Cvičení 12 – opakovací cvičení

Příklad 1.: Jednou za semestr studenti pomocí bodů hodnotí přednášející. Známe bodové hodnocení deseti náhodně vybraných přednášejících loni a letos, přičemž hodnoty v tabulce jsme získali zprůměrováním bodového hodnocení všech studentů, kteří se do ankety zapojili:

Číslo učitele	1	2	3	4	5	6	7	8	9	10
Body loni	932	906	943	907	893	870	889	902	866	887
Body letos	933	923	942	909	908	893	890	900	870	895

Naším úkolem je na hladině významnosti 0,05 posoudit, zda úroveň hodnocení přednášejících je stejná loni i letos.
Výsledek: Úloha vede na párový t-test, testujeme $H_0: \mu_1 - \mu_2 = 0$ proti $H_1: \mu_1 - \mu_2 \neq 0$.
Realizace testové statistiky: $t_0 = -2{,}4893$, kritický obor $W = \left(-\infty; -2{,}2622\right\rangle \cup \left\langle 2{,}2622; \infty\right)$.
Protože $t_0 \in W$, H_0 zamítáme na hladině významnosti 0,05. S rizikem omylu nejvýše 0,05 jsme prokázali, že úroveň hodnocení přednášejících loni a letos se liší. Odpovídající p-hodnota je 0,0345, což je menší než 0,05 a H_0 opět zamítáme na hladině významnosti 0,05.

Příklad 2.: Známe údaje o počtu kusů prodaného zboží určitého druhu ve dvou prodejnách, v první prodejně v šesti po sobě následujících týdnech, ve druhé v pěti týdnech:
prodejna č. 1: 62, 54, 55, 60, 53, 58
prodejna č. 2: 52, 56, 49, 50, 51.
Na hladině významnosti 0,05 ověřte, zda rozdíly v počtu prodaných kusů zboží v obou prodejnách jsou pouze náhodné.
Výsledek:
Úloha vede na dvouvýběrový t-test. Testujeme $H_0: \mu_1 - \mu_2 = 0$ proti $H_1: \mu_1 - \mu_2 \neq 0$.
Předtím je ovšem zapotřebí provést test shody rozptylů, tj. testujeme $H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$ proti $H_1: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$. Protože realizace testové statistiky 1,7534 nepatří do kritického oboru W = $\left(0; 01354\right\rangle \cup \left\langle 9{,}3645; \infty\right)$, nelze na hladině významnosti 0,05 zamítnout hypotézu o shodě rozptylů.
Protože testová statistika dvouvýběrového t-testu se realizuje hodnotou 2,771, která patří do kritického oboru $\left(-\infty; -2{,}2622\right\rangle \cup \left\langle 2{,}2622; \infty\right)$, H_0 zamítáme na hladině významnosti 0,05.
Tedy rozdíl mezi prodejnami je statisticky významný, není způsoben pouze náhodnými vlivy.

Příklad 3.: Výkon 18 gymnastek byl ohodnocen stanovením jejich pořadí od nejlepší (pořadí 1) po nejslabší (pořadí 18). V hodnocené skupině bylo 11 žákyň trenérky A a 7 žákyň trenérky B. V tabulce je uvedeno pořadí žákyň obou trenérek:

A	1	4	5	7	8	10	11	13	14	16	17
B	2	3	6	9	12	15	18				

Na hladině významnosti 0,05 testujte hypotézu, že výukové metody obou trenérek jsou stejně účinné proti oboustranné alternativě.

Výsledek: Testová statistika dvouvýběrového Wilcoxonova testu se realizuje hodnotou 37, kritická hodnota je 16, nulovou hypotézu tedy nezamítáme na hladině významnosti 0,05. Neprokázal se rozdíl ve výukových metodách obou trenérek.

Příklad 4.: Měřením délky deseti válečků byly získány hodnoty (v mm): 5,38 5,36 5,35 5,40 5,41 5,34 5,29 5,43 5,42 5,32. Těchto deset hodnot považujeme za realizace náhodného výběru rozsahu 10 z normálního rozložení $N(\mu, \sigma^2)$.

a) Sestrojte 99% interval spolehlivosti pro neznámou střední hodnotu μ.
b) Sestrojte 99% interval spolehlivosti pro neznámou směrodatnou odchylku σ.

Výsledek:
Ad a) 5,3248 mm < μ < 5,4152 mm s pravděpodobností aspoň 0,99
Ad b) 0,0272 mm < σ < 0,1002 mm s pravděpodobností aspoň 0,99.

Příklad 5.: Firma vyrábějící elektronické přístroje chce uvést na trh nový DVD přehrávač. Marketingový tým mj. zkoumá, jak zákazníci různých věkových kategorií hodnotí tento nový přehrávač. Jsou k dispozici data od 68 zákazníků (datový soubor dvd.sta). Ti byli rozděleni podle věku do 6 skupin (18-24, 25-31, …., 53-59) – proměnná A. Bodové hodnocení přehrávače je obsaženo v proměnné X.
Na hladině významnosti 0,05 máme testovat hypotézu, že rozdíly v bodovém hodnocení různými věkovými skupinami zákazníků jsou způsobeny pouze náhodnými vlivy. V případě zamítnutí nulové hypotézy je třeba identifikovat, které dvojice skupin se liší na hladině významnosti 0,05.
Výsledek: Úloha vede na analýzu rozptylu jednoduchého třídění. Hypotézu o normalitě jednotlivých náhodných výběrů nezamítáme na hladině významnosti 0,05 ani v jednom případě. Rovněž nezamítáme na hladině významnosti 0,05 hypotézu o homogenitě rozptylů. Na hladině významnosti 0,05 se však prokázal rozdíl mezi jednotlivými středními hodnotami ($F_A = 4{,}6$, $p = 0{,}0012$). Výsledek Scheffého metody ukazuje, že na hladině významnosti 0,05 se liší věkové skupiny 39 - 45 let a 46 – 52 let a dále 39 - 45 let a 53 – 59 let. Rozdíly mezi ostatními dvojicemi věkových skupin nejsou prokazatelné na hladině významnosti 0,05.

Příklad 6.: Při kontrole pěti balíčků cukru o deklarované hmotnosti 1000 g byly zjištěny tyto odchylky: -1, 2, -2, 3, 1. Považujeme je za realizace náhodného výběru rozsahu 5 z rozložení $N(\mu, \sigma^2)$. Sestrojte 95% interval spolehlivosti pro neznámou směrodatnou odchylku σ.
Výsledek: 1,24 < σ < 5,96 s pravděpodobností aspoň 0,95.

Příklad 7.: Na automatické balicí lince se sledoval počet zastavení činnosti automatu v průběhu osmihodinové směny. Jsou k dispozici tyto výsledky:

hodina	1	2	3	4	5	6	7	8
počet zastavení	16	17	19	16	24	19	17	16

Na asymptotické hladině významnosti 0,05 testujte hypotézu, že počet zastavení automatu se řídí rovnoměrným diskrétním rozložením.
Výsledek:
Podmínky dobré aproximace jsou splněny, protože $np_j \geq 5$ pro všechna j.
Testová statistika: K = 52/18 = 2,91. Kritický obor: $W = \langle \chi^2_{0,95}(7), \infty) = \langle 14{,}1, \infty)$. Protože testová statistika se nerealizuje v kritickém oboru, na asymptotické hladině významnosti 0,05 nezamítáme hypotézu o rovnoměrném diskrétním rozložení.