9 Analýza rozptylu jednoduchého třídění

Příklad 9.1. V jisté továrně se měřil čas, který potřeboval každý ze tří dělníků k uskutečnění téhož pracovního úkonu. Čas v minutách:

1.dělník:	3.6	3.8	3.7	3.5		
2.dělník:	4.3	3.9	4.2	3.9	4.4	4.7
3.dělník:	4.2	4.5	4.0	4.1	4.5	4.4

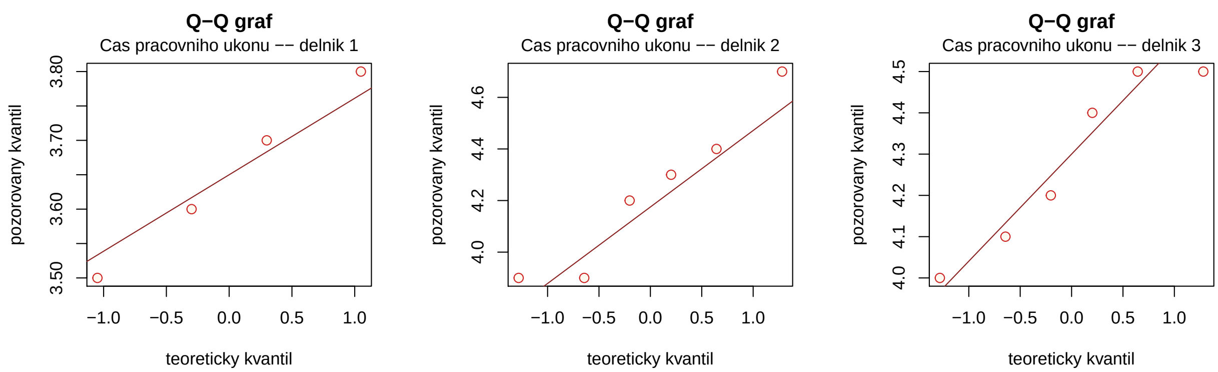
Na hladině významnosti $\alpha = 0.05$ testujte hypotézu, že výkony těchto tří dělníků jsou stejné. Zamítnete-li nulovou hypotézu, určete, výkony kterých dělníků se liší na dané hladině významnosti $\alpha = 0.05$.

Testování normality

1. H_0 : ....
2. H_1 : ....

Na testování normality všech tří výběrů použijeme kvůli jejich malým rozsahům .... test.

```
## [1] "Delnik 1: 0.9719"
## [1] "Delnik 2: 0.5819"
## [1] "Delnik 3: 0.3313"
```



Protože ve všech třech případech je p-hodnota testu .... než α = ..., nulovou hypotézu o normalitě časů všech tří dělníků .... na hladině významnosti .... Všechny tři výběry tedy .... z normálního rozdělení.

Test homogenity rozptylů

1. H_0 : ....
2. H_1 : ....

Jelikož náhodné výběry pochází z normálního rozdělení, na testování hypotézy o shodě rozptylů všech tří výběrů použijeme .... test.

```
## Test Statistic
##       1.514205
## [1] 0.2563563
```

Testovací statistika .... testu nabývá hodnoty ...., odpovídající p-hodnota = ... je .... než α = ..., tedy na hladině významnosti .... hypotézu o shodě rozptylů σ_1^2, σ_2^2 a σ_3^2 ....

Test o shodě středních hodnot:

1. H_0 : ....
2. H_1 : ....

```
##              Df Sum Sq Mean Sq F value  Pr(>F)
## factor(ID)   2 1.1177  0.5589   9.665 0.00268 **
## Residuals   13 0.7517  0.0578
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## [1] 3.805565
```

Skupinový součet čtverců S_A = ..., počet stupňů volnosti f_A = ..., reziduální součet čtverců S_E = ..., počet stupňů volnosti f_E = ..., testovací statistika F_A = ... p-hodnota = ..., nulovou hypotézu o shodě středních hodnot tedy .... na hladině významnosti ....

