Marie Turčičová

Intervaly spolehlivosti

Odhady pro daný parametr lze rozdělit do dvou skupin - bodové a intervalové. Příklady bodových odhadů:

parametr	**bodový odhad**
střední hodnota ($\mathrm{E}\,X, \mu$)	výběrový průměr ($\bar{X}$)
rozptyl ($\mathrm{var}\,X, \sigma^2$)	výběrový rozptyl (S_X^2)
směrodatná odchylka ($\mathrm{sd}\,X, \sigma$)	výběrová směrodatná odchylka (S_X)
$\vdots$	$\vdots$

kde $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ a $S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})$, jako obvykle.
Bodový odhad ale může být poměrně nepřesný, a proto se často uvádí i odhad intervalový, kdy parametr neodhadujeme jedním číslem, ale nějakých intervalem, v němž by daný parametr měl s velkou pravděpodobností ležet.

Obecný tvar a interpretace

Interval spolehlivosti (někdy též nazývaný konfidenční interval, angl. confidence interval) je vždy tvaru:

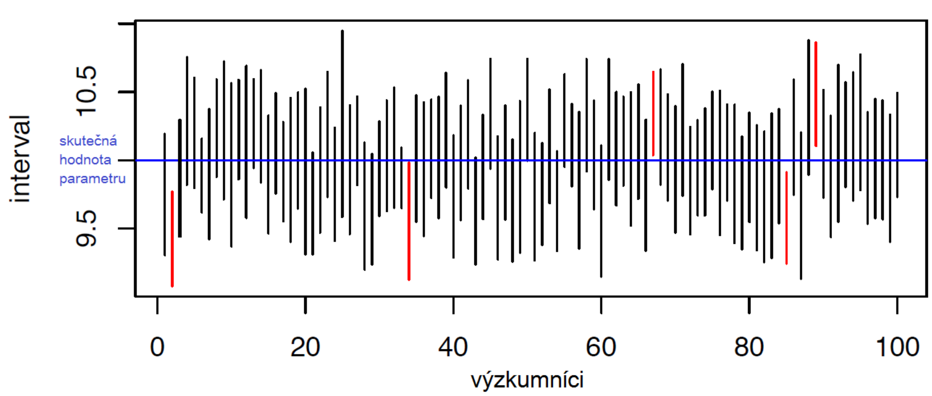
(bodový odhad - chyba odhadu · vhodný kvantil , bodový odhad + chyba odhadu · vhodný kvantil)

Spolehlivost intervalu se značí $(1 - \alpha)$ a obvykle nabývá hodnot 0.95, 0.99, 0.90 apod. Hovoříme pak o 95% intervalu spolehlivosti, nebo také o intervalu se spolehlivostí 95 %.
Spolehlivost vyjadřuje, s jakou pravděpodobností daný interval pokrývá skutečnou (neznámou) hodnotu parametru. Např. 95% interval spolehlivosti obsahuje skutečnou hodnotu parametru s pravděpodobností 0.95. Označíme-li si dolní mez intervalu jako d a horní mez jako h, pak tedy platí

$$P\left[parametr \in (d, h)\right] = 0.95$$

(Pozor! V žádném případě se nejedná o interval, ve kterém by měly ležet hodnoty sledovaného znaku u 95 % populace, či 95 % hodnot našich dat!!)
Uvědomte si, že meze intervalu jsou náhodné (závisejí na konkrétním náhodném výběru!), zatímco parametr je pevné (akorát neznámé) číslo.

Spolehlivost si lze také představit následovně. Mějme 100 výzkumníku, kteří se snaží intervalově odhadnout nějaký parametr. Každý z výzkumníků si vytvoří svůj náhodný výběr (všichni však o stejném rozsahu), tj. náhodně vybere n jedinců z populace, a na jeho základě spočte 95% interval spolehlivosti. Každý z výzkumníků dostane interval trochu odlišný, neboť jak už bylo řečeno, konkrétní meze intervalu závisejí na náhodném výběru. No a 95 % výzkumníků dostane interval, který obsahuje skutečnou hodnotu parametru. 5 % výzkumníků dostane interval, který se do této hodnoty netrefí (viz obrázek).



Konstrukce intervalu pro střední hodnotu

Konkrétní tvar intervalu spolehlivosti závisí samozřejmě na tom, jaký parametr odhadujeme a také na předpokladu o pravděpodobnostním rozdělení daného znaku v populaci. My si ukážeme odvození intervalového odhadu pro střední hodnotu normálně rozděleného výběru.

Předpokládejme tedy, že máme náhodný výběr $X_1, X_2, \ldots, X_n$ (na náhodně vybraných n jedincích z populace jsme změřili hodnotu nějakého znaku). Předpokládejme dále, že daný znak má v populaci normální rozdělení se střední hodnotou μ a rozptylem σ^2, tedy $X_i \sim N(\mu, \sigma^2)$ pro každé $i = 1, \ldots, n$. Bodovým odhadem μ by samozřejmě bylo $\bar{X}$. Pokusme se nyní zkonstruovat odhad intervalový o spolehlivosti 95 %.

