

MA012 Statistika II

cvičení 1

Ondřej Pokora (pokora@math.muni.cz)

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

(podzim 2015)



Příklad 1

Projděte si **Příklad 1** z přednášky. Data jsou uložena v souboru **brambory.csv**, ukázkové řešení naleznete ve skriptu **reseni-01-1.R**.

Příklad 2

Projděte si **Příklad 2** z přednášky. Data jsou dostupná v souboru **data01.csv**, ukázkové řešení naleznete ve skriptu **reseni-01-2.R**.

Příklad 3

Jsou známy měsíční tržby (v tisících Kč) tři prodavačů za dobu půl roku. Data jsou dostupná v souboru **prodavaci.csv**.

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.

[Na hladině významnosti 0,05 se liší tržby prodavačů 1, 3 a 2, 3.]

Příklad 4

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.

[$n = 31$, $a = 5$, $S_A = 12$, $f_A = 26$, $F_{0,95}(4, 26) = 2,7426$ Protože $f_A \geq F_{0,95}(4, 26)$, H_0 zamítáme na hladině významnosti 0,05.]

Příklad 5

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

^aproměnná definovaná v R

[Možno jít na stupeň 7.]

Příklad 6

U 126 podniků řepařské oblasti v České Republice byl sledován hektarový výnos cukrovky ve vztahu ke spotřebě průmyslových hnojiv.

Data jsou uložena v souboru **cukrovka.csv** ve 4 sloupcích:

1 dolní hranice spotřeby K_2O (kg/ha)

2 horní hranice spotřeby K_2O (kg/ha)

3 četnosti

4 průměrné výnosy cukrovky (q/ha)

a) odhadněte parametry regresní funkce tvaru

$$y = \beta_0 + \beta_1 x$$

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

Za hodnoty nezávisle proměnné volte střed intervalu.

b) Porovnejte vhodnost použitých regresních modelů pomocí analýzy rozptylu.

[Kvadratický model je významný.]

Příklad 7

*Klinickou studií byla sledována závislost mezi dietou a dobou za kterou dojde ke koagulaci (sražení) krve. Výzkum bym proveden na 24 zvířatech krmených čtyřmi různými dietami (označených kategoriemi A–D) Výsledky jsou uloženy v souboru **srazlivost.csv** .*

Na hladině významnosti 0,05 testujte hypotézu, že střední hodnoty doby srážlivosti jsou stejné pro všechny testované diety. Pokud nulovou hypotézu zamítnete, zjistěte, u kterých dvojic diet je srážlivost signifikantně odlišná.

[Doba srážlivosti se liší v závislosti na podávané dietě. Tukeyova metoda ukazuje významný rozdíl pro dvojice diet A-B, A-C, B-D, C-D.]

Příklad 8

Tři druhy myši byly testovány na agresivitu dle svého chování v bludišti. Každá myš byla umístěna do středu bludiště tvaru čtverce o straně délky 1 m, rozděleného na 49 stejných čtverečků. Myš se pokoušela o únik z tohoto bludiště a za čas 5 minut se sledoval počet čtverečků, které myš překonala. Data jsou dostupná v souboru `mysi.csv`. Počet překonaných čtverečků je pro každý druh myši zaznamenán v samostatném sloupci, u třetí druhu myši poslední měření chybí (NA).

Na hladině významnosti 0,05 testujte hypotézu, že střední hodnoty počtu překonaných čtverečků jsou stejné pro všechny tři druhy testovaných myši. testované diety. Pokud nulovou hypotézu zamítnete, zjistěte, u kterých dvojic se hodnoty významně liší.