po výkladu tzv. *limitních vět* a jejich souvislostí v některé z následujících přednášek.

2.24. Náhodným pokusem je hod *spravedlivou* kostkou. Náhodnou veličinou $\mathbb{X}$ je číslo, které na kostce padlo při tomto hodu. Vypočtěte střední hodnotu náhodné veličiny $\mathbb{X}$.

Řešení: Možné hodnoty $\mathbb{X}$ jsou čísla padlá na kostce, je tedy množina všech hodnot náhodné veličiny $\mathbb{X}$ rovna $M = \{1, 2, 3, 4, 5, 6\}$. Předpokládáme, že kostka je *spravedlivá*, tj. všechny hodnoty padají se stejnou pravděpodobností, a platí

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}.$$

Dosazením do vzorce (2.2) dostaneme

$$\begin{aligned} E(\mathbb{X}) = \mu = \sum_{x_i \in M} x_i \cdot P(x_i) &= x_1 \cdot P(x_1) + x_2 \cdot P(x_2) + \ldots + x_6 \cdot P(x_6) \\ &= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \\ &= \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} \\ &= \frac{21}{6} = \frac{7}{2} = 3,5. \end{aligned}$$

Střední hodnota vyšla $\mu = 3,5$, což je v souladu s naším očekáváním. Při dostatečném počtu opakování hodů totiž můžeme očekávat přibližně stejný počet jedniček, dvojek, trojek, ..., šestek a průměr z těchto hodnot bude blízký číslu 3,5 (vypočtěte si průměr z hodnot $1, 2, 3, 4, 5, 6$).

Podobně jako pro první obecný moment náhodné veličiny máme speciální označení i pro druhý obecný moment náhodné veličiny, viz definice .

Definice 2.2.5. Pro diskrétní náhodnou veličinu $\mathbb{X}$ s možnými hodnotami $x_1, x_2, \ldots, x_n$ a jejich pravděpodobnostmi $p_i = P(x_i) = P(\mathbb{X} = x_i)$ definujeme druhý obecný moment $E(\mathbb{X}^2)$ vztahem

$$E(\mathbb{X}^2) = \sum x_i^2 \cdot P(x_i). \tag{2.3}$$

Kromě obecných momentů pracujeme i s tzv. centrovanými momenty náhodné veličiny. Připomeneme obecnou definici centrovaného momentu k-tého řádu a pak se budeme věnovat nejdůležitějšímu z centrovaných momentů – centrovanému momentu druhého řádu – tzv. *rozptylu*, pro který používáme označení $D(\mathbb{X})$ nebo σ^2.

Definice 2.2.6. Pro diskrétní náhodnou veličinu $\mathbb{X}$ s možnými hodnotami $x_1, x_2, \ldots, x_n$ a jejich pravděpodobnosti $p_i = P(x_i) = P(\mathbb{X} = x_i)$ definujeme k-tý centrovaný moment vztahem

$$m_c^k = \sum_{x_i \in M} \left[x_i - E(\mathbb{X})\right]^k \cdot P(x_i) = E\Big(\left[\mathbb{X} - E(\mathbb{X})\right]^k\Big). \tag{2.4}$$

Z Definice 2.2.6 plyne, že při výpočtu centrovaných momentů již musíme znát střední hodnotu $E(\mathbb{X})$. Každou hodnotu x pak snížíme o hodnotu $E(\mathbb{X})$, tento rozdíl umocníme na k-tou mocninu, tuto mocninu vynásobíme pravděpodobností hodnoty x. Tyto operace provedeme se všemi hodnotami x a výsledky sečteme. Konkrétní způsob výpočtu si ukážeme na případě druhého centrovaného momentu – rozptylu.

Definice 2.2.7. Pro diskrétní náhodnou veličinu $\mathbb{X}$ s možnými hodnotami $x_1, x_2, \ldots, x_n$ a jejich pravděpodobnosti $p_i = P(x_i) = P(\mathbb{X} = x_i)$ definujeme rozptyl $D(\mathbb{X})$, resp. σ^2 vztahem

$$m_c^2 = D(\mathbb{X}) = \sigma^2 = \sum_{x_i \in M} \big[x_i - E(\mathbb{X})\big]^2 \cdot P(x_i) = E\Big(\big[\mathbb{X} - E(\mathbb{X})\big]^2\Big). \tag{2.5}$$

2.25. Náhodným pokusem je hod *spravedlivou* kostkou. Náhodnou veličinou $\mathbb{X}$ je číslo, které na kostce padlo při tomto hodu. Vypočtěte rozptyl zadané náhodné veličiny.

Řešení: Navážeme na řešení Příkladu 2.24, ve kterém jsme pro tuto náhodnou veličinu vypočítali střední hodnotu $E(\mathbb{X}) = 3,5$. Ve vzorci (2.5) tak místo výrazu $E(\mathbb{X})$ můžeme uvažovat hodnotu $3,5$. Tím dostaneme

$$\begin{aligned}
D(\mathbb{X}) = \sigma^2 &= \sum_{x_i \in M} \big[x_i - E(\mathbb{X})\big]^2 \cdot P(x_i) \\
&= \sum_{x_i \in M} \big[x_i - 3,5)\big]^2 \cdot P(x_i) \\
&= (1-3,5)^2 \cdot \frac{1}{6} + (2-3,5)^2 \cdot \frac{1}{6} + (3-3,5)^2 \cdot \frac{1}{6} + (4-3,5)^2 \cdot \frac{1}{6} + \\
&\quad + (5-3,5)^2 \cdot \frac{1}{6} + (6-3,5)^2 \cdot \frac{1}{6} \\
&= (-2,5)^2 \cdot \frac{1}{6} + (-1,5)^2 \cdot \frac{1}{6} + (-0,5)^2 \cdot \frac{1}{6} + (0,5)^2 \cdot \frac{1}{6} + (1,5)^2 \cdot \frac{1}{6} + (2,5)^2 \cdot \frac{1}{6} \\
&= 6,25 \cdot \frac{1}{6} + 2,25 \cdot \frac{1}{6} + 0,25 \cdot \frac{1}{6} + 0,25 \cdot \frac{1}{6} + 2,25 \cdot \frac{1}{6} + 6,25 \cdot \frac{1}{6} \\
&= \frac{6,25}{6} + \frac{2,25}{6} + \frac{0,25}{6} + \frac{0,25}{6} + \frac{2,25}{6} + \frac{6,25}{6} = \frac{17,5}{6} = \frac{35}{12} = 2,91\overline{6} \doteq 2,92.
\end{aligned}$$

Výpočet rozptylu lze zjednodušit použitím následující věty.

Věta 2.2.8. *Pro diskrétní náhodnou veličinu $\mathbb{X}$ se střední hodnotou $E(\mathbb{X})$ a druhým obecným momentem $E(\mathbb{X}^2)$ lze rozptyl $D(\mathbb{X})$ vypočítat pomocí vztahu*

$$D(\mathbb{X}) = \sigma^2 = E(\mathbb{X}^2) - [E(\mathbb{X})]^2. \tag{2.6}$$

Použití Věty 2.2.8 si ukážeme novým výpočtem rozptylu náhodné veličiny ze zadání z Příkladu 2.25.

2.26. Náhodným pokusem je hod *spravedlivou* kostkou. Náhodnou veličinou $\mathbb{X}$ je číslo, které na kostce padlo při tomto hodu. Vypočtěte rozptyl zadané náhodné veličiny pomocí vzorce (2.6).

Řešení: Již víme, že zadaná náhodná veličina má střední hodnotu $E(\mathbb{X}) = 3,5$. Pro dosazení do vzorce (2.6) budeme potřebovat druhý obecný moment. Podle vzorce (2.3) platí

$$\begin{aligned}
E(\mathbb{X}^2) &= \sum_{x_i \in M} x_i^2 \cdot P(x_i) \\
&= x_1^2 \cdot P(x_1) + x_2^2 \cdot P(x_2) + x_3^2 \cdot P(x_3) + x_4^2 \cdot P(x_4) + x_5^2 \cdot P(x_5) + x_6^2 \cdot P(x_6) \\
&= 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} \\
&= \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + \frac{36}{6} = \frac{91}{6}.
\end{aligned}$$

Nyní již víme, že $E(\mathbb{X}) = 3,5 = \frac{7}{2} = \frac{21}{6}$ a $E(\mathbb{X}^2) = \frac{91}{6}$ a můžeme dosadit do vzorce (2.6).

$$D(\mathbb{X}) = \sigma^2 = E(\mathbb{X}^2) - [E(\mathbb{X})]^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{546}{36} - \frac{441}{36} = \frac{105}{36} = \frac{35}{12}.$$

Výsledek $D(\mathbb{X})$ vyšel v obou případech (Příklad 2.25 i Příklad 2.26) stejně a hodnota rozptylu je rovna $D(\mathbb{X}) = \frac{35}{12}$.

Další významnou statistickou veličinou je *směrodatná odchylka* náhodné veličiny. Její definice vychází z veličiny rozptyl, viz následující definice.

Definice 2.2.9. *Směrodatná odchylka* σ náhodné veličiny $\mathbb{X}$ je definována jako odmocnina z rozptylu této náhodné veličiny, tedy

$$\sigma = \sqrt{D(\mathbb{X})}. \tag{2.7}$$

2.27. Náhodným pokusem je hod *spravedlivou* kostkou. Náhodnou veličinou $\mathbb{X}$ je číslo, které na kostce padlo při tomto hodu. Vypočtěte směrodatnou odchylku zadané náhodné veličiny.

Řešení: Vyjdeme z již známých výsledků. Víme, že náhodná veličina $\mathbb{X}$ má rozptyl $D(\mathbb{X}) = \frac{35}{12}$. Použitím vzorce (2.7) dostaneme

$$\sigma = \sqrt{D(\mathbb{X})} = \sqrt{\frac{35}{12}} = \sqrt{2,91\overline{6}} \doteq 1,71.$$

Směrodatná odchylka náhodné veličiny $\mathbb{X}$ je rovna $\sigma \doteq 1,71$.

2.28. Obsluha nápojového automatu potřebuje zjistit, jak často má doplňovat jistý druh nápoje v daném automatu. Na základě minulých údajů o prodeji odhadla obsluha automatu následující pravděpodobnostní rozdělení veličiny $\mathbb{X}$, kde náhodná veličina $\mathbb{X}$ znamená počet lahví daného

x	15	16	17	18	19	20	$\sum$
$P(x)$	0.10	0.15	0.30	0.20	0.15	0.10	1

Tabulka 2.5: Pravděpodobnostní funkce náhodné veličiny $\mathbb{X}$: počet prodaných lahví denně

nápoje prodaných za jeden den, viz Tabulka 2.5. Vypočtěte střední hodnotu počtu prodaných lahví nápoje za jeden den a rozptyl této veličiny.

Řešení: Nejprve vypočteme střední hodnotu. Jsou známy hodnoty x i jejich pravděpodobnosti $P(x)$, proto můžeme přímo dosadit do vzorce (2.2).

$$\begin{aligned} E(\mathbb{X}) &= \sum_{x_i \in M} x_i \cdot P(x_i) \\ &= 15 \cdot 0,10 + 16 \cdot 0,15 + 17 \cdot 0,30 + 18 \cdot 0,20 + 19 \cdot 0,15 + 20 \cdot 0,10 \\ &= 17,45 \end{aligned}$$

Střední hodnota počtu prodaných lahví denně vyšla rovna 17,45 lahví/den.

Rozptyl náhodné veličiny vypočteme ze vzorce (2.5). K tomu potřebujeme vypočítat hodnotu $E(\mathbb{X}^2)$. K výpočtu $E(\mathbb{X}^2)$ použijeme rovnici (2.3). Tím dostaneme

$$\begin{aligned} E(\mathbb{X}^2) &= \sum_{x_i \in M} x_i^2 \cdot P(x_i) \\ &= 15^2 \cdot 0,10 + 16^2 \cdot 0,15 + 17^2 \cdot 0,30 + 18^2 \cdot 0,20 + 19^2 \cdot 0,15 + 20^2 \cdot 0,10 \\ &= 306,55 \end{aligned}$$

Nyní již můžeme dosadit do vzorce (2.5). Je

$$D(\mathbb{X}) = E(\mathbb{X}^2) - [E(\mathbb{X})]^2 = 306.55 - (17.45)^2 = 2,0475.$$

náhodné veličiny $\mathbb{X}$ je roven přibližně $D(\mathbb{X}) = 2,05$.

Ze známého rozptylu náhodné veličiny můžeme vypočítat také směrodatnou odchylku náhodné veličiny

$$\sigma = \sqrt{D(\mathbb{X})} = \sqrt{2,0475} \doteq 1,431.$$

Směrodatná odchylka náhodné veličiny $\mathbb{X}$ je rovna přibližně $\sigma = 1,43$.

Komentář k výsledkům úlohy Můžeme si samozřejmě klást otázku, k čemu nám vypočtené hodnoty „jsou“, tj. jaké je použití hodnot $E(\mathbb{X})$, $D(\mathbb{X})$, resp. σ.

Význam střední hodnoty $E(\mathbb{X})$, resp. μ, spočívá v odhadu, jaké je průměrné množství prodaných lahví za den. Pak můžeme přibližně odvozovat počty prodaných lahví za dva dny ($2 \cdot \mu$), za týden ($7 \cdot \mu$), měsíc ($30 \cdot \mu$) atd. Odtud pravděpodobně plyne ekvivalentní název *očekávaná hodnota*. O tom, jak blízko budou tyto odhady skutečným realizacím náhodné veličiny nám podávají informaci hodnoty $D(\mathbb{X})$, resp. σ. Asi tušíte, že čím nižší budou obě hodnoty $D(\mathbb{X})$ a σ, tím více se skutečné realizace náhodné veličiny za dva dny, za týden, měsíc atd. budou blížit vypočteným hodnotám. Více se k tématu dozvíte po výkladu tzv. limitních vět a jejich souvislostí.

2.29. Předpokládejme hypotetickou situaci, při které hráč sází při ruletě 100 Kč stále na černou barvu. Náhodná veličina $\mathbb{X}$ představuje velikost jeho výhry v každé kole hry. Vypočtěte střední hodnotu $E(\mathbb{X})$, rozptyl $D(\mathbb{X})$ a směrodatnou odchylku σ této náhodné veličiny. Interpretujte výslednou hodnotu střední hodnoty této náhodné veličiny.

Řešení: Připomeňme relevantní informace k hře *ruleta*. Hru řídí krupiér, který roztočí hrací kolo s očíslovanými přihrádkami a v protisměru otáčení kola do něj vhodí hrací kuličku. Ta obíhá po obvodu kola až se její rychlost třením sníží natolik, že kulička spadne do některé z očíslovaných přihrádek v prostřední části hracího kola. Těchto přihrádek je (v tzv. evropské verzi rulety) celkem 37 a jsou očíslovány čísly od 0 do 36. Přihrádka s číslem 0 má zelenou barvu. Přihrádky s čísly od 1 do 36 mají víceméně střídavě červenou a černou barvu; podle toho, jakou barvu má přihrádka, do které padla kulička, říkáme, že padla červená, černá nebo zelená barva. Upřesněme, že červených přihrádek je 18, černých přihrádek je také 18.

Jednou ze sázek, které hráč provádí, může být sázka na barvu, která padne. V takovém případě hráč položí na hrací plán obnos symbolizující výši sázky na pole s červenou nebo černou barvou (poznámka: na zelenou barvu se nesází – této situaci odpovídá sázka na padnutí čísla 0). Pokud hráč uhodne barvu, krupiér mu vrátí jeho vklad a navíc přidá výhru ve výši hráčova vkladu. Pokud hráč barvu neuhodl, krupiér si vezme z hracího plánu hráčův vklad a věnuje mu maximálně lítostivý úsměv. Ve skutečnosti navíc hra probíhá tak, že hráči místo s penězi hrají pomocí žetonů s jistou peněžní hodnotou a krupiér má profesionálně neutrální výraz ve tváři.

Nyní již máme dostatek informací k tomu, abychom mohli úlohu řešit. Připomeňme, že náhodná veličina $\mathbb{X}$ představuje velikost výhry hráče v každé kole hry.

Uvažujme situaci, kdy hráč má před zahájením hry hotovost ve výši 1 000 Kč. Vsadí 100 Kč, v tu chvíli má u sebe hotovost ve výši 900 Kč. Pokud vyhraje, krupiér mu vrátí vklad 100 Kč a navrch přidá dalších 100 Kč jako výhru. V takové situaci má hráč po vyplacení sázek hotovost ve výši 1 100 Kč a jeho výhra činí 100 Kč. Pokud hráč prohraje, krupiér si vezme jeho sázku a nic mu nedá. V takové situaci má hráč po vyplacení sázek hotovost 900 Kč a jeho „výhra“ činí -100 Kč.

Možné hodnoty náhodné veličiny $\mathbb{X}$ proto jsou -100, nebo 100. Předpokládáme, že ruleta je spravedlivá, tj. každé číslo má stejnou šanci na padnutí. Hráč vyhraje 100 Kč, pokud padne některé z 18 černých čísel z celkového počtu 37 čísel. Pravděpodobnost výhry je rovna

$$P(\mathbb{X} = 100) = \frac{18}{37}.$$

Hráč prohraje, pokud mu padne některé ze zbývajících 19 čísel (1 zelené a 18 červených). Pravděpodobnost prohry, tj. výhry -100 Kč, je

$$P(\mathbb{X} = -100) = \frac{19}{37}.$$

Pravděpodobnostní funkce tedy má předpis

$$P(x) = \begin{cases} \frac{19}{37}, & \ldots \quad x = -100, \\ \frac{18}{37}, & \ldots \quad x = 100. \end{cases}$$

Známe-li pravděpodobnostní funkci náhodné veličiny, můžeme vypočítat její charakteristiky.

$$\begin{aligned} E(\mathbb{X}) &= \sum_{x_i \in M} x_i \cdot P(x_i) = x_1 \cdot P(x_1) + x_2 \cdot P(x_2) = (-100) \cdot \frac{19}{37} + 100 \cdot \frac{18}{37} \\ &= -\frac{1900}{37} + \frac{1800}{37} = -\frac{100}{37} \doteq -2,7. \end{aligned}$$

Střední hodnota (očekávaná hodnota) nám vyšla $E(\mathbb{X}) = -2,7$. Tuto hodnotu můžeme interpretovat jako částku, kterou hráč průměrně v každém kole hry „vyhrává“. Jinými slovy, hráč průměrně při každém kole hry ztrácí 2,7 Kč. Další výpočty jsou následující.

$$E(\mathbb{X}^2) = \sum_{x_i \in M} x_i^2 \cdot P(x_i) = (-100)^2 \cdot \frac{19}{37} + 100^2 \cdot \frac{18}{37} = \frac{190\,000}{37} + \frac{180\,000}{37} = \frac{370\,000}{37} = 10\,000$$

$$D(\mathbb{X}) = E(\mathbb{X}^2) - \big[E(\mathbb{X})\big]^2 = 10\,000 - (-2,7)^2 = 10\,000 - 7,29 = 9\,992,71$$

$$\sigma = \sqrt{D(\mathbb{X})} = \sqrt{9\,992,71} \doteq 99,96$$

Směrodatná odchylka vyšla násobně větší než je střední hodnota. To ukazuje na značnou variabilitu možných výsledků a díky tomu se může stát, že občas některý hráč „skončí v plusu“. Je nutné však mít na paměti, že častěji vyhrává kasino.

2.2.4 Shrnutí kapitoly - úlohy k samostatnému řešení

2.30. Určete, které z následujících náhodných veličin jsou diskrétní, resp. spojité.

- Počet vadných výrobků v balení 24 ks.
- Počet bodů, které získá FK Teplice na konci soutěžního ročníku.
- Doba trvání cvičení ze statistiky.
- Počet zákazníků, kteří přijdou k pokladně v supermarketu během jedné hodiny.
- Hmotnost dataprojektoru v této místnosti.
- Počet správně určených odpovědí v tomto cvičení.
- Doba strávená čekáním „na zelenou u semaforu“ od doby zastavení u semaforu po naskočení zelené na semaforu.

2.31. Náhodná veličina $\mathbb{X}$ představuje počet padlých šestek při hodu třemi kostkami. Určete pravděpodobnostní funkci a distribuční funkci této náhodné veličiny.

2.32. Náhodná veličina $\mathbb{X}$ představuje počet padlých šestek při hodu třemi kostkami. Stanovte střední hodnotu $E(\mathbb{X})$, rozptyl $D(\mathbb{X})$ a směrodatnou odchylku σ této náhodné veličiny $\mathbb{X}$.

2.33. Náhodná veličina $\mathbb{X}$ představuje počet novinových titulů, které si náhodně vybraný zákazník zakoupí v novinovém stánku. Pravděpodobnostní rozdělení $\mathbb{X}$ je uvedeno v tabulce. Určete pravděpodobnost, že počet zakoupených titulů u náhodně vybraného zákazníka je

x	0	1	2	3	4
$P(x)$	0.3	0.5	0.15	0.1	0.05

- větší než dva,
- nejvýše roven dvěma,
- alespoň dva,

- větší než jedna a menší než tři.

2.34. Vypočtěte střední hodnotu $E(\mathbb{X})$, rozptyl $D(\mathbb{X})$ a směrodatnou odchylku σ náhodné veličiny z Příkladu 2.32.

2.35. Náhodná veličina představuje počet hodů kostkou do doby, než padne šestka. Zapište pravděpodobnostní funkci této veličiny a vypočtěte střední hodnotu $E(\mathbb{X})$, rozptyl $D(\mathbb{X})$ a směrodatnou odchylku σ této náhodné veličiny.

