Intervalové odhady parametrů

Petr Pošík

Části dokumentu jsou převzaty (i doslovně)
z *Mirko Navara: Pravděpodobnost a matematická statistika*,
`https://cw.felk.cvut.cz/lib/exe/fetch.php/courses/a6m33ssl/pms_print.pdf`
s laskavým svolením autora.

Intervalové odhady

Dosud jsme skutečnou hodnotu parametru θ nahrazovali *bodovým odhadem* $\widehat{\Theta}$ (což je náhodná veličina).

Nyní budeme hledat **intervalový odhad**, tzv. **interval spolehlivosti** I, což je minimální interval takový, že

$$P[\theta \in I] \geq 1 - \alpha,$$

kde

- $\alpha \in (0, \frac{1}{2})$ je pravděpodobnost, že meze intervalu I budou překročeny
- $1 - \alpha$ je **koeficient spolehlivosti**.

Obvykle hledáme **horní**, resp. **dolní jednostranný odhad**, kdy

$$I = (-\infty, q_{\widehat{\Theta}}(1-\alpha)\rangle, \text{ resp. } I = \langle q_{\widehat{\Theta}}(\alpha), \infty),$$

nebo **(symetrický) oboustranný odhad**

$$I = \left\langle q_{\widehat{\Theta}}\left(\frac{\alpha}{2}\right), q_{\widehat{\Theta}}\left(1 - \frac{\alpha}{2}\right)\right\rangle.$$

K tomu potřebujeme znát rozdělení odhadu $\widehat{\Theta}$.

Normální rozdělení: intervalový odhad μ při známém σ^2

Mějme realizaci náhodného výběru $\boldsymbol{x}_n = (x_1, \ldots, x_n)$ z normálního rozdělení $\mathrm{N}(\mu, \sigma^2)$. Rozptyl σ^2 známe, chceme odhadnout μ.

Střední hodnotu μ odhadneme výběrovým průměrem $\overline{X}$ s rozdělením $\mathrm{N}(\mu, \frac{\sigma^2}{n})$.
Pro kvantilovou funkci normálního rozdělení platí:

$$q_{\mathrm{N}(\mu,\sigma^2)}(\alpha) = \mu + q_{\mathrm{N}(0,\sigma^2)}(\alpha) = \mu + \sigma\Phi^{-1}(\alpha)$$

Z definice kvantilové funkce platí pro $\overline{X}$

$$P\left[\overline{X} \in \left(-\infty, q_{\mathrm{N}\left(\mu,\frac{\sigma^2}{n}\right)}(1-\alpha)\right\rangle\right] = P\left[\overline{X} \leq q_{\mathrm{N}\left(\mu,\frac{\sigma^2}{n}\right)}(1-\alpha)\right] = 1 - \alpha =$$

$$= P\left[\overline{X} - q_{\mathrm{N}\left(0,\frac{\sigma^2}{n}\right)}(1-\alpha) \leq \mu\right], \text{ což stačí v době počítačů, a dále}$$

$$= P\left[\overline{X} - \frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\alpha) \leq \mu\right], \text{ což je nutné pro hledání v tabulkách.}$$

Pro dolní, horní, resp. oboustranný intervalový odhad pak dostáváme

$$\left\langle \overline{X} - \frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\alpha), \infty\right), \qquad \left(-\infty, \overline{X} + \frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\alpha)\right\rangle,$$

$$\text{resp. } \left\langle \overline{X} - \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right), \overline{X} + \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right\rangle.$$

Při výpočtu pak nahradíme výběrový průměr $\overline{X}$ jeho realizací $\bar{x}$.

Souvislost rozsahu výběru, max. chyby odhadu a spolehlivosti

Ve vztahu pro oboustranný interval spolehlivosti

$$P\left[\bar{X} - \Delta \le \mu \le \bar{X} + \Delta\right] = 1 - \alpha,$$

kde $\Delta = \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$, vystupují následující proměnné:

- koeficient spolehlivosti $P = 1 - \alpha$, tj. míra naší důvěry ve výrok o poloze μ,
- chyba Δ odhadu střední hodnoty μ (pro oboustranný interval spolehlivosti $|I| = 2\Delta$), tj. nejistota při určení hodnoty parametru μ, a
- rozsah výběru n.

