4 Diskrétní náhodné veličiny

4.1 Binomické rozdělení $\mathrm{Bin}(N, p)$

- Bernoulliho pokusy $X_1, \ldots, X_N$:
  - $X_i = 1$ … událost nastala; $X_i = 0$… událost nenastala; $i = 1, \ldots, N$.
  - $\Pr(X_i = 1) = p$
  - $\Pr(X_i = 0) = 1 - p$
- Binomické rozdělení:
  - X… počet událostí v posloupnosti N nezávislých Bernoulliho pokusů, přičemž pravděpodobnost nastání události v každém pokusu je vyjádřena parametrem p.
  - $\sum_{i=1}^{N} X_i = X \sim \mathrm{Bin}(N, p)$.
  - $\boldsymbol{\theta} = (N, p)$
  - pravděpodobnostní funkce:

$$p(x) = \binom{N}{x} p^x (1-p)^{N-x} \qquad x = 0, 1, \ldots, N;$$

  - vlastnosti: $\mathrm{E}[X] = Np$; $\mathrm{Var}[X] = Np(1-p)$
  - dbinom(x, N, p), pbinom(x, N, p)

Dataset: Počet chlapců v rodinách s 12 dětmi
V rámci studie poměru pohlaví u lidí z roku 1889 bylo na základě záznamů z nemocnic v Sasku zaznamenáno rozdělení počtu chlapců v čtrnáctičlenných rodinách. Mezi $M = 6115$ rodinami s $N = 12$ dětmi byla pozorována početnost chlapců. Údaje ze studie jsou uvedeny v následující tabulce.

n	0	1	2	3	4	5	6	7	8	9	10	11	12	$\sum$
$m_{observed}$	3	24	104	286	670	1033	1343	1112	829	478	181	45	7	6115

Příklad 4.1. Výpočet parametru p binomického rozdělení
Předpokládejme, že náhodná veličina X popisující počet chlapců v rodinách s dvanácti dětmi pochází z binomického rozdělení s parametrem $N = 12$. Vypočítejte odhad pravděpodobnosti výskytu chlapců v rodinách s dvanácti dětmi.

Řešení příkladu 4.1
Pravděpodobnost p výskytu chlapců v rodinách s dvanácti dětmi odhadneme pomocí vzorce

$$\widehat{p} = \frac{\text{počet narozených chlapců}}{\text{celkový počet narozených dětí}} = \frac{\sum_{n=0}^{N} n m_{observed}}{NM}. \tag{4.1}$$

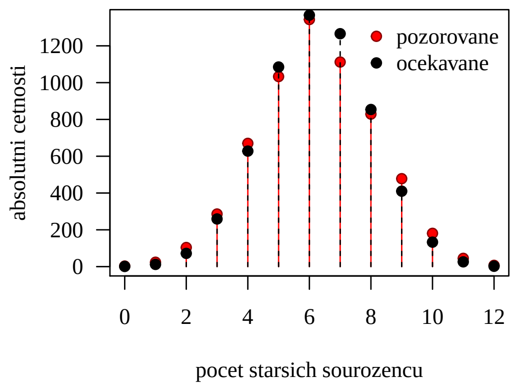
```
[1] 0.5192
```

Interpretace výsledků: Pravděpodobnost výskytu chlapců v rodinách s dvanácti dětmi je .... (............................ %). ★

Příklad 4.2. Pozorované a očekávané početnosti v binomickém rozdělení
Za předpokladu, že počet chlapců v rodinách s dvanácti dětmi pochází z binomického rozdělení s parametry $N =$ .... a $p =$ .... odhadněte očekávané početnosti chlapců v rodinách s dvanácti dětmi a porovnejte je s pozorovanými početnostmi.

Řešení příkladu 4.2

```
      0  1   2   3   4    5    6    7   8   9  10  11  12
m.obs 3 24 104 286 670 1033 1343 1112 829 478 181 45   7
m.exp 1 12  72 259 628 1085 1367 1266 854 410 133 26   2
```



★

Příklad 4.3. Výpočet pravděpodobností za předpokladu binomického rozdělení
Za předpokladu, že náhodná veličina X popisující počet chlapců v rodinách s dvanácti dětmi pochází z binomického rozdělení s parametry $N =$ .... a $p =$ .... vypočítejte pravděpodobnost, že v rodině s dvanácti dětmi bude

a. právě devět chlapců,

b. nejvýše čtyři chlapci,

c. alespoň osm chlapců,

d. čtyři, pět, šest, nebo sedm chlapců.

