NMAI059 Pravděpodobnost a statistika 1
5. přednáška

Robert Šámal

Přehled

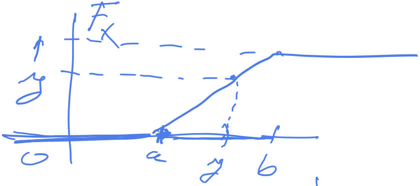
Shrnutí, co už víme

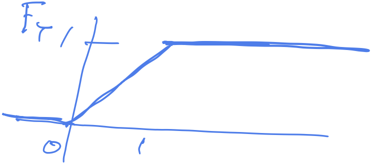
- $\boxed{F_X(x) := P(X \leq x)}$ distribuční funkce, existuje vždy, je neklesající, zprava spojitá, $F_X(-\infty) = 0$, $F_X(+\infty) = 1$
- $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$ pro tzv. spojité n.v. Pro ty dále platí:
- $P(a \leq X \leq b) = \int_a^b f_X(t)dt$, spec. tedy $P(X = x) = 0$ pro každé $x \in \mathbb{R}$
- obecněji: $P(X \in A) = \int_A f_X(t)dt$, kdykoli umíme přes množinu A integrovat
- spec. tedy $\int_{-\infty}^{\infty} f_X(t)dt = 1$
- $\mathbb{E}(X) = \int_{-\infty}^{\infty} t\, f_X(t)dt$ ← def analog. $\sum t\, p_X(t)$ pro diskr. a. v.
- $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(t) f_X(t)dt$ ← věta LOTUS
- Pokud je hustota spojitá, tak navíc platí: $f_X = F_X'$ (základní věta kalkulu).

Uniformní rozdělení

- N.v. X má uniformní rozdělení na intervalu $[a, b]$, píšeme $X \sim U(a, b)$, pokud $f_X(x) = 1/(b-a)$ pro $x \in [a, b]$ a $f_X(x) = 0$ jinak.

$Y = \frac{X-a}{b-a}$ $\quad X = a + (b-a)Y$





$F_Y(y) = P(Y \le y) = P\left(\frac{X-a}{b-a} \le y\right) = P(X \le a + y(b-a))$

$= F_X(a + y(b-a)) = y$

$y \in [0,1]$

Universalita uniformního rozdělení

Věta

Nechť F je rostoucí spojitá funkce s $\lim_{x\to-\infty} F(x) = 0$ *a* $\lim_{x\to+\infty} F(x) = 1$.

1. *Nechť $U \sim U(0,1)$ a $X = F^{-1}(U)$. Pak X má distribuční funkci F.*
2. *Nechť X je n.v. s distribuční funkcí $F = F_X$. Pak $F(X) \sim U(0,1)$.*

$F = \frac{1}{\pi} \operatorname{arctg} x + \frac{1}{2}$

$X \dashrightarrow F(X) = \frac{1}{\pi} \operatorname{arctg} X + \frac{1}{2}$

$\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{F} \mathbb{R}$

$F(X(\omega)) = P(X \le X(\omega))$

$\cancel{F(X)}$

$F(x) = P(X \le x)$

$F(X) = P(X \le X) = 1$??

$\boxed{\text{Dk}}$ (1) $U \sim U(0,1)$ $\quad X = F^{-1}(U)$ $\quad F^{-1}$ def. na $(0,1)$

F rostoucí

$F(X) \in (0,1)$

$$F_X(x) = P(X \le x) = P(F^{-1}(U) \le x)$$

$$= P(U \le F(x)) = F_U(F(x)) = F(x)$$

$\Rightarrow F_X = F$

(2) X má d.f. F, $\quad Y = F(X)$ $\quad$ chceme $Y \sim U(0,1)$

(F rost.)

$$F_Y(y) = P(Y \le y) = P(F(X) \le y)$$

$y \in (0,1)$

$$= P(X \le F^{-1}(y)) = F(F^{-1}(y)) = y$$

$F_Y(y) = 0$ pro $y \le 0$, $\quad F_Y(y) = 1$ pro $y \ge 1$

Rozptyl spojité n.v.

1. moment

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x\, f_X(x) dx$$

def.

