Testy

Pavel Provinský

19. listopadu 2013

Test a intervalový odhad

Testy a intervalové odhady - jsou vlastně to samé. Jiný je jen úhel pohledu. Lze přecházet od jednoho k druhému. Například:

Při odvozování intervalového odhadu pro střední hodnotu (známý rozptyl) dospějeme k tomu, že jakási statistika z má nomované normální rozdělení:

$$z = (\overline{x} - \mu) \cdot \frac{\sqrt{n}}{\sigma} \sim N(0,1).$$

Pokud chceme např. oboustranný 95% - intervalový odhad, získáváme odtud ihned nerovnici:

$$z_{0,025} < (\overline{x} - \mu) \cdot \frac{\sqrt{n}}{\sigma} < z_{0,975},$$

kde $z_{0,025}$ a $z_{0,975}$ jsou příslušné kvantily.

Pokud zůstaneme na této úrovni a spočteme výraz uprostřed a testujeme, jestli opravdu padne mezi příslušné kvantily nebo ne, provedeme test.

Pokud z nerovnice vyjádříme nějakou veličinu, např. μ, provedeme intervalový odhad:

$$\overline{x} + \frac{\sigma}{\sqrt{n}} \cdot z_{0,025} < \mu < \overline{x} + \frac{\sigma}{\sqrt{n}} \cdot z_{0,975}.$$

respektive:

$$\mu \in \overline{x} \pm \frac{\sigma}{\sqrt{n}} \cdot z_{0,975}.$$

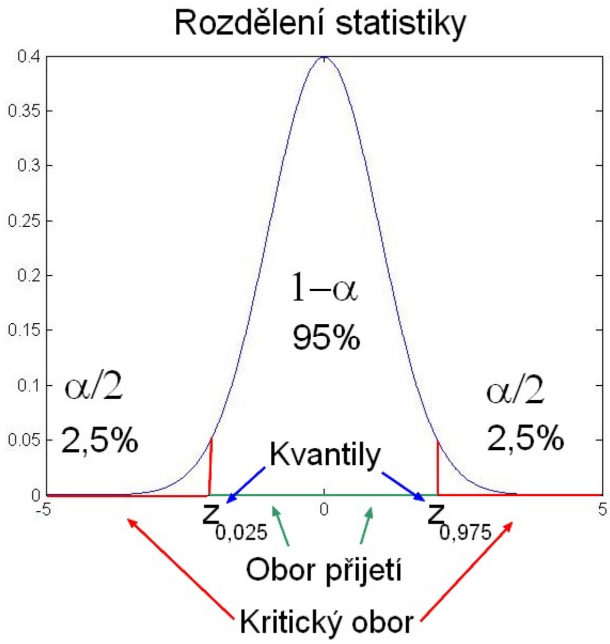
Pojmy

- Nulová hypotéza
- Alternativní hypotéza
- Testová statistika a její rozdělení
- Hladina významnosti
- Obor přijetí
- Kritický obor
- p-hodnota

- pravostranný, levostranný a oboustranný test =pravostranný, levostranný a oboustranný obor přijetí

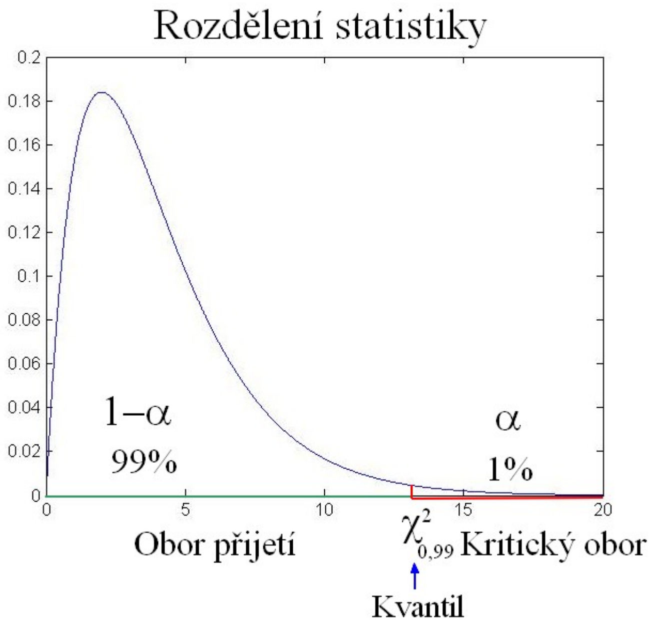
Příklad:

