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)

Hustota: $\varphi(x) = \begin{cases} \lambda e^{-\lambda x} \text{ pro } x > 0 \\ 0 \text{ pro } x \le 0 \end{cases}$, distribuční funkce: $\Phi(x) = \begin{cases} 1 - e^{-\lambda x} \text{ pro } x > 0 \\ 0 \text{ pro } x \le 0 \end{cases}$

Nechť náhodná veličina X udává měsíční příjem náhodně vybraného zaměstnance. Předpokládejme, že $X \sim Ex(\lambda)$. Podle údajů Českého statistického úřadu dosáhla průměrná hrubá mzda v ČR za 1. až 3. čtvrtletí roku 2007 hodnoty 21 119 Kč.

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

Řešení: $E(X) = \int_0^\infty x\lambda e^{-\lambda x} dx = \ldots = \frac{1}{\lambda} = 21119 \Rightarrow \lambda = \frac{1}{21119} = 0{,}00004735$

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: $K_{0,50}(X) = -\frac{1}{\lambda}\ln\frac{1}{2} = \frac{\ln 2}{\lambda} = 21119 \cdot \ln 2 = 14639$

Znamená to, že aspoň polovina osob má průměrnou hrubou mzdu nejvýše 14 639 Kč a aspoň polovina osob má průměrnou hrubou mzdu aspoň 14 639 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 a vytvořte histogram vygenerovaných příjmů.
r = exprnd(20000,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.

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

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