Cvičení 3 s návodem

Příklady na využití exponenciálního rozložení

Příklad 1.: Doba do ukončení opravy v opravně obuvi je náhodná veličina, která se řídí exponenciálním rozložením se střední hodnotou 3 dny. Jaká je pravděpodobnost, že oprava bude ukončena do dvou dnů?
Řešení: X ~ Ex(1/3),

$$P(X \le 2) = \int_0^2 \frac{1}{3} e^{-\frac{x}{3}} dx = \left[-e^{-\frac{x}{3}} \right]_0^2 = 1 - e^{-\frac{2}{3}} = 0{,}4866$$

V MATLABu: p = expcdf(2,3)

Příklad 2.: Životnost žárovky má exponenciální rozložení se střední hodnotou 600 h. Jaká je pravděpodobnost, že žárovka bude svítit dalších aspoň 200 h, jestliže již svítila aspoň 800 h?
Řešení: X ~ Ex(1/600), podle věty 3.2 dostáváme

$$P(X \ge 800 + 200 / X \ge 800) = P(X \ge 200) = 1 - P(X \le 200) + P(X = 200)$$

$$= 1 - \int_0^{200} \frac{1}{600} e^{-\frac{x}{600}} dx = 1 - \left[-e^{-\frac{x}{600}} \right]_0^{200} = e^{-\frac{200}{600}} = e^{-\frac{1}{3}} = 0{,}7165$$

V MATLABu: p = 1- expcdf(200,600)

Příklad 3.: Náhodné doby života dvou součástek jsou stochasticky nezávislé náhodné veličiny, přičemž $X_i \sim Ex(\lambda_i)$, i = 1, 2. Střední hodnota doby života první součástky je 2 roky, druhé součástky 3 roky. Jaká je pravděpodobnost, že druhá součástka přežije první?
Řešení:
Podle věty 3.13 dostáváme:

$$P(X_2 > X_1) = \frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{3}} = \frac{\frac{1}{2}}{\frac{5}{6}} = 0{,}6$$

Příklad 4.: Doba (v hodinách), která uplyne mezi dvěma naléhavými příjmy v jisté nemocnici, se řídí exponenciálním rozložením se střední hodnotou 2 h. Jaká je pravděpodobnost, že uplyne více než 5 h bez naléhavého příjmu?
Řešení: X ~ Ex(1/2),

$$P(X > 5) = 1 - P(X \le 5) = 1 - \int_0^5 \frac{1}{2} e^{-\frac{x}{2}} dx = 1 - \left[-e^{-\frac{x}{2}} \right]_0^5 = e^{-2{,}5} = 0{,}082$$

V MATLABu: p = 1- expcdf(5,2)

Příklad 5.: Zkoumá se funkce dvou nezávisle na sobě pracujících přístrojů. Doba bezporuchové funkce i-tého přístroje je náhodná veličina $X_i \sim Ex(\lambda_i)$, i = 1, 2. Jaká je pravděpodobnost, že za dobu $t_0 > 0$ a) ani jeden přístroj neselže, b) selže aspoň jeden přístroj?
Řešení:
ad a)

$$P(X_1 > t_0 \wedge X_2 > t_0) = P(X_1 > t_0) P(X_2 > t_0) = [1 - P(X_1 \le t_0)][1 - P(X_2 \le t_0)] =$$

$$= [1 - \Phi_1(t_0)][1 - \Phi_2(t_0)] = e^{-\lambda_1 t_0} e^{-\lambda_2 t_0} = e^{-t_0(\lambda_1 + \lambda_2)}$$

ad b)

$$P(X_1 \le t_0 \vee X_2 \le t_0) = 1 - e^{-t_0(\lambda_1 + \lambda_2)}$$

Příklad 6.: Najděte 5. percentil náhodné veličiny X ~ Ex(0,1).