1) Předpokládejme, že σ^2 je známé.

Platí, že mají-li $X_i \sim N(\mu, \sigma^2)$, tak potom

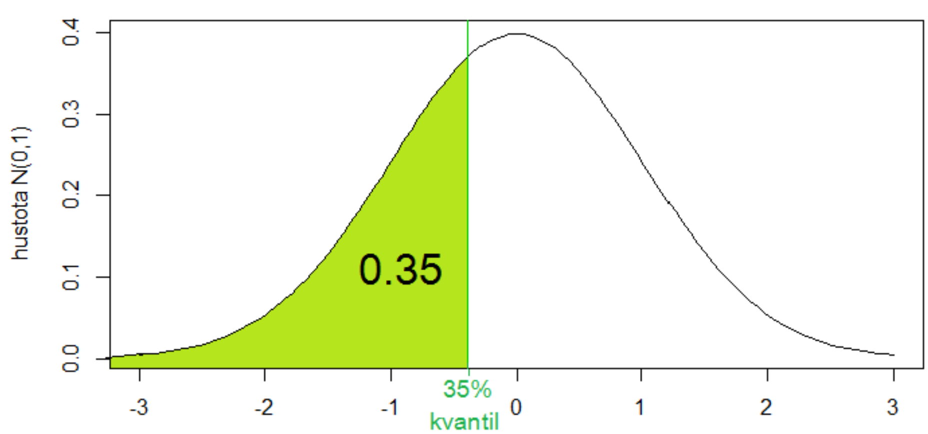
$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}),$$

což lze ekvivalentně zapsat jako

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1). \tag{1}$$

Je to vlastně z-skór pro průměr :-) Odtud již budeme schopní interval spolehlivosti zkonstruovat.

Ještě předtím si ale připomeňme význam kvantilu: Při práci s daty byl k% kvantil takové číslo, že k % dat bylo menších nebo rovno tomuto číslu. Tedy například 35% hodnot dané proměnné bylo menších nebo rovno příslušnému 35% kvantilu. V řeči pravděpodobnosti, veličina, která se řídí daným rozdělením, bude s pravděpodobností 0.35 menší nebo rovna 35% kvantilu tohoto rozdělení. Pokud by daná veličina měla například normované normální rozdělení N(0,1), lze si to graficky představit takto:



Nyní se vraťme k naší úloze. My hledáme interval (d, h) takový, že

$$P[\mu \in (d, h)] = 0.95.$$

Z (1) a významu pojmu kvantil ale víme, že

$$P\left[\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq \text{qnorm}(0.95)\right] = 0.95 \tag{2}$$

95% kvantil rozdělení $N(0, 1)$ jsem kvůli názornosti označila "qnorm(0.95)", častější je ale značení $u(0.95)$ nebo $z(0.95)$, popř. $u_{0.95}$ nebo $z_{0.95}$.

Výraz uvnitř (2) tedy stačí jen upravit do tvaru "$\mu \in (d,h)$"a je hotovo. Tak tedy:

$$P\left[\bar{X} - \mu \le \frac{\sigma}{\sqrt{n}}\,\mathrm{qnorm}(0.95)\right] = 0.95$$

$$P\left[-\mu \le -\bar{X} + \frac{\sigma}{\sqrt{n}}\,\mathrm{qnorm}(0.95)\right] = 0.95$$

$$P\left[\mu \ge \bar{X} - \frac{\sigma}{\sqrt{n}}\,\mathrm{qnorm}(0.95)\right] = 0.95.$$

(Nezapomeňte, že při násobení nerovnosti záporným číslem se nerovnost mění v opačnou.) Ekvivalentně lze poslední vzorec zapsat jako:

$$P\left[\mu \in \left(\bar{X} - \frac{\sigma}{\sqrt{n}}\,\mathrm{qnorm}(0.95)\ ,\ \infty\right)\right] = 0.95.$$

Interval

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} \cdot \mathrm{qnorm}(0.95)\ ,\ \infty\right) \tag{3}$$

je tedy hledaný 95% interval spolehlivosti pro naše μ a říká se mu *dolní*, nebo též *levostranný*.

U výrazu (2) můžete namítnout, že stejně tak platí

$$P\left[\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \ge \mathrm{qnorm}(0.05)\right] = 0.95.$$

Po analogických úpravách bychom pak dostali interval

$$\left(-\infty\ ,\ \bar{X} - \frac{\sigma}{\sqrt{n}} \cdot \mathrm{qnorm}(0.05)\right).$$

Využijeme-li toho, že $\mathrm{qnorm}(0.05) = -\,\mathrm{qnorm}(0.95)$ (neboť normované normální rozdělení je symetrické kolem 0), pak lze tento interval přepsat do běžnějšího tvaru:

$$\left(-\infty\ ,\ \bar{X} + \frac{\sigma}{\sqrt{n}} \cdot \mathrm{qnorm}(0.95)\right). \tag{4}$$

Tomuto intervalu se říká *pravostranný* nebo též *horní* 95% interval spolehlivosti.

