

Vícerozměrné t-testy, výsledky neřešených příkladů

Zdeňka Geršlová

2023-04-04

Příklad 1 Jednovýběrový Hotellingův t-test

V souboru lrm-foot.txt máme k dispozici antropometrické údaje mladých dospělých lidí (převážně studentů VŠ z Brna a Ostravy). Známe pohlaví (sex), délku chodidla (foot.L) a tělesnou výšku (pbody.H). Oba rozměry jsou měřeny v milimetrech. Zaměřte se pouze na ženy a otestujte hypotézu, že vektor středních hodnot délky chodidla a tělesné výšky

je roven $(\mu_1, \mu_2) = (250, 1679)$ oproti alternativní hypotéze $(\mu_1, \mu_2) \neq (250, 1679)$.

- Nejprve vyberte z datového souboru pouze řádky odpovídající ženám.
- Vypočítejte vektor výběrových průměrů a výběrovou varianční matici pro ženy.

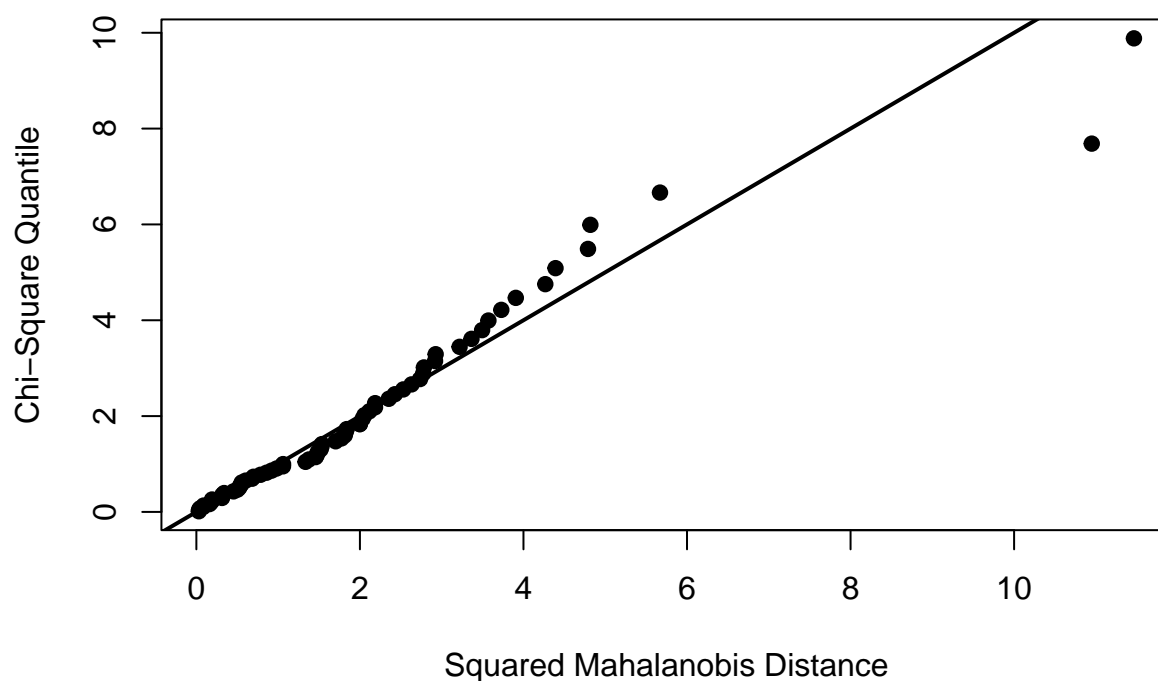
```
##      foot.L      body.H
## 244.3714 1670.7714
```

```
##          foot.L      body.H
## foot.L 130.7586 612.9847
## body.H 612.9847 5079.3093
```

- Na hladině významnosti 0.05 proveďte jednovýběrový Hotellingův test. Nezapomeňte overit předpoklady!

```
##          Test          HZ      p value MVN
## 1 Henze-Zirkler 0.3346412 0.8782532 YES
```

Chi-Square Q-Q Plot



```
##           Test           Statistic           p value Result
## 1 Mardia Skewness  5.10685849124616  0.276507834545936   YES
## 2 Mardia Kurtosis  0.302237733490838  0.76247084127667   YES
## 3           MVN              <NA>              <NA>      YES
```

