Dvouvýběrové testy

Komentované řešení pomocí programu R

Ústav matematiky
Fakulta chemicko inženýrská
Vysoká škola chemicko-technologická v Praze

Popis vstupních dat

Vstupní data pro úlohu **(A)** se nacházejí v souboru

"glukoza.csv".

V prvním sloupci "pred" je hladina glukózy v krvi (v $mmol/l$) před užíváním léku a v druhém sloupci "po" hodnoty hladiny glukózy po dvou týdnech užívání testovaného léku.
Vstupní data pro úlohu **(B)** se nacházejí v souborech

"skupinaI.csv", "skupinaII.csv".

SkupinaI.csv obsahuje údaje o první, 46 členné skupině, jíž byl v první fázi dvojitého slepého pokusu podáván nově testovaný lék (první sloupec) a ve fázi druhé lék tradiční (druhý sloupec). Pro 49 člennou 2. skupinu je tomu přesně naopak.

Načtení vstupních dat

Předpokládejme, že data jsou uložena na disku F ve složce

Aplikovana_statistika.

Načtení vstupních dat do pracovního objektu **glukoza** a vypsání na obrazovku provedeme příkazem

```
glukoza<-read.table("f:\\Aplikovana_statistika\\glukoza.txt", header=TRUE)
glukoza
```

```
   pred  po
1   6.7 5.6
2   5.7 6.1
3   6.5 5.5
4   6.3 5.5
5   6.6 5.4
6   6.3 5.5
7   5.9 5.7
8   6.2 5.8
9   6.4 5.2
...
```

Souhrnné charakteristiky souboru

Podívejme se nejprve na některé souhrnné charakteristiky našich dat pomocí příkazu summary. Ten nám vypíše minimum, maximum, všechny kvartily a medián

```
summary(glukoza)
      pred             po
 Min.   :5.10   Min.   :4.900
 1st Qu.:5.90   1st Qu.:5.500
 Median :6.10   Median :5.650
 Mean   :6.15   Mean   :5.673
 3rd Qu.:6.50   3rd Qu.:5.900
 Max.   :6.90   Max.   :6.500
```

Součástí příkazu summary nejsou výběrové rozptyly, takže si necháme vypsat zvlášť kovarianční matici příkazem var

```
var(glukoza)

            pred          po
pred  0.13745763 -0.01745763
po   -0.01745763  0.12436158
```

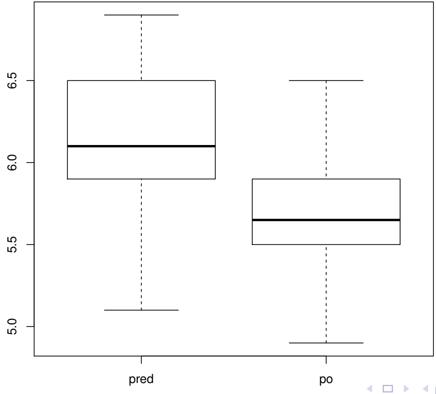
Předpokladem pro párový test je, že výběr pochází z dvourozměrného normálního rozdělení. Práci s tímto typem rozdělení jsme se na přednáškách nevěnovali. Tento předpoklad může být nahrazen předpokladem, že rozdíly mezi měřeními před a po aplikaci léků u jendotlivých pacientů má normální rozdělení. Definujme si tedy rozdíl mezi daty před a po a pro tento rozdíl budeme zkoumat normalitu.

```
glukoza_rozdil<-glukoza$pred-glukoza$po)
glukoza_rozdil
 [1]  1.1 -0.4  1.0  0.8  1.2  0.8  0.2  0.4  1.2  0.8  0.6 -0.5  0.2  0.1  0.3
[16]  0.7  1.2 -0.1  0.5  0.6 -0.2  0.4  0.4 -0.1  1.1  1.0  0.5  1.2  0.2  0.1
[31]  0.2  0.8  0.3  0.0  0.0  0.0  0.2  0.7  0.4  1.5  0.3  0.2  1.1 -0.5  0.4
[46] -0.2 -0.6  1.3  0.3  1.0  0.7  0.9  0.9 -0.9  0.4  0.5  1.6  0.3  1.2  0.3
```

