NMAI059 Pravděpodobnost a statistika 1
10. přednáška

Robert Šámal

bodové odhady

Přehled

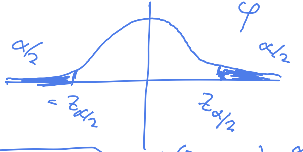
Intervalové odhady

Testování hypotéz

Intervalové odhady normální náhodné veličiny

Věta

$X_1, \dots, X_n$ *je náhodný výběr z* $N(\vartheta, \sigma^2)$.
σ *známe*, ϑ *chceme určit*, $\alpha \in (0,1)$.
Nechť $\Phi(z_{\alpha/2}) = 1 - \alpha/2$. $\hat{\Theta}_n = \bar{X}_n$.



$P(Z > z_{\alpha/2}) = \frac{\alpha}{2}$

$Z \sim N(0,1)$

$$C_n := [\hat{\Theta}_n - z_{\alpha/2}\sigma/\sqrt{n}, \hat{\Theta}_n + z_{\alpha/2}\sigma/\sqrt{n}]$$

$\bar{X}_n$

Pak $P(C_n \ni \vartheta) = 1 - \alpha$.

Důkaz.

$\bar{X}_n \sim N(\vartheta, \frac{\sigma^2}{n})$

(X, Y normální $\Rightarrow$ $X+Y$ normální)

($E\bar{X}_n = \vartheta$, var $\bar{X}_n = \frac{\sigma^2}{n}$ pokud $X_1, \dots$ nezáv.)

$Z = \frac{\bar{X}_n - \vartheta}{\sigma/\sqrt{n}} \sim N(0,1)$

$\boxed{P(|\bar{X}_n - \vartheta| \geq z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) = \alpha}$

$P(Z \geq z_{\alpha/2}) = \frac{\alpha}{2}$ & $P(Z \leq -z_{\alpha/2}) = \frac{\alpha}{2}$, $P(|Z| \geq z_{\alpha/2}) = \alpha$

∎

Intervalové odhady pomocí CLV

Věta

$X_1, \ldots, X_n$ je náhodný výběr z rozdělení se střední hodnotou ϑ, rozptylem σ^2.
σ známe, ϑ chceme určit, $\alpha \in (0,1)$.
Nechť $\Phi(z_{\alpha/2}) = 1 - \alpha/2$.

$\hat{\Theta}_n = \overline{X}_n$

$$C_n := [\hat{\Theta}_n - z_{\alpha/2}\sigma/\sqrt{n}, \hat{\Theta}_n + z_{\alpha/2}\sigma/\sqrt{n}]$$

Pak $P(C_n \ni \vartheta)$ se limitně blíží $1-\alpha$.

Dk $Z_n = \dfrac{\overline{X}_n - \vartheta}{\sigma/\sqrt{n}}$, $Z_n \xrightarrow{d} N(0,1)$ (CLV)

$\lim_{n\to\infty} P(|Z_n| \geq z_{\alpha/2}) = \alpha$

zbytek jako minule.

Studentovo rozdělení

- $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ ... výběrový průměr
- $\widehat{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ ... výběrový rozptyl

$\widehat{S}_n = \sqrt{\widehat{S}_n^2} = \sqrt{\frac{1}{n-1}\sum \ldots}$

n je malé

- Nechť $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

Gosset u Guinessova pivovaru

- Pak $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- *Studentovo t-rozdělení s $n-1$ stupni volnosti* je rozdělení n.v.

$$\boxed{\frac{\bar{X}_n - \mu}{\widehat{S}_n/\sqrt{n}}}$$

nezávisí na μ

na σ

- Distribuční funkce Ψ_{n-1} (v tabulkách ...)