Kritický obor má tvar W=.. Protože ...., hypotézu H_0 o shodě středních hodnot .... na hladině významnosti ....

Metoda mnohonásobného porovnávání

Jelikož jsme nulovou hypotézu o shodě středních hodnot ...., chceme nyní zjistit, které dvojice středních hodnot se od sebe významně liší. Stanovíme nulové a alternativní hypotézy pro dvojice středních hodnot

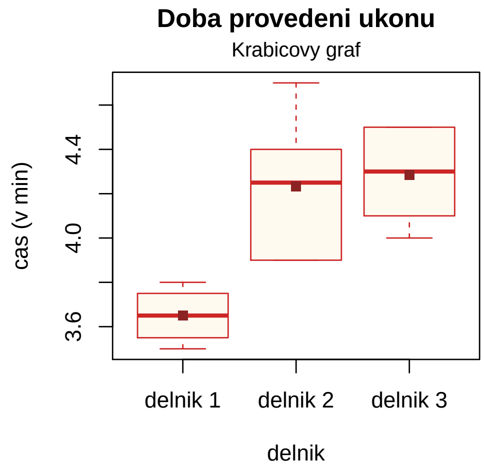
- H_{01} : .... oproti H_{11} : ....
- H_{02} : .... oproti H_{12} : ....
- H_{03} : .... oproti H_{13} : ....

Protože v každé skupině máme různý počet pozorování, použujeme na mnohonásobné porovnávání .... metodu.

```
## $L
##           delnik 1  delnik 2  delnik 3
## delnik 1 0.0000000 0.5833333 0.6333333
## delnik 2 0.5833333 0.0000000 0.0500000
## delnik 3 0.6333333 0.0500000 0.0000000
##
## $R
##           delnik 1  delnik 2  delnik 3
## delnik 1 0.4690839 0.4282131 0.4282131
## delnik 2 0.4282131 0.3830054 0.3830054
## delnik 3 0.4282131 0.3830054 0.3830054
```

Porovnáním pravé a levé strany .... metody vidíme, že na hladině významnosti .... zamítáme nulovou hypotézu o shodě středních hodnot $\mu_{........}$ a $\mu_{........}$ a středních hodnot $\mu_{........}$ a $\mu_{........}$. Výsledek testování tedy ukazuje, že na hladině významnosti .... se liší výkony dělníků .... a ...., naopak výkony dělníků .... se neliší.

Krabicový graf



Příklad 9.2. Z archivních materiálů (Schmidt, 1888) máme k dispozici původní kraniometrické údaje o výšce horní části tváře mužů z pěti populací: bantuské (13 jedinců), čínské (18 jedinců), malajské (69 jedinců), německé (19 jedinců) a peruánské (44 jedinců). Na hladině významnosti $\alpha = 0.05$ otestujte nulovou hypotézu H_0, že výška horní části tváře je stejná pro všechny populace. V případě zamítnutí nulové hypotézy zjistěte, které populace se významně liší ve výšce horní části tváře.

