

CVIČENÍ LINEÁRNÍ STATISTICKÉ MODELY I – PODZIM 2013 – ZADÁNÍ 2. SÉRIE ÚLOH

Úkol ([1]): 29. února 1992 byl v *Singapore Straits Times* uveřejněn inzerát na prodej 48 dámských prstenů opatřených diamanty. Kromě fotografií byla u jednotlivých prstenů uvedena hmotnost diamantu (v karátech = 0,2 g) a cena (v USD).

- ▷ Datový soubor si zkopírujte z `/erko/M5120/data/diamonds.txt` ze serveru `bart`.
- ▷ Zkoumejte závislost ceny prstenu, Y , na hmotnosti diamantu, x , pomocí lineárního regresního modelu s níže uvedenými čtyřmi funkcemi pro hmotnosti diamantů od 0 do 0,5 karátů.
- ▷ Proložte data regresní přímkou v obecné poloze. Obchodníkovi s klenoty by se výsledek jistě nelíbil, proč?
- ▷ Proložte data funkcí $y = \beta_1 x + \beta_2 x^2$. Co je to za funkci? Je tento model reálnější než předchozí?
- ▷ Proložte exponenciální funkci $y = a e^{b x + c x^2}$.
- ▷ Proložte mocninnou funkci $y = a x^d$.
- ▷ Nakreslete grafy všech čtyř odhadnutých regresních funkcí spolu s daty do jednoho obrázku, pro hmotnosti diamantů od 0 do 0,5 karátů (příp. pro širší interval), nezapomeňte na popisky.
- ▷ Pro každou regresní funkci uveďte (např. ve formě tabulky) tyto výsledky: funkční zápis odhadnuté regresní funkce, S_e, s^2, R^2, \bar{R}^2 . Přidejte informace o celkovém F-testu a testech významnosti parametrů v jednotlivých funkcích. Pro testy je důležitý předpoklad rozdělení pravděpodobnosti veličiny Y , o jaká rozdělení u jednotlivých funkcí jde?
- ▷ Porovnejte výsledky. Který model byste vybrali jako nejlepší? Své rozhodnutí doprovodte komentářem.
- ▷ Uveďte také použité tvary příkazů pro řešení lineárních modelů, tj. řádky s příkazem `lm` v R , příp. analogickým příkazem v jiném použitém softwaru.

[1] Chu, Singfat (1996). Diamond Ring Pricing Using Linear Regression, *Journal of Statistics Education* 4 (3).