2.3 Důležitá rozdělení diskrétní náhodné veličiny

2.3.1 Binomické rozdělení

Nyní se budeme věnovat jednomu z nejdůležitějších typů diskrétní náhodné veličiny – náhodné veličině s *binomickým rozdělením* pravděpodobnosti. Před vyslovením definice uvedeme několik ilustračních příkladů, které by měly pomoci rozeznat tento typ náhodné veličiny.

Na úvod uvažujme následující situaci. Náhodným pokusem je hod kostkou a náhodný jev A nastane, pokud při tomto hodu padne šestka. Snadno odvodíme, že pravděpodobnost jevu je rovna $P(A) = \frac{1}{6}$. Nyní provedeme 10 hodů kostkou a při každém hodu sledujeme, zda nastal jev A, resp. kolikrát jev A během oněch 10 hodů nastal, tj. kolik šestek padlo během deseti hodů kostkou. Může se stát, že šestka nepadla ani jednou (tj. počet případů, kdy nastal jev A je roven nule), může se také stát, že šestka během deseti pokusů padla například dvakrát (pak je počet případů, kdy nastal jev A roven dvěma). Pro těchto 10 hodů kostkou tak můžeme zavést náhodnou veličinu $\mathbb{X}$, jejíž hodnoty odpovídají počtu padlých šestek při deseti hodech, tj. odpovídají počtu případů, ve kterých nastal jev A. Možné hodnoty této náhodné veličiny jsou $M = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

Podobně můžeme uvažovat situaci, při které házíme mincí a sledujeme, zda padne rub či líc. Pokud padne rub, řekneme, že nastal jev A. Předpokládejme, že rub i líc mají stejnou šanci padnout, pak je pravděpodobnost padnutí rubu při jednom hodu rovna $P(A) = \frac{1}{2}$. Nyní budeme mincí házet dvacetkrát a budeme sledovat, kolikrát nastal jev A, tj. kolikrát padl rub. Opět můžeme zavést náhodnou veličinu $\mathbb{X}$, jejíž hodnoty říkají, kolikrát nastal jev A v sérii 20 náhodných pokusů, tj. kolik rubů padlo při dvaceti hodech mincí. Možné hodnoty této náhodné veličiny jsou $M = \{0, 1, 2, 3, 4, 5, \ldots, 19, 20\}$.

Předpokládejme, že jistý biatlonista při každém výstřelu zasáhne terč s pravděpodobností $p = 0,9$; tj. dlouhodobě má úspěšnost zásahu terče 90 %. Tento biatlonista přijede na střelnici a střílí pětkrát na terč. Sledujeme, kolikrát se při těchto pěti pokusech trefil do terče. Náhodnou veličinou bude počet úspěšných zásahů při pěti výstřelech na terč. Možné hodnoty této náhodné veličiny jsou $M = \{0, 1, 2, 3, 4, 5\}$.

Všechny uvedené situace mají společný rys, který popíšeme v následující definici.

Definice 2.3.1. Předpokládejme, že při náhodném pokusu sledujeme jistý náhodný jev A, jehož pravděpodobnost označíme symbolem p. Je tedy $P(A) = p$. Uskutečníme-li tento náhodný pokus n-krát, pak počet případů, kdy nastane jev A je náhodnou veličinou $\mathbb{X}$ s tzv. *binomickým rozdělením pravděpodobnosti.* Tuto náhodnou veličinu označujeme symbolem $\mathrm{Bi}(n, p)$, kde n a p jsou parametry tohoto rozdělení (n představuje počet opakování pokusu, p je pravděpodobnost nastání jevu A při jednom opakování pokusu).

Možné hodnoty náhodné veličiny jsou $M = \{0, 1, 2, \ldots, n\}$. Pravděpodobnost, že náhodná věličina $\mathbb{X}$ nabude hodnotu x, tj. pravděpodobnost, že při n opakováních pokusu nastane jev A právě x-krát, je rovna:

$$P(\mathbb{X} = x) = P(x) = \binom{n}{x} \cdot p^x (1-p)^{n-x}. \tag{2.8}$$

Pro střední hodnotu a rozptyl náhodné veličiny s binomickým rozdělením pravděpodobnosti platí:

$$E(\mathbb{X}) = n \cdot p, \qquad D(\mathbb{X}) = n \cdot p \cdot (1-p). \tag{2.9}$$

2.36. Uvažujme běžný balíček s 32 hracími kartami (kde jsou 4 esa.) Z balíčku vybereme kartu, podíváme se, zda je to eso, a pak kartu vrátíme do balíčku a balíček poctivě zamícháme. Takto vybereme a vrátíme kartu celkem pětkrát. Náhodnou veličinou $\mathbb{X}$ je počet případů, kdy vybranou kartou bylo eso. Vypočtěte hodnoty

a) $P(\mathbb{X} = 0) = P(0)$

b) $P(\mathbb{X} = 2) = P(2)$

Dále vypočtěte střední hodnotu $E(\mathbb{X})$ a rozptyl $D(\mathbb{X})$.

Řešení: Ze zadání plyne, že $\mathbb{X}$ je náhodná veličina s binomickým rozdělením pravděpodobnosti – jev A nastane, pokud vytáhneme eso; náhodný pokus opakujeme celkem pětkrát a pravděpodobnost jevu A je při každém opakování výběru stejná (jedná se o *výběr s vracením*). Je tedy

$$\begin{aligned} A \quad & \ldots \quad \text{náhodný jev } \textit{výběr esa} \\ \mathbb{X} \quad & \ldots \quad \text{náhodná veličina měřící } \textit{kolikrát nastal jev } A \\ n = 5 \quad & \ldots \quad \text{počet opakování pokusu} \\ P(A) = p = \tfrac{4}{32} = \tfrac{1}{8} \quad & \ldots \quad \text{pravděpodobnost výběru esa při každém tahu karty} \end{aligned}$$

Dosazením do rovnice (2.8) dostaneme

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

Pro případ $x=0$

$$\begin{aligned}P(\mathbb{X}=0)=P(0)&=\binom{5}{0}\cdot\left(\frac{1}{8}\right)^0\left(1-\frac{1}{8}\right)^{5-0}\\&=\frac{5!}{0!(5-0)!}\cdot\left(\frac{1}{8}\right)^0\left(\frac{7}{8}\right)^5\\&=1\cdot 1\cdot 0,875^5\\&\doteq 0,513\end{aligned}$$

Pro případ $x=2$

$$\begin{aligned}P(\mathbb{X}=2)=P(2)&=\binom{5}{2}\cdot\left(\frac{1}{8}\right)^2\left(1-\frac{1}{8}\right)^{5-2}\\&=\frac{5!}{2!(5-2)!}\cdot\left(\frac{1}{8}\right)^2\left(\frac{7}{8}\right)^3\\&=\frac{5!}{2!\cdot 3!}\cdot 0,125^2\cdot 0,875^3\\&=10\cdot 0,015625\cdot 0,669921875\\&\doteq 0,105\end{aligned}$$

Pravděpodobnost, že při výběru pěti karet s vracením nevytáhneme eso, je rovna přibližně $0,512$. Pravděpodobnost, že při výběru pěti s vracením vytáhneme eso dvakrát, je rovna přibližně $0,105$.

Pro střední hodnotu a rozptyl platí vztahy (2.9).

$$\begin{aligned}E(\mathbb{X})&=n\cdot p=5\cdot\frac{1}{8}=0,625\\D(\mathbb{X})&=n\cdot p\cdot(1-p)=5\cdot\frac{1}{8}\cdot\left(1-\frac{1}{8}\right)=5\cdot\frac{1}{8}\cdot\frac{7}{8}=\frac{35}{64}\doteq 0,55\end{aligned}$$

Poznamenejme, že hodnoty $E(\mathbb{X})$, resp. $D(\mathbb{X})$ lze vypočítat i podle vzorců (2.2.4), resp. (2.5), které jsme zavedli v Kapitole 2.2.3. Nicméně, zde použité vzorce jsou významně jednodušší k použití – není nutné vypočítávat pravděpodobnosti ke každé hodnotě $\mathbb{X}$.

2.37. Studenti píší písemku ve formě testu s 20 otázkami. U každé otázky jsou nabídnuty 4 možné odpovědi, z nichž právě jedna je správná. Jaká je pravděpodobnost, že student, který naprosto náhodně tipuje odpovědi, odpoví správně

a) právě tři otázky?

b) nejvýše tři otázky?

c) alespoň tři otázky?

Řešení: Odpověď na každou otázku představuje podle zadání náhodný pokus s možnými výsledky: student uhodne správnou odpověď, resp. student neuhodne správnou odpověď. Označme náhodným jevem A situaci, kdy student uhádne správnou odpověď. Pravděpodobnost jevu A je

$p = \frac{1}{4}$. Odpovídání na 20 otázek představuje sérii 20 pokusů, při kterých sledujeme, zda nastal jev A – student vybral správnou odpověď. Počet správně zodpovězených otázek představuje náhodnou veličinu s binomickým rozdělením pravděpodobnosti s parametry $n = 20$ a $p = \frac{1}{4} = 0,25$, tedy $\mathbb{X} \approx \text{Bi}\left(20, \frac{1}{4}\right)$. Dosazením do vzorce (2.8) dostaneme odpovědi na zadané otázky.

a) Student, který náhodně tipuje odpovědi, odpoví správně právě tři otázky pro $\mathbb{X} = 3$.

$$\begin{aligned}
P(\mathbb{X} = 3) = P(3) &= \binom{20}{3} \cdot 0,25^3 \cdot (1 - 0,25)^{20-3} \\
&= \frac{20!}{3! \cdot (20-3)!} \cdot 0,25^3 \cdot 0,75^{17} = \frac{20!}{3! \cdot 17!} \cdot 0,25^3 \cdot 0,75^{17} \\
&\doteq \frac{18 \cdot 19 \cdot 20}{1 \cdot 2 \cdot 3} \cdot 0,015625 \cdot 0,007517 \doteq 1140 \cdot 0,00011745 \doteq 0,134
\end{aligned}$$

b) Student, který náhodně tipuje odpovědi, odpoví správně nejvýše tři otázky pro $\mathbb{X} \leq 3$.

$$\begin{aligned}
P(\mathbb{X} \leq 3) &= P(\mathbb{X} = 0) + P(\mathbb{X} = 1) + P(\mathbb{X} = 2) + P(\mathbb{X} = 3) = P(0) + P(1) + P(2) + P(3) \\
&= \binom{20}{0} \cdot 0,25^0 \cdot (1 - 0,25)^{20-0} + \binom{20}{1} \cdot 0,25^1 \cdot (1 - 0,25)^{20-1} + \\
&\quad + \binom{20}{2} \cdot 0,25^2 \cdot (1 - 0,25)^{20-2} + \binom{20}{3} \cdot 0,25^3 \cdot (1 - 0,25)^{20-3} \\
&= \frac{20!}{0! \cdot (20-0)!} \cdot 0,25^0 \cdot 0,75^{20} + \frac{20!}{1! \cdot (20-1)!} \cdot 0,25^1 \cdot 0,75^{19} + \\
&\quad + \frac{20!}{2! \cdot (20-2)!} \cdot 0,25^2 \cdot 0,75^{18} + \frac{20!}{3! \cdot (20-3)!} \cdot 0,25^3 \cdot 0,75^{17} \\
&= 0,75^{20} + 20 \cdot 0,25 \cdot 0,75^{19} + 190 \cdot 0,25^2 + 1140 \cdot 0,25^3 \cdot 0,75^{17} \\
&\doteq 0,003 + 0,021 + 0,067 + 0,134 = 0,225
\end{aligned}$$

c) Student, který náhodně tipuje odpovědi, odpoví správně alespoň tři otázky pro $\mathbb{X} \geq 3$.

$$\begin{aligned}
P(\mathbb{X} \geq 3) &= P(\mathbb{X} = 3) + P(\mathbb{X} = 4) + P(\mathbb{X} = 5) + \ldots + P(\mathbb{X} = 19) + P(\mathbb{X} = 20) \\
&= P(3) + P(4) + P(5) + P(6) + P(7) + P(8) + \ldots + P(19) + P(20)
\end{aligned}$$

Takový výpočet je neúměrně náročný, využijeme zákon o rozdělení pravděpodobnosti náhodné veličiny, viz Věta 2.1.4 na straně 22. Všechny možné hodnoty náhodné veličiny $\mathbb{X}$ jsou v tomto případě $M = \{0, 1, 2, \ldots, 19, 20\}$ a součet jejich pravděpodobností je podle uvedené věty roven jedné, tedy

$$P(0) + P(1) + P(2) + P(3) + P(4) + P(5) + \ldots + P(19) + P(20) = 1.$$

Z toho plyne vztah

$$P(3) + P(4) + P(5) + P(6) + P(7) + \ldots + P(19) + P(20) = 1 - \big(P(0) + P(1) + P(2)\big).$$

Z toho důvodu můžeme hledanou pravděpodobnost vypočítat jednodušeji takto:

$$P(\mathbb{X} \geq 3) = 1 - P(\mathbb{X} \leq 2) = 1 - \big(P(0) + P(1) + P(2)\big) = 1 - (0,003 + 0,021 + 0,067) = 0,909,$$

kde hodnoty $P(0)$, $P(1)$, $P(2)$ jsou použity z řešení části b).

Z předchozích příkladů je zřejmé, že výpočty pravděpodobností u náhodné veličiny s binomickým rozdělením jsou poněkud rozsáhlejší. V příloze č. 1 na straně 50 je popsán postup, jak lze tyto výpočty provést pomocí programu *Excel* z balíku Microsoft Office. Popsaný postup pak funguje stejně i v případě dalších tabulkových procesorů např. v *Tabulkách* dostupných mezi aplikacemi Googlu atd.

Úlohy k samostatnému řešení

2.38. Jaká je pravděpodobnost, že při 10 hodech mincí padne líc

a) právě dvakrát?

b) alespoň jednou?

c) nejvýše třikrát?

2.39. Podíl nezaměstnaných v populaci je 20 %. Vypočtěte pravděpodobnost, že mezi 10 náhodně vybranými osobami budou nejvýše dva nezaměstnaní.

a) nejvýše dva nezaměstnaní.

b) alespoň dva nezaměstnaní.

2.3.2 Hypergeometrické rozdělení

Hypergeometrické rozdělení je základním pravděpodobnostním rozdělením při výběru bez vracení, kdy provedeme náhodný výběr a vybranou jednotku nevracíme do základního souboru.

2.40. Určitý typ součástek je dodáván v sériích po 50 kusech. Při přejímací kontrole je z každé série náhodně vybráno 5 výrobků ke kontrole. Tato kontrola je přitom prováděna tak, že výrobek je podroben destrukční zkoušce. Víme, že v sérii je deset vadných kusů. Potom počet vadných kusů mezi pěti vybranými výrobky je náhodnou veličinou s možnými hodnotami $M = \{0, 1, 2, 3, 4, 5\}$.

2.41. V balíčku 32 karet jsou 4 esa. Hrajeme hru, při které si z balíčku vezmeme do ruky 6 různých karet. Potom počet es v ruce je náhodnou veličinou s možnými hodnotami $M = \{0, 1, 2, 3, 4\}$.

2.42. V balíčku 32 karet je 8 srdcových karet. Hrajeme hru, při které je v talonu odhozeno již 26 karet a ve hře zbývá pouze 6 karet. Potom počet srdcových karet odhozených v talonu je náhodnou veličinou s možnými hodnotami $M = \{2, 3, 4, 5, 6, 7, 8\}$.

Společným rysem uvedených úloh je, že máme nějaký základní soubor prvků (výrobky v sérii, balíček karet) a některé z prvků tohoto souboru mají nějakou vlastnost (být vadný výrobek, být eso) a zbylé prvky základního souboru tuto vlastnost nemají (výrobky bez vady, „ne-esové“ karty). Provedeme výběr několika prvků ze základního souboru. V něm mohou být jak prvky s uvažovanou vlastností, tak i prvky bez této vlastnosti. Náhodnou veličinou je počet prvků s danou vlastností v provedeném výběru.

Definice 2.3.2. Mějme situaci, kdy v základním souboru o N prvcích jich má M určitou vlastnost a zbylých $N - M$ prvků tuto vlastnost nemá. Postupně ze souboru vybereme n prvků, z nichž žádný nevracíme zpět. Počet prvků se sledovanou vlastností mezi n vybranými prvky je náhodnou veličinou $\mathbb{X}$ s tzv. *hypergeometrickým rozdělením pravděpodobnosti.* Pro pravděpodobnostní funkci veličiny $\mathbb{X}$ platí

$$P(\mathbb{X} = x) = P(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \qquad \text{pro } \max\{n - N + M, 0\} \leq x \leq \min\{M, n\}, \tag{2.10}$$

Pro střední hodnotu a rozptyl náhodné veličiny s hypergeometrickým rozdělením pravděpodobnosti platí:

$$E(\mathbb{X}) = n \cdot \frac{M}{N} \qquad D(\mathbb{X}) = n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{N-n}{N-1} \tag{2.11}$$

Připomeňme, že ve vzorci (2.10) je

- N rozsah základního souboru
- M počet prvků s danou vlastností v základním souboru
- n počet prvků vybraných ze základního souboru
- x počet prvků z výběru s uvažovanou vlastností

Důvody podmínek ve vzorci (2.10) omezujících hodnotu x, jak zdola tak shora, jsou naznačeny v úvodních úlohách. V Příkladu 2.40 byla nejvyšší možnou hodnotou x velikost výběru (vadných výrobků bylo více než je velikost výběru). V Příkladu 2.41 bylo nejvyšší možnou hodnotou x množství prvků s danou vlastností v základním souboru (es bylo méně než je velikost výběru). V Příkladu 2.42 bylo nejnižší možnou hodnotou x číslo 2, neboť v balíčku karet je celkem 8 srdcových es a ve hře zbývá posledních 6 karet. To znamená, že minimálně 2 srdcové karty jsou již odhozeny do talonu. Tyto okolnosti vždy musíme uvážit, když stanovujeme možný rozsah hodnot náhodné veličiny s hypergeometrickým rozdělením pravděpodobnosti.

2.43. V balíčku 32 karet jsou 4 esa. Jaká je pravděpodobnost, že při náhodném výběru (bez vracení) 6 karet budou mezi vybranými kartami

a) právě 2 esa,

b) nejvýše dvě esa,

c) alespoň dvě esa.

Řešení: V zadané úloze má základní soubor (balíček karet) celkem 32 prvků – je $N = 32$. Velikost výběru je rovna 6 kartám – je $n = 6$. Sledovanou vlastností je „být esem“ – tuto vlastnost mají celkem čtyři karty – je $M = 4$. Hodnoty x se pak liší podle konkrétního zadání.

a) Ve výběru budou právě dvě esa pro $\mathbb{X} = 2$. Je tedy

$$P(\mathbb{X} = x) = P(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$P(\mathbb{X} = 2) = P(2) = \frac{\binom{4}{2}\binom{32-4}{6-2}}{\binom{32}{6}} = \frac{\binom{4}{2}\binom{28}{4}}{\binom{32}{6}} = \frac{\frac{4!}{2!\cdot(4-2)!}\cdot\frac{28!}{4!\cdot(28-4)!}}{\frac{32!}{6!\cdot(32-6)!}} = \frac{\frac{4!}{2!\cdot 2!}\cdot\frac{28!}{4!\cdot 24!}}{\frac{32!}{6!\cdot 26!}}$$

$$= \frac{\frac{3\cdot 4}{1\cdot 2}\cdot\frac{25\cdot 26\cdot 27\cdot 28}{1\cdot 2\cdot 3\cdot 4}}{\frac{27\cdot 28\cdot 29\cdot 30\cdot 31\cdot 32}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6}} = \frac{(3\cdot 2)\cdot(25\cdot 13\cdot 9\cdot 7)}{9\cdot 28\cdot 29\cdot 31\cdot 4} = \frac{122\,850}{906\,192} \doteq 0,136$$

Provedený výpočet je rozepsán podrobně, aby bylo vidět, jakým způsobem byl uskutečněn. Krácení ve zlomku v posledním řádku je samozřejmě možné provést i jiným způsobem. Pokud používané kalkulačky či SW umožní rovnou vypočítat vzniklá kombinační čísla, je vhodné tuto možnost využít. Pravděpodobnost, že ve výběru šesti karet budou právě dvě esa, je rovna přibližně 0,136.

b) Ve výběru budou nejvýše dvě esa pro $\mathbb{X} \leq 2$, tj. buď žádné eso nebo jedno eso nebo dvě esa. Je tedy

$$P(\mathbb{X} \leq 2) = P(\mathbb{X} = 0) + P(\mathbb{X} = 1) + P(\mathbb{X} = 2) = P(0) + P(1) + P(2)$$

$$= \frac{\binom{4}{0}\binom{32-4}{6-0}}{\binom{32}{6}} + \frac{\binom{4}{1}\binom{32-4}{6-1}}{\binom{32}{6}} + \frac{\binom{4}{2}\binom{32-4}{6-2}}{\binom{32}{6}} = \frac{\binom{4}{0}\binom{28}{6}}{\binom{32}{6}} + \frac{\binom{4}{1}\binom{28}{5}}{\binom{32}{6}} + \frac{\binom{4}{2}\binom{28}{4}}{\binom{32}{6}}$$

$$= \frac{\frac{4!}{0!\cdot 4!}\cdot\frac{28!}{6!\cdot 22!}}{\frac{32!}{6!\cdot 26!}} + \frac{\frac{4!}{1!\cdot 3!}\cdot\frac{28!}{5!\cdot 23!}}{\frac{32!}{6!\cdot 26!}} + \frac{\frac{4!}{2!\cdot 2!}\cdot\frac{28!}{4!\cdot 24!}}{\frac{32!}{6!\cdot 26!}} \doteq 0,4157 + 0,4338 + 0,1356 = 0,9851$$

Pravděpodobnost, že ve výběru šesti karet budou nejvýše dvě esa, je rovna přibližně 0,985.

c) Ve výběru budou alespoň dvě esa, pokud je $\mathbb{X} \geq 2$, tj. ve výběru budou dvě, tři nebo čtyři esa. (Víc es než čtyři ve výběru být nemůže, protože jich tam víc není.)

$$P(\mathbb{X} \geq 2) = P(\mathbb{X} = 2) + P(\mathbb{X} = 3) + P(\mathbb{X} = 4) = P(2) + P(3) + P(4)$$

Výpočet bychom mohli provést analogicky k předchozím výpočtů. Můžeme však také využít již zjištěné výsledky z předchozích podúloh, neboť opačným jevem k *alespoň dvě esa* je jev *nejvýše jedno eso.*

$$P(\mathbb{X} \geq 2) = 1 - P(\mathbb{X} \leq 1) = 1 - \big(P(\mathbb{X} = 0) + P(\mathbb{X} = 1)\big) = 1 - \big(P(0) + P(1)\big)$$
$$= 1 - (0,4157 + 0,4338) = 0,1505$$

Pravděpodobnost, že ve výběru šesti karet budou alespoň dvě esa, je rovna přibližně 0,151. Při výpočtu jsme použili vypočtené pravděpodobnosti z podúlohy b).