2. moment

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2\, f_X(x) dx$$

LOTUS

Označíme-li $\mu = \mathbb{E}(X)$, tak

$$var(X) = \int_{-\infty}^{\infty} (x-\mu)^2\, f_X(x) dx = \underbrace{\mathbb{E}(X^2) - (\mathbb{E}(X))^2}_{\text{většinou}}$$

def $\mathbb{E}(X-\mu)^2$

$\mathbb{E}(aX) = a\,\mathbb{E}X$

$\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y$

Přehled

Spojité náhodné veličiny

Konkrétní spojitá rozdělení a jejich parametry

Náhodné vektory

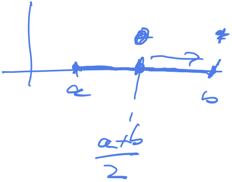
Uniformní rozdělení

- N.v. X má uniformní rozdělení na intervalu $[a, b]$, píšeme $X \sim U(a, b)$, pokud $f_X(x) = 1/(b-a)$ pro $x \in [a, b]$ a $f_X(x) = 0$ jinak.

$$EX = \int_a^b x \frac{1}{b-a} = \left[\frac{x^2}{2}\right]_a^b \frac{1}{b-a} = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

$$\operatorname{Var} X = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a} = \int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} y^2 \frac{1}{b-a}$$

$$y = x - \frac{a+b}{2}$$



$$= \left[\frac{y^3}{3}\right]_{-\frac{b-a}{2}}^{\frac{b-a}{2}} \frac{1}{b-a} = \frac{2(b-a)^3}{3 \cdot 8} \frac{1}{b-a} = \frac{(b-a)^2}{12}$$

$$\sigma_X = \sqrt{\operatorname{Var} X} = \frac{1}{\sqrt{12}} (b-a) = \frac{b-a}{2} \frac{1}{\sqrt{3}}$$

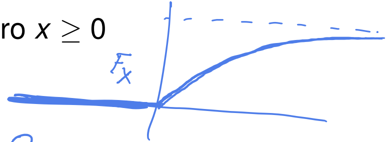
Exponenciální rozdělení

$$\mathrm{Var}(X) = EX^2 - (EX)^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

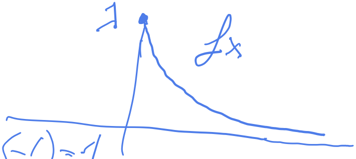
$$\boxed{\lambda > 0}$$

$$F_X(x) = \begin{cases} 0 & \text{pro } x \leq 0 \\ 1 - e^{-\lambda x} & \text{pro } x \geq 0 \end{cases}$$

$x \leq 0$



$$f_X(x) = \begin{cases} 0 & \\ F_X' = \lambda e^{-\lambda x} & x \geq 0 \end{cases}$$



$$\int f_X = \int_0^\infty \lambda e^{-\lambda x} = \left[-e^{-\lambda x}\right]_0^\infty = 0 - (-1) = 1$$

$$EX = \int_0^\infty \underset{u}{x}\, \underset{v'}{\lambda e^{-\lambda x}} = \left[x\left(-e^{-\lambda x}\right)\right]_0^\infty - \int_0^\infty 1 \cdot \left(-e^{-\lambda x}\right) = \int_0^\infty e^{-\lambda x} = \frac{1}{\lambda}$$

$$EX^2 = \int_0^\infty x^2 \lambda e^{-\lambda x} = \left[x^2\left(-e^{-\lambda x}\right)\right]_0^\infty + \int_0^\infty 2x e^{-\lambda x} = \frac{2}{\lambda} \cdot \int x \lambda e^{-\lambda x} = \frac{2}{\lambda^2}$$

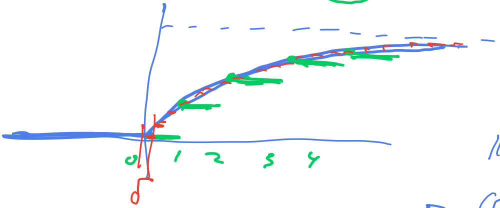
Souvislost *Exp* a *Geom*

$Y \sim Geom(p)$

$P(Y > k) = (1-p)^k \qquad k \in N$

$F_Y(k) = P(Y \le k) = 1 - (1-p)^k = 1 - e^{-\lambda k}$

$F_Y(y) = 0 \qquad y < 0 \qquad e^{-\lambda} = 1-p$



$F_X(\delta y)$

$\parallel$

$Y_\delta = \delta Y$

$\delta \to 0$

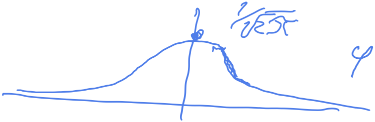
$F_{Y_\delta}(\delta y) = F_Y(y) = 1 - (1-p)^y$

