
CVIČENÍ 14.5.2013

NEPARAMETRICKÉ TESTY Uvažujte data `studenti2013.csv`. Jednotlivé hodnoty jsou odděleny čárkou, chybějící hodnoty jednoduše nejsou vyplněny (není pro ně žádný speciální znak). Význam proměnných z dat již znáte.

1. Načtěte si a prohlédněte si data. Kolik máme pozorování?
2. V rámci prohlížení dat zjistěte následující údaje:
 - Odhadněte procentuální zastoupení jedináčků mezi studenty prvního ročníku.
 - Kolik studentů v souboru má narozeniny v květnu?
 - Kolik let je nejstaršímu studentovi?
 - Vytvořte si graf procentuálního zastoupení mužů a žen.
 - Kolik studentů má velikost bot alespoň 41? Jedná pouze o muže nebo i o ženy?
 - Odhadněte korelační koeficient mezi hmotností a výškou. Nechte si také vykreslit vhodný obrázek pro tuto situaci.
3. Podívejte se, jak je to s normalitou následujících veličin: výška, hmotnost, rozdíl věku rodičů (otec minus matka). Podívejte se vždy na
 - histogram,
 - Q-Q graf,
 - Shapirův - Wilkův test.
4. Otestujte, zda je typická výška studentů rovna 170 cm. Jaký test použijete?
5. Pro srovnání můžeme provést ještě jednovýběrový Wilcoxonův test. V posledním řádku skriptu proto změňte `t.test` na `wilcox.test` a potvrďte. Jak se liší výsledky?
6. Otestujte, zda je střední hmotnost studentů vyšší než 60 kg. Který test zvolíte a proč?
7. Otestujte zda jsou otcové skutečně starší než matky.
Statistics → Nonparametric tests → Paired-samples Wilcoxon test
Podobným způsobem by se provádět i dvouvýběrový Wilcoxonův test.
8. Ověřte, zda jsou otcové o více než 2 roky starší než matky.
9. Otestujte, zda jsou studenti vyšší než studentky.
 - Nejprve zhodnoťte situaci pomocí vhodného obrázku.
 - Ověřte normalitu v obou souborech a dle toho zvolte vhodný test.
 - Proveďte test a interpretujte výsledek.
10. Podobně zjistěte, zda mají studenti typicky větší hmotnost než studentky.

SAMOSTATNÁ PRÁCE

1. Odhadněte procentuální zastoupení jedináčků pro Pražáky a mimopražské zvlášť.
2. Spočítejte 95% interval spolehlivosti pro typickou výšku dívek.
3. Spočítejte aktuální věk studentů. Otestujte, zda je pravda, že typický věk studentů na přednášce je 20 let.
4. Porovnejte testem věk otce a věk matky pouze pro studenty z Prahy.
5. Zjistěte, zda jsou matky chlapců starší než matky dívek.
6. Porovnejte statisticky výšku studentů z Prahy a mimopražských.
7. Dále zjistěte, zda je statisticky významný rozdíl mezi hmotností studentů z Prahy a těch s bydlištěm mimo Prahu.