Testování hypotéz II

Úkol z minula:

$X_1, \dots, X_{25}$ náh. výběr z $N(\mu, \sigma^2)$

$H_0: \mu = 2$

$H_1: \mu < 2$

testová statistika $T = \frac{\bar{X} - 2}{S}\sqrt{n} \overset{H_0}{\sim} t(n-1)$

$W = (-\infty, -t_{1-\alpha}(n-1))$

pro naše data:

$t = \frac{1,99 - 2}{0,1}\sqrt{25} = -0,5$

$t \notin W \Rightarrow$ nezamítáme H_0

$W = (-\infty, -t_{0,95}(24)) = (-\infty; -1,71)$

Nepodařilo se nám prokázat, že nás šidí.

1.) (Výška prvňáčků)

Vybrali jsme 10 dívek a 15 chlapců a změřili jejich výšku:

průměrná ~~...~~ výška dívek je 120,56 a přísl. výb. rozptyl 9,14,

chlapců 124,13 8,95.

Můžeme na hl. významnosti 5 % tvrdit, že chlapci jsou v průměru vyšší než dívky?

X_i = výška i-té dívky

Y_i = -||- chlapce

neznámý, stejný v obou výběrech

$X_1, \dots, X_{10}$ náh. výběr z $N(\mu_1, \sigma^2)$

$Y_1, \dots, Y_{15}$ náh. výběr z $N(\mu_2, \sigma^2)$

vzájemně nezávislé

$H_0: \mu_1 = \mu_2 \qquad \mu_1 - \mu_2 = 0$

$H_1: \mu_1 < \mu_2 \qquad \mu_1 - \mu_2 < 0$

• pivotová statistika pro odhad $\mu_1 - \mu_2$ je $\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{S_*^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$.

• testová statistika $$T = \frac{\bar{X} - \bar{Y}}{\sqrt{S_*^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{H_0}{\sim} t(n_1 + n_2 - 2).$$

Kritický obor $W = (-\infty, -t_{1-\alpha}(n_1+n_2-2))$.

$\alpha = 0{,}05$

$t = -2{,}91$

$W = (-\infty, -1{,}71)$

$\left.\begin{matrix} t = -2{,}91 \\ W = (-\infty, -1{,}71) \end{matrix}\right\} t \in W \Rightarrow$ zamítáme H_0

Výše popsaný test se nazývá dvouvýběrový t-test.

Je předpoklad shodnosti rozptylů v obou výběrech opodstatněný?

obecnější model:

obecně různé, neznámé

$\left.\begin{matrix} X_1, \dots, X_{10} \text{ náh. výběr z } N(\mu_1, \sigma_1^2) \\ Y_1, \dots, Y_{15} \text{ náh. výběr z } N(\mu_2, \sigma_2^2) \end{matrix}\right\}$ vzájemně nezávislé

$H_0: \sigma_1^2 = \sigma_2^2$ $\qquad \frac{\sigma_1^2}{\sigma_2^2} = 1$

$H_1: \sigma_1^2 \neq \sigma_2^2$ $\qquad \frac{\sigma_1^2}{\sigma_2^2} \neq 1$

pivotová statistika pro odhad $\frac{\sigma_1^2}{\sigma_2^2}$ je $\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$

• testová statistika je $T = \frac{S_1^2}{S_2^2} \overset{H_0}{\sim} F(n_1 - 1, n_2 - 1)$,

$$W = \left(0, F_{\frac{\alpha}{2}}(n_1-1, n_2-1)\right) \cup \left(F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)\right).$$

pro naše data je

$T = 1{,}02$

$\alpha = 0{,}05$ (není-li uvedeno, bere se $\alpha = 0{,}05$)

$W = (0, F_{0,025}(9,14)) \cup (F_{0,975}(9,14)) = (0; 0{,}26) \cup (3{,}21; \infty)$.

$T \notin W \Rightarrow H_0$ nezamítáme $\Rightarrow$ předpoklad shodnosti rozptylů je reálný.

2.) Na 9 pokusných polích byly zjišťovány výnosy staré a nové odrůdy kukuřice. Výsledky jsou uvedeny v násl. tabulce:

pole	A	B	C	D	E	F	G	H	I
stará odrůda	4,23	5,09	4,55	5,04	5,31	5,45	4,76	5,57	4,97
nová odrůda	5,83	7,05	6,00	6,31	4,86	5,22	6,44	4,71	6,09

Liší se výnosy obou druhů odrůd? Volte $\alpha = 0,05$.

X_i = výnos staré odrůdy na i-tém poli

Y_i = nové -||-

$$\left.\begin{matrix} X_1, \dots, X_9 \\ Y_1, \dots, Y_9 \end{matrix}\right\} \text{ vzájemně } \underline{\text{ZÁVISLÉ!}}$$

Definujme $Z_i = Y_i - X_i$ ... rozdíl výnosů na i-tém poli

$Z_1, \dots, Z_9$ náh. výběr z $N(\mu, \sigma^2)$

$H_0: \mu = 0$

$H_1: \mu \neq 0$

• testová statistika $T = \dfrac{\bar{Z} - 0}{S}\sqrt{n}$

kritický obor $W = (-\infty, -t_{1-\frac{\alpha}{2}}(n-1)) \cup (t_{1-\frac{\alpha}{2}}(n-1), \infty)$

data:

$z_i = (1,60; 1,96; 1,45; 1,27; -0,45; -0,23; 1,68; -0,86; 1,12)$

$\bar{z} = 0,838$

$s^2 = 1,1086$

$t = 2,39$

$W = (-\infty, -t_{0,975}(8)) \cup (t_{0,975}(8), \infty) = (-\infty; -2,306) \cup (2,306; \infty)$.

$t \in W \Rightarrow$ H_0 zamítáme $\Rightarrow$ Podařilo se nám prokázat rozdíl.

Výše popsaný test se nazývá párový t-test.

3.) Ze 40 hodů mincí nám padlo celkem 22 líců. Můžeme tvrdit, že je mince falešná?

$X_i \begin{cases} 1 & \dots \text{ v } i\text{-tém hodu padl líc} \\ 0 & \dots \text{ jinak} \end{cases}$

$X_1, \dots, X_{40}$ náh. výběr z $A(p)$ $\quad p =$ pravděpodobnost padnutí líce

$H_0: p = \frac{1}{2}$

$H_1: p \neq \frac{1}{2}$

pivotová statistika: $\dfrac{\overline{X} - p}{\sqrt{p(1-p)}} \sqrt{n}$

testová statistika $T = \dfrac{\overline{X} - \frac{1}{2}}{\sqrt{\frac{1}{2}\left(1 - \frac{1}{2}\right)}} \sqrt{n} \overset{H_0}{\approx} N(0,1)$

kritický obor $W = \left(-\infty, -u_{1-\frac{\alpha}{2}}\right) \cup \left(u_{1-\frac{\alpha}{2}}, \infty\right)$.

$\bar{x} = \frac{22}{40} = 0{,}55$

$\left.\begin{array}{l} t = 0{,}6324 \\ \alpha = 0{,}05 \\ W = (-\infty, -1{,}96) \cup (1{,}96; \infty) \end{array}\right\} \; t \notin W \Rightarrow H_0$ nezamítáme.