NMAI059 Pravděpodobnost a statistika 1
6. přednáška

Robert Šámal

Přehled

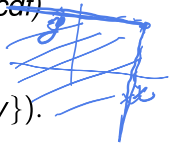
Sdružená distribuční funkce (Joint cdf)

(1D případ: $P(X \in (a,b]) = F_X(b) - F_X(a)$)

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

(x, y)



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

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

$F(b,d) - F(a,d) - F(b,c) + F(a,c)$

Sdružená hustota (Joint pdf)

Fubiniho věta

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

$\iint_{\mathbb{R}^2} f_{X,Y} = 1$

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s,t)\,dt\,ds.$$

$= \int_{t=-\infty}^{y} \int_{s=-\infty}^{x} f(s,t)\, ds\,dt$

- Pak nazýváme n.v. X, Y sdruženě 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:

$$P((X,Y) \in S) = \int_S f_{X,Y}(x,y)\,dx\,dy.$$

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

$f_{X,Y} = \begin{cases} c \\ 0 \end{cases}$

$c = \dfrac{1}{\text{obsah kruhu}}$

LOTUS

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

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

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

$\iint f = 1$

$g(x,y) = ax + by + c$

$E\,g(X,Y) = \iint (ax + by + c) f(x,y) = \iint a x f(x,y) + \iint b y f(x,y) + \underbrace{\iint c f(x,y)}_{= c}$

$a \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y) = a \int x \int f(x,y) = a \int x f_X(x) = a\,EX$

$\sim P(X = x)$ $\quad f_X(x)$?

Nezávislost spojitých náhodných veličin

$\{X = x\}$ a $\{Y = y\}$ -- pro disk. n.v.

Definice

Libovolné náhodné veličiny nazveme nezávislé (independent), pokud jevy $\{X \le x\}$ a $\{Y \le y\}$ jsou nezávislé pro libovolná $x, y \in \mathbb{R}$. Ekvivalentně,

Pr. průniku

$$P(X \le x, Y \le y) = P(X \le x)P(Y \le y),$$

$$\boxed{F_{X,Y}(x,y) = F_X(x)F_Y(y)}$$

$\boxed{DK \Uparrow}$ $\int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\, ds\, dt$

$= \iint f_X(s) f_Y(t) \;\text{?}\; = \int_{-\infty}^{x} f_X(s)\, ds \cdot \int_{-\infty}^{y} f_Y(t)\, dt = F_X(x) \cdot F_Y(y)$

Věta

Nechť X, Y mají sdruženou hustotu ~~f_X, f_Y~~ $f_{X,Y}$. Následující tvrzení jsou ekvivalentní:

- *X, Y jsou nezávislé* $\Uparrow \Downarrow$
- $f_{X,Y}(x,y) = f_X(x) f_Y(y)$ $\forall x, y \in \mathbb{R}$

$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$

$\frac{\partial^2}{\partial x \partial y} F = \left(\frac{\partial}{\partial x} F_X(x)\right)\left(\frac{\partial}{\partial y} F_Y(y)\right)$ (f_X, f_Y)

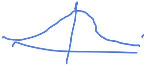
Vícerozměrné normální rozdělení

φ .. hust. $N(0,1)$

$\int_{\mathbb{R}^n} f = 1$

$= \int_{\mathbb{R}} \varphi \cdot \int_{\mathbb{R}} \varphi \ldots = 1$

- $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$



- $f(t_1, \ldots, t_n) = \varphi(t_1)\varphi(t_2)\cdots\varphi(t_n) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{t_1^2+\cdots+t_n^2}{2}} > 0$

f je hustota

- $f(t_1, \ldots, t_n) = (2\pi)^{-n/2} e^{-r^2/2}$, kde $r^2 = t_1^2 + \cdots + t_n^2$
  radiálně symetrická funkce

$Z = (Z_1, \ldots, Z_n)$

n-rozm. norm. rozd.
Gaussovo rozd.

