

## Využití MATLABu při práci s Poissonovým rozložením

### 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  $\lambda$  je střední hodnota počtu těchto událostí. Pokud sledujeme náhodnou veličinu, která udává počet událostí v intervalu délky  $t$ , pak uvedená náhodná veličina má rozložení  $Po(t\lambda)$ .

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

Pravděpodobnostní funkce:

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

Distribuční funkce:

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

Aproximace binomického rozložení pomocí Poissonova rozložení:

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

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

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  $\lambda$  jsou:

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

MATLAB počítá meze  $100(1-\alpha)\%$  přibližného empirického intervalu spolehlivosti pro střední hodnotu  $\lambda$  podle vzorce:

$$d = \frac{1}{2n} \chi^2_{\alpha/2}(2nm), \quad h = \frac{1}{2n} \chi^2_{1-\alpha/2}(2(nm+1))$$

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

```
x=[0:10]';
```

```
pf=poisspdf(x,2);
```

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

```
df=poisscdf(x,2);
```

```
figure
```

```
stairs(x,df)
```

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

Jedno z možných řešení:

```
hold on
```

```
for i=1:(length(x)-2) plot([i,i],[0,1],'w'); end
```

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 meze 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,  $X \sim \text{Po}(2)$ ,  $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) =$

$$1 - \frac{2^0}{0!} e^{-2} = 0,8647.$$

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,  $X \sim \text{Po}(t\lambda)$ , kde  $t = 1/15$  a  $\lambda = 15$ , tedy  $X \sim \text{Po}(1)$ .

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,  $X \sim \text{Bi}(5000, 0,001)$ . Při splnění podmínek dobré aproximace lze rozložení veličiny X aproximovat rozložením Po(5).

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

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  $\theta = 0,1$  ilustруйте aproximaci binomického rozložení  $Bi(n, \theta)$  Poissonovým rozložení  $Po(n\theta)$ . 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]
```

**Příklad 5.:** Na výrobní lince se zhruba každé dvě hodiny vyskytne porucha. Sestrojte tabulku uvádějící, s jakou pravděpodobností se na této lince během osmihodinové pracovní směny nevyskytne žádná porucha, vyskytne jedna porucha, vyskytnou dvě poruchy atd. až vyskytne deset poruch. S jakou pravděpodobností nastane více než deset poruch? Určete nejpravděpodobnější počet poruch během osmihodinové směny.

**Řešení:**

$X$  - počet poruch během osmi hodin,  $X \sim Po(t\lambda)$ , kde  $t = 2$  a  $\lambda = 4$ , tedy  $X \sim Po(4)$ .

$$P(X = 0) = \pi(0) = \frac{4^0}{0!} e^{-4} = 0,0183$$

$$P(X = 1) = \pi(1) = \frac{4^1}{1!} e^{-4} = 0,0733$$

.....

$$P(X = 10) = \pi(10) = \frac{4^{10}}{10!} e^{-4} = 0,0053$$

Přehledný zápis všech výsledků viz následující tabulka (včetně kumulovaných hodnot čili distribuční funkce). Všimněte si, že ačkoli jsme samozřejmě nevyčerpali všechny možné počty poruch (těch je nekonečně mnoho), dosahuje součet všech vypočtených pravděpodobností téměř 1. Najdeme ho jako hodnotu distribuční funkce  $\Phi(10)=0,9972$  a jde o pravděpodobnost, že počet poruch bude nejvýše deset. Celková pravděpodobnost, že by byl počet poruch naopak větší než deset, je tedy  $1-\Phi(10)$  a činí pouze necelá tři promile ( $1-0,9972 = 0,0028$ ).

n	0	1	2	3	4	5	6	7	8	9	10
$\pi(n)$	0,0183	0,0733	0,1465	0,1954	0,1954	0,1563	0,1042	0,0595	0,0298	0,0132	0,0053
$\Phi(n)$	0,0183	0,0916	0,2381	0,4335	0,6288	0,7851	0,8893	0,9489	0,9786	0,9919	0,9972

Nejpravděpodobnější počet poruch během osmihodinové směny je tři až čtyři (hledáme modus, jde o bimodální případ).

V MATLABu:

```
x=[0:10]';
pf=poisspdf(x,4);
df=poisscdf(x,4);
[x pf df]
```

Grafické znázornění:

plot(x,pf,'o')  
stairs(x,df)

**Příklad 6.:** Semena rostlin určitého druhu jsou znečištěna malým množstvím plevel. Je známo, že na jedné jednotce plochy vyrostou po osetí v průměru 4 rostliny plevel. Vypočítejte pravděpodobnost, že na dané jednotce plochy:

- a) Nebude žádný plevel,
- b) Vyrostou nejvýše 3 rostliny plevel,
- c) Vyrostou aspoň 5, ale nejvýše 7 rostlin plevel.

**Výsledek:** ad a) 0,0183, ad b) 0,4335, ad c) 0,32.