$1-p = e^{-\lambda\delta}$

$(1-p)^y = e^{-\lambda\delta y} \Leftrightarrow$

$p = 1 - e^{-\lambda\delta}$

Normální rozdělení



- $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$
- $\Phi(x)$ – primitivní funkce k φ
- *Standardní normální rozdělení* $N(0,1)$ má hustotu φ a distribuční funkci Φ.
- Pokud $Z \sim N(0,1)$, tak $\mathbb{E}(Z) = 0$, $var(Z) = 1$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \text{nelze zjednodušit}$$

sudé fce

$$EZ = \int_{-\infty}^{\infty} \underbrace{z\,\varphi(z)}_{\text{lichá fce}} = 0$$

$$var Z = EZ^2 = \int_{-\infty}^{\infty} z^2 \varphi(z) = \int_{-\infty}^{\infty} z \cdot \frac{1}{\sqrt{2\pi}} z e^{-z^2/2} = \frac{\left[z\left(-e^{-z^2/2}\right)\right]_{-\infty}^{\infty}}{\sqrt{2\pi}} + \int_{-\infty}^{\infty} 1 \cdot \varphi(z) = 1$$

$$= 0$$

$$\int_{-\infty}^{\infty} \varphi(z) = 1$$

$$\int_{-\infty}^{\infty} e^{-x^2/2} = c = \sqrt{2\pi}$$

Normální rozdělení

- Pro $\mu, \sigma \in \mathbb{R}$, $\sigma > 0$ položme $X = \mu + \sigma \cdot Z$, kde $Z \sim N(0,1)$.
- Píšeme $X \sim N(\mu, \sigma^2)$ – obecné normální rozdělení.
- Normální rozdělení $N(\mu, \sigma^2)$ má hustotu $\frac{1}{\sigma}\varphi(\frac{x-\mu}{\sigma})$.
- Pokud $X_i \sim N(\mu_i, \sigma_i^2)$ a $X_1, \dots, X_k$ jsou n.n.v., pak

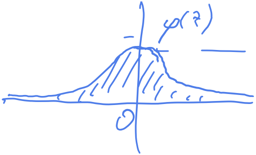
$$X_1 + \dots + X_k \sim N(\mu, \sigma^2).$$

$\mu = \mu_1 + \dots + \mu_k$

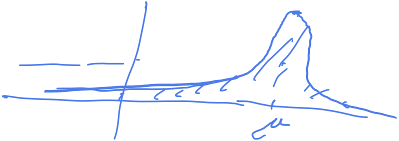
$\sigma^2 = \sigma_1^2 + \dots + \sigma_k^2$

$\mathbb{E}X = \mathbb{E}(\mu + \sigma Z) = \mu + \sigma \cdot \mathbb{E}Z = \mu$

$\operatorname{var} X = \operatorname{var}(\mu + \sigma Z) = \sigma^2 \operatorname{var} Z = \sigma^2$



$\sigma = \frac{1}{2}$



Normální rozdělení

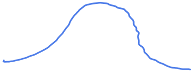
- Pravidlo 3σ (68–95–99.7 rule)
- CLT   CENTR. LIM. VĚTA

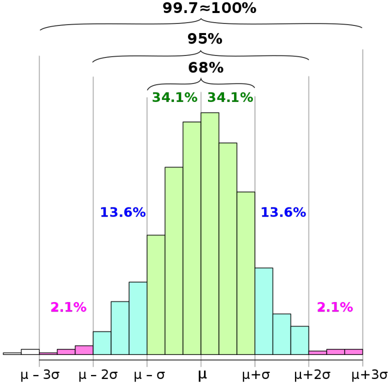
$X_1 \dots X_k$ stejné rozděl.
n.n.v.

$X_1 + \dots + X_k \doteq$ normální

$X_i = 0, 1$ s pstí $\frac{1}{2}$

$\binom{n}{k}$





(Obrázek z Wikipedie, autor Melikamp.)

