5 Pravděpodobnostní funkce, hustoty a distribuční funkce, výpočet pravděpodobností pomocí distribučních funkcí

- experiment složený z opakování jednoho náhodného pokusu stále dokola
- provedu pokus → hodím kostkou → nastane jev
- pravděpodobnost = jak velká je naděje, že daný jev nastane
- náhodná veličina X ... zobrazení ze základního prostoru do množiny reálných čísel
  - podle dat dělíme náhodné veličiny na diskrétní a spojité
- diskrétní
  - pstní fce $p(x)$
    - $P(X = x)$
  - distribuční fce $F(x)$
    - $P(X \leq x)$

5.1 Binomické rozdělení $\mathrm{Bin}(n, \theta)$

- $X \sim \mathrm{Bin}(n, \theta)$
- X... počet úspěchů v posloupnosti n nezávislých opakovaných pokusů, přičemž pravděpodobnost úspěchu v každém pokusu je vyjádřena pomocí parametru θ.

Příklad 5.1. Nakreslete graf pravděpodobnostní funkce a distribuční funkce náhodné veličiny $X \sim \mathrm{Bin}(6, 0.5)$.

```
#graf pravdepodobnostni funkce
x  <- 0:6
px <- dbinom(x, size=6, prob=0.5)
plot(x, px, type='h', main='X~Bin(6,0.5)', ylab='pravdepodobnostni funkce', xlab=
   'x')
points(x, px, col='red', pch=19, cex=0.8)

#graf distribucni funkce
Fx <- pbinom(x, size=6, prob=0.5)
n  <- length(Fx)

plot(x, Fx, type='n', main='X~Bin(6,0.5)', ylab='distribucni funkce',
     xlab='x', xlim=c(-1,n), ylim=c(0,1))
segments(x, Fx, x+1, Fx)
arrows(0,  0, -1, 0, length=0.1)
arrows(n-1,1,  n, 1, length=0.1)
points(x, Fx, col='red', pch=19, cex=0.8)
points(x, c(0, Fx[-n]), col='red', bg='white', pch=21, cex=0.8)
```

5.2 Poissonovo rozdělení Po(λ)

- $X \ldots$ udává počet událostí, které nastanou v jednotkovém časovém intervalu (resp. v jednotkové oblasti), přičemž k událostem dochází náhodně, jednotlivě a vzájemně nezávisle. Parametr $\lambda > 0$ je střední počet těchto událostí.
- počet zákazníků, kteří vešli do pekárny za jeden den
- počet tramvají, které přijeli na zastávku za půl hodiny
- $X \sim \mathrm{Po}(\lambda)$.
- pstní fce

$$p(x) = \begin{cases} \dfrac{\lambda^x}{x!} e^{-\lambda} & \text{pro x=0,1,\ldots;} \\ 0 & \text{jinak.} \end{cases} \tag{1}$$

- $\lambda > 0$; $EX = \lambda$; $DX = \lambda$
- dpois(x, lambda)
- ppois(x, lambda)

Příklad 5.3. Při provozu balicího automatu vznikají během směny náhodné poruchy, které se řídí rozdělením Po(2). Jaká je pravděpodobnost, že během směny dojde k alespoň jedné poruše?

```
1-dpois(0, lambda=2)
1-ppois(0, lambda=2)
```

5.3 Spojitá náhodná veličina

- nabývají lib. hodnoty z daného intervalu
- změříme výšku člověka:
  - pst, vybraný člověk bude měřit 140-145 cm $\Pr(140\,cm < X < 145\,cm)$
  - pst, vybraný člověk bude vyšší než 170 cm $\Pr(X > 170\,cm)$
- hustota $f(x)$
  - pravděpodobnost realizace X v libovolném intervalu I se dá vyjádřit jako plocha pod křivkou hustoty $f(x)$
  - $$\Pr(X \in I) = \int_{x \in I} f(x)dx, \tag{2}$$
  - $\Pr(X = x) = 0 \ldots$ plocha pod bodem je čára a obsah je tedy 0.
  - $f(x) \geq 0$; plocha pod celou křivkou hustoty = 1
  - Gaussova křivka hustoty

- distribuční funkce $F(x)$
  - analogie distribuční funkce v diskrétním případě
  - $F(x) = \Pr(X \leq x) \Pr(X < x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$
  - $F(x)$ ... pst, že náhodná veličina X nepřekročí hodnotu x

  - $\Pr(X \geq x) = \Pr(X > x) = 1 - \Pr(X \leq x) = 1 - \Pr(X < x)$.
- spojité náhodné veličiny se řídí spojitým rozdělením
  - Exponenciální $Ex(\lambda)$
  - Normální $N(\mu, \sigma^2)$
  - Standardizované normální $N(0, 1)$
  - odvozeniny od $N(0, 1)$
    - Chi-kvadrátové $\chi^2(n)$
    - Studentovo $t(n)$
    - Fisherovo-Snedecorovo $F(n_1, n_2)$

5.4 Exponenciální rozdělení Exp(λ)

- X... udává dobu čekání na příchod nějaké události, která se může dostavit každým okamžikem se stejnou šancí bez ohledu na dosud pročekanou dobu. Přitom $\frac{1}{\lambda}$ vyjadřuje střední hodnotu doby čekání.
- doba potřebná k obsloužení zákazníka
- doba během níž přijede tramvaj na zastávku bez ohledu na dobu, kdy přijela poslední tramvaj
- náhodná veličina $X \sim Ex(\lambda)$
- hustota

$$f(x) = \begin{cases} \lambda e^{-x\lambda} & \text{pro } x > 0; \\ 0 & \text{jinak.} \end{cases} \tag{3}$$

