NEX1

Kajetán Poliak, Adam Novotný

18 října 2019

NEX - první domácí úkol

1. část

- Namerte data: experiment nastavte tak, aby byl znahodneny !!!!! Poradi mereni si poznamenejte pro dalsi kontrolu pripadne zavislosti na poradi mereni. Jednotlive lidi ve skupine berte jako ruzne operatory experimentu (blokujte). Pocet replikaci u jednoho cloveka a jednoho casu vemte 1. V protokolu krom popisu experimentu diskutujte i jen promenne, ktere mohli mit na mereni vliv.

Experimetu se účastnili dva operátoři. Každý z operátorů se podrobil devíti měření (tečkování do 3 párů kruhů pomocí dominantní, nedominantní a obou rukou), kde každé meření trvalo 10s. Pořadí měření bylo znáhodněno pomocí generátoru náhodné posloupnosti z www.random.org. Jako první se vygenerovala náhodné posloupnost celých čísel od 1 do 9, která určovala pořadí meření. Následně se pomocí náhodné posloupnosti od 1 do 3 přiřadila jedna z úrovní proměnné ruka ke každému ze tří papírů obsahujícímu tři různé velikosti kruhu. Na meření mohlo mít vliv jak rozpoložení operátora (různé polohy sedu při experimentu, psychický nátlak okolí,..), tak nepřesnosti meření času nebo různé psací potřeby operátorů.

2. část

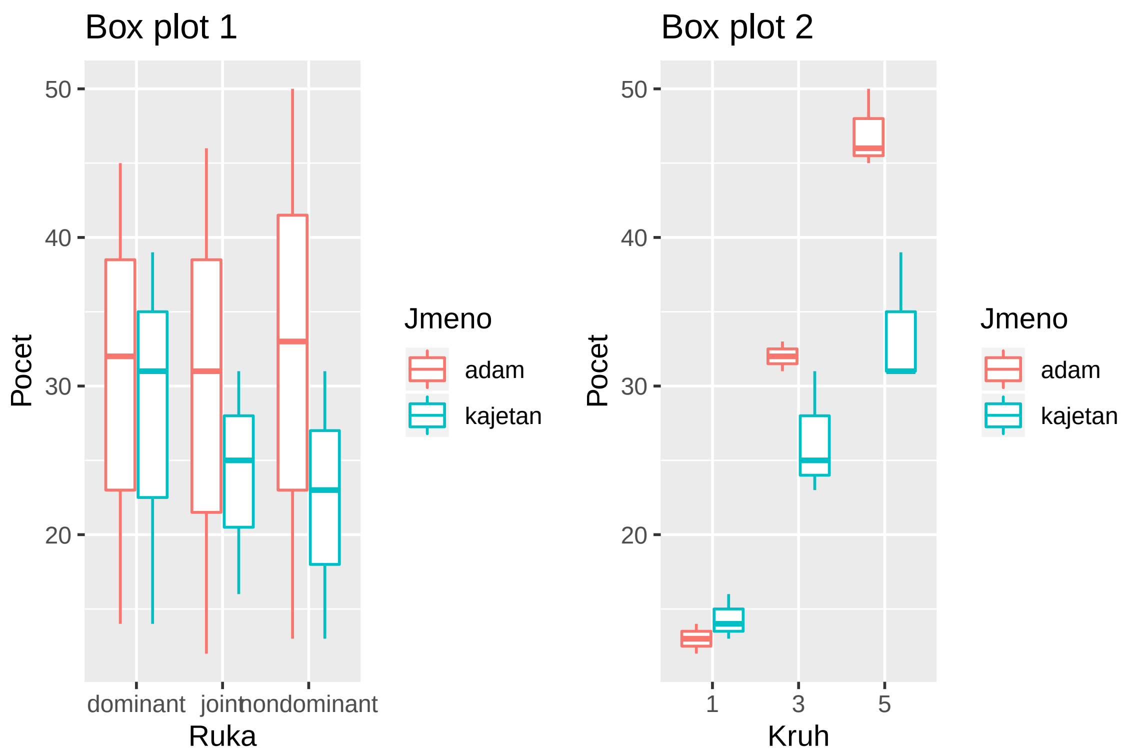
- Spoctete zakladni statistiky (mean, median a sd pro jednotlive faktory - velikost, ruka, operator) Zobrazte namerena data (box plot, interaction plot, effects plot, . . .) a okomentujte je co z danych obrazku muzeme pred samotnou analyzou rici o vysledku?

```
# ZÁKLADÍ STATISTIKY
characteristics
```