Řešení příkladu 4.3

```
[1] 0.067
```
5

```
[1] 0.1589
```
6

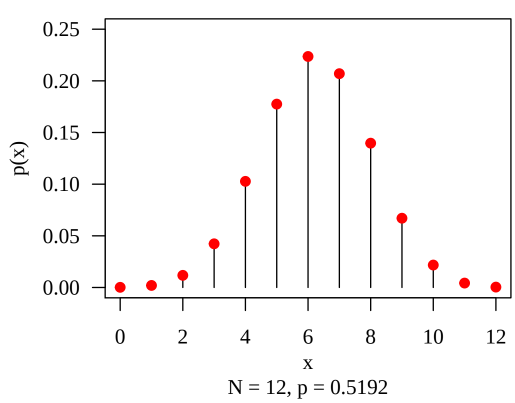
```
[1] 0.2331
```
7

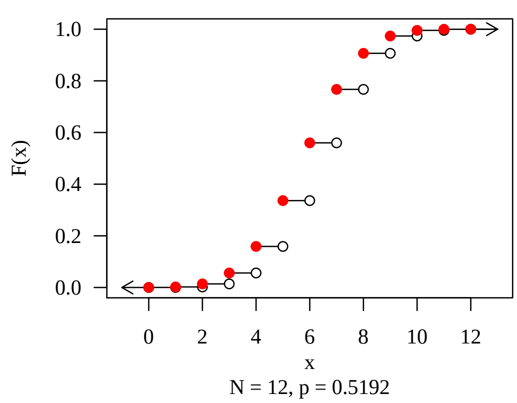
```
[1] 0.7108
```
8

Interpretace výsledků: Pravděpodobnost, že v rodině bude právě devět chlapců, je ....%. Pravděpodobnost, že v rodině budou nejvýše čtyři chlapci, je ....%. Pravděpodobnost, že v rodině bude alespoň osm chlapců, je ....%. Pravděpodobnost, že v rodině bude čtyři, pět, šest, nebo sedm chlapců, je ....%. ★

Příklad 4.4. Graf pravděpodobnostní a distribuční funkce binomického rozdělení
Nakreslete graf pravděpodobnostní funkce a graf distribuční funkce binomického rozdělení $\mathrm{Bin}(N, p)$, kde $N = 12$ a $p = 0.5192$.

Řešení příkladu 4.4





★

4.2 Poissonovo rozdělení Poiss(λ)

- X ... počet událostí, které nastanou v jednotkovém časovém intervalu, přičemž k událostem dochází náhodně, jednotlivě a vzájemně nezávisle. Střední počet těchto událostí je vyjádřen parametrem $\lambda > 0$.
- $X \sim \text{Poiss}(\lambda)$
- $\theta = \lambda$
- pravděpodobnostní funkce:
$$p(x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad x = 0, 1, \ldots;$$
- vlastnosti: $E[X] = \lambda$; $\text{Var}[X] = \lambda$
- dpois(x, lambda), ppois(x, lambda)

Příklad 4.5. Výpočet parametru λ Poissonova rozdělení
Načtete datový soubor 17-anova-newborns-2.txt a odstraňte z něj neznámá pozorování. Zaměřte se na znak $X =$*počet starších sourozenců* novorozence. Za předpokladu, že náhodná veličina X popisující počet starších sourozenců novorozence pochází z Poissonova rozdělení parametrem λ odhadněte střední hodnotu počtu starších sourozenců λ.

Řešení příkladu 4.5
Střední hodnotu počtu starších sourozenců odhadneme pomocí vzorce

$$\lambda = \frac{\text{počet starších sourozenců}}{\text{počet novorozenců}} = \frac{\sum_{i=1}^{N} x_i}{N}. \tag{4.2}$$

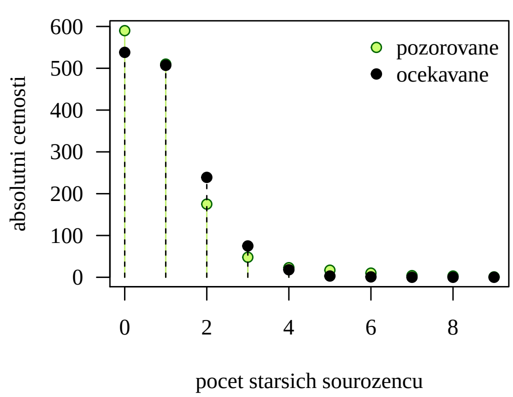
```
[1] 0.9427951
```

Interpetace výsledků: Střední hodnota počtu starších sourozenců novorozenců v datovém souboru $\lambda =$.........
.... ★

