Dvouvýběrový t-test řešené příklady

Příklad 1

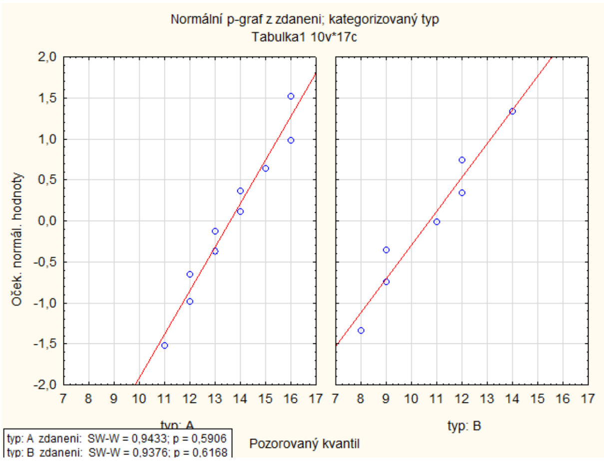
V 10 zemích bylo v roce 2005 zavedeno regresivní zdanění (A). v dalších 7 zemích bylo v tom samém roce zavedeno progresivní zdanění (B).

A: 12; 14; 12; 15; 16; 13; 14; 11; 13; 16;
B: 14; 12; 12; 9; 11; 8; 9;

Na hladině významnosti 5 % testujte hypotézu, že výnosy z daní jsou v zemích stejné. Předpoklad shody rozptylů pro využití nepárového t-testu je v datech splněn.

Řešení 1

Nejprve stanovíme nulovou a alternativní hypotézu: $H_0 : \mu_A = \mu_B$, $H_1 : \mu_A \neq \mu_B$.
V prvním kroku řešení příkladů ověříme předpoklady testu, tedy že oba naše výběry pocházejí z normálního rozdělení, dále bychomo ověřili předpoklad shody rozptylů, ale ze zadání víme, že je splněn. Pro ověření normality použijeme Shapiro-Wilkův test, p-hodnoty vyšly postupně 0,59 a 0,61, tedy budeme nadále předpokládat, že data pochází z normálního rozdělení. Graficky bychom mohli normalitu ověřit například pomocí N-P plotu.



Poté si sestrojíme tabulku:

skupina	velikost výběru	výběrový průměr	výběrová směrodatná odchylka
1	$n_1 = 10$	$\bar{x}_1 = 13,6$	$s_1 = 1,71$
2	$n_2 = 7$	$\bar{x}_2 = 10,7$	$s_2 = 2,14$

Následně vypočítáme rozdíl průměrů, $\bar{x}_1 - \bar{x}_2 = 1,71 - 2,14 = -0,43$ a sdruženou výběrovou směrodatnou odchylku jako:

$$s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{(10-1)1,71^2+(7-1)2,14^2}{10+7-2}} = 1,89$$

z ní dále vypočítáme směrodatnou chybu v rozdílech mezi průměry

$$SE(\bar{x}_1-\bar{x}_2) = s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = 1,89\sqrt{\frac{1}{10}+\frac{1}{7}} = 0,93.$$

Mezi-výpočty dosadíme do vzorce pro výpočet T-statistiky

$$t_0 = \frac{\bar{x}_1-\bar{x}_2}{SE(\bar{x}_1-\bar{x}_2)} = \frac{13,6-10,7}{0,93} = 3,11.$$

V posledním kroku použijeme tabulky pro t-rozdělení, abychom porovnali hodnotu T-statistiky s kritickým oborem využívajícím t rozdělení:

$$W = \left(-\infty; -t_{1-\frac{\alpha}{2}}(n_1+n_2-2)\right\rangle \cup \left\langle t_{1-\frac{\alpha}{2}}(n_1+n_2-2); \infty\right)$$
$$W = (-\infty; -t_{0,975}(15)\rangle \cup \langle t_{0,975}(15); \infty)$$
$$W = (-\infty; -2,1314\rangle \cup \langle 2,1314; \infty)$$

Jelikož vidíme, že realizace testového kritéria spadá do kritického oboru ($t_0 \in W$), pak nulovou hypotézu zamítáme. Jinými slovy, na hladině významnosti pěti procent jsme prokázali, že se průměrný výnos z daní se v zemích lišil.

Příklad 2

Sledujeme vývoje akcií dvou firem. Provedli jsme náhodný výběr rozsahu 20 u akcií firmy A, které se řídí normálním rozložením $N(4;3)$ a náhodný výběr hodnot akcií firmy B rozsahu 15, které se řídí normálním rozložením $N(5;5)$. Oba dva výběry jsou nezávislé. Jaká je pravděpodobnost, že rozdíl průměrů hodnot akcií A-B bude záporný?

Řešení

$$P(M_1-M_2<0) = P\left(\frac{M_1-M_2-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} < \frac{0-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\right) =$$
$$P\left(U < \frac{0-(4-5)}{\sqrt{3/20+5/15}}\right) = P(U<1,61) = \Phi(1,61) = 0,9463$$