2.44. Určitý typ součástek je dodáván v sériích po 50 kusech. Při přejímací kontrole je z každé série náhodně vybráno 5 výrobků. Série je přijata, jestliže mezi kontrolovanými výrobky není žádný zmetek. Jaká je pravděpodobnost, že série bude přijata, jestliže obsahuje 10 zmetků? Kontrola je přitom prováděna tak, že výrobek je podroben destrukční zkoušce.

Řešení: Začneme posledním údajem – to, že výrobek je podroben destrukční zkoušce, znamená, že po výběru ke zkoušce jej zničíme a nemůžeme jej vrátit do základního souboru. Proto se jedná o výběr bez vracení. V zadané úloze má základní soubor (všechny výrobky v jedné sérii) celkem 50 prvků – je $N = 50$. Velikost výběru je rovna 5 výrobkům – tolik jich je vybráno ke kontrole – je $n = 5$. Sledovanou vlastností je „být vadným výrobkem" – tuto vlastnost má celkem deset výrobků – je $M = 10$. Otázka zní s jakou pravděpodobností bude série přijata – přitom série je přijata, pokud je v kontrolním výběru 0 vadných kusů, proto chceme vypočítat pravděpodobnost, že hodnota náhodné veličiny je rovna nule – tj. $x = 0$.

$$P(\mathbb{X} = x) = P(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$P(\mathbb{X} = 0) = P(0) = \frac{\binom{10}{0}\binom{50-10}{5-0}}{\binom{50}{5}} = \frac{\binom{10}{0}\binom{40}{5}}{\binom{50}{5}} = \frac{\frac{10!}{0!\cdot(10-0)!}\cdot\frac{40!}{5!\cdot(40-5)!}}{\frac{50!}{5!\cdot(50-5)!}} = \frac{\frac{10!}{0!\cdot 10!}\cdot\frac{40!}{5!\cdot 35!}}{\frac{50!}{5!\cdot 45!}}$$

$$= \frac{40!\cdot 45!}{35!\cdot 50!} = \frac{36\cdot 37\cdot 38\cdot 39\cdot 40}{46\cdot 47\cdot 48\cdot 49\cdot 50} = \frac{78\,960\,960}{254\,251\,200} \doteq 0,311.$$

Pravděpodobnost, že série je přijata, je rovna přibližně 0,311.

2.45. Ve hře Sportka se losuje šest čísel z celkového počtu 49 čísel. Hra probíhá tak, že sázející před losováním výherních čísel zatrhne na sázkovém tiketu šest čísel - ta představují vsazená čísla. Pokud všech šest vsazených čísel na sázkovém tiketu odpovídá všem šesti vylosovaným číslům, sázející vyhrává první pořadí ve hře. Vypočtěte, jaká je pravděpodobnost výhry prvního pořadí ve hře Sportka.

Řešení: Vylosovaná čísla se do osudí nevrací, stejně tak sázející musí zatrhnout šest různých čísel – jedná se tedy o výběr bez vracení. V zadané úloze má základní soubor (všechna čísla v osudí) celkem 49 prvků – je $N = 49$. Velikost výběru je rovna 6 vsazeným číslům na sázkovém tiketu – je $n = 6$. Sledovanou vlastností je „být výherním číslem" – tuto vlastnost má celkem šest vylosovaných čísel – je $M = 6$. Náhodnou veličinou je počet vylosovaných výherních čísel v hráčově výběru. Otázka zní s jakou pravděpodobností bude v hráčově výběru všech šest čísel s vlastností výherní číslo? Je tedy $x = 6$.

$$P(\mathbb{X} = x) = P(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$P(\mathbb{X} = 6) = P(6) = \frac{\binom{6}{6}\binom{49-6}{6-6}}{\binom{49}{6}} = \frac{\binom{6}{6}\binom{43}{0}}{\binom{49}{6}} = \frac{\frac{6!}{6!\cdot(6-6)!}\cdot\frac{43!}{0!\cdot(43-0)!}}{\frac{49!}{6!\cdot(49-6)!}} = \frac{\frac{6!}{6!\cdot 0!}\cdot\frac{43!}{0!\cdot 43!}}{\frac{49!}{6!\cdot 43!}}$$

$$= \frac{6!\cdot 43!}{49!} = \frac{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6}{44\cdot 45\cdot 46\cdot 47\cdot 48\cdot 49} = \frac{720}{10\,068\,347\,520} = \frac{1}{13\,983\,816} \doteq 7,15\cdot 10^{-8}.$$

Pravděpodobnost výhry prvního pořadí ve hře sportka je přibližně 0,000 000 071 5, což není mnoho. Podle předposledního údaje je zřejmé, že v průměru vyhrává jedna sázka ze 14 milionů sázek. Pro uvážení, jak malá vlastně tato pravděpodobnost je, si můžete přečíst

Úlohy k samostatnému řešení

2.46. Ve skupině 40 osob umí 3 lidé španělsky. Jaká je pravděpodobnost, že při náhodném výběru 5 osob z této skupiny bude alespoň jeden z nich mluvit španělsky?

2.47. V aplikaci Spotify mám sestavený playlist, ve které je celkem 6 českých a 24 zahraničních písní. Jaká je pravděpodobnost, že při přehrávání playlistu v náhodném pořadí zazní tři různé české písně hned za sebou? Jaký bude průměrný počet českých písní v sérii pěti písní za sebou?

2.48. Výhra třetího pořadí ve hře Sportka nastane, dokáže-li sázející správně uhodnout právě pět čísel ze šesti losovaných výherních čísel. Jaká je pravděpodobnost výhry třetího pořadí ve hře Sportka?

2.3.3 Vztah binomického a hypergeometrického rozdělení

Nyní se budeme věnovat aproximaci (nahrazení jisté hodnoty pomocí blízké, podobné hodnoty) náhodné veličiny s hypergeometrickým rozdělením pomocí náhodné veličiny s binomickým rozdělením, viz videopřednáška *Diskrétní náhodná veličina 3* od času 15:45.

V některých případech je náročné či přímo prakticky nemožné vypočítat pravděpodobnosti jednotlivých hodnot náhodné veličiny s hypergeometrickým rozdělením. Názornou ukázkou může být výpočet následující úlohy.

2.49. Rybáři chtějí určit přibližný počet ryb v rybníku (ponechme stranou použitou metodiku a předpokládejme, že v rybníku žije celkem 50 000 ryb). Za tím účelem rybáři vylovili z rybníku 1 000 ryb, označili je a pustili zpět do vody. Za týden se rybáři vrátili a vylovili 2 000 ryb. Vypočtěte pravděpodobnost, že ve znovuvyloveném vzorku bude právě 50 ryb označených během původního výlovu.

Řešení: Jedná se o výběr bez vracení (jedna ryba nemůže být během druhého výlovu v síti dvakrát). Celkem je v rybníku $N = 50\,000$ ryb, vlastnost „být označenou rybou z prvního výlovu“ má $M = 1\,000$ ryb, počet znovuvylovených ryb, tj. velikost výběru je $n = 2\,000$ ryb. Náhodnou veličinou je *počet označených ryb ve druhém výlovu*, přičemž hodnota, pro kterou chceme vypočítat pravděpodobnost, je $x = 40$. Jedná se o tedy o náhodnou veličinu s hypergeometrickým rozdělením.

Výpočet je následující:

$$P(\mathbb{X} = x) = P(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$P(\mathbb{X} = 100) = P(100) = \frac{\binom{1\,000}{40}\binom{50\,000-1\,000}{2\,000-40}}{\binom{50\,000}{2\,000}} = \frac{\binom{1\,000}{40}\binom{49\,000}{1\,960}}{\binom{50\,000}{2\,000}} = \frac{\frac{1\,000!}{40!\cdot 960!}\cdot\frac{49\,000!}{1\,960!\cdot 47\,040!}}{\frac{50\,000!}{2\,000!\cdot 48\,000!}}.$$

Zde počítané faktoriály i po zkrácení ve zlomcích představují tak obrovská čísla, že je prakticky nemožné je počítat.

Nyní přichází ke slovu užitečnost výše zmíněné možnosti aproximace. Opět použijeme ilustrační příklad.

2.50. V regionu žije 100 000 obyvatel, z nichž je 5 000 nezaměstnaných. Jaká je pravděpodobnost, že mezi 10 náhodně vybranými různými osobami v daném regionu je právě jeden nezaměstnaný?

Řešení: Uvažujme náhodnou veličinu *počet nezaměstnaných mezi deseti vybranými osobami.* V principu se jedná o náhodnou veličinu s hypergeometrickým rozdělením ($N = 100\,000$, $M =$

5 000, $n = 10$, $x = 1$). Při výpočtu pravděpodobnosti pomocí hypergeometrického rozdělení dostaneme:

$$P(1) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} = \frac{\binom{5\,000}{1}\binom{100\,000-5\,000}{10-1}}{\binom{100\,000}{10}} = \frac{\binom{5\,000}{1}\binom{95\,000}{9}}{\binom{100\,000}{10}} = 0,315147097.$$

Na úlohu se lze ovšem podívat také tím způsobem, že náhodným jevem A je výběr nezaměstnaného. Pokud je mezi 100 000 lidmi právě 5 000 nezaměstnaných, pak pravděpodobnost výběru nezaměstnané osoby při jednom výběru je

$$P(A) = \frac{5\,000}{100\,000} = 0,05.$$

Nyní desetkrát konáme náhodný pokus (náhodně vybíráme osobu z populace) a nahodnou veličinu definuijeme jako počet případů, kdy nastal jev A. To odpovídá definici náhodné veličiny s binomickým rozdělením pravděpodobnosti. Výpočet potom vypadá následujícím způsobem.

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

$$P(\mathbb{X} = 1) = P(1) = \binom{10}{1} \cdot 0,05^1 \cdot (1-0,05)^{10-1} = 10 \cdot 0,05 \cdot (0,95)^9 = 0,315124705$$

Porovnáním výsledků plynoucích z hypergeometrického a binomického rozdělení vidíme, že získané výsledky se liší až na pátém desetinném místě, takže při zaokrouhlení na tři desetinná místa (což se často děje) rozdíl nepoznáme – z obou přístupů dostaneme výsledek, že hledaná pravděpodobnost je rovna 0,315. Mimochodem – pro ostatní hodnoty obou náhodných veličin jsou příslušné pravděpodobnosti uvedeny v Tabulce 2.6.

x	hypergeom. rozdělení $P(x)$	binomické rozdělení $P(x)$
0	0,598722758	0,598736939
1	0,315147097	0,315124705
2	0,074631459	0,074634799
3	0,010471172	0,010475059
4	0,000963932	0,000963932
5	0,000608344	0,000060935
6	0,000002665	0,000002672
7	0,000000080	0,000000080
8	0,000000002	0,000000002
9	0,000000000	0,000000000
10	0,000000000	0,000000000

Tabulka 2.6: Porovnání hodnot pravděpodobnosti u hypergeometrického a binomického rozdělení

Otázkou samozřejmě je, za jakých okolností lze výše uvedenou záměnu použít. Odpověď je následující.

Věta 2.3.3. *Je-li velikost výběru n malá v porovnání s velikostí základního souboru N, (tj. $n/N < 0.05$), můžeme místo hypergeometrického rozdělení použít binomické rozdělení. Parametr p příslušného binomického rozdělení určíme ze vztahu*

$$p = \frac{M}{N}. \tag{2.12}$$

Podle uvedeného návodu vypočteme pravděpodobnost z Příkladu 2.49.

2.51. Ověřte předpoklady Věty 2.3.3, a pokud budou splněny, vypočtěte pravděpodobnost z Příkladu 2.49 pomocí binomického rozdělení.

Řešení: Při řešení Příkladu 2.49 jsme odvodili následující parametry hypergeometrického rozdělení – $N = 50\,000$, $M = 1\,000$ a $n = 2\,000$. Abychom mohli rozumně nahradit toto rozdělení binomickým rozdělením, měla by být podle Věty 2.3.3 splněna nerovnost $n/N < 0,05$. Zde je

$$\frac{n}{N} = \frac{2\,000}{50\,000} = 0,04 < 0,05,$$

proto je podmínka věty splněna.

Parametr p binomického rozdělení vypočteme ze vztahu 2.12. Je

$$p = \frac{M}{N} = \frac{1\,000}{50\,000} = 0,02.$$

Dosazením do vzorce 2.8 na straně 35 dostaneme:

$$P(\mathbb{X} = 100) = \binom{2\,000}{50} \cdot 0,02^{50}(1 - 0,02)^{2\,000-50} = \frac{2\,000!}{50! \cdot 1950!} \cdot 0,02^{50} \cdot 0,98^{1950} = \text{ajajaj}$$

I v tomto případě je výpočet příliš náročný, budeme si muset pomoci jiným způsobem – ukáže se, že tou pomocí je náhodná veličina s tzv. Poissonovým rozdělením pravděpodobnosti.

2.3.4 Poissonovo rozdělení

Náhodné veličině s Poissonovým rozdělením pravděpodobnosti se věnuje videopřednáška dr. Šimsové *Diskrétní náhodná veličina 2* od času 38:48. Připomeňme, že poissonovskými událostmi rozumíme takové události, které přicházejí v čase, přičemž platí:

- v jednom okamžiku může nastat nejvýše jedna událost,
- události přicházejí nezávisle na sobě (počty vzniklých událostí v disjunktních časových intervalech jsou nezávislé),
- pravděpodobnost, že událost nastane v intervalu $(t, t + h)$ závisí na h, ale nikoliv na t. Jinými slovy – pravděpodobnost výskytu události závisí na délce intervalu, nikoliv na tom, kdy časový interval začíná.

2.52. Po silnici před naší školou projede v době trvání cvičení průměrně 100 automobilů za hodinu. Náhodnou veličinou je počet projíždějích automobilů během jedné hodiny. Možné hodnoty náhodné veličiny tvoří množinu $M = \{0,\ 1,\ 2,\ 3,\ \ldots\}$.

2.53. Na zákaznickou linku firmy ABC se každou minutu obrací průměrně 62 zákazníků. Náhodnou veličinou je počet volajících zákazníků za jednu minutu. Možné hodnoty náhodné veličiny tvoří množinu $M = \{0, 1, 2, 3, \ldots\}$.

Definice 2.3.4. Označme průměrný počet výskytů jisté poissonovské události za časovou jednotku číslem λ. Potom počet událostí x, které nastanou za časovou jednotku, je náhodnou veličinou $\mathbb{X}$ s tzv. Poissonovým rozdělením pravděpodobnosti, jejíž pravděpodobnostní funkce má předpis

$$P(\mathbb{X} = x) = P(x) = \frac{\lambda^x}{x!} \cdot \mathrm{e}^{-\lambda}.$$

Pro střední hodnotu a rozptyl náhodné veličiny s poissonovým rozdělením pravděpodobnosti platí:

$$E(\mathbb{X}) = \lambda \qquad D(\mathbb{X}) = \lambda$$

2.54. K pokladně v supermarketu přistoupí průměrně 5 zákazníků během deseti minut.

a) Jaká je pravděpodobnost, že během následujících 10 minut přijdou právě 3 zákazníci?

b) Jaká je pravděpodobnost, že během následujích dvou minut nepřijde ani jeden zákazník?

Řešení:

a) Při řešení příkladu budeme předpokládat, že příchod zákazníka k pokladně je poissonovský proces. Potom náhodná veličina *počet zákazníků, kteří přišli k pokladně během deseti minut* má poissonovské rozdělení pravděpodobnosti s parametrem $\lambda = 5$ a pro pravděpodobnost, že během následujících 10 minut přijdou právě 3 zákazníci, platí

$$P(\mathbb{X} = x) = P(x) = \frac{\lambda^x}{x!} \cdot \mathrm{e}^{-\lambda}$$
$$P(\mathbb{X} = 3) = P(3) = \frac{5^3}{3!} \cdot \mathrm{e}^{-5} = 0,14038$$

Pravděpodobnost, že během následujících 10 minut přijdou právě 3 zákazníci je rovna přibližně $p = 0,14$.

b) Při řešení příkladu budeme opět předpokládat, že příchod zákazníka k pokladně je poissonovský proces. Za náhodnou veličinu budeme uvažovat *počet zákazníků, kteří přišli k pokladně během dvou minut*. Pak tato veličina má poissonovské rozdělení pravděpodobnosti s parametrem $\lambda = 1$ (když za 10 minut přijde v průměru 5 zákazníků, tak za 2 minuty v průměru přijde 1 zákazník; proto je $\lambda = 1$) a pravděpodobnost, že během následujících dvou minut nepřijde ani jeden zákazník (tj. přijde nula zákazníků), platí

$$P(\mathbb{X} = x) = P(x) = \frac{\lambda^x}{x!} \cdot \mathrm{e}^{-\lambda}$$
$$P(\mathbb{X} = 0) = P(0) = \frac{1^0}{0!} \cdot \mathrm{e}^{-1} = 0,3679$$

Pravděpodobnost, že během následujích dvou minut nepřijde ani jeden zákazník je rovna přibližně $p = 0,37$.

2.55. Na výrobní lince dojde k poruše průměrně jednou za dvě hodiny. Jaká je pravděpodobnost, že během osmihodinové směny dojde na výrobní lince

a) ke dvěma poruchám?

b) k nejvýše dvěma poruchám?

Řešení: Při řešení příkladu budeme předpokládat, že událost, kdy na výrobní lince dojde k poruše, je poissonovský proces. Za náhodnou veličinu budeme uvažovat *počet poruch na výrobní lince za osm hodin provozu.* Pak tato veličina má poissonovské rozdělení pravděpodobnosti s parametrem $\lambda = 4$ (když za dvě hodiny dojde v průměru k jedné poruše, tak za osm hodin v průměru nastanou čtyři poruchy; proto je $\lambda = 4$).

a) Pravděpodobnost, že během osmihodinové směny dojde na výrobní lince ke dvěma poruchám, odpovídá $P(\mathbb{X} = 2)$.

$$P(\mathbb{X} = x) = P(x) = \frac{\lambda^x}{x!} \cdot \mathrm{e}^{-\lambda}$$
$$P(\mathbb{X} = 2) = P(2) = \frac{4^2}{2!} \cdot \mathrm{e}^{-4} = 0,1465$$

Pravděpodobnost, že během osmihodinové směny dojde na výrobní lince ke dvěma poruchám, je rovna přibližně $p = 0,15$.

b) Pravděpodobnost, že během osmihodinové směny dojde na výrobní lince k nejvýše dvěma poruchám, odpovídá $P(\mathbb{X} \leq 2)$.

$$P(\mathbb{X} \leq 2) = P(0) + P(1) + P(2) = \frac{4^0}{0!} \cdot \mathrm{e}^{-4} + \frac{4^1}{1!} \cdot \mathrm{e}^{-4} + \frac{4^2}{2!} \cdot \mathrm{e}^{-4} = 0,2381$$

Pravděpodobnost, že během osmihodinové směny dojde na výrobní lince k nejvýše dvěma poruchám, je rovna přibližně $p = 0,24$.

Poissonovo rozdělení popisuje počty událostí i v jiných než časových jednotkách, např. v jednotkách délky, obsahu atd.

2.56. V pruhu látky se vyskytuje průměrně 5 kazů na 100 metrů látky. Jaká je pravděpodobnost, že:

a) v náhodně vybrané roli látky o délce 100 metrů se nacházejí právě 4 kazy?

b) v náhodně vybrané roli látky o délce 200 metrů se nachází právě 6 kazů?

c) v náhodně vybrané roli látky o délce 50 metrů se nachází nejvýše 1 kaz?

Řešení: Při řešení ověříme (zdůvodníme), že výskyt kazů na m^2 látky opravdu můžeme považovat za poissonovský proces. Připomeňme vlastnosti poissonovského procesu.

- V jednom okamžiku může nastat nejvýše jedna událost:
  - na jednom místě může být nejvýše jeden kaz.

