

Využití MATLABu při práci s exponenciálním a Poissonovým rozložením

Základní poznatky o exponenciálním rozložení $Ex(\lambda)$

Náhodná veličina X udává dobu čekání na příchod nějaké události, která se může dostavit každým okamžikem se stejnou šancí bez ohledu na dosud pročekanou dobu. Přitom $1/\lambda$ vyjadřuje střední dobu čekání.

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

$$\text{kvantilová funkce: } \Phi^{-1}(\alpha) = -\frac{1}{\lambda} \ln(1 - \alpha), \text{ kde } 0 < \alpha < 1.$$

$$\text{Střední hodnota: } E(X) = \frac{1}{\lambda}, \text{ rozptyl: } D(X) = \frac{1}{\lambda^2}.$$

Interval spolehlivosti pro střední hodnotu: Necht' X_1, \dots, X_n je náhodný výběr z rozložení $Ex(\lambda)$ a necht' m je realizace výběrového průměru. Pak meze $100(1-\alpha)\%$ přibližného

$$\text{empirického intervalu spolehlivosti pro } E(X) = \frac{1}{\lambda} \text{ jsou: } d = \frac{2nm}{\chi^2_{1-\alpha/2}(2n)}, h = \frac{2nm}{\chi^2_{\alpha/2}(2n)}$$

Pozor, funkce v MATLABu pro práci s exponenciálním rozložením vyžadují zadávat převrácenou hodnotu parametru λ .

a) Kreslení grafu hustoty a distribuční funkce rozložení $Ex(1/2)$

```
x=[0:0.01:10]';  
f=expPDF(x,2);  
plot(x,f)  
df=expCDF(x,2);  
figure  
plot(x,df)
```

b) Kreslení grafu kvantilové funkce rozložení $Ex(1/2)$

```
alfa=[0.01:0.01:0.99]';  
kf=expinv(alfa,2);  
plot(alfa,kv)
```

c) Generování 100 realizací náhodné veličiny s rozložením $Ex(1/2)$ a kreslení histogramu s 10 třídicími intervaly

```
r=exprnd(2,100,1);  
hist(r)
```

d) Odhad střední hodnoty a meze intervalu spolehlivosti pro střední hodnotu na základě proměnné r

Hodnoty uložené v proměnné r považujeme za realizace náhodného výběru rozsahu 100 z rozložení $Ex(1/2)$

```
[m,meze]=expfit(r)
```

e) Výpočet střední hodnoty a rozptylu rozložení $Ex(1/2)$

```
[m,v]=expstat(2)
```

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?

$$\text{Řešení: } X \sim \text{Ex}(1/3), P(X \leq 2) = \int_0^2 \frac{1}{3} e^{-\frac{x}{3}} dx = \left[-e^{-\frac{x}{3}} \right]_0^2 = 1 - e^{-\frac{2}{3}} = 0,4866$$

V MATLABu: $p = \text{expcdf}(2,3)$

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: $X \sim \text{Ex}(1/600)$,

$$\begin{aligned} P(X \geq 800 + 200 / X \geq 800) &= P(X \geq 200) = 1 - P(X \leq 200) + P(X = 200) = 1 - \int_0^{200} \frac{1}{600} e^{-\frac{x}{600}} dx = \\ &= 1 - \left[-e^{-\frac{x}{600}} \right]_0^{200} = e^{-\frac{200}{600}} = e^{-\frac{1}{3}} = 0,7165 \end{aligned}$$

V MATLABu: $p = 1 - \text{expcdf}(200,600)$

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

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

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?

$$\text{Řešení: } X \sim \text{Ex}(1/2), P(X > 5) = 1 - P(X \leq 5) = 1 - \int_0^5 \frac{1}{2} e^{-\frac{x}{2}} dx = 1 - \left[-e^{-\frac{x}{2}} \right]_0^5 = e^{-2,5} = 0,082$$

V MATLABu: $p = 1 - \text{expcdf}(5,2)$

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

Řešení: ad a)

$$P(X_1 > t_0 \wedge X_2 > t_0) = P(X_1 > t_0)P(X_2 > t_0) = [1 - P(X_1 \leq t_0)][1 - P(X_2 \leq t_0)] =$$
$$= [1 - \Phi_1(t_0)][1 - \Phi_2(t_0)] = e^{-\lambda_1 t_0} e^{-\lambda_2 t_0} = e^{-t_0(\lambda_1 + \lambda_2)}$$

$$\text{ad b) } P(X_1 \leq t_0 \vee X_2 \leq t_0) = 1 - e^{-t_0(\lambda_1 + \lambda_2)}$$

Příklad 6.: Najděte 5. percentil náhodné veličiny $X \sim \text{Ex}(0,1)$

$$\text{Řešení: } 0,05 = \Phi(K_{0,05}(X)) = 1 - \exp(-0,1K_{0,05}(X)) \Rightarrow K_{0,05}(X) = -10 \ln 0,95 = 0,5129$$

V MATLABu: $K = \text{expinv}(0.05,10)$

Základní poznatky o Poissonově rozložení $Po(\lambda)$

Náhodná veličina $X \sim Po(\lambda)$ udává počet událostí, které nastanou v jednotkovém časovém intervalu případně v jednotkové oblasti, jestliže k událostem dochází náhodně, jednotlivě a vzájemně nezávisle. Parametr λ je střední hodnota počtu těchto událostí. Pokud sledujeme náhodnou veličinu, která udává počet událostí v intervalu délky t , pak uvedená náhodná veličina má rozložení $Po(t\lambda)$.