- lze použít k jednoduchému generování bodu na n-rozměrné sféře $\quad Z/\|Z\|$
- první souřadnice je $N(0, 1)$, $\quad Z_2 \sim N(0,1), \ldots, Z_n \sim N(0,1)$
- tudíž skal. součin s libovolným jednotkovým vektorem je $N(0, 1)$

$u \in \mathbb{R}^n$, $\|u\| = 1$ ..... u je první vektor báze

$\sum u_i^2 = \|u\| = 1$

$\langle Z, u \rangle =$ souř. $Z \sim N(0,1)$

→ $Z_1, \ldots, Z_n$ jsou n.n.v.

→ Z vygen. jako n-tici $N(0,1)$ nez. n.v.

$\langle Z, u \rangle = \left(\sum u_i Z_i\right) \sim N(0,1)$

Vícerozměrné normální rozdělení obecné

- Obecněji můžeme vzít náhodný vektor s hustotou $c \cdot e^{Q(t)}$, kde $c > 0$ je vhodná konstanta a $Q(t)$ je obecná kvadratická funkce.

$$Q(t) = -\frac{t_1^2 + \dots + t_4^2}{2}$$

n^2 koef.

- Používá se ve strojovém učení.
- Souřadnice nejsou nezávislé!

Součet spojitých n.v.

$\mathbb{E}Z = \mathbb{E}X + \mathbb{E}Y$

Věta

Nechť spojité X, Y jsou n.n.v. Pak $Z = X + Y$ je také spojitá n.v. a její hustotu dostaneme jako konvoluci funkcí f_X, f_Y, neboli

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx.$$

Dk vynech., obdoba diskr. případu

$$p_Z(z) = \sum_{x \in \mathrm{Im}\, X} p_X(x)\, p_Y(z-x)$$

Podmiňování

Definice

X je n.v. na $(\Omega, \mathcal{F}, P)$, $B \in \mathcal{F}$.

$$F_{X|B}(x) := \boxed{P(X \le x \mid B)}$$

$$= \frac{P(X \le x \,\&\, B)}{P(B)}$$

$$= \int_{-\infty}^{x} f_{X|B}(s)$$

K tomu přísluší hustotní funkce $f_{X|B}$.

Věta

$B_1, B_2, \ldots$ rozklad Ω

$$F_X(x) = \sum_i F_{X|B_i}(x) P(B_i)$$ -- věta o rozboru případů

$$f_X(x) = \sum_i f_{X|B_i}(x) P(B_i)$$

der.

pro p. v. n. d.

Důkaz – spec. případ byl na cvičení (dva algoritmy).

Podmíněná hustota $B=\{Y=y\}$

chceme : $f_{X|B}$ .... $P(B)=0$!

Definice

Pro spojité n.v. X, Y definujeme podmíněnou hustotu (conditional pdf) předpisem

$$\int_{x=-\infty}^{\infty} \boxed{f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}}$$

pokud je $f_Y(y) > 0$, jinak ji nedefinujeme.

- připomeňme, že $f_Y(y) = \int_{x\in\mathbb{R}} f_{X,Y}(x,y)dx$
- pro fixované y je $f_{X|Y}(x|y)$ hustota

Podmíněná, sdružená a marginální hustota

Věta

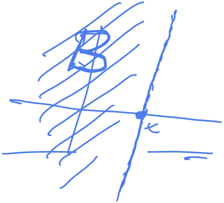
$$f_{X,Y}(x,y) = f_Y(y) f_{X|Y}(x|y)$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y) f_Y(y) dy$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy$$

$$F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{X,Y}(s,t)\, ds\, dt$$

$$= \int_B f_{X,Y}(s,t) = P((X,Y) \in B)$$



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

Spojitá Bayesova věta

$P(A/B) \rightsquigarrow P(B/A)$

$f_{X/Y} \rightsquigarrow f_{Y/X}$

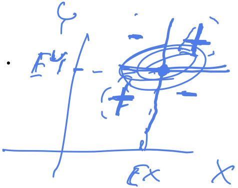
Přehled

Kovariance

Definice

Pro n.v. X, Y definujeme jejich kovarianci (covariance) předpisem

$$cov(X,Y) = \mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y)).$$



Věta

$$cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$EXY = EX \cdot EY$ .... dk jsou pro diskr. n.n.v.

