

Řešení vzorové počítačové části zkoušky z Výpočetní statistiky

Ve dvou laboratořích byly pořízeny nezávislé náhodné výběry rozsahu 17 z populace laboratorních myší, které byly zasaženy polychlorovanými bifenyly. Proměnná X udává velikost jater uhynulého zvířete a proměnná Y udává aktivitu jaterního enzymu.

Výběr číslo 1: (14290, 5177), (13459, 5840), (5586, 1732), (19593, 6908), (15193, 5389), (14741, 5683), (12055, 5073), (19265, 4101), (7908, 4584), (11451, 4666), (14379, 5164), (16236, 5178), (20032, 8475), (15875, 5284), (12724, 3548), (16829, 2393), (13998, 5155)

Výběr číslo 2: (19728, 7130), (12136, 4855), (20484, 2357), (19187, 4339), (24462, 3786), (4434, 991), (4089, 1936), (8707, 1641), (24229, 5467), (13739, 2451), (5365, 1517), (4810, 832), (18995, 3194), (12460, 4292), (20146, 3597), (10163, 3124), (13998, 1552)

a) V obou výběrech Lilieforsovou variantou Kolmogorovova – Smirnovova testu ověřte na hladině významnosti 0,05 normalitu proměnných X a Y. (Uveďte hodnotu testové statistiky, p-hodnotu a rozhodnutí o nulové hypotéze).

První výběr:

proměnná X: $d = 0,11546$, $p > 0,2$, proměnná Y: $d = 0,17587$, $p < 0,2$

Druhý výběr:

proměnná X: $d = 0,17742$, $p < 0,15$, proměnná Y: $d = 0,1214$, $p > 0,2$

Ani v jednom případě hypotézu o normalitě nezamítáme na hladině významnosti 0,05.

b) Vypočítejte průměry a směrodatné odchylky velikosti jater uhynulých myší v obou výběrech a na hladině významnosti 0,05 testujte hypotézu, že střední hodnota velikosti jater uhynulých myší je v obou výběrech shodná.

První výběr:

Průměr $m_1 = 14330,24$

směrodatná odchylka $s_1 = 3810,846$

Druhý výběr:

Průměr $m_2 = 13948,94$

směrodatná odchylka $s_2 = 6963,882$

Test hypotézy o shodě rozptylů.

Testová statistika F-testu: $F = 3,339335$,

počet stupňů volnosti čitatele = 16, počet stupňů volnosti jmenovatele = 16,

p-hodnota = 0,020962,

rozhodnutí o hypotéze o shodě rozptylů: na hladině významnosti 0,05 zamítáme hypotézu o shodě rozptylů.

Testová statistika Levenova testu: $F = 8,72554$,

počet stupňů volnosti čitatele = 1, počet stupňů volnosti jmenovatele = 32,

p-hodnota = 0,005841,

rozhodnutí o hypotéze o shodě rozptylů na hladině významnosti 0,05 zamítáme hypotézu o shodě rozptylů.

Při testování hypotézy o shodě středních hodnot musíme použít variantu dvouvýběrového t-testu se separovanými odhady rozptylů.

Test hypotézy o shodě středních hodnot.

Testová statistika dvouvýběrového t-testu: $t = 0,198039$,

počet stupňů volnosti = 24,79412,

p-hodnota = 0,844629,

rozhodnutí o hypotéze o shodě středních hodnot: hypotézu o shodě středních hodnot nezamítáme na hladině významnosti 0,05.

c) Variabilitu velikosti jater a aktivity jaterního enzymu v prvním a druhém výběru posuďte pomocí koeficientů variace. Který výběr vykazuje vyšší variabilitu velikosti jater a který výběr vykazuje vyšší variabilitu aktivit jaterního enzymu?

První výběr:

$$\text{koeficient variace velikosti jater: } \frac{3810,846}{14330,24} = 0,2659,$$

$$\text{koeficient variace aktivity jaterního enzymu: } \frac{1538,911}{4961,765} = 0,310154$$

Druhý výběr:

$$\text{koeficient variace velikosti jater: } \frac{6963,882}{13948,94} = 0,499241,$$

$$\text{koeficient variace aktivity jaterního enzymu: } \frac{1725,368}{3119,647} = 0,553065$$

Komentář: V obou případech jsou koeficienty variace vyšší ve druhém výběru.

d) V obou výběrech vypočtete hodnotu výběrového koeficientu korelace, sestrojte asymptotický 95% interval spolehlivosti pro ρ a na hladině významnosti 0,05 testujte hypotézu o nezávislosti veličin X, Y.

První výběr:

$$R_{12} = 0,574337,$$

$$\text{testová statistika} = 2,717254,$$

$$\text{počet stupňů volnosti} = 15,$$

$$\text{p-hodnota} = 0,015899,$$

rozhodnutí o hypotéze o nezávislosti: na hladině významnosti 0,05 zamítáme hypotézu o nezávislosti veličin X a Y.

$$\text{dolní mez 95\% asymptotického intervalu spolehlivosti pro } \rho: 0,129418$$

$$\text{horní mez 95\% asymptotického intervalu spolehlivosti pro } \rho: 0,826754$$

Druhý výběr:

$$R_{12} = 0,669384,$$

$$\text{testová statistika} = 3,489639,$$

$$\text{počet stupňů volnosti} = 15,$$

$$\text{p-hodnota} = 0,0003293,$$

rozhodnutí o hypotéze o nezávislosti: na hladině významnosti 0,05 zamítáme hypotézu o nezávislosti veličin X a Y.

$$\text{dolní mez 95\% asymptotického intervalu spolehlivosti pro } \rho: 0,278268$$

$$\text{horní mez 95\% asymptotického intervalu spolehlivosti pro } \rho: 0,87009$$

e) Na hladině významnosti 0,05 testujte hypotézu, že koeficienty korelace obou rozložení jsou stejné.

$$\text{Testujeme } H_0: \rho = \rho^* \text{ proti } H_1: \rho \neq \rho^*.$$

$$\text{p-hodnota} = 0,6834,$$

rozhodnutí o hypotéze o shodě korelačních koeficientů: na hladině významnosti 0,05 se neprokázal rozdíl mezi korelačními koeficienty.