- $\lambda > 0$; $E(X) = \frac{1}{\lambda}, D(X) = \frac{1}{\lambda^2}$
- $f(x)$ ... dexp(x, lambda)
- $F(x)$ ... pexp(x, lambda)

Příklad 5.4. Nakreslete graf hustoty a distribuční funkce náhodné veličiny $X \sim \mathrm{Exp}(2)$.

```
#graf hustoty
x  <- seq(from=0, to=2.5, length=512)
fx <- dexp(x, rate=2)
plot(x, fx, type='l', main='X~Exp(2)', xlab='x', col='red')

#graf distribucni funkce
Fx <- pexp(x, rate=2)
plot(x, Fx, main='X~Exp(2)', xlab='x', ylab='distribucni funkce', type='l', col='
   red')
```

Příklad 5.5. Doba do ukončení opravy v opravně obuvi je náhodná veličina, která se řídí exponenciálním rozdělením se střední dobou opravy 3 dny. Jaká je pravděpodobnost, že oprava bude ukončena do dvou dnů?

```
pexp(2, rate=1/3)
```

5.5 Normální rozdělení $N(\mu, \sigma^2)$

- nejpoužívanější ze všech rozložení
- $X \sim N(\mu, \sigma^2)$
- hustota
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}. \tag{4}$$
- $E(X) = \mu$; $D(X) = \sigma^2$.

standardizované normální rozložení

- $\mu = 0$, $\sigma^2 = 0$
- $X \sim N(0, 1)$
- hustota:
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{5}$$
- distribuční funkci $N(0, 1)$ značíme $\Phi(x)$
- normální rozdělení je symetrické okolo nuly $\rightarrow \Phi(-x) = 1 - \Phi(x)$ (platí jen pro $N(0, 1)$.
- $f(x)$ ... dnorm(x, μ, $\sqrt{\sigma^2}$)
- $F(x)$ ... pnorm(x, μ, $\sqrt{\sigma^2}$)

Vlastnosti normálního rozložení

- Pokud $X \sim N(\mu, \sigma^2)$, potom $\dfrac{X-\mu}{\sigma} \sim N(0, 1)$
- Pokud $X \sim N(\mu, \sigma^2)$, potom $Y = a + bX \sim N(a + b\mu, b^2\sigma^2)$

Příklad 5.7. Nakreslete graf hustoty a distribuční funkce náhodné veličiny $X \sim N(-1, 2)$.

```
#graf hustoty
x  <- seq(from=-5, to=3, length=512)
fx <- dnorm(x,  mean=-1, sd=sqrt(2))
plot(x, fx, type='l', main=bquote(paste('X ~ N( ',mu,'=-1, ',sigma^2,'=2 )')),
   xlim=c(-4.5,2.5),
     col='red')

#graf distribucni funkce
Fx <- pnorm(x, mean=-1, sd=sqrt(2))
plot(x, Fx, type='l', main=bquote(paste('X ~ N( ',mu,'=-1, ',sigma^2,'=2 )')),
   xlim=c(-4.5,2.5),
     ylab='distribucni funkce', col='red')
```

Příklad 5.9. Životnost baterie v hodinách je náhodná veličina, která má normální rozdělení se střední hodnotou 300 hodin a směrodatnou odchylkou 35 hodin. Jaká je pravděpodobnost, že náhodně vybraná baterie bude mít životnost

a) alespoň 320 hodin?

b) nejvýše 310 hodin?

a) $\Pr(X \geq 320) = \Pr(X > 320) = 1 - \Pr(X \leq 320) = 1 - \Pr(X < 320) = 1 - \Pr\left(\frac{X-\mu}{\sqrt{\sigma^2}} < \frac{320-\mu}{\sqrt{\sigma^2}}\right) =$
$1 - \Pr\left(Y < \frac{320-300}{35}\right) = 1 - \Phi(0.5714286)$

```
1-pnorm(320, mean=300, sd=35)
1-pnorm((320-300)/35, mean=0, sd=1)
```

b) $\Pr(X \leq 310) = \Pr(X < 310)$

```
pnorm(310, mean=300, sd=35)
```

Pro zajímavost: (Nedělali jsme na hodině)

Příklad 5.11. Nakreslete graf hustoty dvourozměrného standardizovaného normálního rozložení.

```
source('AS-funkce.R')
x  <- seq(from=-3, to=3, length=40)
y  <- seq(from=-3, to=3, length=40)
nx <- length(x)
ny <- length(y)
z  <- matrix(NA, nrow=nx, ncol=ny)
for(i in 1:nx){
  for(j in 1:ny){
    z[i,j]  <- norm2(x[i], y[j], mu1=0, mu2=0, sigma1=1, sigma2=1)
  }
}

color      <- terrain.colors(12)
```

```
stredy      <- (z[-1, -1] + z[-1, -ncol(z)] + z[-nrow(z), -1] + z[-nrow(z), -ncol(
   z)])/4
stredy.col <- cut(stredy, 12)
persp(x, y, z, col = color[stredy.col], phi = 30, theta = -45)
```