$Z \sim N(\mu, \sigma^2)$

$P(|Z-\mu| \le 1\sigma) \doteq 68\%$

$P(|Z-\mu| \le 2\sigma) \doteq 95\%$

$P(|Z-\mu| \le 3\sigma) \doteq 99.7\%$

Cauchyho rozdělení

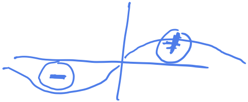
- *Cauchyho rozdělení:* hustota $f(x) = \frac{1}{\pi(1+x^2)}$
- nemá střední hodnotu!!!

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} = \left[\operatorname{arctg} x\right]_{-\infty}^{\infty} = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi$$

$$\int f = 1$$

$$\int x \cdot f(x) = \int \frac{x}{\pi(1+x^2)} = \left[\frac{1}{2\pi} \ln(1+x^2)\right]_{-\infty}^{\infty}$$

$$= \infty - \infty$$



Gamma rozdělení

- *Gamma*(w, λ), *gamma rozdělení s parametry* $w > 0$ *a* $\lambda > 0$ má hustotu

$$f(x) = \begin{cases} 0 & \text{pro } x \leq 0 \\ \frac{1}{\Gamma(w)} \lambda^w x^{w-1} e^{-\lambda x} & \text{pro } x \geq 0 \end{cases}$$

kde $\Gamma(w) = (w-1)! = \int_0^\infty x^{w-1} e^{-x} dx$.
Pro $w = 1$ dostáváme znovu exponenciální rozdělení.
Pokud sečteme n n.n.v. s rozdělením $Exp(\lambda)$, dostaneme n.v. $\sim Gamma(n, \lambda)$.
Modeluje mj. životnost součástky.

Další stručněji

- *Beta*(s, t) – beta rozdělení s parametry $s, t > 0$ má hustotu definovanou na $(0, 1)$ předpisem
$$f(t) = \frac{1}{B(s,t)} x^{s-1}(1-x)^{t-1},$$
přičemž *normovací konstanta* $B(s, t)$ je definovaná tak, aby vznikla hustota.
- χ^2 rozdělení s n stupni volnosti (χ^2_n) je jiné jméno $\Gamma(\frac{1}{2}n, \frac{1}{2})$. Je to rozdělení $Z_1 + \cdots + Z_n$, kde $Z_i \sim N(0, 1)$ jsou n.n.v.
- Student t-distribution
- atd. atd.

Přehled

Spojité náhodné veličiny

Konkrétní spojitá rozdělení a jejich parametry

Náhodné vektory

Sdružená distribuční funkce (Joint cdf)

Definice

Pro n.v. X, Y na pravděpodobnostním prostoru $(\Omega, \mathcal{F}, P)$ definujeme jejich sdruženou distribuční funkci (joint cdf) $F_{X,Y} : \mathbb{R}^2 \to [0,1]$ předpisem

$$F_{X,Y}(x,y) = P(\{\omega \in \Omega : X(\omega) \leq x \& Y(\omega) \leq y).$$

- Mohli bychom definovat i pro více než dvě n.v. $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$.
- Můžeme odsud odvodit pravděpodobnost obdélníku: $P(X \in (a,b] \& Y \in (c,d]) =$

Sdružená hustota (Joint pdf)

- Často můžeme sdruženou distribuční funkci psát jako integrál pomocí nezáporné funkce $f_{X,Y}$

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x,y) dxdy.$$

- Pak nazýváme n.v. X, Y sdružené spojité. Funkce $f_{X,Y}$ je jejich *sdružená hustota.*
- Jako u jednorozměrného případu může být $f_{X,Y} > 1$.
- Stejně jako u jednorozměrného případu můžeme pak pomocí hustoty vyjádřit i další pravděpodobnosti:

$$P((X,Y) \in S) = \int_{S} f_{X,Y}(x,y) dxdy.$$

- $f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$

LOTUS

- Stejně jako v diskrétním případu platí pro střední hodnotu funkce dvou n.v.

$$\mathbb{E}(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy.$$

- A stejně jako v diskrétním případu odsud odvodíme $\mathbb{E}(aX + bY + c) = a \cdot \mathbb{E}(X) + b \cdot \mathbb{E}(Y) + c$.

Vícerozměrné normální rozdělení