- Události přicházejí nezávisle na sobě (počty vzniklých událostí v disjunktních časových intervalech jsou nezávislé):
  - Kazy se vyskytují nezávisle na sobě (počty kazů v různých pruzích látky jsou nezávislé ve smyslu, že počet kazů v předchozím m^2 neříká nic o tom, jaký počet kazů bude v následujícím m^2 látky).
- Pravděpodobnost, že událost nastane v intervalu $(t, t+h)$ závisí na h, ale nikoliv na t. Jinými slovy – pravděpodobnost výskytu události závisí na délce intervalu, nikoliv na tom, kdy časový interval začíná.
  - Počet kazů v m^2 nezávisí na tom, od kolikátého metru látky začínáme počet kazů měřit, ale pouze na tom, jaká je délka toho pruhu látky.

Nyní se již můžeme zaměřit na řešení konkrétních podúloh. Při jejich řešení budeme předpokládat, že výskyt kazů v m^2 látky je poissonovský proces.

a) Za náhodnou veličinu budeme uvažovat *počet kazů v roli látky o délce 100 metrů.* Pak tato veličina má poissonovské rozdělení pravděpodobnosti s parametrem $\lambda = 5$ a příslušnou pravděpodobností je $P(\mathbb{X} = 4)$.

$$P(\mathbb{X} = x) = P(x) = \frac{\lambda^x}{x!} \cdot \mathrm{e}^{-\lambda}$$
$$P(\mathbb{X} = 4) = P(4) = \frac{5^4}{4!} \cdot \mathrm{e}^{-5} = 0,175$$

Pravděpodobnost, že v náhodně vybrané roli látky o délce 100 metrů se nacházejí právě 4 kazy, je rovna $p = 0,175$

b) Za náhodnou veličinu budeme uvažovat *počet kazů v roli látky o délce 200 metrů.* Pak tato veličina má poissonovské rozdělení pravděpodobnosti s parametrem $\lambda = 10$ (jestliže ve 100 metrech látky je průměrně 5 kazů, pak v 200 metrech jich můžeme průměrně očekávat 10; proto je $\lambda = 10$) a příslušnou pravděpodobností je $P(\mathbb{X} = 6)$.

$$P(\mathbb{X} = x) = P(x) = \frac{\lambda^x}{x!} \cdot \mathrm{e}^{-\lambda}$$
$$P(\mathbb{X} = 6) = P(6) = \frac{10^6}{6!} \cdot \mathrm{e}^{-10} = 0,063$$

Pravděpodobnost, že v náhodně vybrané roli látky o délce 200 metrů se nachází právě 6 kazů, je rovna $p = 0,063$.

c) Za náhodnou veličinu budeme uvažovat *počet kazů v roli látky o délce 50 metrů.* Pak tato veličina má poissonovské rozdělení pravděpodobnosti s parametrem $\lambda = 2,5$ (jestliže ve 100 metrech látky je průměrně 5 kazů, pak v 50 metrech jich můžeme průměrně očekávat 2,5; proto je $\lambda = 2,5$) a příslušnou pravděpodobností je $P(\mathbb{X} \leq 1)$.

$$P(\mathbb{X} \leq 1) = P(0) + P(1) = \frac{2,5^0}{0!} \cdot \mathrm{e}^{-2,5} + \frac{2,5^1}{1!} \cdot \mathrm{e}^{-2,5} = 0,287$$

Pravděpodobnost, že v náhodně vybrané roli látky o délce 50 metrů se nachází nejvýše jeden kaz, je rovna $p = 0,287$.

2.57. Redaktor časopisu ze zkušenosti ví, že v dodaných rukopisech jsou průměrně 3 chyby na jedné tiskové straně. Jaká je pravděpodobnost, že v dodaném článku o rozsahu 5 stran bude méně než 10 chyb?

Řešení: Ze zadání plyne, že na 5 stranách textu bude průměrně 15 chyb. Proto bude v zadaném případě $\lambda = 15$.

$$\begin{aligned} P(\mathbb{X} < 10) &= P(0) + P(1) + P(2) + \ldots + P(8) + P(9) \\ &= \frac{15^0}{0!} \cdot \mathrm{e}^{-15} + \frac{15^1}{1!} \cdot \mathrm{e}^{-15} + \frac{15^2}{2!} \cdot \mathrm{e}^{-15} + \ldots + \frac{15^8}{8!} \cdot \mathrm{e}^{-15} + \frac{15^9}{9!} \cdot \mathrm{e}^{-15} \\ &= \mathrm{e}^{-15} \cdot \left(\frac{15^0}{0!} + \frac{15^1}{1!} + \frac{15^2}{2!} + \ldots + \frac{15^8}{8!} + \frac{15^9}{9!} \right) \\ &\doteq 0,0699 \end{aligned}$$

Pravděpodobnost, že v článku o rozsahu 5 stran bude méně než 10 tiskových chyb, je přibližně rovna $p = 0,07$.

2.58. Předpokládejme, že pravděpodobnost, že se 35letý muž nedožije následujícího roku je 0,01. Roční pojistné této věkové skupiny činí 2 500 Kč. V případě úmrtí pojišťovna jednorázově vyplatí 100 000 Kč. Jaká je pravděpodobnost, že zisk z 1 000 pojištěných mužů ve věku 35 let bude alespoň 2 000 000 Kč?

Řešení: Ze zadání plyne, že pojišťovna získá z pojištění 1 000 mužů celkem 2 500 000 Kč. Pokud má mít z pojištění zisk alespoň 2 000 000 Kč, potom nesmí vyplatit více než 500 000 Kč, tj. nesmí dojít k úmrtí více než 5 mužů. Dále ze zadání plyne, že v průměru zemře během jednoho roku jeden muž ze 100, tj. 10 mužů z 1 000. Je tedy $\lambda = 10$ a hledáme pravděpodobnost $P(\mathbb{X} \leq 5)$.

$$\begin{aligned} P(\mathbb{X} \leq 5) &= P(0) + P(1) + P(2) + P(3) + P(4) + P(5) \\ &= \frac{10^0}{0!} \cdot \mathrm{e}^{-10} + \frac{10^1}{1!} \cdot \mathrm{e}^{-10} + \frac{10^2}{2!} \cdot \mathrm{e}^{-10} + \frac{10^3}{3!} \cdot \mathrm{e}^{-10} + \frac{10^4}{4!} \cdot \mathrm{e}^{-10} + \frac{10^5}{5!} \cdot \mathrm{e}^{-10} \\ &= \mathrm{e}^{-10} \cdot \left(\frac{10^0}{0!} + \frac{10^1}{1!} + \frac{10^2}{2!} + \frac{10^3}{3!} + \frac{10^4}{4!} + \frac{10^5}{5!} \right) \\ &\doteq 0,067 \end{aligned}$$

Pravděpodobnost, že zisk pojišťovny z daného pojištění bude alespoň 2 000 000 Kč, je přibližně rovna $p = 0,067$.

2.3.5 Aproximace binomického a hypergeometrického rozdělení Poissonovým rozdělením

Definice 2.3.5. Pro dostatečně velké n (prakticky pro $n > 30$) a pro malou pravděpodobnost (prakticky pro $p \leq 0,1$), lze binomické, resp. hypergeometrické rozdělení aproximovat Poissonovým rozdělením, kde $\lambda = np$, resp. $\lambda = nM/N$.

2.59. V regionu s 1 000 000 obyvatel v produktivním věku je 40 000 nezaměstnaných. V dotazníkovém šetření vybereme 100 obyvatel z regionu. Jaká je pravděpodobnost toho, že mezi nimi bude právě 8 nezaměstnaných.

Řešení:

Jde o výběr bez vracení, tedy o hypergeometrické rozdělení s parametry $N = 1\,000\,000$, $M = 40\,000$, $n = 100$ a $x = 8$. Je

$$P(\mathbb{X} = 8) = P(8) = \frac{\binom{40\,000}{8} \cdot \binom{960\,000}{92}}{\binom{1\,000\,000}{100}} = 0,0285.$$

Protože ale jde o „velké osudí“, je možné to počítat pomocí binomického rozdělení s parametry $n = 100$ a $p = 40\,000/1\,000\,000 = 0,04$. Tak dostaneme

$$P(\mathbb{X} = 8) = P(8) = \binom{100}{8}(0,04)^8(0,96)^{92} = 0,0285.$$

Pomocí nahodné veličiny s poissonovým rozdělením dostaneme ($\lambda = np = 100 \cdot 0,04 = 4$)

$$P(\mathbb{X} = 8) = P(8) = \frac{4^8 \cdot e^{-4}}{8!} = 0,0298.$$

2.4 Úlohy k samostatné práci

2.60. Na zákaznickou linku zavolá průměrně 12 zákazníků za hodinu. S jakou pravděpodobností nezavolá během následující čtvrthodiny ani jeden zákazník?

2.61. V rybníku je 50 000 ryb. Rybáři chytí 1 000 ryb, označí je a vrátí do rybníka. Pak vyloví 20 ryb. Jaká je pravděpodobnost, že mezi nimi budou právě dvě označené ryby?

2.62. Hráč hodí šesti kostkami najednou. Jaká je pravděpodobnost, že na kostkách při hodu padla alespoň jedna jednička?

2.63. Amatérský střelec trefí při jednom výstřelu černý střed terče s pravděpodobností $p = 0.65$. Jaká je pravděpodobnost, že se střelec trefí při 10 výstřelech do černého středu terče nejvýše devětkrát?

2.64. Na plese bylo prodáno 150 losů do tomboly. V tombole je deset cen. Jaká je pravděpodobnost, že soutěžící, který si zakoupil 10 losů, vyhraje více než jednu cenu?

2.65. Předpokládejme, že pravděpodobnost narození chlapce je stejná jako pravděpodobnost narození děvčete. Vypočtěte jaká je pravděpodobnost, že v náhodně vybrané rodině se třemi dětmi je víc synů než dcer.

2.66. V extraligovém play-off se utkaly týmy A a B. Předpokládejme, že tým A je lepší a pravděpodobnost jeho výhry činí $p = 0,6$ (jedná se o play-off, a není tedy možná remíza na konci zápasu). Jaká je pravděpodobnost výhry slabšího týmu v celé sérii, hraje-li se tato na 4 vítězné zápasy?

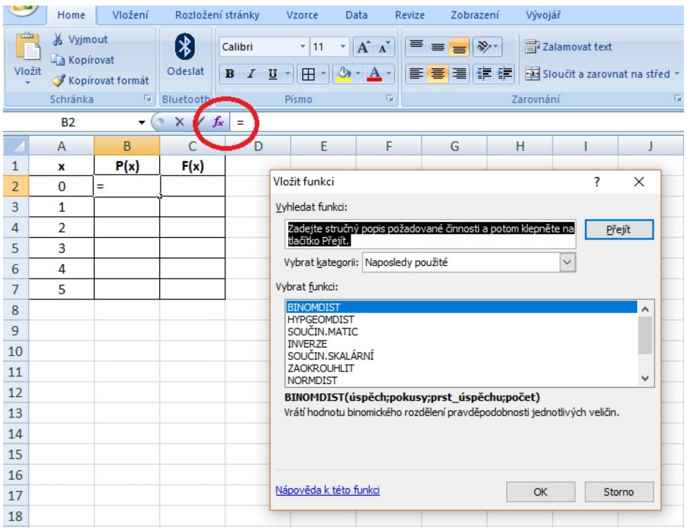
2.67. V balíčku 32 karet je osm karet červené barvy. Hráč dostane při rozdání 5 karet. Jaká je pravděpodobnost, že všechny budou mít červenou barvu?

Příloha č. 1: Výpočet náhodné veličiny s binomickým rozdělením pomocí Excelu

V této příloze si popíšeme postup při určování pravděpodobnosti hodnot náhodné veličiny s binomickým rozdělením. Postup si ukážeme na poslední z ilustračních úloh v úvodu kapitoly o binomickém rozdělení.

2.68. Předpokládejme, že jistý biatlonista při každém výstřelu zasáhne terč s pravděpodobností $p = 0,9$; tj. dlouhodobě má úspěšnost zásahu terče 90 %. Tento biatlonista přijede na střelnici a střílí pětkrát na terč. Sledujeme, kolikrát se při těchto pěti pokusech trefil do terče. Náhodnou veličinou bude počet úspěšných zásahů při pěti výstřelech na terč. Možné hodnoty této náhodné veličiny jsou $M = \{0, 1, 2, 3, 4, 5\}$. Určete pravděpodobnosti všech těchto hodnot.

Řešení: Ze zadání plyne, že náhodná veličina *počet úspěšných zásahů terče při pěti výstřelech* má binomické rozdělení pravděpodobnosti. (Jev A – zásah terče při výstřelu – má při každém opakování stejnou pravděpodobnost a pokus opakujeme pětkrát.) Popišme si postup výpočtu v programu Excel. Postupovat budeme tak, že vytvoříme základní tabulku, zapíšeme možné hodnoty náhodné veličiny a necháme program spočítat pravděpodobnosti těchto hodnot. Možné hodnoty jsou 0, 1, 2, 3, 4, 5; tyto napíšeme do jednotlivých buněk pod sebe. Do druhého sloupce vložíme



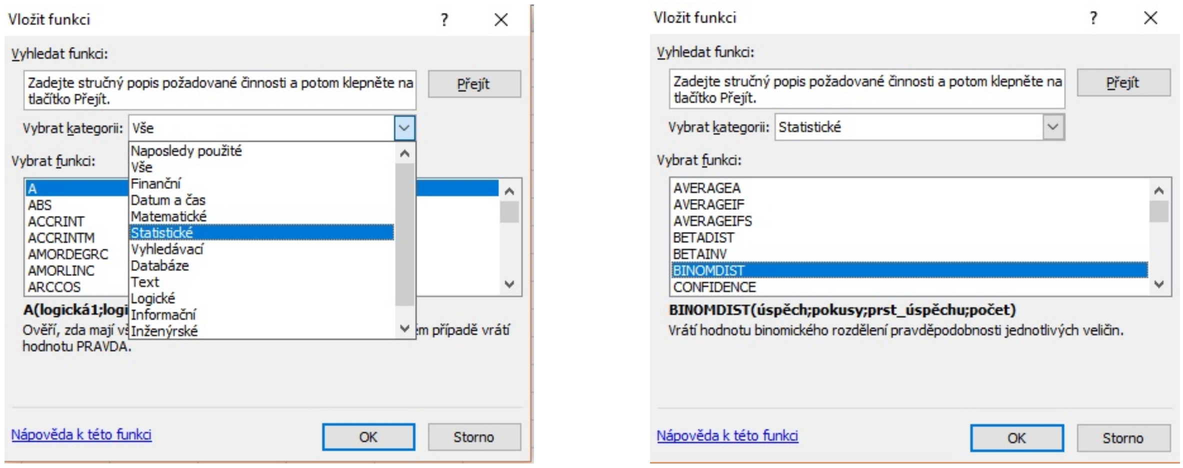
Obrázek 2.3: Vytvoření základní tabulky - zadání hodnot náhodné veličiny a vyvolání okna *Vložit funkci*

pravděpodobnosti příslušných hodnot následujícím způsobem. Označíme (vybereme) buňku vedle první hodnoty x, kliknutím na symbol f_x vyvoláme okno *Vložit funkci*, viz Obrázek 2.3. Vybereme kategorii statistických funkcí a z nabídky vybereme funkci BINOMDIST, viz Obrázek 2.4.

Funkce BINOMDIST má celkem čtyři argumenty k vyplnění, viz Obrázek 2.5.

Prvním je tzv. *Úspěch* (nápověda – počet úspěšných pokusů) - obecně se jedná o hodnotu náhodné veličiny, pro kterou počítáme pravděpodobnost. V našem případě navíc hodnoty náhodné veličiny odpovídají počtu úspěšných pokusů biatlonisty, takže označení je ve shodě s naší intuicí. Do políčka kliknutím na buňku A2 zadáme relativní odkaz na buňku, ve které se nachází hodnota x (odkaz lze zadat i ručně vypsáním).

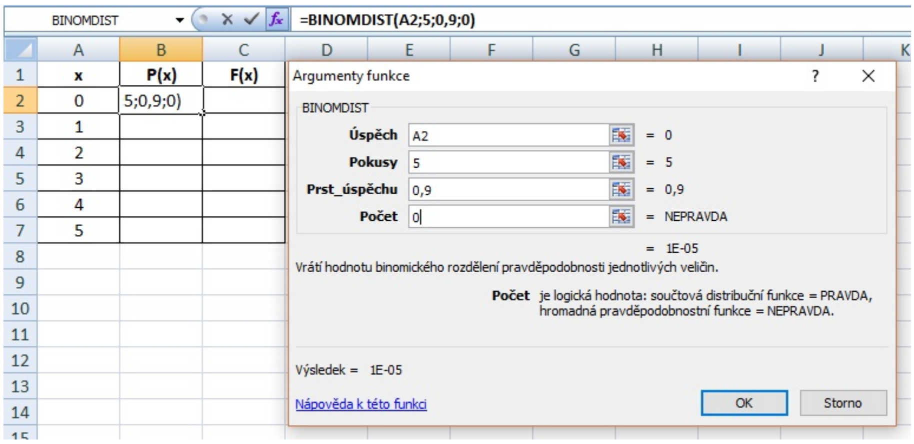
Druhý argument má název *Pokusy* a znamená počet opakování náhodného pokusu. V našem případě se jednalo o pět výstřelů, je tedy doplněno číslo 5.



Obrázek 2.4: Výběr kategorie statistických funkcí a funkce BINOMDIST

Třetí argument odkazuje na pravděpodobnost jevu A. Jmenuje se sice *Prst-úspěchu*, ale tím se nenecháme zaskočit. Pravděpodobnost úspěšného zásahu při jednom výstřelu je 0,9, proto do třetího políčka zadáme toto číslo.

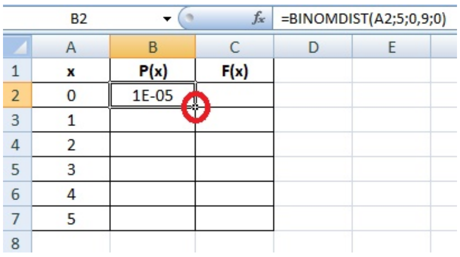
Poslední argument *Počet* určuje, zda-li chceme vypočítat pravděpodobnost dané hodnoty x, tj. hodnotu $P(\mathbb{X} = x)$, nebo hodnotu distribuční funkce pro dané x, tj. $F(x) = P(\mathbb{X} \leq x)$. V prvním případě zadáme hodnotu 0, ve druhém případě zadáme hodnotu 1. Zde chceme vypočítat hodnotu pravděpodobnosti, zadali jsme číslo 0.



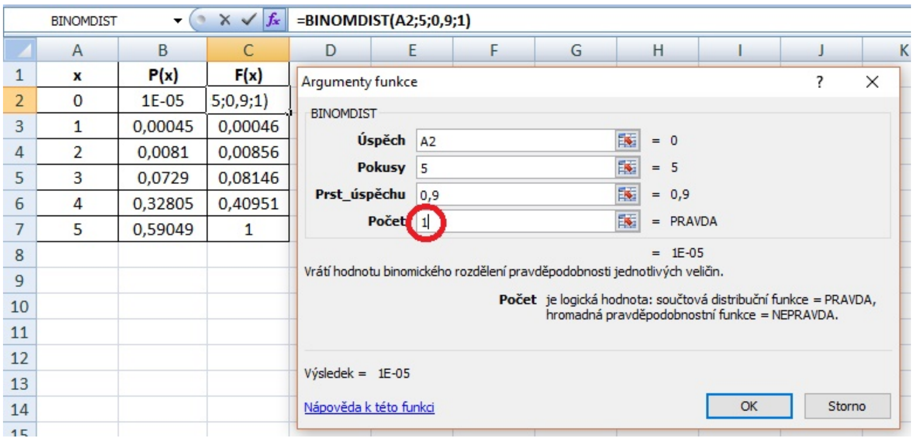
Obrázek 2.5: Zadání hodnot funkce BINOMDIST

Vyplněné okno potvrdíme kliknutím na tlačítko OK. V buňce A2 se nám zobrazí výraz 1E-05, či něco podobného. Toto je zápis v exponenciálním tvaru a znamená číslo $1 \cdot 10^{-5}$, tj. 0,00001 – což znamená prakticky hodnotu nula. Nyní najedeme kurzorem na dolní pravý roh buňky A2. Kurzor se změní ze šipky na znaménko plus (+), klikneme na tento dolní roh a tím se vzorec nakopíruje do ostatních buněk s tím, že se ve vzorci vždy změní odkaz na hodnotu x – proto jsme ji v prvním kroku zadávali pomocí relativního odkazu, viz Obrázek 2.6.

Místo pravděpodobností jednotlivých hodnot bychom mohli také chtít vypočítat hodnoty distribuční funkce. Postup je skoro stejný, pouze při vyplňování posledního argumentu zadáme hodnotu 1, viz Obrázek 2.7.



Obrázek 2.6: Pro nakopírování vzorce do ostatních buněk klikněte na dolní pravý roh buňky s vytvořeným vzorcem



Obrázek 2.7: Hodnoty distribuční funkce získáme zadáním hodnoty 1 do posledního řádku okna

Z výsledků v tabulce můžeme vyčíst mnoho informací, viz Obrázek 2.7. Například je zde uvedena informace, že

- pravděpodobnost sestřelení právě dvou terčů je rovna $0,0081$,
- pravděpodobnost sestřelení všech pěti terčů je rovna $0,59049$,
- nejvyšší pravděpodobnost má zásah všech pěti terčů,
- pravděpodobnost zásahu nejvýše dvou terčů je rovna $0,00856$,
- pravděpodobnost zásahu nejvýše čtyř terčů je rovna $0,40951$.

Zde naleznete videonávod k výpočtu.