Rozdělení testové statistiky pro oboustranný test s hladinou významnosti 5%. Rozdělení statistiky je v tomto případě normální $N(0,1)$:



Příklad:

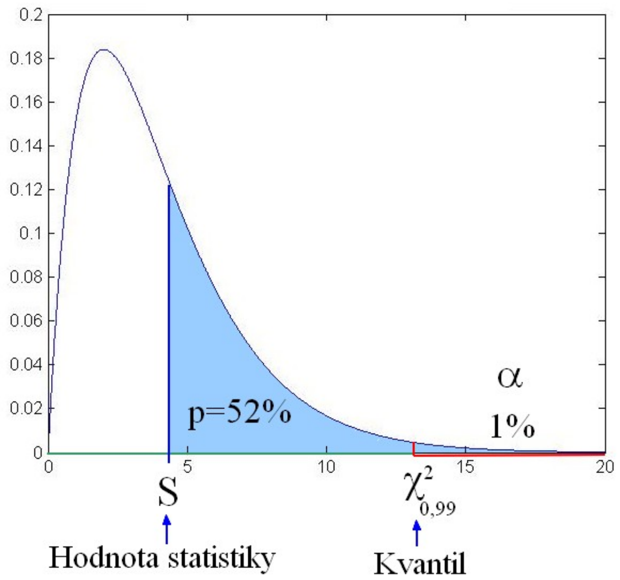
Rozdělení testové statistiky pro pravostranný test s hladinou významnosti 1%. Rozdělení statistiky je v tomto případě Chí-kvadrát se 4 stupni volnosti $Chi^2(4)$:



Příklad:

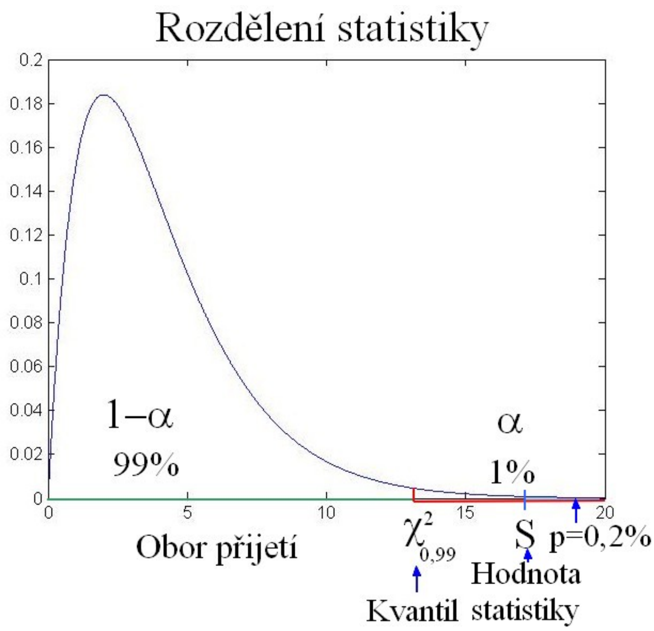
Pokud nulovou hypotézu zamítneme, pak p-hodnota znamená pravděpodobnost, že jsme se v tomto zmýlili a nulová hypotéza přesto platí.

Zakreslení p-hodnoty pro statistiku, která padla do oboru přijetí. p-hodnota je větší než hladina významnosti ($p > \alpha$), nulovou hypotézu tedy nezamítáme:



Příklad:

Zakreslení p-hodnoty pro statistiku, která padla do kritického oboru. p-hodnota je menší než hladina významnosti ($p < \alpha$), nulovou hypotézu tedy zamítáme:



Některé testy

Test střední hodnoty μ, znám rozptyl σ^2

Statistika:

$$z = (\overline{x} - \mu) \cdot \frac{\sqrt{n}}{\sigma}.$$

Rozdělení statistiky:

$$z \sim N(0, 1).$$

Směrování testu dle alternativní hypotézy, možné jsou všechny tři varianty.