Tyto proměnné spolu souvisí následujícím způsobem:

P konstatntní	$n \nearrow \iff \Delta \searrow$
Δ konstatntní	$n \nearrow \iff P \nearrow$
n konstatntní	$P \nearrow \iff \Delta \nearrow$

Výpočet potřebného rozsahu výběru

Pro *oboustranný interval spolehlivosti* $I = \langle \bar{X} - \Delta, \bar{X} + \Delta \rangle$ platí, že $P\left[\bar{X} - \Delta \le \mu \le \bar{X} + \Delta\right] \ge 1 - \alpha$, kde $\Delta = \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$.

Stanovíme-li maximální přípustnou chybu odhadu $\Delta_{\max}$, můžeme určit **potřebný rozsah výběru** $n_{\min}$ při požadované spolehlivosti $1 - \alpha$:

$$\Delta_{\max} \ge \Delta = \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right),$$

$$n \ge \left(\frac{\sigma}{\Delta}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)^2 = n_{\min}.$$

Pro *jednostranné intervaly spolehlivosti* lze postupovat obdobně, např. pro horní intervalový odhad $I = (-\infty, \bar{X}\rangle$ platí, že $P\left[\mu \le \bar{X} + \Delta\right] \ge 1 - \alpha$, kde $\Delta = \frac{\sigma}{\sqrt{n}}\Phi^{-1}(1 - \alpha)$.

Stanovíme-li maximální přípustnou chybu odhadu $\Delta_{\max}$, můžeme určit **potřebný rozsah výběru** $n_{\min}$ při požadované spolehlivosti $1 - \alpha$:

$$\Delta_{\max} \ge \Delta = \frac{\sigma}{\sqrt{n}}\Phi^{-1}(1 - \alpha),$$

$$n \ge \left(\frac{\sigma}{\Delta}\Phi^{-1}(1 - \alpha)\right)^2 = n_{\min}.$$

Normální rozdělení: intervalový odhad μ při neznámém σ^2

Střední hodnotu μ odhadneme výběrovým průměrem $\bar{X}$ s rozdělením $\mathrm{N}(\mu, \frac{\sigma^2}{n})$.
Rozptyl σ^2 odhadneme výběrovým rozptylem S_X^2; $\frac{(n-1)S_X^2}{\sigma^2}$ má rozdělení $\chi^2(n-1)$.

Použití *odhadu* rozptylu místo skutečné hodnoty rozptylu je *další zdroj neurčitosti*; abychom zachovali koeficient spolehlivosti $1-\alpha$, musíme rozšířit interval spolehlivosti, tj. použít kvantily jiného rozdělení:

Studentovo t-rozdělení s $n-1$ stupni volnosti, $t(n-1)$.

Z toho vyplývají následující intervalové odhady:

$$\left\langle \bar{X} - \frac{S_X}{\sqrt{n}} q_{t(n-1)}(1-\alpha), \infty \right), \qquad \left(-\infty, \bar{X} + \frac{S_X}{\sqrt{n}} q_{t(n-1)}(1-\alpha) \right\rangle,$$
$$\left\langle \bar{X} - \frac{S_X}{\sqrt{n}} q_{t(n-1)}\left(1-\frac{\alpha}{2}\right), \bar{X} + \frac{S_X}{\sqrt{n}} q_{t(n-1)}\left(1-\frac{\alpha}{2}\right) \right\rangle.$$

Při výpočtu nahradíme výběrový průměr $\bar{X}$ jeho realizací $\bar{x}$ a výběrovou směrodatnou odchylku S_X její realizací s_X.