Grafické výstupy - Boxplot

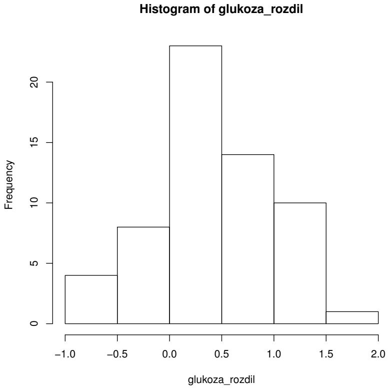
Boxplot si můžeme vykreslit pro data před i po zvlášť, neboť nám slouží k sumárním statistikám nikoliv testovánńormality.

```
boxplot(glukoza)
```



Grafické výstupy - Histogramy

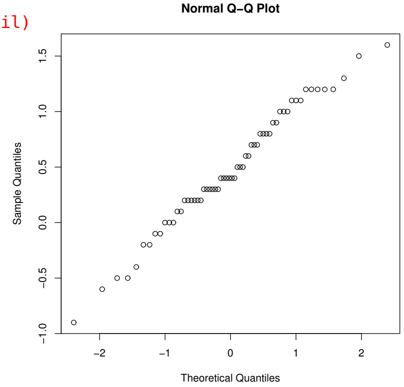
```
hist(glukoza_rozdil)
```



Jsou-li data normální, kopírují vrcholky sloupců přibližne Gaussovu křivku.

Grafické výstupy - Q-Q plot

```
qqnorm(glukoza_rozdil)
```



Normalita dat by znamenala, že body grafu by se měly ležet v přímce, což se zdá být přibližně splněno.

Test normality dat

```
shapiro.test(glukoza_rozdil)
        Shapiro-Wilk normality test

data:  glukoza_rozdil
W = 0.98372, p-value = 0.6039
```

P-hodnota testu je, jak vidíme poměrně vysoká. Hypotézu o "nenormalitě" dat bychom zamítli až na hladině přes 60 %. To svědčí ve prospěch normality rozdílu, což ovšem podporuje i vzhled Q-Q grafu.

Párový test

Protože bychom rádi prokázali, že lék má průkazné pozitivní vliv na hladinu glukózy v krvi, budeme testovat oproti jednostranné alternativě, že střední hodnota rozdílu "pred-po" je vetší než nula. To můžeme udělat příkazem

```
t.test(glukoza$pred,glukoza$po,alternative = "greater",paired = TRUE,
                   var.equal = FALSE,conf.level = 0.95)
```

```
data:  glukoza$pred and glukoza$po
t = 6.7781, df = 59, p-value = 3.209e-09
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.3591474       Inf
sample estimates:
mean of the differences
              0.4766667
```

Interpretace výsledku testu

Ve výpisu vidíme, že testová statistika má hodnotu $t = 6,7781$, rozdíl výběrových průměrů je 0,4766667 a jednostranný 95 % interval spolehlivosti je $[0,3591474; \infty)$. Protože interval spolehlivosti neobsahuje nulu, můžeme na hladině 5 % zamítnout hypotézu o shodě středních hodnot. Na hladině 5 % jsme tedy statisticky význmně prokázali účinost léku na snížení hladiny glukózy. Podstatné pro nás je, že p-hodnota testu je na úrovni miliardtin, tedy velmi blízká nule, takže bychom zamítali i na velmi velmi nízkých hladinách.

Pro úlohu (B) budeme muset nejprve načíst nová data. Předtím (pokud pokračujeme ve stejné konzoli) si pro jistotu vymažeme všechny stávající objekty.

```
rm(list=ls(all=TRUE))
```

Teď si můžeme načíst nová data:

```
skupinaI<-read.table("F:\\Aplikovana_statistika\\dvouvyberove_testy\\
skupinaI.txt", header=TRUE)
skupinaII<-read.table("F:\\Aplikovana_statistika\\dvouvyberove_testy\\
skupinaII.txt", header=TRUE)
```

Pro přehlednost si pojmenujeme jednotlivé sloupce:

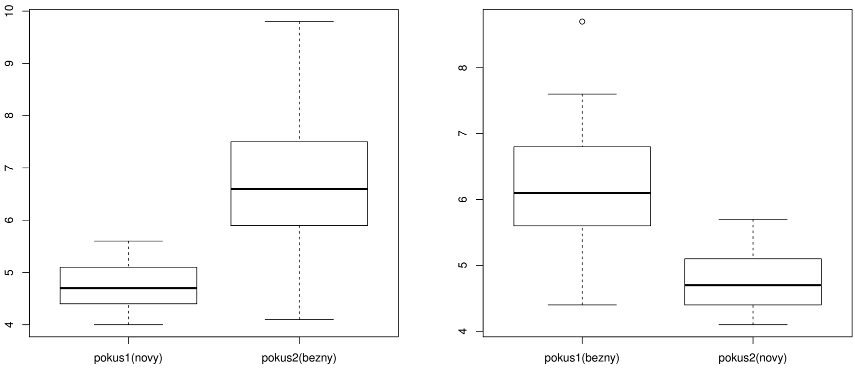
```
names(skupinaI)=c("pokus1novy","pokus2bezny")
names(skupinaII)=c("pokus1bezny","pokus2novy")
```

Sloupec "novy" odpovídá novému léku, sloupec "bezny", běžně používanému.

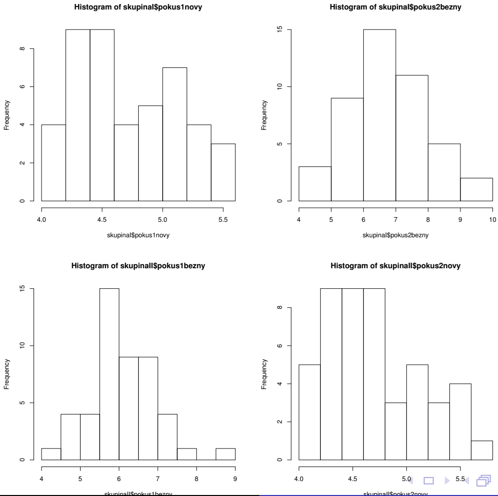
Základním předpokladem pro použítí dvouvýběrového testu je opět normalita dat. Provedeme proto obdobné grafické výstupy jako v předchozím párovém testu, tj.

```
boxplot(skupinaI)
boxplot(skupinaII)

hist(skupinaI$pokus1novy)
hist(skupinaI$pokus2bezny)

hist(skupinaII$pokus1bezny)
hist(skupinaII$pokus2novy)

qqnorm(skupinaI$pokus1novy)
qqnorm(skupinaI$pokus2bezny)

qqnorm(skupinaII$pokus1bezny)
qqnorm(skupinaII$pokus2novy)
```

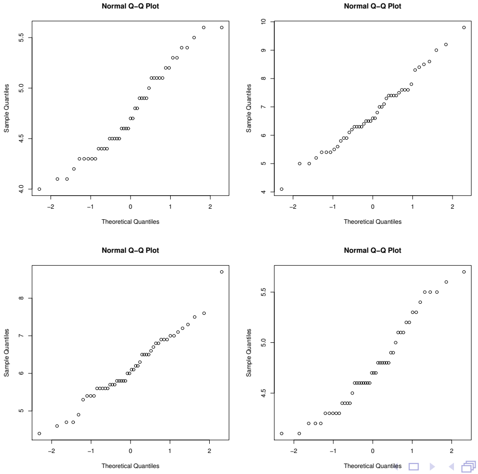
Boxploty


4
5
6
7
8
9
10
pokus1(novy)
pokus2(bezny)
4
5
6
7
8
pokus1(bezny)
pokus2(novy)

Histogramy


Histogram of skupinaI$pokus1novy
Frequency
skupinaI$pokus1novy
Histogram of skupinaI$pokus2bezny
Frequency
skupinaI$pokus2bezny
Histogram of skupinaII$pokus1bezny
Frequency
Histogram of skupinaII$pokus2novy
Frequency

Q-Q grafy


Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles
Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles
Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles
Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles

Test normality-skupina I

```
shapiro.test(skupinaI$pokus1novy)

        Shapiro-Wilk normality test

data:  skupinaI$pokus1novy
W = 0.95921, p-value = 0.114

shapiro.test(skupinaI$pokus2bezny)

        Shapiro-Wilk normality test

data:  skupinaI$pokus2bezny
W = 0.98431, p-value = 0.7935
```

U obou pokusů je p-hodnota vyšší než všechny standardně používané hladiny testu, u druhého pokusu je dokonce výrazně vysoká. Normalitu dat tedy nemůžeme zamítnout. Přesto test o shodě středních hodnot budeme považovat za asymptotický.