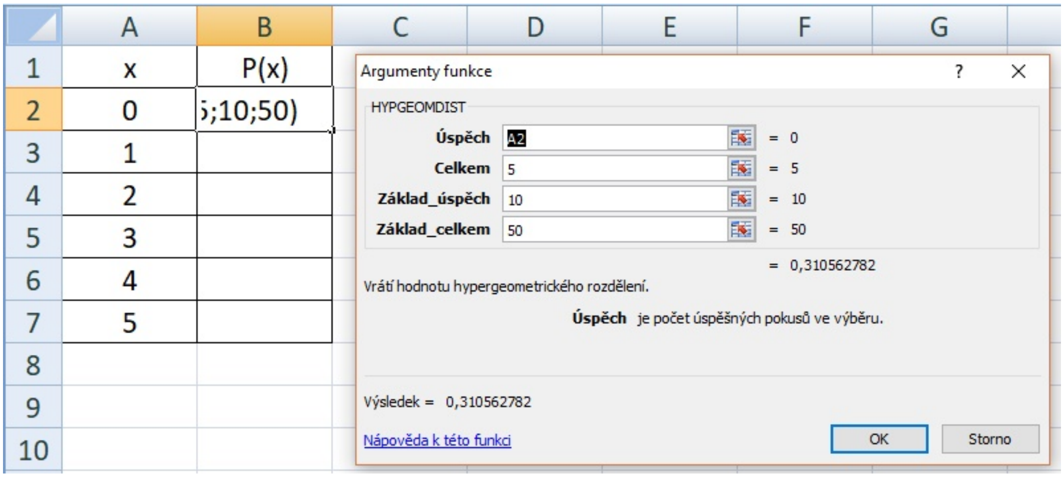
Příloha č. 2: Výpočet náhodné veličiny s hypergeometrickým rozdělením pomocí Excelu

V této příloze si popíšeme postup při určování pravděpodobnosti hodnot náhodné veličiny s hypergeometrickým rozdělením pravděpodobnosti. Postup si ukážeme na Příkladu 2.40 na straně 38

2.69. Určitý typ součástek je dodáván v sériích po 50 kusech. Při přejímací kontrole je z každé série náhodně vybráno 5 výrobků ke kontrole. Tato kontrola je přitom prováděna tak, že výrobek je podroben destrukční zkoušce. Víme, že v sérii je deset vadných kusů. Potom počet vadných kusů mezi pěti vybranými výrobky je náhodnou veličinou s možnými hodnotami $M = \{0, 1, 2, 3, 4, 5\}$.

Řešení: Ze zadání plyne, že náhodná veličina *počet vadných kusů mezi pěti vybranými výrobky* má hypergeometrické rozdělení pravděpodobnosti, kde $N = 50$, $M = 10$, $n = 5$. Popišme si postup výpočtu v programu Excel. Začátek je stejný jako v příloze č. 1. na straně 50. Vytvoříme základní tabulku, zapíšeme možné hodnoty náhodné veličiny a necháme program spočítat pravděpodobnosti těchto hodnot. Možné hodnoty jsou 0, 1, 2, 3, 4, 5; tyto napíšeme do jednotlivých buněk pod sebe. Do druhého sloupce vložíme pravděpodobnosti příslušných hodnot následujícím způsobem. Označíme (vybereme) buňku vedle první hodnoty x, kliknutím na symbol f_x vyvoláme okno *Vložit funkci*, vybereme kategorii statistických funkcí a z nabídky funkcí vybereme položku HYPGEOMDIST.

Funkce HYPGEOMDIST má celkem čtyři argumenty k vyplnění, viz Obrázek 2.8.



Obrázek 2.8: Zadání hodnot funkce HYPGEOMDIST

Prvním je tzv. *Úspěch* (nápověda – počet úspěšných pokusů ve výběru) - obecně se jedná o hodnotu náhodné veličiny, pro kterou počítáme pravděpodobnost. Do políčka kliknutím na buňku A2 zadáme relativní odkaz na buňku, ve které se nachází hodnota x (odkaz lze zadat i ručně vypsáním).

Druhý argument má název *Celkem* (nápověda – velikost výběru) a znamená počet opakování náhodného pokusu. V našem případě se jednalo o výběr pěti výrobků, je tedy doplněno číslo 5.

Třetí argument se jmenuje *Základ-úspěch*, (nápověda – počet úspěšných pokusů v základním souboru) a představuje počet prvků s danou vlastností v základním souboru. V základní souboru je celkem 10 vadných výrobků, zadáme číslo 10.

Poslední argument *Základ-celkem* (nápověda – velikost základního souboru) je roven počtu všech prvků v základním souboru. Série obsahuje 50 výrobků, prot zadáme číslo 50.

Vyplněné okno potvrdíme kliknutím na tlačítko OK. V buňce A2 se objevila příslušná hodnota pravděpodobnosti. Nyní vzorec zkopírujeme do ostatních buněk ke zbývajícím hodnotám náhodné veličiny např. tak, že kursorem najedeme na dolní pravý roh buňky A2. Kurzor se změní ze šipky na znaménko plus (+), klikneme na tento dolní roh a tím se vzorec nakopíruje do ostatních buněk.

Z výsledků v tabulce můžeme vyčíst mnoho informací, viz Obrázek 2.9. Například je zde uvedena informace, že

	A	B
1	x	P(x)
2	0	0,31056
3	1	0,43134
4	2	0,20984
5	3	0,04418
6	4	0,00396
7	5	0,00012
8		

Obrázek 2.9: Výsledky příkladu vypočtené pomocí HYPGEOMDIST

- pravděpodobnost výskytu dvou vadných výrobků ve výběru je rovna přibližně $0,21$,
- pravděpodobnost výskytu čtyř vadných výrobků ve výběru je rovna přibližně $0,004$,
- nejvyšší pravděpodobnost má výskyt jednoho vadného výrobku.

Zde naleznete videonávod k výpočtu.

Příloha č. 3: Výpočet náhodné veličiny s poissonovým rozdělením pomocí Excelu

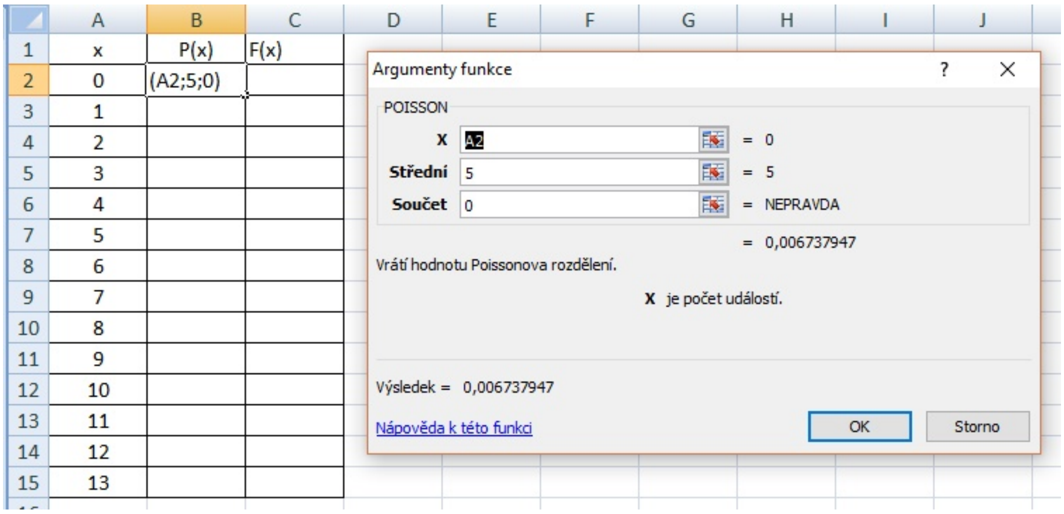
V této příloze si popíšeme postup při určování pravděpodobnosti hodnot náhodné veličiny s poissonovým rozdělením pravděpodobnosti. Postup si ukážeme na Příkladu 2.54 na straně 45 a vypočteme pravděpodobnosti pro některé další hodnoty.

2.70. K pokladně v supermarketu přistoupí průměrně 5 zákazníků během deseti minut. Určeme pravděpodobnosti několika prvních hodnot náhodné veličiny *počet zákazníků u pokladny během deseti minut.*

Řešení: Ze zadání plyne, že náhodná veličina *počet zákazníků u pokladny během deseti minut* má poissonovo rozdělení pravděpodobnosti, kde $\lambda = 5$. Začátek bude opět podobný jako v předešlých přílohách. Vytvoříme základní tabulku, zapíšeme několik prvních hodnot náhodné veličiny (obecně není počet zákazníků u pokladny shora ohraničen) a necháme program spočítat pravděpodobnosti těchto hodnot. Možné hodnoty budeme uvažovat $0, 1, 2, 3, 4, 5\ldots, 12, 13$; tyto napíšeme do jednotlivých buněk pod sebe.

Do druhého sloupce vložíme pravděpodobnosti příslušných hodnot následujícím způsobem. Označíme (vybereme) buňku vedle první hodnoty x (v našem případě se jedná o buňku A2), kliknutím na symbol f_x vyvoláme okno *Vložit funkci*, vybereme kategorii statistických funkcí a z nabídky funkcí vybereme položku POISSON.

Funkce POISSON má celkem tři argumenty k vyplnění, viz Obrázek 2.10.



Obrázek 2.10: Zadání hodnot funkce HYPGEOMDIST

První argument mná označení X (nápověda – počet událostí) - opět se jedná o hodnotu náhodné veličiny, pro kterou počítáme pravděpodobnost. Do políčka kliknutím na buňku A2 zadáme relativní odkaz na buňku, ve které se nachází hodnota x (odkaz lze zadat i ručně vypsáním).

Druhý argument má název *Střední* (nápověda – předpokládaná číselná hodnota, kladné číslo) a znamená střední (očekávanou) hodnotu náhodné veličiny. V našem případě v průměru chodí 5 zákazníků za 10 minut, je tedy doplněno číslo 5.

Poslední argument *Součet* určuje, zda-li chceme vypočítat pravděpodobnost dané hodnoty x, tj. hodnotu $P(\mathbb{X} = x)$, nebo hodnotu distribuční funkce pro dané x, tj. $F(x) = P(\mathbb{X} \leq x)$.

V prvním případě zadáme hodnotu 0, ve druhém případě zadáme hodnotu 1. Zde chceme vypočítat hodnotu pravděpodobnosti, zadali jsme číslo 0.

Vyplněné okno potvrdíme kliknutím na tlačítko OK. V buňce A2 se objevila příslušná hodnota pravděpodobnosti. Nyní vzorec zkopírujeme do ostatních buněk ke zbývajícím hodnotám náhodné veličiny např. tak, že kursorem najedeme na dolní pravý roh buňky A2. Kurzor se změní ze šipky na znaménko plus (+), klikneme na tento dolní roh a tím se vzorec nakopíruje do ostatních buněk.

Místo pravděpodobností jednotlivých hodnot bychom mohli také chtít vypočítat hodnoty distribuční funkce. Postup je skoro stejný, pouze při vyplňování posledního argumentu zadáme hodnotu 1.

Z výsledků v tabulce opět vyčteme mnoho informací, viz Obrázek 2.11. Například je zde uvedena informace, že

	A	B	C
1	x	P(x)	F(x)
2	0	0,00674	0,00674
3	1	0,03369	0,04043
4	2	0,08422	0,12465
5	3	0,14037	0,26503
6	4	0,17547	0,44049
7	5	0,17547	0,61596
8	6	0,14622	0,76218
9	7	0,10444	0,86663
10	8	0,06528	0,93191
11	9	0,03627	0,96817
12	10	0,01813	0,9863
13	11	0,00824	0,99455
14	12	0,00343	0,99798
15	13	0,00132	0,9993

Obrázek 2.11: Výsledky příkladu vypočtené pomocí HYPGEOMDIST

- pravděpodobnost příchodu tří zákazníků během deseti minut je rovna přibližně $0,14037$,
- pravděpodobnost příchodu nejvýše tří zákazníků během deseti minut je rovna přibližně $0,26503$,
- nejvyšší pravděpodobnost (modální hodnotu) má příchod čtyř nebo pěti zákazníků během deseti minut.

Zde naleznete videonávod k výpočtu.

Kapitola 3

Spojitá náhodná veličina

V této kapitole se budeme zabývat pojmem *spojitá náhodná veličina*, vlastnostmi spojité náhodné veličiny a některými důležitými rozděleními spojité náhodné veličiny. Budeme přitom vycházet z videopřednášek v Moodlu, které připravila přednášející dr. Šimsová - tzn. že byste tato videa měli mít shlédnuta, než začnete číst tento text. Po prostudování kapitoly byste měli být schopni:

- rozumět pojmu *spojitá náhodná veličina*,
- rozumět pojmu *hustota pravděpodobnosti* spojité náhodné veličiny,
- rozumět pojmu *distribuční funkce* náhodné veličiny,
- pracovat s *charakteristikami* spojité náhodné veličiny, mezi které patří zejména:
  - první obecný moment, tedy *střední hodnota* spojité náhodné veličiny,
  - druhý obecný moment spojité náhodné veličiny,
  - druhý centrovaný moment, neboli *rozptyl* spojité náhodné veličiny,
  - *směrodatná odchylka* spojité náhodné veličiny,
  - modální hodnota, neboli *modus* náhodné veličiny,
  - *p% kvantil* náhodné veličiny,
- pracovat s pojmem náhodná veličina s *rovnoměrným rozdělením* pravděpodobnosti,
- pracovat s pojmem náhodná veličina s *exponenciálním rozdělením* pravděpodobnosti,
- pracovat s pojmem náhodná veličina s *normálním rozdělením* pravděpodobnosti,

V úvodní kapitole o náhodné veličině jsme se naučili rozlišovat mezi diskrétní a spojitou náhodnou veličinou. Připomeňme hlavní rozdíl mezi oběma typy náhodných veličin. Ten spočívá v hodnotách, které může nabývat příslušná náhodná veličina. Diskrétní náhodná veličina mohla nabývat konečný i nekonečný počet hodnot, nicméně v případě nekonečného počtu hodnot bylo těchto hodnot spočetně mnoho. Zmínili jsme zákon rozdělení pravděpodobnosti diskrétní náhodné veličiny, který vyžadoval, aby součet pravděpodobností všech hodnot diskrétní náhodné veličiny byl roven jedné.

Nepovinná vsuvka o nekonečnu – pouze pro matematicky otrlé čtenáře Naše zkušenost s nekonečnem je omezená, a tak nás mohou některé jeho vlastnosti překvapit. Jednou z takových situací může být případ, kdy má veličina nekonečně mnoho hodnot, každá z těchto hodnot má (alespoň teoreticky) nenulovou pravděpodobnost, a přitom součet těchto pravděpodobností (tj. součet nekonečně mnoha kladných čísel) je roven jedné. V matematické teorii o nekonečných řadách čísel se lze seznámit s podmínkami, za kterých je součet nekonečně mnoha kladných čísel roven konečně velkému číslu.

Ostatně, toto by vás nemělo překvapit. Uvažujme například číslo $\frac{1}{3} = 0,\overline{3}$. Číslo

$$\begin{aligned} 0,\overline{3} &= 0,333\,333\,333\,33\ldots \\ &= 0 + 0,3 + 0,03 + 0,003 + 0,000\,3 + 0,000\,03 + \ldots \\ &= \frac{3}{10} + \frac{3}{100} + \frac{3}{1\,000} + \frac{3}{10\,000} + \frac{3}{100\,000} + \ldots \\ &= \sum_{n=1}^{\infty} \frac{3}{10^n} \end{aligned}$$

je jistě menší než jedna a přitom je zřejmé, že vzniká jako součet nekonečně mnoho kladných čísel. Principem konečného součtu je, že tato čísla se zmenšují dostatečně rychle, aby jejich součet vedl ke konečně velkému číslu.

Vlastnost, že součet nekonečně mnoha kladných čísel (pravděpodobností) může být roven jedné, však mají pouze ta čísla, kterých je nanejvýše spočetně mnoho, tj. množiny hodnot diskrétní náhodné veličiny. Intervaly reálných čísel mají nespočetné množství prvků a součet nespočetně mnoha kladných hodnot nebude konečně velký. To ale znamená, že pokud hodnoty náhodné veličiny tvoří interval reálných čísel, musí být pravděpodobnosti těchto hodnot rovny nule (záporné být nemohou z principu; kladné být nemohou, neboť jejich součet by nebyl konečně velký – takže zbývá pouze nulová hodnota pravděpodobnosti). Tedy pravděpodobnost každé hodnoty z intervalu reálných čísel je rovna nule.

Ostatně, ani toto by nás nemělo překvapit. Možná si vzpomenete, že spojitou náhodnou veličinou je například doba čekání na nějakou událost. Pravděpodobnost, že na tuto událost budeme čekat přesně 4,584 651 694 215 362 948 321 362 749 362 215 594 36 sekundy je prostě nulová. Je proto dobré zamyslet se nad významem následujícího tvrzení. Pravděpodobnost nemožného jevu je rovna nule, na druhou stranu obrácené tvrzení neplatí – to, že nějaký jev má pravděpodobnost rovnu nule neznamená, že se jedná o nemožný jev. Při náhodném pokusu se spojitou náhodnou veličinou vždy nějaká hodnota nastane, přestože pravděpodobnost její události je rovna nule. *Konec vsuvky o nekonečnu.*

Definice 3.0.1. Náhodnou veličinu, která může nabývat libovolnou hodnotu z konečného nebo nekonečného intervalu reálných čísel, nazýváme *spojitá náhodná veličina.*

V nepovinné vsuvce jsme vysvětlili, že u spojité náhodné veličiny nemá smysl pracovat s pravděpodobnostní funkcí, neboť všechny její hodnoty jsou rovny nule. Z toho důvodu nejdříve začneme definicí pojmu distribuční funkce.

Definice 3.0.2. Distribuční funkce náhodné veličiny $\mathbb{X}$ je pro každé reálné číslo x definována vztahem

$$F(x) = P(\mathbb{X} \leq x). \tag{3.1}$$

Hodnota distribuční funkce v bodě x tedy znamená pravděpodobnost, že náhodná veličina $\mathbb{X}$ má hodnotu nejvýše x, tj. pravděpodobnost, že náhodná veličina $\mathbb{X}$ nabude některou z hodnot od $-\infty$ do x.

Z úvodního textu o pravděpodobnosti byste si měli pamatovat, že pravděpodobnost sjednocení vzájemně neslučitelných jevů je rovna součtu jednotlivých pravděpodobností těchto jevů; tedy, jsou-li jevy A_1, A_2, ..., A_n vzájemně neslučitelné, potom platí

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Z této rovnosti vyplývá několik závěrů.

- Snadno nahlédneme platnost rovnosti $(-\infty, x\rangle = (-\infty, x) \cup \{x\}$. Jevy $\mathbb{X} < x$ a $\mathbb{X} = x$ jsou jistě neslučitelné (hodnota náhodné veličiny $\mathbb{X}$, která je výsledkem konkrétního náhodného pokusu, je prostě buď menší než dané x, nebo je rovna x, nebo je větší než x, ale nemůže být současně menší než x a současně rovna x). Proto je

  $$P(\mathbb{X} \leq x) = P\big((\mathbb{X} < x) \cup (\mathbb{X} = x)\big) = P\big(\mathbb{X} < x\big) + P(\mathbb{X} = x) = P\big(\mathbb{X} < x\big) + 0 = P\big(\mathbb{X} < x\big),$$

  kde rovnost $P(\mathbb{X} = x) = 0$ plyne ze znění nepovinné matematické vsuvky. Tím jsme zdůvodnili platnost rovnosti $P(\mathbb{X} < x) = P(\mathbb{X} \leq x)$, kde $\mathbb{X}$ je spojitá náhodná veličina. Tuto rovnost budeme občas využívat, proto by vás neměla překvapit případná „nedůslednost“ v práci s (ne)ostrou nerovností $\mathbb{X} < x$, resp. $\mathbb{X} \leq x$, neboť pravděpodovnosti obou jevů se rovnají. Z toho důvodu se také můžete v některých případech setkat s definicí distribuční funkce ve tvaru $F(x) = P(\mathbb{X} < x)$. Tato poznámka ukazuje, že takto vyslovená definice není v rozporu s naší definicí (v případě spojité náhodné veličiny).

  Poznamenejme, že analogicky můžeme pro spojitou náhodnou veličinu $\mathbb{X}$ zdůvodnit rovnost $P(x < \mathbb{X}) = P(x \leq \mathbb{X})$.