Studentovo rozdělení

Studentovo t-rozdělení s η stupni volnosti, $t(\eta)$ je rozdělení náhodné veličiny

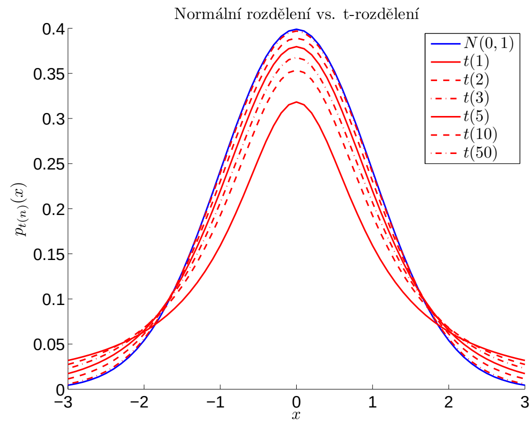
$$\frac{U}{\sqrt{\frac{V}{\eta}}},$$

kde

- U má rozdělení $\mathrm{N}(0,1)$,
- V má rozdělení $\chi^2(\eta)$ a
- U a V jsou nezávislé.

Pro velký počet stupňů volnosti se nahrazuje normálním rozdělením.

Hustota:



V případě odhadu střední hodnoty při neznámém rozptylu

$$U = \frac{\sqrt{n}}{\sigma}(\bar{X}-\mu) \text{ má } \mathrm{N}(0,1),$$
$$V = \frac{(n-1)S_X^2}{\sigma^2} \text{ má } \chi^2(n-1), \qquad \eta = n-1,$$
$$\frac{U}{\sqrt{\frac{V}{\eta}}} = \frac{\frac{\sqrt{n}}{\sigma}(\bar{X}-\mu)}{\sqrt{\frac{S_X^2}{\sigma^2}}} = \frac{\sqrt{n}}{S_X}(\bar{X}-\mu) \text{ má } t(n-1), \text{ z čehož vyplývají odhady na předchozím slidu.}$$

Příklad: intervalové odhady μ při neznámém σ^2

Zadání: Vliv alkoholismu matky na inteligenci dítěte. Nalezeno 6 žen, které byly v době těhotenství chronickými alkoholičkami. IQ jejich dětí bylo změřeno v 7 letech: $n = 6$, $\bar{x} = 78$, $\sum_{i=1}^{n}(x_i - \bar{x})^2 = 1805$. Vypočtěte horní a oboustranný 95% interval spolehlivosti pro střední hodnotu rozdělení IQ dětí alkoholiček.

Řešení: Předpokládáme, že výběr alkoholiček byl náhodný. Protože výběrový rozptyl odhadujeme a rozsah výběru je malý, použijeme kvantily Studentova rozdělení.

Pro výběrovou směrodatnou odchylku a oboustranný interval spolehlivosti platí

$$s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} = \sqrt{\frac{1805}{5}} = 19$$

$$\widehat{\mu} = \bar{x} \pm \frac{s_x}{\sqrt{n}} q_{t(n-1)}\left(1 - \frac{\alpha}{2}\right) = 78 \pm \frac{19}{\sqrt{6}} 2.57 = 78 \pm 19.94,$$

a pro horní odhad

$$\widehat{\mu} \leq \bar{x} + \frac{s_x}{\sqrt{n}} q_{t(n-1)}(1-\alpha) = 78 + \frac{19}{\sqrt{6}} 2.02 = 78 + 15.63 = 93.63.$$

Závěry:

1. S pravděpodobností 95 % střední hodnota μ rozdělení IQ dětí alkoholiček leží v $\langle 58.06, 97.94\rangle$.
2. S pravděpodobností 95 % střední hodnota μ rozdělení IQ dětí alkoholiček nepřekročí 93.96.

Intervalový odhad rozptylu

Rozptyl σ^2 odhadneme výběrovým rozptylem S_X^2; n.v. $\frac{(n-1)S_X^2}{\sigma^2}$ má rozdělení $\chi^2(n-1)$.

$$P\left[\frac{(n-1)S_X^2}{\sigma^2} \in (-\infty, q_{\chi^2(n-1)}(1-\alpha)\rangle\right] = 1 - \alpha =$$

$$= P\left[\frac{(n-1)S_X^2}{\sigma^2} \leq q_{\chi^2(n-1)}(1-\alpha)\right] = P\left[\frac{(n-1)S_X^2}{q_{\chi^2(n-1)}(1-\alpha)} \leq \sigma^2\right] =$$

$$= P\left[\sigma^2 \in \left\langle \frac{(n-1)S_X^2}{q_{\chi^2(n-1)}(1-\alpha)}, \infty\right)\right]$$

Dostali jsme *dolní* odhad, ostatní obdobně.