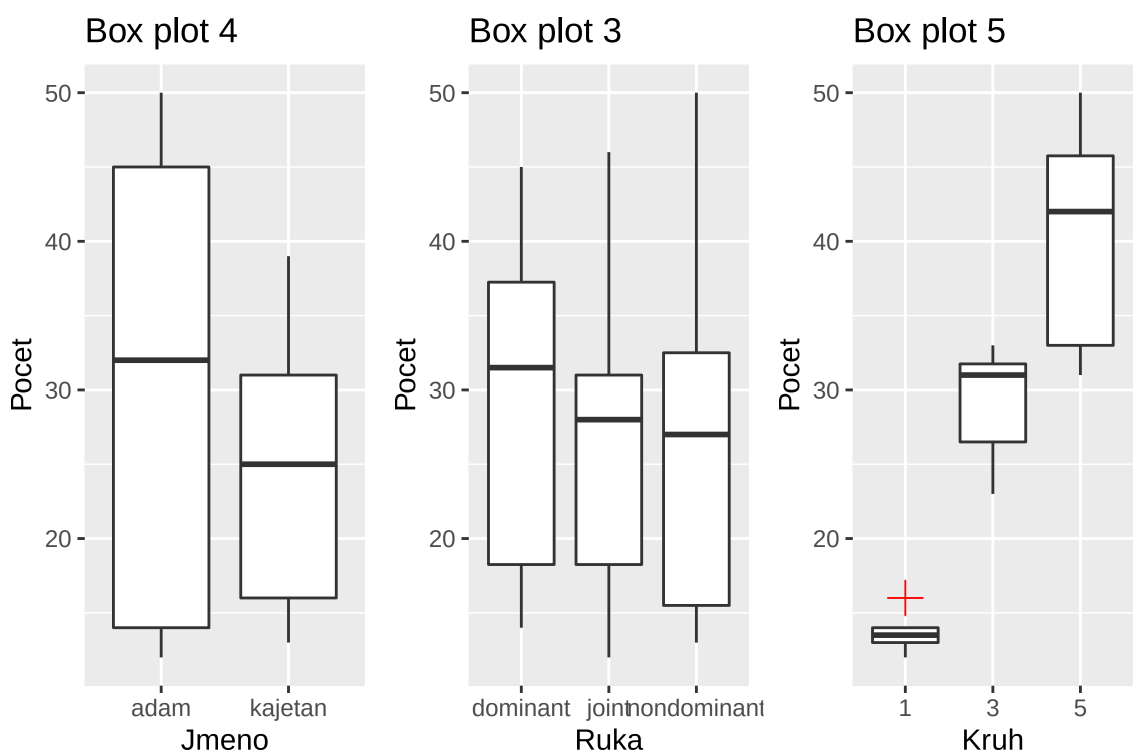
```
##             adam   kajetan dominant nondominant    joint     1 cm
## Min.    12.00000 13.000000 14.00000    13.00000 12.00000 12.00000
## 1st Qu. 14.00000 16.000000 18.25000    15.50000 18.25000 13.00000
## Median  32.00000 25.000000 31.50000    27.00000 28.00000 13.50000
## Mean    30.66667 24.777778 29.16667    27.16667 26.83333 13.66667
## 3rd Qu. 45.00000 31.000000 37.25000    32.50000 31.00000 14.00000
## Max.    50.00000 39.000000 45.00000    50.00000 46.00000 16.00000
## Sd.     14.83240  9.038498 12.79714    14.06295 12.18879  1.36626
##              3 cm      5 cm
## Min.    23.000000 31.000000
## 1st Qu. 26.500000 33.000000
## Median  31.000000 42.000000
## Mean    29.166667 40.333333
## 3rd Qu. 31.750000 45.750000
## Max.    33.000000 50.000000
## Sd.      4.119061  8.041559
```

Tabulka zobrazuje základní statistiky experimentu. Vidíme, že operátoři mají podobná minima ale maximum Adama je o celých 11 vyšší. Stejně tak je vyšší střední hodnota počtu teček Adama. Maxima u proměnných dominant, nondominant a joint tedy souvisí s operátorem Adam. Podle očekávání vidíme, že hodnoty maxim i minim jsou vyšší se zvětšujícím se poloměrem kruhu. Směrodatná odchylka u proměnných adam, kajetán, dominant, nondominant a joint dosahují poměrně vysokých hodnot, to je vysvětlitlé tím, že počet teček nabýval pro každý kruh se radikálně lišil. Nejnižší střední hodnotu z různých úrovní proměnné ruka má joint.

```r
#BOX-PLOTS
grid.arrange(p1,p2,ncol=2)
```