$$\text{Pravděpodobnostní funkce: } \pi(x) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & \text{pro } x = 0, 1, 2, \dots \\ 0 & \text{jinak} \end{cases}$$

$$\text{Distribuční funkce: } \Phi(x) = \sum_{t=0}^x \frac{\lambda^t}{t!} e^{-\lambda}$$

$$\text{Střední hodnota: } E(X) = \lambda, \text{ rozptyl: } D(X) = \lambda$$

Aproximace binomického rozložení pomocí Poissonova rozložení:

Nechť náhodná veličina $X \sim Bi(n, \nu)$. Za předpokladu, že $n \geq 30$ a $\nu \leq 0,1$, lze pravděpodobnostní funkci této náhodné veličiny uspokojivě aproximovat pravděpodobnostní funkcí rozložení $Po(n\nu)$:

$$P(X = x) \approx \frac{(n\nu)^x}{x!} e^{-n\nu}$$

Interval spolehlivosti pro střední hodnotu: Nechť X_1, \dots, X_n je náhodný výběr z rozložení $Po(\lambda)$ a nechť m je realizace výběrového průměru. Pak meze $100(1-\alpha)\%$ přibližného empirického intervalu spolehlivosti pro střední hodnotu λ jsou:

$$d = m - \sqrt{\frac{m}{n}} u_{1-\alpha/2}, \quad h = m + \sqrt{\frac{m}{n}} u_{1-\alpha/2}$$

a) Kreslení grafu pravděpodobnostní a distribuční funkce rozložení $Po(2)$

`x=[0:10]'`;

`pf=poisspdf(x,2);`

`plot((x,pf,'o')`

`df=poisscdfx,2);`

`figure`

`stairs(x,df)`

(Samostatný úkol: Jak nakreslit graf distribuční funkce bez svislých čar?)

b) Generování 100 realizací náhodné veličiny s rozložením $Po(2)$ a kreslení histogramu

`r=poissrnd(2,100,1);`

`hist(r,x)`

c) Odhad střední hodnoty a výpočet mezí intervalu spolehlivosti pro střední hodnotu na základě proměnné r

Hodnoty uložené v proměnné r považujeme za realizace náhodného výběru rozsahu 100 z rozložení $Po(2)$

`[m,meze]=poissfit(r)`

d) Výpočet střední hodnoty a rozptylu rozložení $Po(2)$

`[m,v]=poisstat(2)`

Příklady na využití Poissonova rozložení:

Příklad 1.: Při provozu balicího automatu vznikají během směny náhodné poruchy, které se řídí rozložením $Po(2)$. Jaká je pravděpodobnost, že během směny dojde k aspoň jedné poruše?

Řešení: X – počet poruch během směny, $X \sim Po(2)$, $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) =$
 $= 1 - \frac{2^0}{0!} e^{-2} = 0,8647.$

V MATLABu: $p = 1 - \text{poisspdf}(0,2)$

Příklad 2.: Telefonní ústředna zapojí během hodiny průměrně 15 hovorů. Jaká je pravděpodobnost, že během 4 minut ústředna zapojí a) právě jeden hovor, b) aspoň dva hovory?

Řešení: X – počet zapojených hovorů během 4 minut = 1/15 hodiny, $X \sim Po(t\lambda)$, kde $t = 1/15$ a $\lambda = 15$, tedy $X \sim Po(1)$.

ad a) $P(X = 1) = e^{-1} = 0,36788$,

ad b) $P(X \geq 2) = 1 - P(X \leq 1) = 1 - \pi(0) - \pi(1) = 1 - 2e^{-1} = 1 - 2 \cdot 0,36788 = 0,264242$

V MATLABu: a) $p = \text{poisspdf}(1,1)$, b) $p = 1 - \text{poisscdf}(1,1)$

Příklad 3.: Ze zkušenosti víme, že při správné obsluze stroje je v průměru 0,1% výrobků zmetkových. Ke stroji nastoupil nový pracovník. Za týden vyrobil 5 000 kusů, z nichž 11 bylo zmetkových. Lze takto vysoký počet zmetků vysvětlit působením náhodných vlivů?

Řešení: Budeme počítat pravděpodobnost, že pracovník vyrobil aspoň 11 zmetků za předpokladu, že stroj je obsluhován správně.

X – počet vyrobených zmetků za týden, $X \sim Bi(5000, 0,01)$. Při splnění podmínek dobré aproximace lze rozložení veličiny X aproximovat rozložením $Po(5)$.

$$P(X \geq 11) = 1 - P(X \leq 10) = 1 - \sum_{t=0}^{10} \frac{5^t}{5!} e^{-5} = 1 - 0,9863 = 0,0137.$$

Je zřejmé, že nový pracovník nepracuje správně.

V MATLABu: $p = 1 - \text{poisscdf}(10,5)$

Přesný výpočet v MATLABu: $p = 1 - \text{binocdf}(10,5000,0.001)$

Příklad 4.: Pro $n = 30$ a $v = 0,1$ ilustруйте aproximaci binomického rozložení $Bi(n, v)$ Poissonovým rozložením $Po(nv)$. Vypočtené hodnoty obou pravděpodobnostních funkcí v bodech $x = 0, 1, \dots, 30$ zapište do tabulky.

Řešení:

```
x=[0:1:30]';
```

```
pf1=binopdf(x,30,0.01);
```

```
pf2=poisspdf(x,0.3);
```

```
[x pf1 pf2]
```