Řešení:

$$0{,}05 = \Phi\left(K_{0,05}(X)\right) = 1 - \exp\left(-0{,}1 K_{0,05}(X)\right) \Rightarrow K_{0,05}(X) =$$

$$= -10 \ln 0{,}95 = 0{,}5129$$

V MATLABu: K = expinv(0.05,10)

Využití exponenciálního rozložení při analýze příjmů

Úvod do problému: Je známo, že příjmy obyvatelstva ve společnosti jsou rozděleny nerovnoměrně. Jako první zkoumal toto rozdělení italský inženýr Vilfredo Pareto na konci 19. století. Zjistil, že příjmy lze modelovat mocninnou funkcí. V dalších letech se ukázalo, že tento tzv. Paretův zákon platí jen pro 5 % nejbohatších lidí. Příjmy ostatních 95 % obyvatel lze modelovat pomocí exponenciálního rozložení. (Proč to tak je? To je vysvětleno v článku F. Slaniny, Vesmír č. 9, rok 2001)
Nechť náhodná veličina X udává měsíční příjem náhodně vybraného zaměstnance. Předpokládejme, že X ~ Ex(λ). Podle údajů Českého statistického úřadu dosáhla průměrná hrubá mzda v ČR ve 4. čtvrtletí roku 2010 hodnoty 25 752 Kč.

Úkol 1.: Zjistěte parametr λ pro náhodnou veličinu X.

Řešení: $E(X) = \int_0^\infty x\lambda e^{-\lambda x}\,dx = \ldots = \frac{1}{\lambda} = 25752 \Rightarrow \lambda = \frac{1}{25752} = 0{,}00003883$

Úkol 2.: Odvoďte obecný vzorec pro výpočet α-kvantilu náhodné veličiny X a pak vyjádřete medián náhodné veličiny X.
Co lze říci o vztahu střední hodnoty a mediánu?

Řešení: $\alpha = \Phi(K_\alpha(X)) = 1 - e^{-\lambda K_\alpha(X)} \Rightarrow K_\alpha(X) = -\frac{1}{\lambda}\ln(1-\alpha)$

Výpočet mediánu: $K_{0,50}(X) = -\frac{1}{\lambda}\ln\frac{1}{2} = \frac{\ln 2}{\lambda} = 25752 \cdot \ln 2 = 17850$
Znamená to, že aspoň polovina osob má průměrnou hrubou mzdu nejvýše 17 850 Kč a aspoň polovina osob má průměrnou hrubou mzdu aspoň 17 850 Kč.
Protože exponenciální rozložení je rozložení s kladnou šikmostí (lze spočítat, že šikmost = 2), bude medián vždy menší než střední hodnota.

Úkol 3.: Kolik procent zaměstnanců má podprůměrnou hrubou mzdu?

Řešení: $P\left(X < \frac{1}{\lambda}\right) = \int_0^{\frac{1}{\lambda}} \lambda e^{-\lambda x}\,dx = 1 - e^{-1} = 0{,}6321$
Znamená to, že téměř 2/3 zaměstnanců nedosáhnou na průměrnou mzdu. Průměr tedy není vhodnou charakteristikou střední úrovně mezd.

Práce se systémem MATLAB

Úkol 1.: Pomocí funkce exprnd náhodně vygenerujte příjmy n = 1000, 10 000 a 100 000 osob (střední hodnotu volte 25 752) a vytvořte histogram vygenerovaných příjmů.

```
r = exprnd(25752,n,1);
hist(r)
```

Úkol 2.: Vypočtěte průměrný příjem a vypočtěte medián příjmů.

```
m = mean(r)
x50 = median(r)
```

Zjištěné hodnoty porovnejte s teoretickými hodnotami: střední hodnota = 25 752 Kč, medián = 17 850 Kč.

Úkol 3.: Zjistěte, kolik procent osob bude mít podprůměrné příjmy.

```
pocet=0;
pocet=sum(r<m);
procento=100*pocet/n
```
Zjištěnou hodnotu porovnejte s teoretickou hodnotou 63,2%.