Pozn: Mardiov test se vyhodnocuje společně, takže i v případě, že dojde k zamítnutí jen u šikmosti nebo jen u špičatosti, zamítá Mardiov test hypotézu, že data pocházejí z vícerozměrného norm. rozdělení. Alternativou je HZ test pomocí knihovny mvnTest a funkce HZ.test() na data:

```
##           Henze-Zirkler test for Multivariate Normality
##
## data : data
##
## HZ           : 0.3286551
## p-value      : 0.8880806
##
## Result  : Data are multivariate normal (sig.level = 0.05)
```

Jednovýběrový Hotellingův test

```
##
## Hotelling's one sample T2-test
##
## data: data
## T.2 = 13.513, df1 = 2, df2 = 68, p-value = 1.145e-05
## alternative hypothesis: true location is not equal to c(250,1679)
```

- V případě zamítnutí nulové hypotézy proveďte jednorozměrné t-testy. Pomocí nich zjistíte, kvůli kterým složkám vektoru došlo k zamítnutí vícerozměrné hypotézy. Nezapomeňte, že je potřeba upravit hladinu významnosti.

Pro výpočet Bonferroniho korigované hladiny významnosti podělíme zvolenou hladinu významnosti (tj. 0.05) počtem testovaných hypotéz (v našem případě 2). Pozor! U jednorozměrných t-testů tedy budeme zamítat hypotézu v případě, že p-hodnota bude menší než korigovaná hladina významnosti (tj. 0.025). Pro jednorozměrné t-testy použijeme funkci `t.test`:

```
## [1] 0.025

##
## One Sample t-test
##
## data: data$foot.L
## t = -4.1182, df = 69, p-value = 0.0001043
## alternative hypothesis: true mean is not equal to 250
## 95 percent confidence interval:
## 241.6449 247.0980
## sample estimates:
## mean of x
## 244.3714

##
## One Sample t-test
##
## data: data$body.H
## t = -0.96599, df = 69, p-value = 0.3374
## alternative hypothesis: true mean is not equal to 1679
## 95 percent confidence interval:
## 1653.778 1687.765
## sample estimates:
## mean of x
## 1670.771
```

Příklad 2 Dvouvěbový Hotellingův t-test

V souboru `head.txt` máme k dispozici antropometrická data mladých dospělých lidí (prevážně studentů z Brna a Ostravy). Známe pohlaví (`sex`), tělesnou výšku (`body.H`), délku hlavy (`head.L`), šířku hlavy (`head.W`) a šířku čelisti (`bigo.W`), rozměry jsou v milimetrech. Chceme otestovat hypotézu, že vektor středních hodnot sledovaných proměnných je stejný pro muže a pro ženy.