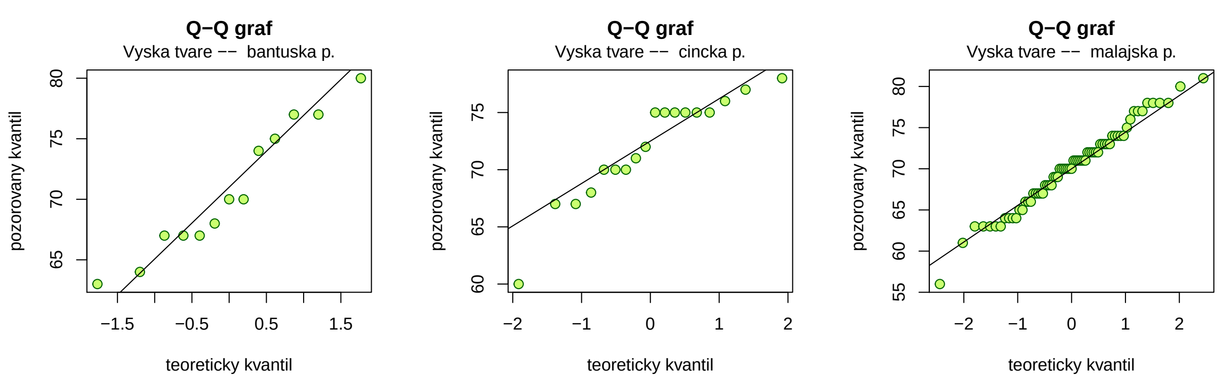
Testování normality

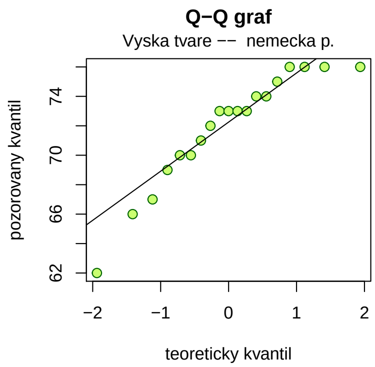
1. H_0 : ....

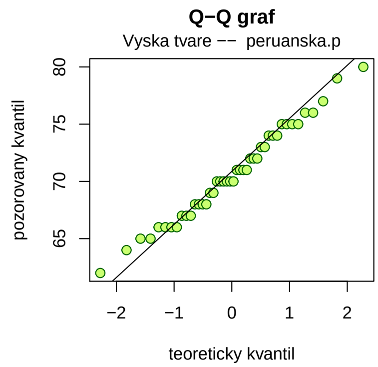
2. H_1 : ....

Rozsah náhodných výběrů pro bantuskou, čínskou a německou populaci je .... 30, proto normalitu otestujeme .... testem. Rozsah náhodného výběru pro malajskou a peruánskou populaci je .... 30, proto normalitu otestujeme .... testem.

```
## [1] "Bantuska p.:  0.4321"
## [1] "Cinska p.:    0.0513"
## [1] "Malajska p.:  0.2703"
## [1] "Nemecka p.:   0.0419"
## [1] "Peruanska p.: 0.6447"
```







V případě bantuské, čínské, malajské a peruánské populace je p-hodnota testu normality .... než $\alpha =$ ...., nulovou hypotézu o normalitě tedy .... na hladině významnosti .... V případě německé populace je p-hodnota .... než α. Porušení normality je však mírné, proto i zde předpokládáme, že data pochází z normálního rozdělení.

Test homogenity rozptylů

1. H_0 : ....
2. H_1 : ....

Jelikož náhodné výběry pochází z normálního rozdělení, na testování hypotézy o shodě rozptylů všech pěti populací použijeme .... test.

```
## Test Statistic
##       0.966642
## [1] 0.4275202
```

Testovací statistika .... testu nabývá hodnoty ...., odpovídající p-hodnota = ... je .... než $\alpha =$ ...., tedy na hladině významnosti .... hypotézu o shodě rozptylů u všech pěti populací ....

Test o shodě středních hodnot:

1. H_0 : ....
2. H_1 : ....

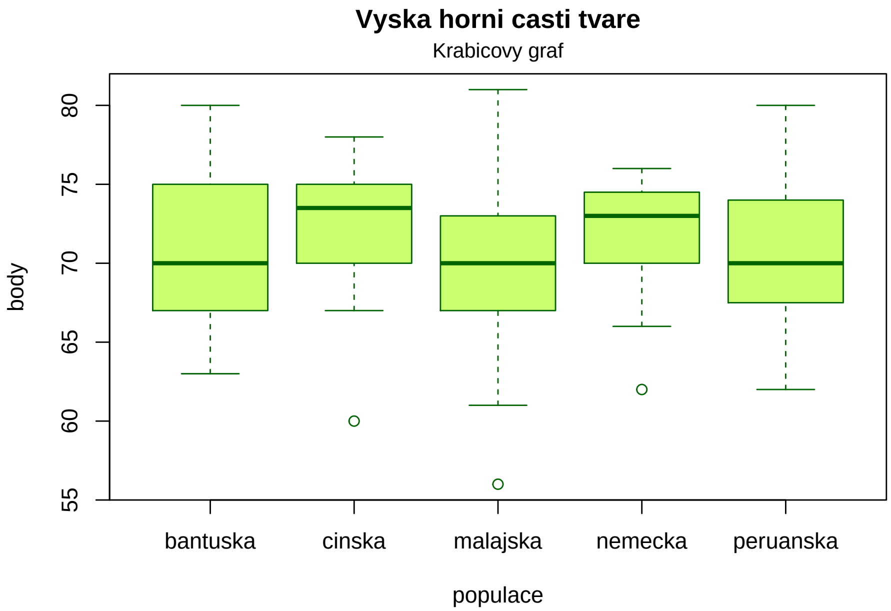
```
##              Df Sum Sq Mean Sq F value Pr(>F)
## factor(ID)    4     80   20.05   0.943  0.441
## Residuals   158   3361   21.27
## [1] 1.278004
```

Skupinový součet čtverců S_A = ..., počet stupňů volnosti f_A = ..., reziduální součet čtverců S_E = ..., počet stupňů volnosti f_E = ..., testovací statistika F_A = ... p-hodnota = ..., nulovou hypotézu o shodě středních hodnot tedy .... na hladině významnosti ....

Kritický obor má tvar W=.. Protože ...., hypotézu H_0 o shodě středních hodnot .... na hladině významnosti ....

V délce horní části tváře tedy .... významný rozdíl.

Krabicový graf



Příklad 9.3. Je dána neúplná tabulka analýzy rozptylu jednoduchého třídění. Na volná místa doplňte chybějící hodnoty a na hladině významnosti 0.05 testujte hypotézu o shodě středních hodnot. Stanovte, jaký je celkový počet pozorování n a kolik úrovní r má faktor A?

zdroj variability	součet čtverců	stupně volnosti	podíl	statistika F
skupinový		2		
reziduální	172			-
celkový	326	17	-	-

```
## [1] 154
## [1] 15
## [1] 77
## [1] 11.467
## [1] 6.715
## [1] 3.68232
## [1] 0.008267191
```

Testovací statistika F nabývá hodnoty .... Kritický obor má tvar W=...................................... .
Protože ...., hypotézu H_0 o shodě středních hodnot .... na hladině významnosti ....

Testovací statistika F nabývá hodnoty ...., p-hodnota = ... Protože p-hodnota= .... je .... než α = ..., hypotézu H_0 o shodě středních hodnot .... na hladině významnosti ....
Celkový počet pozorování je ....; faktor A má celkem .... úrovní.

Příklad 9.4. V rámci studie byly získány čtyři nezávislé náhodné výběry o rozsazích 15, 11, 6, 6 přičemž i-tý výběr pochází z rozdělení $N(\mu_i, \sigma_i^2)$, $i = 1, 2, 3, 4$. Byl vypočten celkový součet čtverců $S_T = 7606$ a reziduální součet čtverců $S_E = 3881$. Na hladině významnosti $\alpha = 0.05$ testujte hypotézu H_0 o shodě středních hodnot. Testování proveďte pomocí kriického oboru a pomocí p-hodnoty.

```
## [1] 10.87778
## [1] 2.882604
## [1] 3.685735e-05
```

zdroj variability	součet čtverců	stupně volnosti	podíl	statistika F
skupinový				
reziduální				-
celkový			-	-

Testovací statistika F nabývá hodnoty .... Kritický obor má tvar W=...................................... . Protože ...., hypotézu H_0 o shodě středních hodnot .... na hladině významnosti ....

Testovací statistika F nabývá hodnoty ...., p-hodnota = ... Protože p-hodnota= .... je .... než $\alpha =$ ...., hypotézu H_0 o shodě středních hodnot .... na hladině významnosti ....