- V tomto odstavci se budeme věnovat důsledkům rovnosti $(-\infty, a\rangle \cup (a, b\rangle = (-\infty, b\rangle$, pro $a \leq b$. Opět snadno nahlédneme, že jevy $\mathbb{X} \leq a$ a $a < \mathbb{X} \leq b$ jsou neslučitelné (hodnota náhodné veličiny $\mathbb{X}$ je prostě buď nejvýše rovna a, nebo je větší než a). Proto je

  $$P(-\infty < \mathbb{X} \leq b) = P\Big((-\infty < \mathbb{X} \leq a) \cup (a < \mathbb{X} \leq b)\Big) = P\big(-\infty < \mathbb{X} \leq a\big) + P\big(a < \mathbb{X} \leq b\big).$$

  Tuto nerovnost můžeme přepsat do tvaru

  $$P(a < \mathbb{X} \leq b) = P(-\infty < \mathbb{X} \leq b) - P(-\infty < \mathbb{X} \leq a) = F(b) - F(a).$$

  Společně s tvrzením v předchozí odrážce tak pro spojitou náhodnou veličinu $\mathbb{X}$ dostáváme následující rovnosti.

  $$\begin{aligned} F(b) - F(a) &= P(a < \mathbb{X} < b) \\ &= P(a \leq \mathbb{X} < b) \\ &= P(a < \mathbb{X} \leq b) \\ &= P(a \leq \mathbb{X} \leq b) \end{aligned}$$

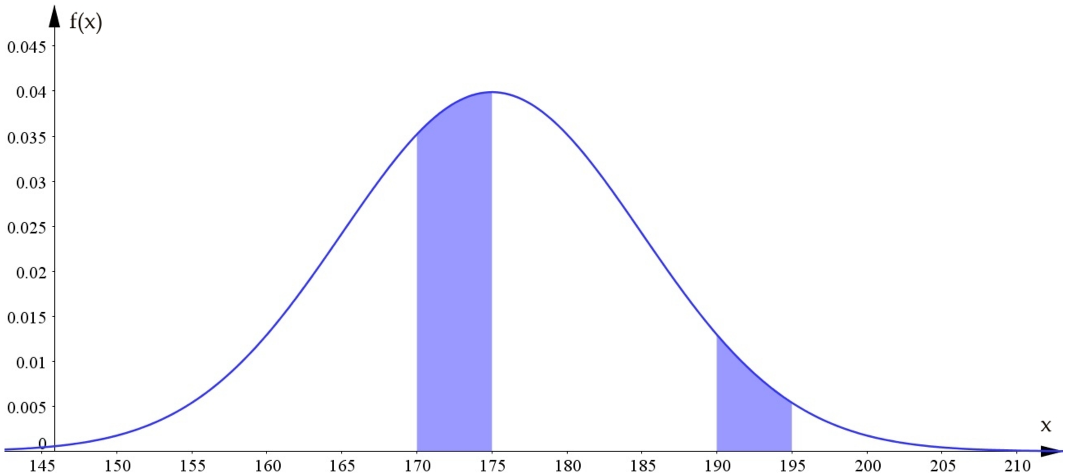
V další části si přiblížíme důležitý pojem *hustota pravděpodobnosti*. K tomuto pojmu nás vede následující představa. I když je pravděpodobnost každé konkrétní hodnoty spojité náhodné veličiny rovna nule, přesto můžeme mít tendenci některé hodnoty náhodné veličiny očekávat častěji než jiné hodnoty.

Uvažujme například náhodnou veličinu, kterou je tělesná výška náhodně vybraného dospělého člověka. Jistě se shodneme na tom, že pravděpodobnost, že náhodně vybraný dospělý člověk bude mít výšku 175 cm, tj. 175,000 000 000 000 000 cm, stejně tak jako 325,154 127 368 145 234 cm, je rovna nule. Nicméně – asi se shodneme na tom, že výška náhodně vybraného dospělého člověka se bude pohybovat spíše v blízkosti čísla 175 cm než v blízkosti 325 cm – a tuto míru přesvědčení bychom potřebovali nějak číselně vyjádřit.

Vytvoříme tedy funkci, jejíž funkční hodnoty budou vyšší u hodnot, v jejichž okolí předpokládáme častější výskyt hodnot náhodné veličiny. To by k přesnému vymezení funkce nestačilo, proto přidáme další podmínku. Tuto funkci sestrojíme tak, aby její funkční hodnoty měly nějaký vztah k pravděpodobnostem souvisejícím s danou náhodnou veličinou. Tento vztah může být zvolen různě – my si zvolíme tu možnost, aby plocha A ohraničená shora grafem této funkce, zdola osou x a zleva, resp. zprava ohraničená svislou přímkou protínající na ose x bod a, resp. bod b odpovídala svým obsahem $S(A)$ pravděpodobnosti, že hodnoty náhodné veličiny $\mathbb{X}$ leží v rozmezí od a do b; tedy, aby platilo

$$S(A) = P(a < \mathbb{X} < b) = P(a \leq \mathbb{X} < b) = P(a < \mathbb{X} \leq b) = P(a \leq \mathbb{X} \leq b) = F(b) - F(a). \quad (3.2)$$

Takto pojatá funkce je již dána jednoznačně. Podívejme se na Obrázek 3.1. Zde je zobrazen



Obrázek 3.1: Hustota pravděpodobnosti

graf funkce $f(x)$ a dále jsou uvedeny dvě různé plochy (modré pruhy), které jsou ohraničeny shora grafem funkce $f(x)$, zdola osou x a zleva a zprava svislými přímkami, které na ose x protínají hodnoty 170 a 175, resp. 190 a 195. Obsahy těchto ploch jsou rovny pravděpodobnostem, že hodnota náhodné veličiny bude ležet v rozmezí od 170 do 175 (levý pruh), resp v rozmezí od 190 do 195 (pravý pruh). Pokud bychom za hodnotu náhodné veličiny považovali výšku náhodně vybraného dospělého člověka, potom by obsahy obou ploch znamenaly pravděpodobnost, že náhodně vybraný dospělý člověk bude mít výšku v rozmezí od 170 do 175 cm (obsah levého pruhu), resp. bude mít výšku v rozmezí od 190 do 195 cm (obsah pravého pruhu). Všimněte si, že oba pruhy mají stejnou šířku 5 cm; levý pruh je vyšší – jeho obsah je větší a to znamená vyšší pravděpodobnost výskytu osob s výškou od 170 do 175 cm oproti osobám s výškou v rozmezí od 190 do 195 cm.

Takto definovaná funkce tedy naznačuje, jaké hodnoty náhodné veličiny můžeme očekávat, a přímo poskytuje návod k výpočtu pravděpodobností pomocí obsahu jisté plochy. Z kurzu matematiky byste si měli pamatovat, že obsah plochy A ohraničené shora grafem funkce $f(x)$ – spojité a nezáporné v intervalu $\langle a, b\rangle$, zdola osou x, zleva a zprava ohraničené svislými přímkami protínajícími na ose x body a a b, lze vypočítat pomocí vzorce

$$S(A) = \int_a^b f(x)\,\mathrm{d}x = F(b) - F(a).$$

Funkce $F(x)$ je tzv. primitivní funkce k funkce $f(x)$ na intervalu (a, b), tj. pro všechna $x \in (a, b)$ je $F'(x) = f(x)$.

My již víme – viz rovnice (3.2), že obsah plochy $S(A)$ odpovídá rozdílu hodnot distribuční funkce $F(b) - F(a)$, proto distribuční funkci $F(x) = P(\mathbb{X} \le x)$ můžeme považovat za primitivní funkci k naší uvažované funkci $f(x)$ a vztah mezi touto funkcí $f(x)$ a distribuční funkcí $F(x)$ je $F'(x) = f(x)$. Nyní již můžeme vyslovit následující definici.

Definice 3.0.3. Je-li $\mathbb{X}$ spojitá náhodná veličina, potom existuje nezáporná funkce $f(x)$ taková, že pro všechna reálná čísla x platí

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t. \tag{3.3}$$

Funkci $f(x)$ nazýváme *hustota pravděpodobnosti* spojité náhodné veličiny $\mathbb{X}$.

Hustota pravděpodobnosti $f(x)$ má tyto vlastnosti:

- $f(x) \ge 0$
- $\displaystyle\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$
- $P(a \le \mathbb{X} \le b) = F(b) - F(a) = \displaystyle\int_a^b f(x)\,\mathrm{d}x$ pro $a \le b$, kde $F(x)$ je distribuční funkce náhodné veličiny $\mathbb{X}$.

Později se seznámíme s některými speciálními typy spojitých náhodných veličin a poznáme, jaký je předpis hustoty pravděpodobnosti v těchto případech.

3.0.1 Číselné charakteristiky spojité náhodné veličiny

Mějme spojitou náhodnou veličinu $\mathbb{X}$ s hustotou pravděpodobnosti $f(x)$, kde M představuje množinu všech možných hodnot x náhodné veličiny $\mathbb{X}$. Potom lze zavést tyto číselné veličiny, které charakterizují spojitou náhodnou veličinu.

Střední hodnota náhodné veličiny $\mathbb{X}$ je definována vztahem

$$E(\mathbb{X}) = \int_M x \cdot f(x)\,\mathrm{d}x. \tag{3.4}$$

Druhý obecný moment náhodné veličiny $\mathbb{X}$ je definován vztahem

$$E(\mathbb{X}^2) = \int_M x^2 \cdot f(x)\,\mathrm{d}x. \tag{3.5}$$

Rozptyl náhodné veličiny $\mathbb{X}$ je definován vztahem

$$D(\mathbb{X}) = E\Big\{\big[\mathbb{X} - E(\mathbb{X})\big]^2\Big\} = \int_M \big[x - E(\mathbb{X})\big]^2 \cdot f(x)\,\mathrm{d}x = E(\mathbb{X}^2) - \big(E(\mathbb{X})\big)^2. \tag{3.6}$$

Směrodatná odchylka náhodné veličiny $\mathbb{X}$ je definována vztahem

$$\sigma = \sqrt{D(\mathbb{X})}. \tag{3.7}$$

Medián $\tilde{x}$ je taková hodnota náhodné veličiny $\mathbb{X}$, pro kterou platí

$$P(\mathbb{X} \le \tilde{x}) = P(\mathbb{X} \ge \tilde{x}) = 0,5. \tag{3.8}$$

Význam uvedených charakteristik je stejný, jako u diskrétní náhodné veličiny. Střední hodnota $E(\mathbb{X})$, někdy též *očekávaná hodnota*, představuje hodnotu, ke které se bude blížit průměr z realizací náhodných pokusů při rostoucím počtu opakování těchto pokusů. Rozptyl, potažmo směrodatná odchylka, představují míru variability výsledků náhodné veličiny. Čím menší je tato variabilita (vzhledem k nezápornosti těchto měr to znamená – čím více se variabilita blíží nule), tím více se výsledky náhodné veličiny blíží střední hodnotě. Jakým způsobem chápat toto přibližování si opět řekneme při aplikacích limitních vět.

Konkrétními výpočty uvedených charakteristik si budeme zabývat při probírání již zmíněných speciálních typů spojitých náhodných veličin. Zde uvedeme pouze jeden příklad na výpočet uvedených charakteristik.

3.1. Je dána funkce

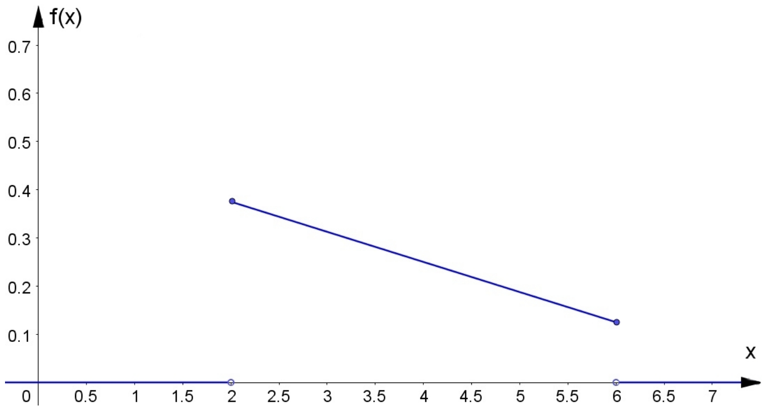
$$f(x) = \begin{cases} \dfrac{1}{2} - \dfrac{x}{16}, & \text{pro } x \in \langle 2,\, 6\rangle, \\ 0, & \text{pro zbývající hodnoty } x. \end{cases}$$

Ověřte, zda uvedená funkce může být hustotou pravděpodobnosti spojité náhodné veličiny. Pokud ano, vypočtěte střední hodnotu, rozptyl a směrodatnou odchylku a medián této náhodné veličiny.

Řešení: Graf funkce je uveden na Obrázku 3.2. Ze znázornění grafu je zřejmé, že náhodná veličina může nabývat pouze hodnoty z intervalu $\langle 2,\, 6\rangle$, přičemž hodnoty v pravostranném okolí čísla 2 můžeme očekávat o něco častěji než hodnoty v levostranném okolí čísla 6 (mají vyšší hustotu pravděpodobnosti).

Nejprve ověříme, zda uvedená funcke může být hustotou pravděpodobnosti spojité náhodné veličiny, tj zda splňuje podmínky $f(x) \ge 0$ pro všechna $x \in \mathbb{R}$ a $\displaystyle\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$. Z grafu funkce je zřejmé, že funkce je nezáporná na celém definičním oboru – první podmínka je tedy splněna. Nyní ověříme druhou podmínku. Vzhledem k tomu, že funkce nabývá nenulových hodnot pouze pro $x \in \langle 2,\, 6\rangle$ lze integrační meze omezit pouze na tuto množinu. Pak platí

$$\int\limits_{-\infty}^{\infty} f(x)\mathrm{d}x = \int\limits_{2}^{6} \left(\frac{1}{2} - \frac{x}{16}\right)\mathrm{d}x = \left[\frac{x}{2} - \frac{x^2}{32}\right]_2^6 = \left(\frac{6}{2} - \frac{36}{32}\right) - \left(\frac{2}{2} - \frac{4}{32}\right) = \frac{6-2}{2} - \frac{36-4}{32} = 1.$$



Obrázek 3.2: Hustota pravděpodobnosti

Druhá podmínka, která představuje zákon rozdělení pravděpodobnosti spojité náhodné veličiny, je také splněna, a funkce $f(x)$ je tedy hustotou pravděpodobnosti nějaké náhodné veličiny.

Nyní vypočteme střední hodnotu této náhodné veličiny. Použijeme přitom vzorec (3.4). Jak bylo zmíněno výše, množina M všech možných hodnot náhodné veličiny je rovna intervalu $\langle 2, 6\rangle$, proto integrační meze ve vzorci (3.4) budou odpovídat tomuto intervalu.

$$E(\mathbb{X}) = \int_M x \cdot f(x)\,\mathrm{d}x = \int\limits_2^6 x \cdot \left(\frac{1}{2} - \frac{x}{16}\right)\,\mathrm{d}x = \int\limits_2^6 \left(\frac{x}{2} - \frac{x^2}{16}\right)\,\mathrm{d}x = \left[\frac{x^2}{4} - \frac{x^3}{48}\right]_2^6$$
$$= \left(\frac{36}{4} - \frac{216}{48}\right) - \left(\frac{4}{4} - \frac{8}{48}\right) = \frac{36-4}{4} - \frac{216-8}{48} = \frac{176}{48} = \frac{11}{3} = 3,\overline{6}$$

Poznámka na okraj Tento výsledek je v souladu s naším očekáváním. Kdyby byly všechny hodnoty od 2 do 6 stejně očekávatelné, střední hodnotou by byl střed intervalu – tedy číslo 4. Nicméně, hodnoty v blízkosti čísla 2 mají vyšší hustotu pravděpodobnosti, jsou tedy očekávatelnější v tom smyslu, že při realizaci náhodných pokusů budou nastávat častěji, proto se průměr z realizací poněkud sníží a bude lehce nižší než 4 – což se stalo.

Dále budeme pokračovat výpočtem rozptylu. K tomu použijeme vzorec na pravé straně rovnice (3.6). Nejdříve proto podle rovnice (3.5) vypočteme druhý obecný moment $E(\mathbb{X}^2)$.

$$E(\mathbb{X}^2) = \int_M x^2 \cdot f(x)\,\mathrm{d}x = \int\limits_2^6 x^2 \cdot \left(\frac{1}{2} - \frac{x}{16}\right)\,\mathrm{d}x = \int\limits_2^6 \left(\frac{x^2}{2} - \frac{x^3}{16}\right)\,\mathrm{d}x = \left[\frac{x^3}{6} - \frac{x^4}{64}\right]_2^6$$
$$= \left(\frac{216}{6} - \frac{1\,296}{64}\right) - \left(\frac{8}{6} - \frac{16}{64}\right) = \frac{216-8}{6} - \frac{1\,296-16}{64} = \frac{44}{3} = 14,\overline{6}$$

Nyní již známe hodnoty $E(\mathbb{X}^2)$ a $E(\mathbb{X})$ a můžeme tedy dosadit do rovnice (3.6).

$$D(\mathbb{X}) = E(\mathbb{X}^2) - \Big(E(\mathbb{X})\Big)^2 = \frac{44}{3} - \left(\frac{11}{3}\right)^2 = \frac{132}{3} - \frac{121}{9} = \frac{11}{9} = 1,\overline{2}$$

Směrodatnou odchylku určíme ze vzorce (3.7).

$$\sigma = \sqrt{D(\mathbb{X})} = \sqrt{\frac{11}{9}} \doteq 1,106$$

Výpočtem jsme zjistili, že rozptyl je roven $D(\mathbb{X}) = 1,\overline{2}$ a směrodatná odchylka má hodnotu $\sigma \doteq 1,106$.

Na závěr příkladu vypočteme medián $\widetilde{x}$ této náhodné veličiny. Pro medián platí rovnost (3.8), proto vzhledem k rovnosti (3.1) můžeme psát $F(\widetilde{x}) = 0,5$. Vyjádříme předpis distribuční funkce a vypočteme hodnotu x, pro kterou je $F(x) = 0,5$ – to bude hledaný medián náhodné veličiny. Předpis distribuční funkce určíme z hustoty pravděpodobnosti pomocí rovnice (3.3). Vzhledem k tomu, že náhodná veličina nabývá nenulové hodnoty pouze pro $\langle 2,\, 6\rangle$, bude pro všechna $x \in (-\infty,\, 2)$ platit $F(x) = 0$ a pro všechna $x \in (6,\, \infty)$ vztah $F(x) = 1$. Níže vypočteme předpis funkce $F(x)$ pro $x \in \langle 2,\, 6\rangle$.

$$F(x) = \int\limits_{-\infty}^{x} f(t)\,\mathrm{d}t = \int\limits_{2}^{x} f(t)\,\mathrm{d}t = \int\limits_{2}^{x} \left(\frac{1}{2} - \frac{t}{16}\right)\,\mathrm{d}t = \left[\frac{t}{2} - \frac{t^2}{32}\right]_2^x = \left(\frac{x}{2} - \frac{x^2}{32}\right) - \left(\frac{2}{2} - \frac{4}{32}\right)$$
$$= \frac{x-2}{2} - \frac{x^2-4}{32} = \frac{16x-32}{32} - \frac{x^2-4}{32} = \frac{-x^2+16x-28}{32}$$

Pro předpis distribuční funkce tedy platí

$$F(x) = \begin{cases} 0, & \text{pro } x < 2, \\ \dfrac{-x^2+16x-28}{32}, & \text{pro } x \in \langle 2,\, 6\rangle, \\ 1, & \text{pro } x > 6. \end{cases}$$

Ptáme se, pro jakou hodnotu x je funkční hodnota této funkce rovna $0,5$.

$$\begin{aligned} F(x) = \frac{-x^2+16x-28}{32} &= 0,5 && \ldots|\cdot 32 \\ -x^2+16x-28 &= 16 && \ldots|-16 \\ -x^2+16x-44 &= 0 && \ldots|\ \textit{nalezneme kořeny kvadratické rovnice} \\ x &= 8 \pm \sqrt{20} && \ldots|\ \textit{bereme kořeny pouze z intervalu } \langle 2,\, 6\rangle \\ x &= 8 - \sqrt{20} \doteq 3,53 && \ldots|\ \textit{což je hodnota hledaného mediánu} \end{aligned}$$

Medián náhodné veličiny $\mathbb{X}$ se zadanou hustotou pravděpodobnosti $f(x)$ je roven $\widetilde{x} \doteq 3,53$. Poznamenejme, že analogicky bychom postupovali při výpočtu ostatních kvantilů – pouze místo hodnoty $0,5$ bychom pracovali s hodnotou p (u $p\%$ kvantilu).

Úlohy k samostatnému řešení

3.2. Je dána funkce $f(x)$, jejíž předpis je roven $f(x) = 0,2$ pro $x \in \langle 2,\, 7\rangle$, resp. $f(x) = 0$ pro $x < 2$ a $x > 7$. Rozhodněte a zdůvodněte, zda takto definovaná funkce může být hustotou pravděpodobnosti náhodné veličiny. Jestliže ano, určete její střední hodnotu, rozptyl a směrodatnou odchylku a medián.

3.3. Je dána funkce $f(x)$, jejíž předpis je roven $f(x) = \frac{3}{4} - \frac{x}{4}$ pro $x \in \langle 0,\, 4\rangle$, resp. $f(x) = 0$ pro $x < 0$ a $x > 4$. Rozhodněte a zdůvodněte, zda takto definovaná funkce může být hustotou pravděpodobnosti náhodné veličiny. Jestliže ano, určete její střední hodnotu, rozptyl a směrodatnou odchylku a medián.

3.4. Je dána funkce $f(x)$, jejíž předpis je roven $f(x) = 3x^2 - 6x + 3$ pro $x \in \langle 0, 1\rangle$, resp. $f(x) = 0$ pro $x < 0$ a $x > 1$. Rozhodněte a zdůvodněte, zda takto definovaná funkce může být hustotou pravděpodobnosti náhodné veličiny. Jestliže ano, určete její střední hodnotu, rozptyl a směrodatnou odchylku a medián.

3.1 Některá důležitá rozdělení spojité náhodné veličiny

3.1.1 Rovnoměrné rozdělení spojité náhodné veličiny

Rovnoměrné rozdělení používáme k modelování situací, kdy obor hodnot náhodné veličiny je oboustranně omezen (je to tedy konečný interval) a není možné předpokládat, že některé hodnoty z tohoto intervalu by se vyskytovaly častěji než jiné. To znamená, že pravděpodobnost, že hodnota náhodné veličiny pochází z libovolného podintervalu od α do β je stejná, jako že pochází z jakéhokoliv jiného podintervalu stejné délky.