$\lim_{n\to\infty} \Psi_{n-1}(t) = \Phi(t)$

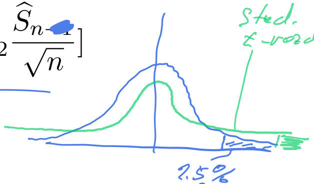
Int. odhady normální n.v. pomocí Studentova t

Věta

$X_1, \dots, X_n$ *je náhodný výběr z* $N(\vartheta, \sigma^2)$.
ϑ *chceme určit,* $\boxed{\sigma \text{ neznáme}}$; $\alpha \in (0,1)$. *Nechť*
$\Psi_{n-1}(z_{\alpha/2}) = 1 - \alpha/2$. $\hat{\Theta}_n = \bar{X}_n$, $\widehat{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$

vyjde větší interval než pro N.R.

$$C_n := [\hat{\Theta}_n - z_{\alpha/2}\frac{\widehat{S}_{n\cancel{-1}}}{\sqrt{n}}, \hat{\Theta}_n + z_{\alpha/2}\frac{\widehat{S}_{n\cancel{-1}}}{\sqrt{n}}]$$



Pak $P(C_n \ni \vartheta) = 1 - \alpha$.

Dk $\quad Z = \dfrac{\bar{X}_n - \vartheta}{\widehat{S}_n/\sqrt{n}}$ má St. t-rozd. s n-1 st. volnosti

$P(|Z| \geq z_{\alpha/2}) = \alpha$

zbytek jako minule ... $P(C_n \not\ni \vartheta) = \alpha$

$\boxtimes$

Přehled

Intervalové odhady

Testování hypotéz

Úvod do testování hypotéz

- Je naše mince spravedlivá?
- Je naše kostka spravedlivá?
- Má vylepšený program kratší dobu běhu než původní?
- Je léčba nemoci metodou X dobrá? (Lepší než placebo, lepší než metoda Y, ...)
- Jsou leváci lepší boxeři?

- dvě hypotézy: H_0, H_1
- H_0 – *nulová hypotéza (null hypothesis)* – značí defaultní, konzervativní model (léčba, mince je spravedlivá)
- H_1 – *alternativní hypotéza (alternative hypothesis)* – značí alternativní model „zajímavost“

Testování hypotéz – ilustrace

- Chceme testovat, zda je mince spravedlivá.
- Hodíme n-krát mincí, orel padne S-krát.
- Pokud je $|S - n/2|$ moc velké, tak mince není spravedlivá.

nezamítneme H_0

$$P(S \le 472) = \sum_{k=0}^{472} \binom{1000}{k} \frac{1}{2^{1000}} \quad \ldots\ldots \text{ distr. fce } Bin(1000, \tfrac{1}{2})$$

$$P(S \ge 528) =$$

$$\le \alpha$$

$$< \alpha/2$$

$$\boxed{500 \pm 31}$$

Testování hypotéz – ilustrace

- Chceme testovat, zda je mince spravedlivá.
- H_0: je spravedlivá
- H_1: není spravedlivá („Vědci objevili, že v kasinu byla použita falešná mince.“)
- Výsledky: zamítneme H_0/nezamítneme H_0 0.05
- Chyba 1. druhu: chybné zamítnutí. Zamítneme H_0, i když platí. Trapas. Chceme se vyvarovat ... P(uděláme trapas) $\leq \alpha$
- Chyba 2. druhu: chybné přijetí. Nezamítneme H_0, ale ona neplatí. Promarněná příležitost.
- Potřebujeme určit k takové, že budeme zamítat H_0 pokud $|S - n/2| > k$.

Pozn: užíváme nezávislost

Testování hypotéz – obecný postup

- Vybereme vhodný statistický model.
- $\boxed{\text{Volíme}}$ *hladinu významnosti (significance level)* α: pravd. chybného zamítnutí H_0. Typicky $\alpha = 0.05$.
- Určíme *testovou statistiku* $S = h(X_1, \dots, X_n)$, kterou budeme určovat z naměřených dat.
- Určíme *kritický obor (rejection region)* – množinu W.
- Naměříme hodnoty $x_1, \dots, x_n$ náh. veličin $X_1, \dots, X_n$.
- Rozhodovací pravidlo: zamítneme H_0 pokud $h(x_1, \dots, x_n) \in W$.