Test normality-skupina II

```
shapiro.test(skupinaII$pokus1bezny)

        Shapiro-Wilk normality test

data:  skupinaII$pokus1bezny
W = 0.97705, p-value = 0.4627

shapiro.test(skupinaII$pokus2novy)

        Shapiro-Wilk normality test

data:  skupinaII$pokus2novy
W = 0.94871, p-value = 0.03552
```

U prvního pokusu je p-hodnota vyšší než všechny standardně používané hladiny testu, u druhého pokusu je bohužel pouze lehce vyšší než 3 %, tedy na hladině 5 % bychom už hypotézu o normalitě dat zamítli. Test o shodě středních hodnot budeme budeme tedy opět považovat za asymptotický.

Chceme-li u prvního pokusu prokázat lepší působení nového léku, budeme testovat oproti jednostranné alternativě, že očekávaná hladina je při užívání nového léku nižší, v tomto případě to znamená "menší než":

```
t.test(skupinaI$pokus1novy, skupinaII$pokus1bezny,alternative = "less",
                    paired = FALSE, var.equal = FALSE,conf.level = 0.95)

        Welch Two Sample t-test

data:  skupinaI$pokus1novy and skupinaII$pokus1bezny
t = -9.8607, df = 70.129, p-value = 3.501e-15
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
    -Inf -1.1518
sample estimates:
mean of x mean of y
 4.755556  6.141667
```

Jak vidíme, p-hodnota je velmi nízká, protomůžeme na všech myslitelných hadinách zamítnout shodu středních hodnot. Lze tedy spolehlivě tvrdit (na základě Pokusu 1), že nový lék je prokazatelně účinnější.

Analogicky u druhého pokusu prokázat lepší působení nového léku, budeme testovat oproti jednostranné alternativě "větší než":

```
t.test(skupinaI$pokus2bezny, skupinaII$pokus2novy,alternative = "greater",
                    paired = FALSE, var.equal = FALSE,conf.level = 0.95)
        Welch Two Sample t-test
data:  skupinaI$pokus2bezny and skupinaII$pokus2novy
t = 10.518, df = 54.115, p-value = 5.403e-15
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 1.707484      Inf
sample estimates:
mean of x mean of y
 6.788889  4.758333
```

Jak vidíme, p-hodnota je velmi nízká, protomúžeme na všech myslitelných hadinách zamítnout shodu středních hodnot. Lze tedy spolehlivě tvrdit (na základě Pokusu 2), že nový lék je prokazatelně účinnější. Tedy celý slepý dvojitý pokus proběhl úspěšně a zavedení novéo léku můžeme spolehlivě doporučit.

Jako dodatek k dvouvýběrovým testům jenom dodáváme návod, jak provést test shody rozptylů. Jak již bylo řečeno v teoretické části, pokud z principu nevíme, že výběry pochází z rozdělení se shodnými rozptyly, **nebudeme** užívat test shody středních hodnot při shodných rozptylech. Tato kapitola je tedy pouze dodatková, abychom věděli, jak v případě potřeby testovat v **R** shodu rozptylů.

Test si procvičíme například na datech pro nový lék získaných v obou skupinách.

```
var.test(skupinaI$pokus1novy, skupinaII$pokus2novy, ratio=1,
                           alternative = "two.sided",conf.level = 0.95)
```

```
        F test to compare two variances
data:  skupinaI$pokus1novy and skupinaII$pokus2novy
F = 1.0088, num df = 44, denom df = 47, p-value = 0.9739
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.5615002 1.8229735
sample estimates:
ratio of variances
          1.008816
```

Jak vidíme, 95 % interval spolehlivosti obsahuje jedničku, proto nemůžeme hypotézu o shodě rozptylu nemůžeme zamítnout. Navíc p-hodnota je velmi blízká jedné, z čehož lze na shodu usuzovat.

Poznámka: Argumentem ratio můžeme změnit typ testovaného podílu (vztahu) mezi rozptyly.