

Příklady k prezentaci na 9. cvičení

Budete-li chtít prezentovat i příkazy v R-studiu, přineste si prosím příslušný kód na flashce, nebo mi ho s předstihem pošlete emailem. Pokud byste v příkladu narazili na nějakou nejasnost, neváhejte se mi ozvat.

Dataset `Heart_rate.csv` obsahuje údaje o 500 jedincích. Během šestitýdenního tréninkového programu u nich byly zkoumány následující veličiny:

- Age = věk
- SleepHours = průměrná doba spánku jedince během studie
- ExerciseFrequency = počet dní v týdnu, kdy jedinec cvičí
- HeartRateBefore = klidová tepová frekvence před začátkem šestitýdenního tréninkového programu (průměr z deseti měření)
- HeartRateAfter = klidová tepová frekvence po skončení šestitýdenního tréninkového programu (průměr z deseti měření)
- MaxHeartRate = maximální tepová frekvence jedince během tréninkového programu

Tabulku si do R můžete načísti příkazem

```
Heart <- read.csv2("Heart_rate.csv", header=TRUE, dec=",")
```

Příklad 1

- Vytvořte veličinu udávající rozdíl klidové tepové frekvence před a po absolvování tréninkového programu. Vypočítejte pro ni základní popisné statistiky a vykreslete krabicový diagram.
- Pomocí vhodných obrázků prozkoumejte, zda lze u této veličiny předpokládat, že se řídí normálním rozdělením.
- Statistický test hypotézy:

H_0 : náš výběr pochází z normálního rozdělení, proti alternativě

H_1 : náš výběr nepochází z normálního rozdělení

se jmenuje Shapiro-Wilkův test a v R ho provedete příkazem `shapiro.test(výběr)`. Uvažujte hladinu $\alpha = 5\%$. Na základě výsledné p-hodnoty rozhodněte, zda lze předpokládat, že data pocházejí z normálního rozdělení. (Pravidlo pro zamítnutí či nezamítnutí H_0 na základě p-hodnoty je stejné jako u t-testu).

Příklad 2

Pomocí vhodného testu rozhodněte, zda tréninkový program pomohl jedincům snížit jejich klidovou tepovou frekvenci. Tj. zajímá nás, zda klidová tepová frekvence po absolvování tréninkového kurzu je ve střední hodnotě statisticky významně nižší než klidová tepová frekvence před kurzem. Uvažujte hladinu $\alpha = 5\%$.

- (a) Vytvořte veličinu udávající rozdíl klidové tepové frekvence před a po absolvování tréninkového programu. Tuto veličinu si označíme jako Z .
- (b) Formulujte nulovou a alternativní hypotézu (za pomoci veličiny Z).
- (c) Platnost nulové hypotézy ověřte pomocí jednovýběrového t-testu. (Ověření předpokladu normality nemusíte provádět, neboť to je obsahem Příkladu 1.) Formulujte svůj závěr.

Jednovýběrový t-test aplikovaný na rozdíly dvou veličin naměřených na témže jedinci se nazývá **párový t-test**.

Příklad 3

Na základě datové tabulky **Heart_rate** ověřte, zda průměrná doba spánku v populaci je rovna 7,5 hodiny, nebo je nižší. Uvažujte hladinu $\alpha = 5\%$.

- (a) Zapište nulovou a alternativní hypotézu.
- (b) Ověřte předpoklady jednovýběrového t-testu.
- (c) Proved'te test zadané hypotézy pomocí jednovýběrového t-testu a svůj závěr zformulujte bez použití výrazů „zamítáme/nezamítáme H_0 “.
- (d) Museli bychom výpočet provést znovu, kdybychom chtěli rozhodnout na hladině $\alpha = 10\%$ místo původních 5%?