- $var(X) = cov(X,X)$
- $cov(X, aY + bZ + c) = a\,cov(X,Y) + b\,cov(X,Z)$
- $cov(X,Y) = 0$ pokud X, Y jsou nezávislé
- ale nejen tehdy

platí pro libov. n.n.v.

Korelace

Definice

Korelace náhodných veličin X, Y je definována předpisem

$$corr(X,Y) = \varrho(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)\, var(Y)}}.$$

- je to přenormovaná kovariance
- $-1 \leq \varrho(X,Y) \leq 1$ (cvič.)
- korelace neznamená příčinnou souvislost!

$X = Y \dots \rho = 1$

$X = -Y \quad \rho = -1$

Rozptyl součtu

Věta

Nechť $X = \sum_{i=1}^{n} X_i$. *Pak*

$$var(X) = \sum_{i=1}^{n}\sum_{j=1}^{n} cov(X_i, X_j) = \sum_{i=1}^{n} var(X_i) + \sum_{i \neq j} cov(X_i, X_j).$$

Spec. jsou-li $X_1, \ldots, X_n$ *nezávislé, pak*

$$var(X) = \sum_{i=1}^{n} var(X_i).$$

označ.: $Y_i = X_i - EX_i$ $\quad Y = \sum Y_i = X - EX$

$= \sum X_i - \sum EX_i$

$cov(X_i, X_j)$

$var\, X = E\,Y^2 = E(X - EX)^2$ $\quad E\left(\sum_i Y_i \cdot \sum_j Y_j\right) = \sum_{i,j} E\, Y_i Y_j$

Rozptyl součtu – ukázka

Problém šatnářky: n klobouků přiřadíme náhodně n lidem.
X_i je indikátor jevu „i-tý člověk dostal svůj klobouk“
$X = \sum_{i=1}^{n} X_i$

Přehled

Náhodné vektory

Zpátky k základům

Limitní věty – zákony velkých čísel

První zákon velkých čísel

Věta

Nechť $X_1, X_2, \dots$ jsou n.n.v. se střední hodnotou μ a rozptylem σ^2. Označme $S_n = (X_1 + \cdots + X_n)/n$. Pak

$$\mathbb{E}\big((S_n - \mu)^2\big) \to 0.$$

Říkáme, že posloupnost průměrů konverguje k μ v L_2-normě (in mean square).

Slabý zákon velkých čísel (weak law of large numbers)

Věta

Nechť $X_1, \dots, X_n$ jsou n.n.v. se střední hodnotou μ a rozptylem σ^2. Označme $S_n = (X_1 + \cdots + X_n)/n$. Pak pro každé $\varepsilon > 0$ platí

$$\lim_{n\to\infty} P(|S_n - \mu| > \varepsilon) = 0.$$

Říkáme, že posloupnost S_n konverguje k μ v pravděpodobnosti (in probability).

Silný zákon velkých čísel (strong law of large numbers)

Věta

Nechť $X_1, \dots, X_n$ jsou n.n.v. se střední hodnotou μ a rozptylem σ^2. Označme $S_n = (X_1 + \cdots + X_n)/n$. Pak pro každé $\varepsilon > 0$ platí

$$\lim_{n\to\infty} S_n = \mu \quad \textit{skoro jistě (tj. s pravděpodobností 1).}$$

Říkáme, že posloupnost S_n konverguje k μ skoro jistě (almost surely).

CLT