

MV011 Statistika I – cvičení 12

Analýza rozptylu

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Brno

jaro 2016



Příklad 1

Je dáno pět nezávislých náhodných výběrů o rozsazích 5, 7, 6, 8, 5, přičemž i -tý výběr pochází z rozložení $N(\mu_i, \sigma^2)$, $i = 1, \dots, 5$. Byl vypočten celkový součet čtverců $S_T = 15$ a reziduální součet čtverců $S_e = 3$.

Na hladině významnosti 0,05 testujte hypotézu o shodě středních hodnot.

Příklad 2

Datový soubor [brambory.csv](#): hmotnost trsu brambor (v kilogramech) u 4 odrůd brambor. Na hladině významnosti 0,05 testujte hypotézu, že střední hodnota hmotnosti trsu brambor nezávisí na odrůdě. Zamítnete-li nulovou hypotézu, zjistěte, které dvojice odrůd se liší na hladině významnosti 0,05.

Příklad 3

Datový soubor [prodavaci.csv](#): měsíční tržby (v Kč) tří prodavačů za dobu půl roku. Na hladině významnosti 0,05 testujte hypotézu, že střední hodnoty tržeb všech tří prodavačů jsou stejné. Pokud zamítneme nulovou hypotézu, zjistěte, tržby kterých dvou prodavačů se liší na hladině významnosti 0,05.

Příklad 4

Pro datový soubor `rada1.csv` uvažujte 4 regresní modely: přímku procházející počátkem, obecnou přímku, parabolu a polynom 3. stupně. Ty spočítejte, výsledky graficky prezentujte a modely statisticky porovnejte, mj. pomocí analýzy rozptylu.

Příklad 5

V proměnné `LakeHuron` v R jsou uloženy údaje o hloubce (ve stopách) jezera Huron měřené v letech 1875–1972. Data proložte polynomem 8. stupně. Pomocí analýzy rozptylu zkoumejte možnosti zmenšení stupně regresního polynomu.

Výsledky

1. $n = 31$, $a = 5$, $S_A = 12$, $F_A = 26$, $F_{0,95}(4, 26) = 2,7426$
Protože $F_A > F_{0,95}(4, 26)$, zamítáme H_0 na hladině významnosti 0,05.
2. H_0 zamítáme, liší se odrůdy brambor 1–3, podle Tukeyovy metody také navíc 1–2.
3. H_0 zamítáme, liší se tržby prodavačů 1–3 a 2–3.
4. Z uvedených modelů je nejvhodnější model s regresní parabolou.
5. Je možné snížit stupeň polynomu na 7.