Příkladem náhodné veličiny s tímto rozdělením je doba čekání na pravidelně se opakující jev. Předpokládejme například, že autobosová linka přijíždí pravidelně na zastávku v desetiminutových intervalech. Náhodnou veličinou je doba, která uplyne od našeho příchodu na zastávku do doby, než přijede první autobus této linky. Předpokládáme přitom, že svůj příchod nijak nesynchronizujeme s jízdním řádem a na zastávku přijdeme nezávisle na časech příjezdu autobusu. Pak tato náhodná veličina má rovnoměrné rozdělení pravděpodobnosti s možnými hodnotami z intervalu $(0, 10)$ minut.

Jiným příkladem je velikost zaokrouhlovací chyby při měření. Dejme tomu, že měříme délku výrobku s přesností 1 mm a zaokrouhlování provádíme podle matematických pravidel. Tedy pokud má předmět skutečnou délku 45,23 mm zapíšeme si délku 45 mm a chyba měření činí 0,23 mm. Pokud má předmět skutečnou délku 148,76 mm zapíšeme si délku 149 mm a chyba měření činí 0,24 mm. Náhodnou veličinou je velikost chyby, které se dopouštíme při každém měření. Pak tato chyba je náhodná veličina s rovnoměrným rozdělením s možnými hodnotami $(0, 0.5)$ milimetrů.

Definice 3.1.1. *Rovnoměrným rozdělením* spojité náhodné veličiny $\mathbb{X}$, která nabývá hodnot $x \in (a, b)$, nazveme takové rozdělení pravděpodobnosti, jehož hustota je

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

Pro $x \in (a, b)$ je

$$F(x) = \int_a^x \frac{1}{b-a}\,\mathrm{d}t = \left[\frac{t}{b-a}\right]_a^x = \frac{x}{b-a} - \frac{a}{b-a} = \frac{x-a}{b-a}.$$

Distribuční funkce je potom popsána rovnicemi

$$F(x) = \begin{cases} 0, & x \le a, \\ \frac{x-a}{b-a}, & x \in (a, b), \\ 1, & x \ge b. \end{cases}$$

Střední hodnotu vypočteme pomocí rovnice (3.4).

$$E(\mathbb{X}) = \int\limits_a^b x \cdot f(x)\,\mathrm{d}x = \int\limits_a^b x \cdot \frac{1}{b-a}\,\mathrm{d}x = \int\limits_a^b \frac{x}{b-a}\,\mathrm{d}x = \left[\frac{x^2}{2(b-a)}\right]_a^b$$
$$= \frac{b^2}{2(b-a)} - \frac{a^2}{2(b-a)} = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}.$$

Pro výpočet druhého obecného momentu použijeme rovnici (3.5).

$$E(\mathbb{X}^2) = \int\limits_a^b x^2 \cdot f(x)\,\mathrm{d}x = \int\limits_a^b \frac{x^2}{b-a}\,\mathrm{d}x = \left[\frac{x^3}{3(b-a)}\right]_a^b$$
$$= \frac{b^3}{2(b-a)} - \frac{a^3}{2(b-a)} = \frac{b^3-a^3}{2(b-a)} = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{a^2+ab+b^2}{3}.$$

Rozptyl určíme pomocí vzorce (3.6).

$$D(\mathbb{X}) = E(\mathbb{X}^2) - (E(\mathbb{X}))^2 = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2$$
$$= \frac{4a^2+4ab+4b^2}{12} - \frac{3a^2+6ab+3b^2}{12} = \frac{a^2-2ab+b^2}{12} = \frac{(a-b)^2}{12}.$$

Řešené úlohy

3.5. Představme si „ideální“ systém městské hromadné dopravy, ve které určitá linka přijíždí na zastávku v naprosto přesném intervalu 10 minut. Náhodnou veličinou $\mathbb{X}$ je *doba čekání* na tuto linku při náhodném příchodu na zastávku.

a) Jaké jsou možné hodnoty této náhodné veličiny?

b) Namalujte a popište graf hustoty pravděpodobnosti této náhodné veličiny.

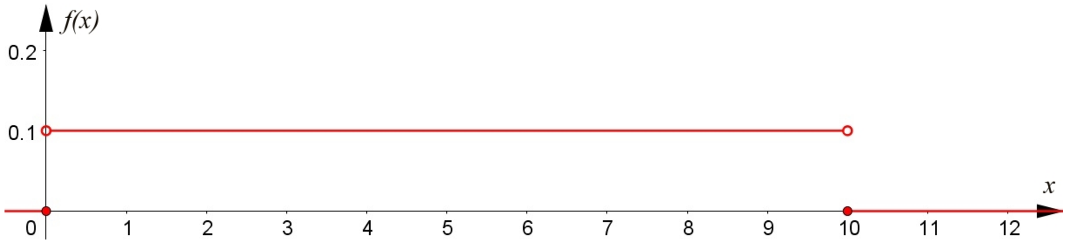
c) Jaká je pravděpodobnost, že na autobus budete čekat od 2 do 7 minut?

d) Jaká je pravděpodobnost, že na autobus budete čekat více než 6 minut?

Řešení:
Jedná se samozřejmě o idealizovanou úlohu, neboť ve skutečnosti nelze zajistit, aby autobus přijížděl přesně každých deset minut. Navíc úloha nepočítá s tím, že autobus se ve stanici chvíli zdrží atd. Nicméně, pokusíme se toto zjednodušení přijmout a úlohu řešit tak, jako kdyby se autobus každých deset minut objevil ve stanici a hned zase „zmizel“, takže ten, kdo přijde bezprostředně po zmizení autobusu, bude na další bus čekat 10 minut.

a) Hodnoty této náhodné veličiny představují všechny myslitelné doby čekání na autobus, tedy všechna $x \in (0,\ 10)$ (v minutách).

b) Graf hustoty pravděpodobnosti náhodné veličiny bude konstantní kladná funkce pro všechna $x \in (0, 10)$. Mimo tento interval nabývá hustota pravděpodobnosti nulové hodnoty. Funkční hodnota hustoty pravděpodobnosti v intervalu (0, 10) je rovna
$$\frac{1}{b-a} = \frac{1}{10-0} = 0,1.$$
Graf naleznete na Obrázku 3.3.



Obrázek 3.3: Hustota pravděpodobnosti v Příkladu 3.5

c) Pravděpodobnost, že na autobus budeme čekat od 2 do 7 minut je rovna pravděpodobnosti, že náhodná veličina bude nabývat hodnoty z intervalu od 2 do 7 minut.

$$P(2 < \mathbb{X} < 7) = F(7) - F(2) = \frac{7-0}{10-0} - \frac{2-0}{10-0} = \frac{7}{10} - \frac{2}{10} = \frac{5}{10} = \frac{1}{2} = 0,5.$$

Stejně tak bylo možné výpočet provést pomocí integrálů (zejména pokud rádi integrujete).

$$P(2 < \mathbb{X} < 7) = \int_2^7 f(x)\,\mathrm{d}x = \int_2^7 \frac{1}{10}\,\mathrm{d}x = \left[\frac{x}{10}\right]_2^7 = \frac{7}{10} - \frac{2}{10} = \frac{5}{10} = 0,5.$$

d) Pravděpodobnost, že na autobus budeme čekat čekat více než 6 minut je rovna pravděpodobnosti, že náhodná veličina bude nabývat hodnoty z intervalu od 6 do 10 minut.

$$P(6 < \mathbb{X} < 10) = F(10) - F(6) = \frac{10-0}{10-0} - \frac{6-0}{10-0} = \frac{10}{10} - \frac{6}{10} = \frac{4}{10} = 0,4.$$

Výpočet bylo možné provést také pomocí opačného jevu.

$$P(\mathbb{X} > 6) = 1 - P(\mathbb{X} \le 6) = 1 - F(6) = 1 - \frac{6-0}{10-0} = 1 - 0,6 = 0,4.$$

Výpočet bylo možné provést i pomocí integrálů (opět zejména pokud rádi integrujete).

$$P(\mathbb{X} > 6) = \int_6^{10} f(x)\,\mathrm{d}x = \int_6^{10} \frac{1}{10}\,\mathrm{d}x = \left[\frac{x}{10}\right]_6^{10} = \frac{10}{10} - \frac{6}{10} = \frac{4}{10} = 0,4.$$

Úlohy k samostatnému řešení

3.6. Pro náhodnou veličinu s rovnoměrným rozdělením platí rovnost $E(\mathbb{X}) = \widetilde{x}$. Zdůvodněte proč.

3.7. Generátor náhodných čísel v PC „vytváří“ náhodná čísla z rozmezí $\langle 0, 1\rangle$. S jakou pravděpodobností bude vylosováno číslo

a) $x = 0,3251$

b) menší než $x = 0,4$

c) číslo větší než $x = 0,7$

d) číslo z rozmezí 0,4 až 0,7?

3.1.2 Exponenciální rozdělení náhodné veličiny

Tato kapitola představuje vzhledem k sdělení dr. Šimsové v rámci videopřednášky text nepovinný ke studiu v tom smyslu, že z něj nebudete zkoušeni ani jinak testováni. Na druhou stranu, řada praktických úloh například z teorie hromadné obsluhy (často příchod požadavku do systému – což může být příchod zákazníka, potřeba jistého dílu ve výrobě atd.) je popsána právě touto náhodnou veličinou a byla by škoda se s tématem alespoň trochu neseznámit.

Časový interval $\mathbb{X}$ mezi dvěma sousedními výskyty událostí v Poissonovském procesu (viz kapitola **??** na straně **??**) je náhodná veličina s tzv. *exponenciálním rozdělením pravděpodobnosti.*

3.8. Průměrná doba mezi průjezdy dvou po sobě jedoucích automobilů po silnici před naší školou v době konání cvičení ze statistiky je 1 minuta. Náhodnou veličinou je doba mezi průjezdy dvou po sobě jedoucích automobilů po silnici před naší školou v době konání cvičení ze statistiky. Možné hodnoty náhodné veličiny teoreticky tvoří interval $M = (0, \infty)$, prakticky je to interval $M = (0, 240)$ (měříme v minutách).

3.9. Průměrná doba mezi dvěma telefonáty na zákaznickou linku firmy ABC je 20 sekund. Náhodnou veličinou je doba mezi dvěma telefonáty. Možné hodnoty náhodné veličiny tvoří interval $M = (0, \infty)$.

Definice 3.1.2. Uvažujme spojitou náhodnou veličinu $\mathbb{X}$, kde $\mathbb{X}$ je doba do nastání události v poissonovském procesu. Pak tato náhodná veličina má *exponenciální rozdělení pravděpodobnosti* $\mathrm{Exp}(\delta)$, a její hustota pravděpodobnosti má předpis

$$f(x) = \begin{cases} \dfrac{1}{\delta}\,\mathrm{e}^{-x/\delta}, & x \in (0,\infty),\ \delta > 0, \\ 0, & \text{jinak.} \end{cases}$$

Parametr δ představuje střední hodnotu této náhodné veličiny. Je tedy $E(\mathbb{X}) = \delta$, také platí $D(\mathbb{X}) = \delta^2$.

Vzhledem k definici můžeme odvodit vztah pro distribuční funkci. Pro všechna $x > 0$ je

$$F(x) = \int_0^x f(t)\,\mathrm{d}t = \int_0^x \frac{1}{\delta}\,\mathrm{e}^{-t/\delta}\,\mathrm{d}t = -\int_0^x -\frac{1}{\delta}\,\mathrm{e}^{-t/\delta}\,\mathrm{d}t = \left| \begin{array}{rcl} -\frac{t}{\delta} & = & s \\ -\frac{1}{\delta}\,\mathrm{d}t & = & \mathrm{d}s \end{array} \right|$$

$$= -\int_0^{-x/\delta} \mathrm{e}^s\,\mathrm{d}s = \Big[-\mathrm{e}^s\Big]_0^{-x/\delta} = \Big(-\mathrm{e}^{-x/\delta}\Big) - \big(-\mathrm{e}^0\big) = 1 - \mathrm{e}^{-x/\delta}.$$

Pokud se vám zdál výpočet obtížný (přeci jenom jsme zde použili substituční metodu v určitém integrálu) nevěšte hlavu, důležité je vnímat výsledek; předpis distribuční funkce náhodné veličiny s exponenciálním rozdělením zní

$$F(x) = \begin{cases} 0, & x \le 0, \\ 1 - \mathrm{e}^{-x/\delta}, & x \in (0,\infty). \end{cases}$$

Řešené úlohy

3.10. V jisté restauraci čekáme na obsloužení průměrně 9 minut. Jaká je pravděpodobnost, že budeme obslouženi do 4 minut?

Řešení:
Náhodnou veličinu *doba čekání od příchodu do restaurace do příchodu obsluhy* můžeme považovat za náhodnou veličinu s exponenciálním rozdělením, kde $\delta = 9$ (v průměru čekáme na obsloužení 9 minut). Pravděpodobnost, že budeme obslouženi do 4 minut je potom rovna pravděpodobnosti, že hodnota náhodné veličiny bude nejvýše rovna 4, tj. počítáme $P(\mathbb{X} \leq 4)$. Z definice distribuční funkce víme, že tato pravděpodobnost odpovídá funkční hodnotě v bodě 4. Je tedy

$$P(\mathbb{X} \leq 4) = F(4) = 1 - \mathrm{e}^{-4/9} \doteq 1 - 0,641 = 0,359.$$

Pravděpodobnost, že budeme obslouženi do 4 minut, je rovna přibližně 0,359.

3.11. Pracovnice call centra přijme hovor průměrně jednou za 6 minut. Jaká je pravděpodobnost, že po dobu, kdy se na dvě minuty vzdálí, nepřijde žádný hovor?

Řešení:
Náhodnou veličinu *doba do příštího telefonátu* můžeme považovat za náhodnou veličinu s exponenciálním rozdělením se střední hodnotou $\delta = 6$ minut. Ze zadání úlohy plyne, že chceme znát pravděpodobnost, že doba do příštího hovoru bude větší než 2 minuty, tj. chceme znát pravděpodobnost, že $\mathbb{X} > 2$.

$$P(\mathbb{X} > 2) = 1 - P(\mathbb{X} \leq 2) = 1 - F(2) = 1 - \left(1 - \mathrm{e}^{-2/6}\right) = \mathrm{e}^{-1/3} = \frac{1}{\sqrt[3]{\mathrm{e}}} \doteq 0,717.$$

Pravděpodobnost, že během následujících dvou minut nepřijde na call centrum hovor k vyřízení, je rovna přibližně 0,717.

Úloha k samostatnému řešení

3.12. Zkouška nového stroje musí probíhat nepřetržitě 24 hodin a je nezbytně nutné, aby po celou tuto dobu byl stroj pod kontrolou diagnostického zařízení. Víme, že diagnostické zařízení má poruchu průměrně jednou za 2 500 hodin. Zjistěte, zda „čas čekání na poruchu diagnostického zařízení“ je s pravděpodobností $p = 0.99$ delší, než čas vymezený na zkoušku.

3.1.3 Normální rozdělení náhodné veličiny

Normální rozdělení pravděpodobnosti má zcela výjimečnou pozici mezi ostatními pravděpodobnostními rozděleními spojité náhodné veličiny. Toto rozdělení je použitelné tam, kde kolísání hodnot náhodné veličiny kolem střední hodnoty je způsobeno velkým počtem nepatrných a nezávislých vlivů.

3.13. Výška dospělého člověka (muže) se pohybuje v průměru okolo 180 cm (ve skutečnosti to bude zřejmě jiné číslo, nicméně pro tuto chvíli to není podstatné). Toto však neznamená, že každý muž měří 180 cm. Někteří muži jsou vyšší, někteří nižší, a pokud bychom za náhodnou veličinu považovali *výšku náhodně vybraného muže*, pak kolísání hodnot této náhodné veličiny kolem čísla 180 cm, tj. různé výšky náhodně vybraných dospělých mužů, je dáno řadou nejrůznějších vlivů – genetickou výbavou, stravou, prodělanými nemocemi, prodělanými úrazy, provozovanými sportovními aktivitami, kvalitou spánku, denní dobou měření výšky a jistě řadou dalších zde neuvedených vlivů. Tedy kolísání hodnot náhodné veličiny okolo střední hodnoty je dáno tím, že tato hodnota je ovlivňována řadou nejrůznějších, na sobě nezávislých, jevů. Tak nějak očekáváme, že většina mužů bude mít výšku v blízkosti čísla 180 cm, a počet mužů klesá tak, jak se jejich výška vzdaluje od oné střední hodnoty.

3.14. Automatická linka plní hladkou mouku do papírového balení. Automat je nastaven tak, aby mouka v jednom balení měla hmotnost 1 kg. V každém balení však není přesně jeden kilogram mouky. Pokud budeme uvažovat náhodnou veličinu *hmotnost mouky v náhodně vybraném balení*, budou se její hodnoty lišit od předepsané hmotnosti. Tyto odchylky mohou být způsobeny nejrůznějšími vlivy – špatné nastavení stroje, měnící se teplota prostředí (tepelná roztažnost změní nastavení stroje), vlhkost vzduchu (ovlivní lepivost mouky), vlhkost mouky (poměr vody ovlivní hmotnost), opotřebení stroje, nečistoty ve stroji, chyba obsluhy atd. Opět můžeme očekávat, že většina balení bude mít hmotnost blízkou hodnotě 1 kg, počet balení s rozdílnou hmotností bude klesat se zvětšující se odchylkou od 1 kg.

V obou výše uvedených příkladech lze zavedenou náhodnou veličinu považovat za náhodnou veličinu s normálním rozdělením pravděpodobnosti. Hustota pravděpodobnosti této náhodné veličiny nabývá maximum ve své střední hodnotě μ, s rostoucí vzdáleností x od μ se hustota symetricky snižuje. Graf hustoty pravděpodobnosti připomíná svým tvarem zvon, proto se grafu hustoty pravděpodobnosti veličiny s normálním rozdělením říká zvonovitá křivka (ang. *bell shaped curve*), či gaussova křivka (podle německého matematika C. F. Gausse (1777-1855).

Definice 3.1.3. Řekneme, že náhodná veličina $\mathbb{X}$ má *normální rozdělení* pravděpodobnosti, tj. $\mathbb{X} \sim \mathrm{No}[\mu,\ \sigma^2]$, jestliže hustota pravděpodobnosti $f(x)$ je popsána vzorcem

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty. \tag{3.9}$$

Normální rozdělení má dva parametry μ a σ^2, kde μ je střední hodnota rozdělení a σ^2 je jeho rozptyl. Je tedy

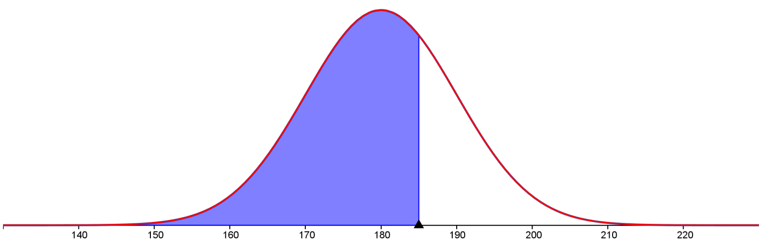
$$E(\mathbb{X}) = \mu, \qquad D(\mathbb{X}) = \sigma^2. \tag{3.10}$$

Distribuční funkce normálního rozdělení má tvar

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\, \mathrm{d}t. \tag{3.11}$$

Výpočet $F(x)$ podle vzorce (3.11) je obtížný, ne-li nemožný. Proto byly hodnoty distribuční funkce tabelovány (vypočteny numerickými metodami a zaneseny do tabulek). Vzhledem k tomu, že náhodná veličina s normálním rozdělením má dva parametry μ a σ^2, museli bychom vytvořit tabulky pro každou kombinaci těchto parametrů – což je značně nepraktické. Proto libovolnou náhodnou veličinu $\mathbb{X} \sim \mathrm{No}[\mu,\ \sigma^2]$ s parametry μ a σ^2 transformujeme na normovanou veličinu $\mathbb{U}$, jejíž parametry jsou $\mu = 0$ a $\sigma^2 = 1$, a to pomocí převodního vztahu

$$\mathbb{U} = \frac{\mathbb{X} - \mu}{\sigma}. \tag{3.12}$$



Obrázek 3.4: Hustota pravděpodobnosti normálního rozdělení $\mathbb{X} \sim \mathrm{No}[180, 10^2]$

Po této transformaci dostaneme tzv. *normované normální rozdělení* $\mathbb{U}$, kde je

$$f(u) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{u^2}{2}}, \qquad F(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

Parametry normovaného normálního rozdělení $\mathbb{U}$ jsou

$$E(\mathbb{U}) = 0, \quad D(\mathbb{U}) = 1.$$

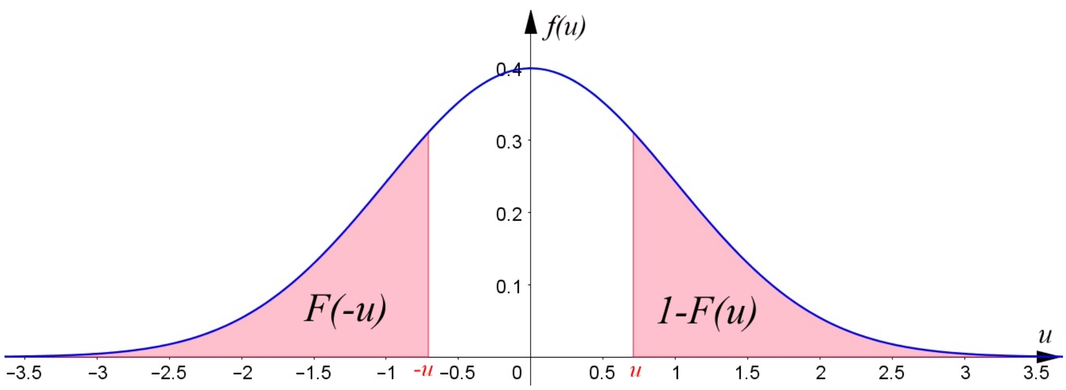
Hodnoty distribuční funkce normovaného normálního rozdělení jsou tabelovány. Protože je rozdělení symetrické okolo nuly, stačí v tabulce uvést hodnoty pravděpodobností pouze pro $u > 0$. Pro hodnoty $u < 0$ lze využít vzorec

$$F(-u) = 1 - F(u), \tag{3.13}$$

který tuto symetrii využívá. Viz Obrázek 3.5, kde jsou naznačeny rovnosti

$$F(-u) = P(\mathbb{U} \leq -u) = P(\mathbb{U} > u) = 1 - P(\mathbb{U} \leq u) = 1 - F(u).$$

Vypočtěme pomocí vzorce (3.13) například hodnotu $F(-1,5)$. Je $F(-1,5) = 1 - F(1,5)$, kde hodnotu $F(1,5)$ již nalezneme v tabulkách distribuční funkce normovaného normálního rozdělení.