- Vypočítejte vektory výběrových průměrů a výběrové varianční matice pro ženy a pro muže.

```
## body.H head.L head.W bigo.W
## 1667.33 185.01 146.92 100.57

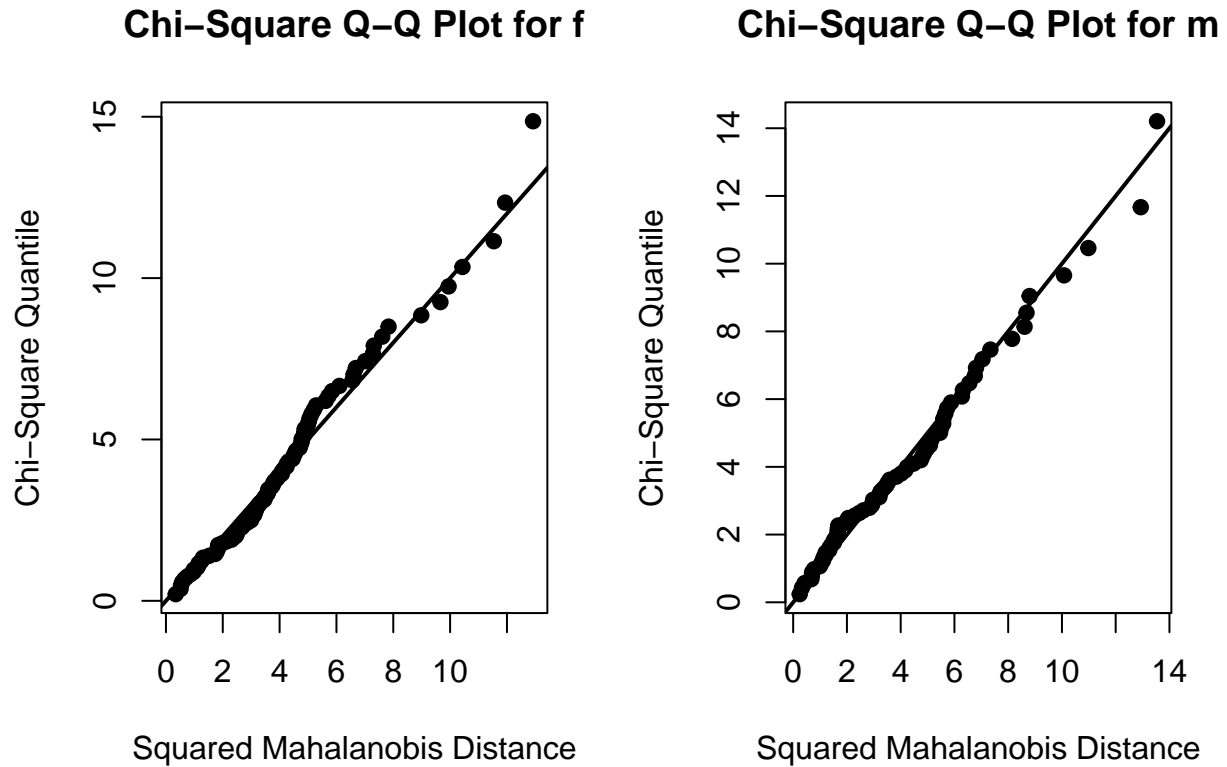
##          body.H    head.L    head.W    bigo.W
## body.H 4516.93040 121.855253 85.976162 42.789798
## head.L 121.85525 42.838283 5.526061 6.943737
## head.W 85.97616 5.526061 28.478384 10.056162
## bigo.W 42.78980 6.943737 10.056162 22.085960

##          body.H    head.L    head.W    bigo.W
## 1789.7200 195.9467 155.6533 107.8133

##          body.H    head.L    head.W    bigo.W
## body.H 3564.85297 123.295676 64.671892 37.717297
```

```
## head.L 123.29568 48.591712 8.008288 7.300721
## head.W 64.67189 8.008288 36.986306 15.880360
## bigo.W 37.71730 7.300721 15.880360 47.234955
```

- Na hladině významnosti 0.05 proveďte dvouvýběrový Hotellingův test. Nezapomeňte ověřit předpoklady!



```
## $f
##           Test           HZ    p value MVN
## 1 Henze-Zirkler 0.8599071 0.2976864 YES
##
## $m
##           Test           HZ    p value MVN
## 1 Henze-Zirkler 0.9063688 0.1357247 YES
##
## Box's M-test for Homogeneity of Covariance Matrices
##
## data: dataS[, 2:5]
## Chi-Sq (approx.) = 17.261, df = 10, p-value = 0.06879
```

Dvourozměrný Hotellingův test:

```
HotellingsT2(dataS[dataS$sex=="f", 2:5], dataS[dataS$sex=="m", 2:5])
```

```
##
## Hotelling's two sample T2-test
##
## data: dataS[dataS$sex == "f", 2:5] and dataS[dataS$sex == "m", 2:5]
## T.2 = 66.309, df1 = 4, df2 = 170, p-value < 2.2e-16
```

```
## alternative hypothesis: true location difference is not equal to c(0,0,0,0)
```

- V případě zamítnutí nulové hypotézy proveďte simultánní testy. Pomocí nich zjistíte, v kterých složkách vektoru se muži a ženy liší. Použijeme vzorec pro simultánní t-testy z přednášky.

```
##          body.H  head.L  head.W  bigo.W
## F          38.3751 27.8000 25.0025 16.8188
## p-hodnota  0.0000  0.0000  0.0000  0.0000
## kvantil    2.4248  2.4248  2.4248  2.4248
```

Poznámka: Výsledné p-hodnoty neuvádíme jako 0, i když tak po zaokrouhlení na určitý počet desetinných míst vycházejí. Je vhodné je uvádět pomocí nerovnosti, např. $p < 0.0001$.

Antropologický závěr: Zamítli jsme shodu vektoru středních hodnot čtyř rozměrů těla mezi pohlavími, všechny čtyři střední hodnoty byly mezipohlavně statisticky významně odlišné i samostatně. Toto je obvyklý stav, který najdeme běžně u zdravé populace. Můžeme pak takto otestované rozměry použít při sledování změn tělesných rozměrů u některých patologických stavů, kdy dochází k ovlivnění pohlavně obvyklé tělesné stavby v důsledku nerovnováh pohlavních hormonů.