

Cvičení 13

V následujících příkladech předpokládejte, že data jsou pořízena jako náhodné výběry z populací sledujících normální rozdělení. K tomuto předpokladu však nepřistupujte nekriticky! Pokud budete přesvědčeni, že zde musíte použít jiný test, učiňte to.

1. Při sledování závislosti těla tloušťě na jeho stáří bylo zváženo 14 tloušťů (viz tabulka 1).

Tabulka 1: Tloušťky tloušťě

Tloušť číslo	Věk tloušťě (roky)	Hmotnost tloušťě (g)
1	1	10
2	1	12
3	1	15
4	2	28
5	2	29
6	2	29
7	3	60
8	3	55
9	4	100
10	4	115
11	4	98
12	5	160
13	5	168
14	5	170

- (a) Vyjádřete funkční závislost mezi sledovanými znaky pomocí paraboly, hyperboly a exponenciální funkce.
 - (b) Pro každý regresní model spočítejte ukazatel těsnosti korelační závislosti a na základě věcného rozboru problému spolu s přihlédnutím k hodnotě příslušného ukazatele vyberte nejvhodnější model. Jaký z nich to je?
 - (c) [Opakování:] Otestujte zda hmotnost tloušťě závisí na jeho stáří (test na shodu středních hodnot). Hladinu významnosti volte $\alpha = 0,01$.
2. Při volejbalovém memoriálu R. Myslíka byly získány antropometrické údaje (viz soubor [memorial.sta](#)), které již znáte např z šestého cvičení.

Pro další úvahy vycházejte z následujících možných závislostí (nezávislá \rightarrow závislá proměnná):

1. Výška \rightarrow Váha
2. Výška \rightarrow Dosah 2 stoj
3. Výška \rightarrow Dosah 1 stoj
4. Výška \rightarrow Abs. rozb.
5. Výška \rightarrow Dosah rozb.
6. Dosah rozb. \rightarrow Abs. rozb.

- (a) Zobrazte korelační pole a i s ohledem na věcnou stránku věci rozhodněte, který regresní model nejlépe vystihuje vzájemnou vazbu mezi sledovanými ukazateli. Modelů můžete uvažovat několik.
 - (b) Vypočítejte parametry regresního modelu(ů).
 - (c) Prostřednictvím testu zjistěte zda je(jsou) modely celkově statisticky významné. Dále pracujte jen s významnými modely.
 - (d) Prostřednictvím testu zjistěte zda lze regresní model(y) zjednodušit. Pokud ano, zjednodušte.
 - (e) Pokud jste mimo jiné uvažovali lineární regresní model, interpretujte hodnotu regresního koeficientu u tohoto modelu.
 - (f) Vypočítejte a interpretujte hodnoty koeficientů respektive indexů korelace a determinace. Na základě těchto hodnot vyberte „nejlepší“ model a dále pracujte jen s ním.
 - (g) Určete jakou(ý) . . . (viz vysvětlovaná proměnná) lze očekávat u volejbalisty vysokého 180 cm, resp. jaký absolutní smečářský výskok bude mít, má-li smečářský dosah 321 cm.
 - (h) Určete 99% interval spolehlivosti pro tyto odhady.
 - (i) Graficky znázorněte empirické a teoretické (tj. odhadnuté na základě zvoleného modelu) hodnoty závisle proměnné. Vypočítejte rezidua.
3. Na základě příkladu o rodinných výdajích z minulého cvičení vyřešte následující úkoly.
- (a) Vysvětlete výdaje za potraviny v závislosti jak na počtu osob v domácnosti, tak příjmu na domácnost. Volte lineární aditivní regresní model.
 - (b) Prostřednictvím testu zjistěte zda je tento model celkově statisticky významný.
 - (c) Prostřednictvím testu zjistěte zda lze regresní model zjednodušit, tj. jsou-li oba regresory a absolutní člen v modelu statisticky významné. Pokud nejsou, zjednodušte.
 - (d) Interpretujte hodnotu(y) regresních koeficientu(ů) tohoto modelu.
 - (e) Odhadněte bodově i intervalově (95% interval spolehlivosti) v závislosti na čistém příjmu domácnosti a počtu osob v domácnosti výdaje za potraviny pro příjem Kč 50 000,- a 4 osoby.
4. Rozhodněte, zda spotřeba automobilu závisí na jeho výkonu motoru, když navíc přihlédnete k hmotnosti a maximální rychlosti (viz soubor [auta.sta](#)), které již znáte např. ze sedmého cvičení.