Obrázek 3.5: Porovnáním obsahů obou ploch zdůvodníme rovnici (3.13)

Nyní již máme dostatečný teoretický základ k provádění výpočtů. Před začátkem procvičování pouze připomeneme obecné tvrzení, platné pro jakoukoliv náhodnou veličinu. Často máme určit pravděpodobnost, že náhodná veličina $\mathbb{X}$ nabude hodnoty z intervalu x_1 až x_2. Z vlastností distribuční funkce plyne

$$P(u_1 < \mathbb{U} < u_2) = F(u_2) - F(u_1). \tag{3.14}$$

Následující příklad slouží k procvičení základní práce s pravděpodobnostmi různých hodnot náhodné veličiny s normovaným normálním rozdělením a tabulkami příslušné distribuční funkce.

3.15. Nechť $\mathbb{U}$ je náhodná veličina s normovaným normálním rozdělením No$[0, 1]$. Určete, jaká je pravděpodobnost, že hodnota $\mathbb{U}$ bude:

a) menší než $1,25$,

b) větší než $2,1$,

c) menší než $-1,5$,

d) větší než $-0,45$,

e) v rozmezí $0,35$ až $1,87$,

f) v rozmezí $-1,35$ až $0,55$.

Řešení: Při výpočtech budeme používat tabulky s hodnotami distribuční funkce normovaného normálního rozdělení, viz příloha tohoto textu na konci dokumentu. Tabulka obsahuje dva sloupce – v levém je uvedena hodnota u, v pravém sloupci se nachází příslušná hodnota distribuční funkce $F(u)$.

a) Pravděpodobnost, že hodnota $\mathbb{U}$ bude menší než $1,25$, odpovídá rovnosti $P(\mathbb{U} < 1,25)$. Z definice distribuční funkce tak použitím vzorce (3.1) dostaneme řešení

$$P(\mathbb{U} < 1,25) = F(1,25) = 0,894350.$$

b) Pravděpodobnost, že hodnota $\mathbb{U}$ bude větší než $2,1$, odpovídá rovnosti $P(\mathbb{U} > 2,1)$. Vzhledem k tomu, že jev $\mathbb{U} > 2,1$ je opačným jevem k jevu $\mathbb{U} \leq 2,1$, platí mezi pravděpodobnostmi vztah $P(\mathbb{U} > 2,1) = 1 - P(\mathbb{U} \leq 2,1)$. Z definice distribuční funkce tak použitím vzorce (3.1) dostaneme řešení

$$P(\mathbb{U} > 2,1) = 1 - P(\mathbb{U} \leq 2,1) = 1 - F(2,1) = 1 - 0,982136 = 0,017864.$$

c) Pravděpodobnost, že hodnota $\mathbb{U}$ bude menší než $-1,5$ odpovídá rovnosti $P(\mathbb{U} < -1,5)$. Ze vzorců (3.1) a (3.13) dostaneme řešení

$$P(\mathbb{U} < -1,5) = F(-1,5) = 1 - F(1,5) = 1 - 0,933193 = 0,066807.$$

d) Pravděpodobnost, že hodnota $\mathbb{U}$ bude větší než $-0,45$ odpovídá rovnosti $P(\mathbb{U} > -0,45)$. Vzhledem k tomu, že jev $\mathbb{U} > -0,45$ je opačným jevem k jevu $\mathbb{U} \leq -0,45$, platí mezi pravděpodobnostmi vztah $P(\mathbb{U} > -0,45) = 1 - P(\mathbb{U} \leq -0,45)$. Ze vzorců (3.1) a (3.13) dostaneme řešení

$$\begin{aligned} P(\mathbb{U} > -0,45) &= 1 - P(\mathbb{U} \leq -0,45) = 1 - F(-0,45) \\ &= 1 - \big(1 - F(0,45)\big) = F(0,45) = 0,673645. \end{aligned}$$

e) Pravděpodobnost, že $\mathbb{U}$ bude v rozmezí $0,35$ až $1,87$ zapíšeme $P(0,35 < \mathbb{U} < 1,87)$. Z rovnice (3.14) dostaneme řešení

$$P(0,35 < \mathbb{U} < 1,87) = F(1,87) - F(0,35) = 0,969258 - 0,636831 = 0,332427.$$

f) Pravděpodobnost, že $\mathbb{U}$ bude v rozmezí $-1,35$ až $0,55$ zapíšeme $P(-1,35 < \mathbb{U} < 0,55)$. Použitím rovnic (3.14) a (3.13) dostaneme řešení

$$\begin{aligned} P(-1,35 < \mathbb{U} < 0,55) &= F(0,55) - F(-1,35) \\ &= F(0,55) - \big(1 - F(1,35)\big) = F(0,55) + F(1,35) - 1 = 0,620332. \end{aligned}$$

Nyní se již můžeme věnovat slovním úlohám.

3.16. Předpokládejme, že životnost baterie je náhodnou veličinou $\mathbb{X}$ s normálním rozdělením pravděpodobnosti s parametry $\mu = 300$ hodin a $\sigma = 35$ hodin. Jaká je pravděpodobnost, že náhodně vybraná baterie bude mít životnost větší než 320 hodin?

Řešení: Je $\mathbb{X} \sim \mathrm{No}[300, 35^2]$, kde $\mathbb{X}$ představuje dobu životnosti baterie. Naším úkolem je zjistit pravděpodobnost, že hodnota $\mathbb{X}$ bude větší než 320 hodin, tedy určit hodnotu $P(\mathbb{X} > 320)$. Abychom mohli používat tabulky s hodnotami distribuční funkce, musíme příslušné hodnoty (meze) veličiny $\mathbb{X}$ převést na odpovídající hodnoty (meze) normované veličiny $\mathbb{U}$. To znamená že musíme vypočítat, která hodnota normované veličiny $\mathbb{U}$ odpovídá číslu 320 u veličiny $\mathbb{X}$. Pomocí rovnice (3.12) dostaneme následující rovnosti.

$$\mathbb{U} = \frac{\mathbb{X} - \mu}{\sigma} = \frac{\mathbb{X} - 300}{35} = \frac{320 - 300}{35} = \frac{20}{35} = \frac{4}{7} \doteq 0,57.$$

Hodnotě $\mathbb{X} = 320$ tedy odpovídá $\mathbb{U} = 0,57$. Z toho důvodu je $P(\mathbb{X} > 320) = P(\mathbb{U} > 0,57)$. Postupem, jak vypočítat hodnotu $P(\mathbb{U}) > 0,57$ jsme se zabývali v Příkladu 3.15 b).

$$P(\mathbb{X} > 320) = P(\mathbb{U} > 0,57) = 1 - P(\mathbb{U} \leq 0,57) = 1 - F(0,57) = 1 - 0,715661 = 0,284339.$$

Pravděpodobnost, že náhodně vybraná baterie vydrží pracovat déle než 320 hodin je přibližně $p = 0,284$. Tento výsledek také můžeme interpretovat tak, že přibližně $28,4$ % těchto baterií vydrží déle než 320 hodin.

3.17. Při výstupní kontrole je součástka uznána za kvalitní, jestliže se její rozměr pohybuje v rozmezí 35 až 37 mm. Měřením se zjistilo, že rozměry součástek tvoří náhodnou veličinu s normálním rozdělením pravděpodobnosti se střední hodnotou $\mu = 36,2$ mm a směrodatnou odchylkou $\sigma = 0,3$ mm. Jaká je pravděpodobnost, že rozměr součástky, náhodně vybrané ke kontrole, bude v požadovaných mezích?

Řešení: Je $\mathbb{X} \sim \mathrm{No}[36.2, 0.3^2]$, kde $\mathbb{X}$ představuje délku náhodně vybraného výrobku. Naším úkolem je zjistit pravděpodobnost, že hodnota $\mathbb{X}$ se bude nacházet v rozmezí $35 < \mathbb{X} < 37$, tj. vypočítat hodnotu $P(35 < \mathbb{X} < 37)$. Opět musíme příslušné hodnoty (meze) veličiny $\mathbb{X}$ převést na odpovídající hodnoty (meze) normované veličiny $\mathbb{U}$. To znamená že musíme vypočítat, které hodnoty u_1, resp. u_2 normované veličiny $\mathbb{U}$ odpovídají číslům $x_1 = 35$, resp. $x_2 = 37$ veličiny $\mathbb{X}$. Pomocí rovnice (3.12) dostaneme následující rovnosti.

$$\mathbb{U} = \frac{\mathbb{X} - 36.2}{0.3}, \qquad u_1 = \frac{35 - 36.2}{0.3} = \frac{-1.2}{0.3} = -4, \qquad u_2 = \frac{37 - 36.2}{0.3} = \frac{0.8}{0.3} \doteq 2.67.$$

Hledaná pravděpodobnost je

$$\begin{aligned} P(35 < \mathbb{X} < 37) &= P(-4 < \mathbb{U} < 2,67) = F(2,67) - F(-4) = F(2,67) - \big(1 - F(4)\big) \\ &= F(2,67) + F(4) - 1 = 0,9962 + 1 - 1 \doteq 0,9962, \end{aligned}$$

kde hodnota $F(4)$ není v tabulkách uvedena, nicméně z trendu je zřejmé, že její hodnota je při zaokrouhlení na čtyři desetinná místa rovna 1.

Pravděpodobnost, že součástka bude v požadovaných mezích činí přibližně $p = 0,996$.

3.18. Výrobce trolejbusů připravuje nový typ trolejbusu. Při návrhu stanovuje různé detaily týkající se trolejbusu. Jedním z nich je výška dveří, která má být taková, aby jimi prošlo (bez nutnosti ohnutí) 95 % lidí (a tedy pouze 5 % lidí bylo vyšších než je výška dveří). Předpokládejme, že výška lidí, kteří používají městskou dopravu, podléhá normálnímu rozdělení s parametry $\mu = 180$ cm a směrodatnou odchylkou $\sigma = 10$ cm. Jaká by tedy měla být výška dveří?

Řešení: Ze zadání plyne, že chceme určit 95% kvantil zadané náhodné veličiny, neboť potřebujeme určit takovou hodnotu x veličiny $\mathbb{X}$, aby platily rovnosti $P(\mathbb{X} \leq x) = 0,95$ a $P(\mathbb{X} > x) = 0,05$. Hledáme tedy hodnotu x, pro kterou je hodnota distribuční funkce rovna 0,95. Pro rozlišení budeme pro označení distribuční funkce normovaného normálního rozdělení používat symbol $\Phi(u)$ a pro označení distribuční funkce nenormované veličiny $\mathbb{X}$ použijeme symbol $F(x)$.

Z tabulek pro normovanou distribuční funkci normálního rozdělení vyčteme, že $\Phi(u) = 0,95$ platí pro $u = 1,65$. Je tedy $\Phi(1,65) = 0,95$, resp. $P(\mathbb{U} \leq 1,65) \doteq 0,95$.

Nyní musíme zjistit, která hodnota x v $\mathbb{X} \sim \mathrm{No}[180, 10^2]$ odpovídá $u = 1,65$ pro $\mathbb{U} \sim N[0,1]$.

$$\mathbb{U} = \frac{\mathbb{X} - \mu}{\sigma} \quad \Rightarrow \quad 1,65 = \frac{x - 180}{10} \quad \Rightarrow \quad 16,5 = x - 180 \quad \Rightarrow \quad x = 196,5$$

Výška dveří trolejbusu by měla být rovna 196,5 cm.

Úlohy k samostatné práci

3.19. Vypočtěte pravděpodobnost, že náhodná veličina s normálním rozdělením pravděpodobnosti $X \sim \mathrm{No}(64; 16)$ nabude hodnotu větší než 58. (Poznámka na okraj – připomeňme, že v takto uvedeném zápisu náhodné veličiny je pod číslem 16 uveden rozptyl $D(\mathbb{X}) = \sigma^2$, nikoliv směrodatná odchylka σ.)

3.20. Vypočtěte pravděpodobnost, že náhodná veličina s normálním rozdělením pravděpodobnosti $X \sim \mathrm{No}(5.3; 1.44)$ nabude hodnotu v rozmezí od 4,7 do 6,5.

3.21. Je dána náhodná veličina s normálním rozdělením pravděpodobnosti $X \sim \mathrm{No}(5.3; 1.44)$. Vypočtěte 30% kvantil této náhodné veličiny.

3.22. Váha automaticky vyráběného výrobku je náhodná veličina s normálním rozdělením s parametry $\mu = 220$ gramů a $\sigma = 4$ gramy. Kolik procent výrobků je těžších než 222 gramy?

Příloha č. 1: Hodnoty distribuční funkce náhodné veličiny s normovaným normálním rozdělením

u	F(u)	u	F(u)	u	F(u)	u	F(u)	u	F(u)	u	F(u)
0	0,5	0,6	0,72575	1,2	0,88493	1,8	0,96407	2,4	0,9918	3	0,99865
0,01	0,50399	0,61	0,72907	1,21	0,88686	1,81	0,96485	2,41	0,99202	3,01	0,99869
0,02	0,50798	0,62	0,73237	1,22	0,88877	1,82	0,96562	2,42	0,99224	3,02	0,99874
0,03	0,51197	0,63	0,73565	1,23	0,89065	1,83	0,96638	2,43	0,99245	3,03	0,99878
0,04	0,51595	0,64	0,73891	1,24	0,89251	1,84	0,96712	2,44	0,99266	3,04	0,99882
0,05	0,51994	0,65	0,74215	1,25	0,89435	1,85	0,96784	2,45	0,99286	3,05	0,99886
0,06	0,52392	0,66	0,74537	1,26	0,89617	1,86	0,96856	2,46	0,99305	3,06	0,99889
0,07	0,5279	0,67	0,74857	1,27	0,89796	1,87	0,96926	2,47	0,99324	3,07	0,99893
0,08	0,53188	0,68	0,75175	1,28	0,89973	1,88	0,96995	2,48	0,99343	3,08	0,99896
0,09	0,53586	0,69	0,7549	1,29	0,90147	1,89	0,97062	2,49	0,99361	3,09	0,999
0,1	0,53983	0,7	0,75804	1,3	0,9032	1,9	0,97128	2,5	0,99379	3,1	0,99903
0,11	0,5438	0,71	0,76115	1,31	0,9049	1,91	0,97193	2,51	0,99396	3,11	0,99906
0,12	0,54776	0,72	0,76424	1,32	0,90658	1,92	0,97257	2,52	0,99413	3,12	0,9991
0,13	0,55172	0,73	0,7673	1,33	0,90824	1,93	0,9732	2,53	0,9943	3,13	0,99913
0,14	0,55567	0,74	0,77035	1,34	0,90988	1,94	0,97381	2,54	0,99446	3,14	0,99916
0,15	0,55962	0,75	0,77337	1,35	0,91149	1,95	0,97441	2,55	0,99461	3,15	0,99918
0,16	0,56356	0,76	0,77637	1,36	0,91309	1,96	0,975	2,56	0,99477	3,16	0,99921
0,17	0,56749	0,77	0,77935	1,37	0,91466	1,97	0,97558	2,57	0,99492	3,17	0,99924
0,18	0,57142	0,78	0,7823	1,38	0,91621	1,98	0,97615	2,58	0,99506	3,18	0,99926
0,19	0,57535	0,79	0,78524	1,39	0,91774	1,99	0,9767	2,59	0,9952	3,19	0,99929
0,2	0,57926	0,8	0,78814	1,4	0,91924	2	0,97725	2,6	0,99534	3,2	0,99931
0,21	0,58317	0,81	0,79103	1,41	0,92073	2,01	0,97778	2,61	0,99547	3,21	0,99934
0,22	0,58706	0,82	0,79389	1,42	0,9222	2,02	0,97831	2,62	0,9956	3,22	0,99936
0,23	0,59095	0,83	0,79673	1,43	0,92364	2,03	0,97882	2,63	0,99573	3,23	0,99938
0,24	0,59483	0,84	0,79955	1,44	0,92507	2,04	0,97932	2,64	0,99585	3,24	0,9994
0,25	0,59871	0,85	0,80234	1,45	0,92647	2,05	0,97982	2,65	0,99598	3,25	0,99942
0,26	0,60257	0,86	0,80511	1,46	0,92785	2,06	0,9803	2,66	0,99609	3,26	0,99944
0,27	0,60642	0,87	0,80785	1,47	0,92922	2,07	0,98077	2,67	0,99621	3,27	0,99946
0,28	0,61026	0,88	0,81057	1,48	0,93056	2,08	0,98124	2,68	0,99632	3,28	0,99948
0,29	0,61409	0,89	0,81327	1,49	0,93189	2,09	0,98169	2,69	0,99643	3,29	0,9995
0,3	0,61791	0,9	0,81594	1,5	0,93319	2,1	0,98214	2,7	0,99653	3,3	0,99952
0,31	0,62172	0,91	0,81859	1,51	0,93448	2,11	0,98257	2,71	0,99664	3,31	0,99953
0,32	0,62552	0,92	0,82121	1,52	0,93574	2,12	0,983	2,72	0,99674	3,32	0,99955
0,33	0,6293	0,93	0,82381	1,53	0,93699	2,13	0,98341	2,73	0,99683	3,33	0,99957
0,34	0,63307	0,94	0,82639	1,54	0,93822	2,14	0,98382	2,74	0,99693	3,34	0,99958
0,35	0,63683	0,95	0,82894	1,55	0,93943	2,15	0,98422	2,75	0,99702	3,35	0,9996
0,36	0,64058	0,96	0,83147	1,56	0,94062	2,16	0,98461	2,76	0,99711	3,36	0,99961
0,37	0,64431	0,97	0,83398	1,57	0,94179	2,17	0,985	2,77	0,9972	3,37	0,99962
0,38	0,64803	0,98	0,83646	1,58	0,94295	2,18	0,98537	2,78	0,99728	3,38	0,99964
0,39	0,65173	0,99	0,83891	1,59	0,94408	2,19	0,98574	2,79	0,99736	3,39	0,99965
0,4	0,65542	1	0,84134	1,6	0,9452	2,2	0,9861	2,8	0,99744	3,4	0,99966
0,41	0,6591	1,01	0,84375	1,61	0,9463	2,21	0,98645	2,81	0,99752	3,41	0,99968
0,42	0,66276	1,02	0,84614	1,62	0,94738	2,22	0,98679	2,82	0,9976	3,42	0,99969
0,43	0,6664	1,03	0,84849	1,63	0,94845	2,23	0,98713	2,83	0,99767	3,43	0,9997
0,44	0,67003	1,04	0,85083	1,64	0,9495	2,24	0,98745	2,84	0,99774	3,44	0,99971
0,45	0,67364	1,05	0,85314	1,65	0,95053	2,25	0,98778	2,85	0,99781	3,45	0,99972
0,46	0,67724	1,06	0,85543	1,66	0,95154	2,26	0,98809	2,86	0,99788	3,46	0,99973
0,47	0,68082	1,07	0,85769	1,67	0,95254	2,27	0,9884	2,87	0,99795	3,47	0,99974
0,48	0,68439	1,08	0,85993	1,68	0,95352	2,28	0,9887	2,88	0,99801	3,48	0,99975
0,49	0,68793	1,09	0,86214	1,69	0,95449	2,29	0,98899	2,89	0,99807	3,49	0,99976
0,5	0,69146	1,1	0,86433	1,7	0,95543	2,3	0,98928	2,9	0,99813	3,5	0,99977
0,51	0,69497	1,11	0,8665	1,71	0,95637	2,31	0,98956	2,91	0,99819	3,51	0,99978
0,52	0,69847	1,12	0,86864	1,72	0,95728	2,32	0,98983	2,92	0,99825	3,52	0,99978
0,53	0,70194	1,13	0,87076	1,73	0,95818	2,33	0,9901	2,93	0,99831	3,53	0,99979
0,54	0,7054	1,14	0,87286	1,74	0,95907	2,34	0,99036	2,94	0,99836	3,54	0,9998
0,55	0,70884	1,15	0,87493	1,75	0,95994	2,35	0,99061	2,95	0,99841	3,55	0,99981
0,56	0,71226	1,16	0,87698	1,76	0,9608	2,36	0,99086	2,96	0,99846	3,56	0,99981
0,57	0,71566	1,17	0,879	1,77	0,96164	2,37	0,99111	2,97	0,99851	3,57	0,99982
0,58	0,71904	1,18	0,881	1,78	0,96246	2,38	0,99134	2,98	0,99856	3,58	0,99983
0,59	0,7224	1,19	0,88298	1,79	0,96327	2,39	0,99158	2,99	0,99861	3,59	0,99983