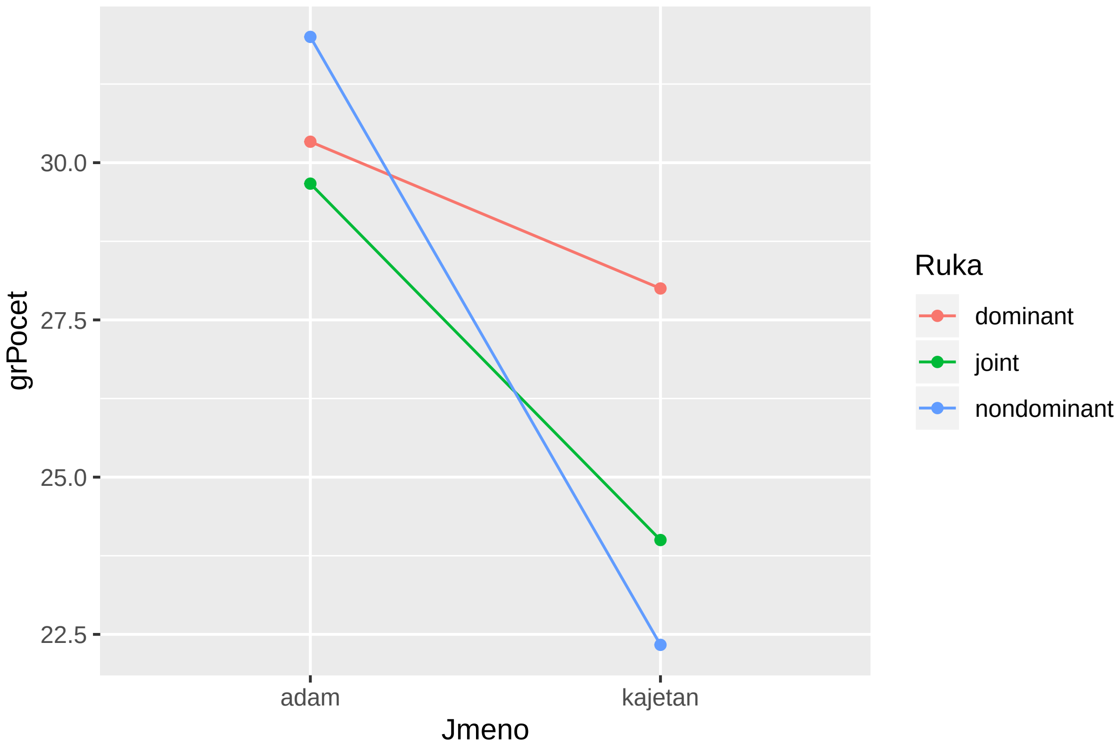



```r
grid.arrange(p4,p3,p5,ncol=3)
```