Příklad:

H_0: Střední hodnota je $\mu = 5$.
$H_A : \mu \neq 5$............oboustranný test.
Známe rozptyl veličiny X_i: $\sigma^2 = 100$.
Volíme $\alpha = 5\%$.
Data: $\overline{x} = 7$, $n = 36$.

$z = (7 - 5) \cdot \frac{\sqrt{36}}{10} = 2 \cdot 0,6 = 1,2$
$z_{1-\frac{\alpha}{2}} = z_{0,975} = 1,96$
$z_{\frac{\alpha}{2}} = z_{0,025} = -1,96$
Statistika padne do oboru přijetí, hypotézu H_0 tedy na hladině významnosti 5% nezamítáme.

Test střední hodnoty μ, neznám rozptyl σ^2

Statistika:

$$t = (\overline{x} - \mu) \cdot \frac{\sqrt{n}}{s}.$$

Rozdělení statistiky: Studentovo rozdělení s $n-1$stupni volnosti.

$$t \sim St\,(n-1)\,.$$

Směrování testu dle alternativní hypotézy, možné jsou všechny tři varianty.

Příklad:

H_0: Střední hodnota je $\mu \leq 5$.
$H_A : \mu > 5$............pravostranný test.
Volíme $\alpha = 5\%$.
Data: $\overline{x} = 6,07$, $s = 2,56$, $n = 29$.

$t = (6,07 - 5) \cdot \frac{\sqrt{29}}{2,56} = 2,25$
$t_{1-\alpha}\,(n-1) = t_{0,95}\,(28) = 1,70$
Statistika padne do kritického oboru, hypotézu H_0 tedy na hladině významnosti 5% zamítáme.

Varianta s výpočtem p-hodnoty (např. v Matlabu):
$p = 0,0164$
$p < \alpha$, hypotézu H_0 tedy na hladině významnosti 5% zamítáme.

Test podílu

Statistika:

$$z = (p - \pi) \cdot \sqrt{\frac{n}{p \cdot (1-p)}}.$$

(Poznámka: Lze najít ještě lepší statistiku, která je ale složitější.)

Rozdělení statistiky:

$$z \sim N(0,1).$$

Směrování testu dle alternativní hypotézy, možné jsou všechny tři varianty.

Příklad:

Je herní kostka poctivá, když jsme při 20 hodech hodili šestku 10-krát, nebo padá šestka častěji?
H_0: Podíl šestky v základním souboru je $\pi \leq \frac{1}{6}$.
$H_A : \pi > \frac{1}{6}$............pravostranný test.
Volíme $\alpha = 1\%$.
Data: $p = \frac{1}{2}$, $n = 20$.

$z = \left(\frac{1}{2} - \frac{1}{6}\right) \cdot \sqrt{\frac{20}{\frac{1}{2} \cdot \left(1 - \frac{1}{2}\right)}} = 2,98$
$z_{1-\alpha} = z_{0,99} = 2,33$

Statistika padne do kritického oboru, hypotézu H_0 tedy na hladině významnosti 1% zamítáme. (Pravděpodobnost, že se mýlíme, je tedy menší než 1%.)

Test rozptylu

Statistika:

$$\chi^2 = \frac{(n-1)\cdot s^2}{\sigma^2}$$

Rozdělení statistiky:

$$\chi^2 \sim Chí^2\,(n-1)\,.$$

Směrování testu dle alternativní hypotézy, možné jsou všechny tři varianty. Nejčastější je pravostranný test, tedy zda není rozptyl příliš velký.

Příklad:

Může být rozptyl základního souboru nejvýše 50, když jsme zjistili na vzorku 20 hodnot výběrový rozptyl $s^2 = 150$?
H_0: Rozptyl základního souboru je $\sigma^2 \leq 50$.
$H_A : \sigma^2 > 50$............pravostranný test.
Volíme $\alpha = 1\%$.
Data: $s^2 = 150$, $n = 20$.