Dolní, horní a oboustranný intervalové odhady rozptylu jsou

$$\left\langle \frac{(n-1)S_X^2}{q_{\chi^2(n-1)}(1-\alpha)}, \infty\right), \qquad \left(-\infty, \frac{(n-1)S_X^2}{q_{\chi^2(n-1)}(\alpha)}\right\rangle,$$

$$\left\langle \frac{(n-1)S_X^2}{q_{\chi^2(n-1)}\left(1-\frac{\alpha}{2}\right)}, \frac{(n-1)S_X^2}{q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right)}\right\rangle.$$

Při výpočtu nahradíme výběrový rozptyl S_X^2 jeho realizací s_x^2.

Alternativní rozdělení: odhad populační pravděpodobnosti q

Mějme náhodný výběr rozsahu n, $\boldsymbol{X}_n = \{X_1, \ldots, X_n\}$, kde $X_i \in \{0, 1\}$, $X_i \sim Ber(q)$.

Populační pravděpodobnost q odhadneme výběrovou relativní četností $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Pro dostatečně velký rozsah výběru ($n > 100$) lze *rozdělení výběrových relativních četností $\overline{X}$* aproximovat normálním rozdělením,

$$\overline{X} \sim \mathrm{N}\left(q, \frac{q(1-q)}{n}\right) \approx \mathrm{N}\left(q, \frac{\overline{X}(1-\overline{X})}{n}\right),$$

a použít intervaly spolehlivosti pro střední hodnotu normálního rozdělení:

$$\left\langle \overline{X} - \sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\Phi^{-1}(1-\alpha), \infty \right), \qquad \left(-\infty, \overline{X} + \sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\Phi^{-1}(1-\alpha) \right\rangle,$$

$$\text{resp. } \left\langle \overline{X} - \sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\Phi^{-1}\left(1-\frac{\alpha}{2}\right), \overline{X} + \sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\Phi^{-1}\left(1-\frac{\alpha}{2}\right) \right\rangle.$$

Při výpočtu pak nahradíme výběrovou relativní četnost $\overline{X}$ její realizací $\bar{x}$.

Při malém rozsahu výběru ($n < 100$) je třeba hledat kvantily rozdělení výběrových relativních četností ve speciálních tabulkách.

Příklad: Odhad populační pravděpodobnosti q

Zadání: V roce 1973 bylo na Berkeley přijato 3700 z 8300 uchazečů můžského pohlaví a 1500 z 4300 uchazečů ženského pohlaví. Vypočtete 95% oboustranné intervaly spolehlivosti ($\alpha = 0.05$) pro pravděpodobnosti přijetí u mužů, q_M a u žen q_Z.

Řešení: Intervaly spolehlivosti pro populační pravděpodobnosti přijetí u mužů a u žen:

$$\bar{x}_M = \frac{3700}{8300} = 0.4458$$

$$\bar{x}_Z = \frac{1500}{4300} = 0.3488$$

$$\widehat{q}_M = \bar{x}_M \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\bar{x}_M(1-\bar{x}_M)}{n_M}} =$$

$$= 0.4458 \pm 1.96\sqrt{\frac{0.4458(1-0.4458)}{8300}} = 44.58 \pm 1.07\%$$

$$\widehat{q}_Z = \bar{x}_Z \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\bar{x}_Z(1-\bar{x}_Z)}{n_Z}} =$$

$$= 0.3488 \pm 1.96\sqrt{\frac{0.3488(1-0.3488)}{4300}} = 34.88 \pm 1.42\%$$

Co můžete říct o rozdílu mezi pravděpodobnostmi přijetí? Je to důkaz pohlavní diskriminace? Odpovědi v příštích přednáškách.

Shrnutí: Odhady parametrů

- Výběrové statistiky (např. $\bar{X}$, S_X^2, apod.), resp. jejich realizace ($\bar{x}$, s_x^2, apod.), se používají k odhadu populačních parametrů (μ, σ^2, apod.). Jsou to bodové odhady.
- Intervalové odhady jsou lepší než bodové odhady v tom směru, že umožňují stanovit rozmezí, v jakém může hodnota odhadovaného populačního parametru ležet, a „míru důvěry“ v tento interval.