Kromě jednostranných intervalů spolehlivosti lze zkonstruovat též interval oboustranný, který se asi používá nejčastěji. Vyjdeme ze vztahu

$$P\left[\mathrm{qnorm}(0.025) \le \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \le \mathrm{qnorm}(0.975)\right] = 0.95.$$

který je založen na jednoduché úvaze, že daná veličina se nachází v prostoru mezi 2.5% a 97.5% kvantilem opět s pravděpodobností 0.95. Úplně stejnými úpravami jako výše (akorát bychom nyní pracovali se dvěma nerovnostmi současně) dostaneme oboustranný 95% interval spolehlivosti pro μ

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} \cdot \mathrm{qnorm}(0.975)\ ,\ \bar{X} + \frac{\sigma}{\sqrt{n}} \cdot \mathrm{qnorm}(0.975)\right). \tag{5}$$

(Opět jsme navíc využili symetrie rozdělení N(0,1), totiž že $\mathrm{qnorm}(0.025) = -\,\mathrm{qnorm}(0.975)$).

2) Je-li σ^2 neznámé.

Je-li σ^2 neznámé, musíme ho odhadnout. Odhadneme ho tradičně pomocí výběrového rozptylu S_X^2. Výsledné intervaly mají pak stejný tvar jako (3), (4), (5), akorát místo σ^2 je tam S_X^2. To, že jsme rozptyl museli odhadnout ale ještě musíme penalizovat tím, že místo kvantilu normovaného normální rozdělení použijeme kvantily t-rozdělení (tzv. Studentova rozdělení), které jsou obecně o trochu větší. Dostáváme tak

- horní (pravostranný) interval:

$$\left(-\infty\ ,\ \bar{X} + \frac{S_X}{\sqrt{n}} \cdot \mathrm{qt}_{n-1}(0.95)\right),$$

- dolní (levostranný) interval:

$$\left(\bar{X} - \frac{S_X}{\sqrt{n}} \cdot \mathrm{qt}_{n-1}(0.95)\ ,\ \infty\right),$$

- oboustranný interval:

$$\left(\bar{X} - \frac{S_X}{\sqrt{n}} \cdot \mathrm{qt}_{n-1}(0.975)\ ,\ \bar{X} + \frac{S_X}{\sqrt{n}} \cdot \mathrm{qt}_{n-1}(0.975)\right),$$

kde $qt_{n-1}(0.95)$ značí 95% kvantil t-rozdělení. Častěji se pro něj používá značení $t_{n-1}(0.95)$. Analogicky $qt_{n-1}(0.975)$.

(Pro fajnšmekry: Intervaly výše by se odvodili stejnými úvahami jako v sekci 1), pouze místo (1) by se pracovalo s

$$\frac{\bar{X} - \mu}{\frac{S_X}{\sqrt{n}}} \sim t_{n-1},$$

tedy že po nahrazení σ jeho výběrových protějškem S už má výraz vlevo t-rozdělení s $n-1$ stupni volnosti.)

Shrnutí

$(1-\alpha)\cdot 100\%$ interval spolehlivosti pro střední hodnotu normálně rozděleného výběru:

	je-li σ^2 známé	je-li σ^2 neznámé
pravostranný	$\left(-\infty\ ,\ \bar{X} + \frac{\sigma}{\sqrt{n}} \cdot \mathrm{qnorm}(1-\alpha)\right)$	$\left(-\infty\ ,\ \bar{X} + \frac{S_X}{\sqrt{n}} \cdot \mathrm{qt}_{n-1}(1-\alpha)\right)$
levostranný	$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} \cdot \mathrm{qnorm}(1-\alpha)\ ,\ \infty\right)$	$\left(\bar{X} - \frac{S_X}{\sqrt{n}} \cdot \mathrm{qt}_{n-1}(1-\alpha)\ ,\ \infty\right)$
oboustranný	$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} \cdot \mathrm{qnorm}(1-\frac{\alpha}{2})\ ,\ \bar{X} + \frac{\sigma}{\sqrt{n}} \cdot \mathrm{qnorm}(1-\frac{\alpha}{2})\right)$	$\left(\bar{X} - \frac{S_X}{\sqrt{n}} \cdot \mathrm{qt}_{n-1}(1-\frac{\alpha}{2})\ ,\ \bar{X} + \frac{S_X}{\sqrt{n}} \cdot \mathrm{qt}_{n-1}(1-\frac{\alpha}{2})\right)$

Poznámka

Intervaly spolehlivosti pro jiné parametry (např. pro rozptyl) se konstruují zcela analogicky, jen je potřeba si rozmyslet rozdělení, jehož kvantily se použijí.