Příklad 4.6. Porovnání pozorovaných a očekávaných početností v Poissonově rozdělení
Za předpokladu, že počet starších sourozenců novorozenců pochází z Poissonova rozdělení s parametrem $\lambda =$...........
.... odhadněte očekávané početnosti starších sourozenců a porovnejte je s pozorovanými početnostmi.

Řešení příkladu 4.6

```
        0   1   2  3  4  5  6 7 8 9
m.obs 590 510 175 48 23 17 10 4 3 1
m.exp 538 507 239 75 18  3  1 0 0 0
```



Obrázek 1: Porovnání pozorovaných a očekávaných početností v Poissonově rozdělení

★

Příklad 4.7. Výpočet pravděpodobností za předpokladu Poissonova rozdělení
Za předpokladu, že data pochází z Poissonova rozdělení s parametrem $\lambda =$ .... určete pravděpodobnost, novorozenec má

a. dva, tři nebo čtyři starší sourozence,

b. alespoň čtyři starší sourozence,

c. nejvýše dva starší sourozence,

d. právě jednoho staršího sourozence.

Řešení příkladu 4.7

```
[1] 0.2403526                                                    13
```

```
[1] 0.01567936                                                   14
```

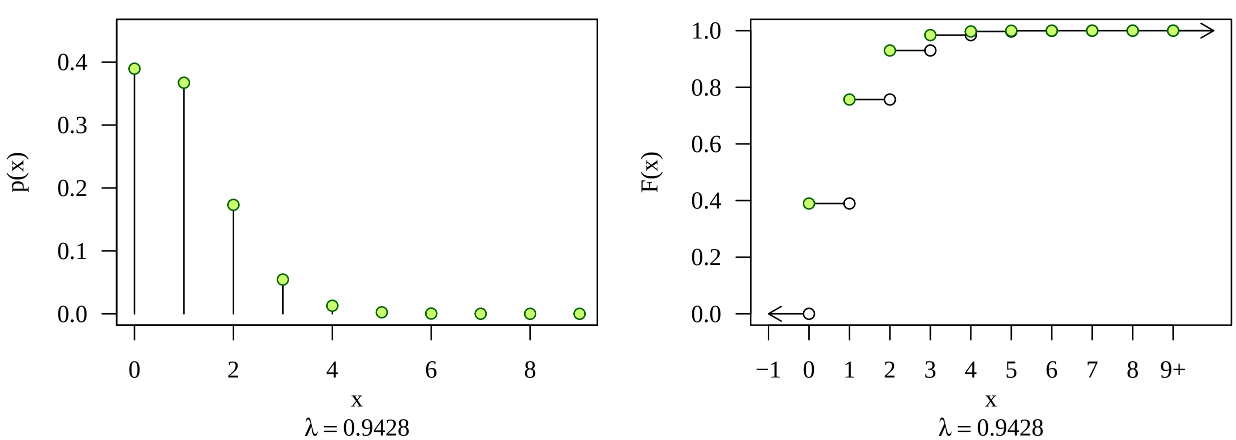
```
[1] 0.9299142                                                    15
```

```
[1] 0.3672541                                                    16
```

Interpretace výsledů: Pravděpodobnost, že novorozenec má dva, tři nebo čtyři starší sourozence je ....%. Pravděpodobnost, že novorozenec má alespoň čtyři starší sourozence je ....%. Pravděpodobnost, že novorozenec má nejvýše dva starší sourozence je ....%. Pravděpodobnost, že novorozenec má jednoho staršího sourozence je ....%. ★

Příklad 4.8. Graf pravděpodobnostní a distribuční funkce Poissonova rozdělení
Nakreslete graf pravděpodobnostní a distribuční funkce Poissonova rozdělení Poiss(0.9428) v hodnotách $x = 0$, 1, 2, 3, 4, 5, 6, 7, 8, a $x \geq 9$.

Řešení příkladu 4.8



Obrázek 2: Pravděpodobnostní a distribuční funkce Poissonova rozdělení

★