$\chi^2 = \frac{(20-1)\cdot 150}{50} = 57$
$\chi^2_{1-\alpha} = \chi^2_{0,99} = 36,2$

Statistika padne do kritického oboru, hypotézu H_0 tedy na hladině významnosti 1% zamítáme, a to ve prospěch alternativní hypotézy, tedy že $\sigma^2 > 50$.

χ^2 test nezávislosti

Statistika:

$$S = \sum \frac{(O_i - E_i)^2}{E_i}$$

Rozdělení statistiky:

$$S \sim Chí^2\left((m-1)\cdot(n-1)\right).$$

Test je pravostranný.

Nulová hypotéza: Veličiny jsou nezávislé.

Příklad:

Závisí barva trička na pohlaví?

H_0 : Barva trička a pohlaví jsou nezávislé veličiny.

Volíme hladinu významnosti $\alpha = 5\%$.

Data:

	bílé	červené	modré	jiné	
muž	15	6	12	22	...O_i
žena	8	11	6	20	

Spočteme E_i a pak statistiku. Nejprve určíme marginální četnosti:

	bílé	červené	modré	jiné	
muž	15	6	12	22	55
žena	8	11	6	20	45
	23	17	18	42	100

Určíme marginální pravděpodobnosti a pravděpodobnosti za předpokladu nezávislosti:

	bílé	červené	modré	jiné	
muž	$\frac{55}{100}\cdot\frac{23}{100}$	$\frac{55}{100}\cdot\frac{17}{100}$	$\frac{55}{100}\cdot\frac{18}{100}$	$\frac{55}{100}\cdot\frac{42}{100}$	$\frac{55}{100}$
žena	$\frac{45}{100}\cdot\frac{23}{100}$	$\frac{45}{100}\cdot\frac{17}{100}$	$\frac{45}{100}\cdot\frac{18}{100}$	$\frac{45}{100}\cdot\frac{42}{100}$	$\frac{45}{100}$
	$\frac{23}{100}$	$\frac{17}{100}$	$\frac{18}{100}$	$\frac{42}{100}$	1

Určíme četnosti za předpokladu nezávislosti:

	bílé	červené	modré	jiné	
muž	$\frac{55}{100}\cdot\frac{23}{100}\cdot 100 = 12,65$	9,35	9,9	23,1	....E_i
žena	$\frac{45}{100}\cdot\frac{23}{100}\cdot 100 = 10,35$	7,65	8,1	18,9	

$S = \sum \frac{(O_i-E_i)^2}{E_i} = 4,74$

$\chi^2_{0,95}\left((n-1)(m-1)\right) = \chi^2_{0,95}\left((4-1)(2-1)\right) = \chi^2_{0,95}(3) = 7,81$

Statistika padla do oboru přijetí, hypotézu H_0 tedy nezamítáme.

Varianta s p-hodnotou (spočtenou např. v Matlabu):

$p = 0,19$

$p > \alpha$, hypotézu H_0 tedy nezamítáme.

Tj. na dané hladině významnosti jsme neprokázali závislost. Pozorované odchylky mohou být pouze náhodné.

χ^2 test dobré shody

Statistika:

$$S = \sum \frac{(O_i - E_i)^2}{E_i}$$

Rozdělení statistiky:

$$S \sim Chí^2\,(n-1)\,.$$

Test je pravostranný.

Nulová hypotéza: Veličina má takové a takové rozdělení.

Příklad:

H_0 : Počet cestujících přepravených v jednotlivé dni v týdnu má rovnoměrné rozdělení.
Volíme hladinu významnosti $\alpha = 5\%$.
Data:

Počet přepravených cestujících:

Po	Út+St	Čt+Pá	So	Ne
615	1281	1121	615	630

...O_i

Spočteme statistiku. Nejprve určíme ideální četnosti za předpokladu rovnoměrného rozdělení:
Celkový počet cestujících: 4262.