- $\alpha = P(h(X) \in W; H_0)$ → není podm. pravd.
- $\beta = P(h(X) \notin W; H_1)$ … *síla testu*

  chyba 2. druhu

  vyjde ze zvolené metody
- často α nevolíme předem, ale spočítáme p-hodnotu: minimální α, pro které bychom H_0 zamítli.

Testování hypotéz – příklad

- $X_1, \ldots, X_n$ náhodný výběr z $N(\vartheta, \sigma^2)$
- $\boxed{\sigma^2 \text{ známe}}$
- $H_0 : \vartheta = 0 \qquad H_1 : \vartheta \neq 0$

X_i = výška i-tého člena

– 171 cm

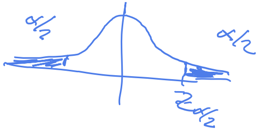
$\alpha = 0.05$

$\Phi^{-1}(1 - \frac{\alpha}{2}) = 1.96$

$$S(X_1, \ldots X_n) = \cancel{\frac{X_1 + \ldots + X_n}{n} = \overline{X}_n}$$

$$= \frac{\overline{X}_n}{\sigma / \sqrt{n}} = \frac{X_1 + \ldots + X_n}{\sigma \sqrt{n}} \sim N(0, 1)$$

$$W = \{ s \in \mathbb{R},\ |s| > z_{\alpha/2} \}$$



zamítnu $H_0 \iff |S| > 1.96$

Testování hypotéz – příklad

$var = \frac{\vartheta_x(1-\vartheta_x)}{}$

- $X_1, \ldots, X_{n_1}$ náhodný výběr z $Ber(\vartheta_X)$
- $Y_1, \ldots, Y_{n_2}$ náhodný výběr z $Ber(\vartheta_Y)$
- $H_0 : \vartheta_X = \vartheta_Y \qquad H_1 : \vartheta_X \neq \vartheta_Y$

$X_i = I$ (i-tý člověk, na kterého jsem op. X, se vyléčil)

$\hat\Theta_x = \frac{X_1 + \ldots + X_{n_1}}{n_1} \qquad \hat\Theta_y = \frac{Y_1 + \ldots + Y_{n_2}}{n_2}$ .... přibližně nez. normální n.v. (CLV)

odhadneme $\vartheta_x - \vartheta_y$ pomocí $Z = \hat\Theta_x - \hat\Theta_y$ norm. se stř. hod. $\vartheta_x - \vartheta_y$

$var(\hat\Theta_x - \hat\Theta_y) = var(\hat\Theta_x) + var(\hat\Theta_y) = \boxed{\frac{\vartheta_x(1-\vartheta_x)}{n_1} + \frac{\vartheta_y(1-\vartheta_y)}{n_2}} = \sigma^2$

pokud H_0: $EZ = \vartheta_x - \vartheta_y = 0 \qquad Z \sim N(0, \sigma^2)$

$\hat\Theta = \frac{\sum X_i + \sum Y_i}{n_1 + n_2}$ aprox. $\sigma^2 = var\, Z$ pomocí $\hat\sigma^2 = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\hat\Theta(1-\hat\Theta)$

$\frac{\hat\Theta_x - \hat\Theta_y}{\hat\sigma} \sim N(0,1)$

zamítneme H_0, pokud $\frac{|\hat\Theta_x - \hat\Theta_y|}{\hat\sigma} > \xi$ $\qquad \Phi(\xi) = 1 - \alpha/2 \qquad (\alpha = 0.05 \quad \xi = 1.96)$

p-hacking

- napřed získáme data, pak v nich hledáme zajímavosti
- když máme dost dat, tak tam nějaké budou „shodou okolností“
- *reprodukovatelnost* – po explorační analýze dat uděláme nezávislý sběr dat a ten analyzujeme konfirmačně
- nebo dopředu náhodně rozdělíme data na část pro tvorbu hypotéz a část pro jejich potvrzení ... jednoduchý případ křížové validace (cross validation)