
CVIČENÍ 7.5.2013

DVOUVÝBĚROVÝ T-TEST Uvažujte stejná data jako na minulé hodině.

1. Zajímáme se o porovnání IQ chlapců a dívek.

- (a) Pomocí popisných statistik si udělejte představu o tom, zda se typické IQ chlapců a dívek liší. `Statistics` → `Summaries` → `Numerical summaries`
- (b) Vykreslete si vhodný ilustrativní obrázek (boxplot).
- (c) Ověřte, zda lze předpokládat, že IQ chlapců a dívek jsou výběry z normálního rozdělení se shodným rozptylem.
 - Číselné porovnání směrodatných odchylek (např.): `Statistics` → `Summaries` → `Table of statistics`. Zde vyberte IQ v závislosti na POHLAVI a zvolte `Standard deviation`.
 - Test normality: Poslední řádek ve `Script Window` přepište tak, že namísto `sd`, `na.rm=TRUE` napíšete `shapiro.test`. Poté dejte `Submit`.
 - Test shody rozptylů lze provést také pomocí tzv. F-testu.
`Statistics` → `Variances` → `Two Variance F-test`.
- (d) Otestujte, zda je typické IQ chlapců stejné jako typické IQ dívek.
 - Uveďte uvažovaný model. Jsou splněny předpoklady tohoto testu?
 - Formulujte nulovou a alternativní hypotézu.
 - Proveďte dvouvýběrový t-test.

`Statistics` → `Means` → `Independent samples t-test`. Zde vhodně doplňte jednotlivá políčka (především správná volba alternativní hypotézy a předpokladu shody rozptylů).

- Interpretujte výsledek.

2. Zjistěte, zda mají chlapci na konci sedmé třídy lepší známky než dívky.

- Nejdříve se podívejte na popisné statistiky a vhodný obrázek.
- Posuďte, jak je to s normalitou průměrné známky z konce sedmé třídy pro dívky a chlapce zvlášť. Zvažte také, zda lze předpokládat shodu rozptylů.
- Odpovězte na zkoumanou otázku pomocí vhodného testu.

3. Ověřte domněnku, podle které mají dívky o více než 0.2 horší průměr v pololetí osmé třídy než chlapci.

4. Zjistěte, zda je statisticky významný rozdíl mezi známkami dětí ze školy č.1 a školy č.2. Berte známku v pololetí osmé třídy.

SAMOSTATNÁ PRÁCE Uvažujte data `Davis.txt`, která obsahují údaje o mužích a ženách, kteří pravidelně navštěvují náhodně vybranou posilovnu. U každé osoby známe jeho/její pohlaví, skutečnou váhu a výšku a váhu a výšku, kterou osoba veřejně uvádí.

- `Pohlavi` pohlaví osoby,
- `Vaha` skutečná (naměřená) hmotnost v kg,
- `Vyska` skutečný (naměřená) výška v cm,
- `RVaha` hmotnost v kg, kterou osoba veřejně uvádí,
- `RVyska` výška v cm, kterou osoba veřejně uvádí.

1. Načtěte si do R Commanderu soubor data a prohlédněte si základní popisné statistiky.
2. Pomocí vhodného testu ověřte, že jsou muži typicky skutečně vyšší než ženy.
 - (a) Nejprve srovnajte výběrové průměry a vykreslete si boxploty.
 - (b) Poté toto tvrzení ověřte statistickým testem. Posuďte, zda jsou splněny předpoklady testu a jako variantu vybereme.
3. Podobně ověřte domněnku, že muži mají typicky o více než 10 kg vyšší hmotnost než ženy.
4. Zajímá nás, zda lidé v posilovně uvádějí jinou hmotnost, než ve skutečnosti mají.
 - (a) Spočtete novou proměnnou, která bude udávat tento rozdíl.
 - (b) Prohlédněte si popisné statistiky a obrázky této proměnné.
 - (c) Posuďte, zda lze rozdělení této proměnné považovat za normální.
 - (d) Otestujte, zda lze na základě daných dat prokázat, že lidé v posilovně uvádějí jinou hmotnost, než skutečně mají.
5. Otestujte, zda muži uvádějí statisticky významně vyšší hmotnost než skutečně mají.
6. Zjistěte, zda ženy uvádějí statisticky významně nižší hmotnost než skutečně mají.
7. Otestujte, zda je statisticky významný rozdíl v rozdílu uváděné a skutečné hmotnosti pro muže a ženy.
8. Otestujte, zda lidé uvádějí statisticky významně větší výšku než skutečně mají.
 - (a) Nejprve proved'te srovnání pro všechny osoby, bez rozlišení pohlaví.
 - (b) Poté proved'te test pro muže a ženy zvlášť.