Po	Út+St	Čt+Pá	So	Ne
$\frac{1}{7} \cdot 4262 = 609$	$\frac{2}{7} \cdot 4262 = 1218$	$\frac{2}{7} \cdot 4262 = 1218$	$\frac{1}{7} \cdot 4262 = 609$	$\frac{1}{7} \cdot 4262 = 609$

...E_i

$S = \sum \frac{(O_i-E_i)^2}{E_i} = 11,83$
$\chi^2_{0,95}\,(n-1) = \chi^2_{0,95}\,(5-1) = \chi^2_{0,95}\,(4) = 9,49$
Statistika padla do kritického oboru, hypotézu H_0 tedy na hladině významnosti 5% zamítáme.
Tj. Počet cestujících není rovnoměrný a pravděpodobnost, že se v tomto mýlíme, je menší než 5%.

Varianta s p-hodnotou (spočtenou např. v Matlabu):
$p = 0,0187$
$p < \alpha$, hypotézu H_0 tedy zamítáme.
Tj. Počet cestujících není rovnoměrný a pravděpodobnost, že se v tomto mýlíme, je 1,87%.

Test mediánu = znaménkový test

Statistika:

$$z = \frac{2b - n}{\sqrt{n}}$$

Poznámka: Parametr b je počet hodnot větších než testovaný medián.

Rozdělení statistiky:

$$z \sim N(0, 1).$$

Test dle alternativní hypotézy.

Nulová hypotéza: Medián základního souboru je tolik a tolik.

Příklad:

H_0 : Polovina lidí v našem podniku má plat nad 16 000. ($\widetilde{x} \geq 16\,000$)
H_A : $\widetilde{x} < 16\,000$............levostranný test.
Volíme hladinu významnosti $\alpha = 5\%$.
Data:

Platy náhodně vybraných deseti lidí v tisících:
17, 14, 11, 25, 15, 14, 18, 11, 12, 10

Označím si hodnoty větší nebo rovny testovanému mediánu a spočtu jejich počet:
17+, 14, 11, 25+, 15, 14, 18+, 11, 12, 10
$b = 3$
Spočteme statistiku:
$z = \frac{2 \cdot 3 - 10}{\sqrt{10}} = -1,26$
Zjistíme příslušný kvantil:
$z_{0,05} = -1,65$
Statistika padla do oboru přijetí, hypotézu H_0 tedy na hladině významnosti 5% nezamítáme.
Tj. Data nejsou v tak příprém rozporu s nulovou hypotézou, abychom ji zamítli.

Test pořadové nezávislosti

Statistika:

$$z = \frac{2b - (n-2)}{\sqrt{n-1}}$$

Poznámka: Parametr b znamená počet nepřerušených sérií dat, která jsou nad resp. pod výběrovým mediánem.

Rozdělení statistiky:

$$z \sim N(0,1).$$

Test levostranný.

Nulová hypotéza: Data jsou pořadově nezávislá.

Příklad:

Máme pocit, že zkoušející se nechává ovlivňovat výsledky předchozích zkoušených. To mu zlepší nebo zhorší náladu a to se pak projeví v jeho aktuálním hodnocení. Je to pravda?

H_0 : Při zkoušení jsou známky nezávislé na předchozích výsledcích.

Volíme hladinu významnosti $\alpha = 5\%$.

Data:

Pořadí známek:

A, B, A C, D, F, F, E, C, A, B, A, D, E, E, C, B, A, A, B, A

Seřadíme hodnoty a určíme výběrový medián:

A, A, A, A, A, A, A, B, B, B, B, C, C, C, D, D, E, E, E, F, F

n=21

medián = B

Poznámka: Hodnotu B budu počítat do aktuální série. Teprve vyšší resp. nižší hodnota přepne sérii.

Série:

+++——+++——++++

Počet sérií:

b=5

Spočteme statistiku:

$z = \frac{2b-(n-2)}{\sqrt{n-1}} = \frac{10-19}{\sqrt{20}} = -2,01$

Zjistíme příslušný kvantil:

$z_{0,05} = -z_{0,95} = -1,64$

Statistika padla do kritického oboru, hypotézu H_0 tedy na hladině významnosti 5% zamítáme.

Tj. Známky jsou pořadově závislé a pravděpodobnost, že se v tomto mýlíme, je menší než 5%.