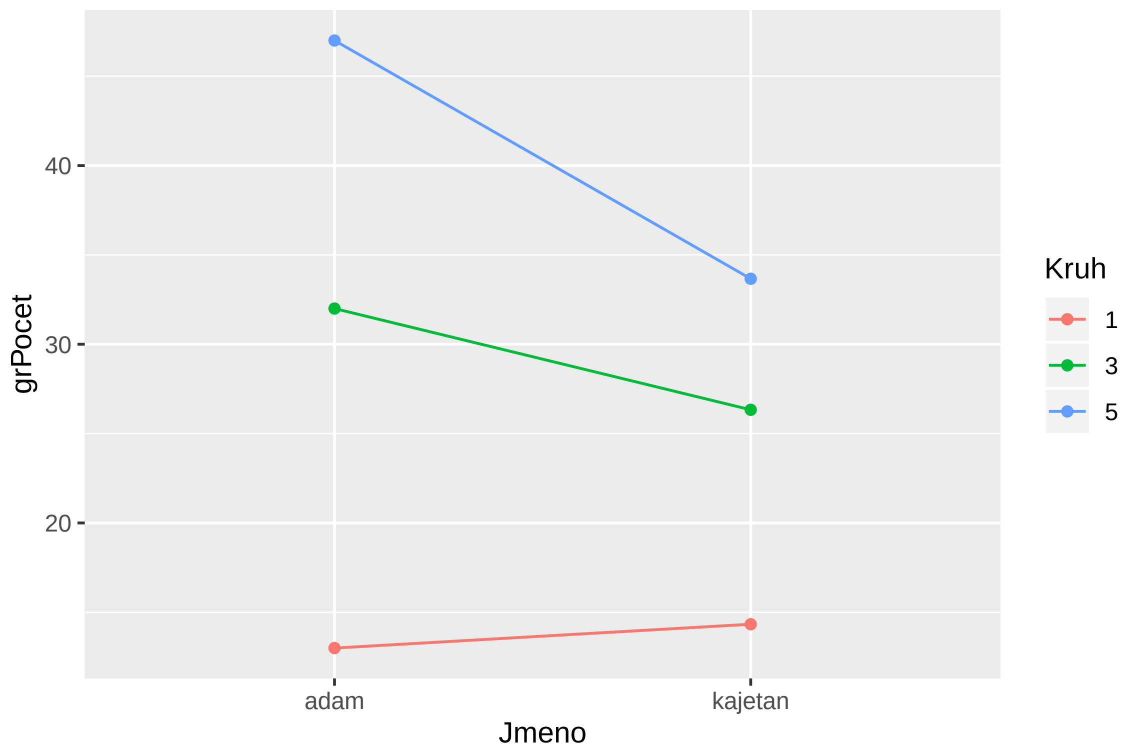


Z boxplotu 1 vidíme, že střední hodnoty u dominant mají operátoři podobné. U nondominant a joint má Adam vyšší střední hodnoty i znatelně větší variance. Boxplot 2 ukazuje, že Adam měl vyšší počet teček všude, kromě nejmenšího kruhu.

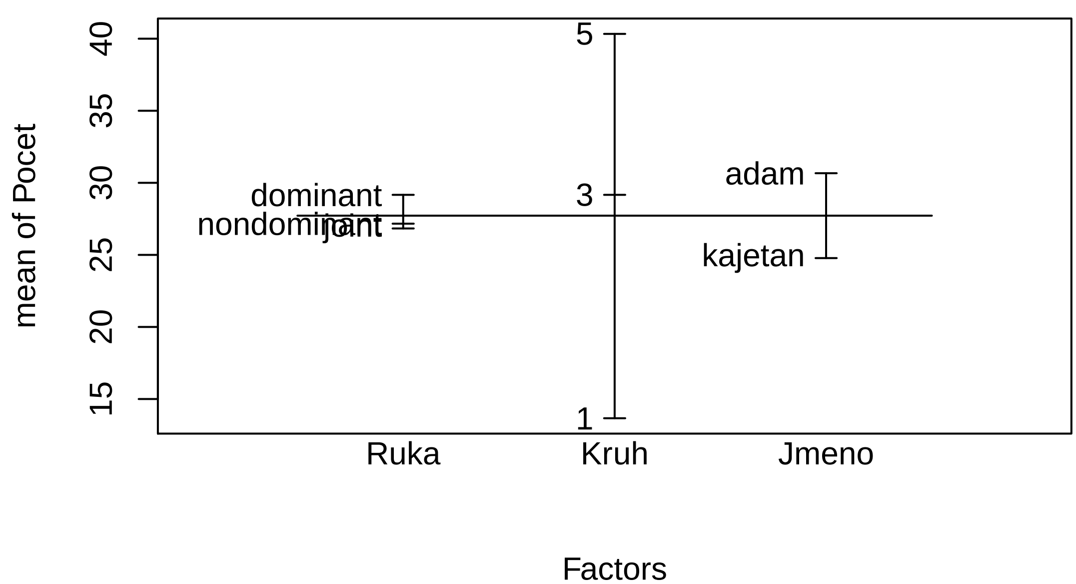
```
#INTERACTION PLOTS
dat2 %>%
  ggplot() +
  aes(x = Jmeno, y = grPocet, color = Ruka) +
  geom_line(aes(group = Ruka)) +
  geom_point()
```




```
dat3 %>%
  ggplot() +
  aes(x = Jmeno, y = grPocet, color = Kruh) +
  geom_line(aes(group = Kruh)) +
  geom_point()
```




```r
#EFFECT PLOTS
plot.design(Pocet~Ruka+Kruh+Jmeno, data = dat)
```

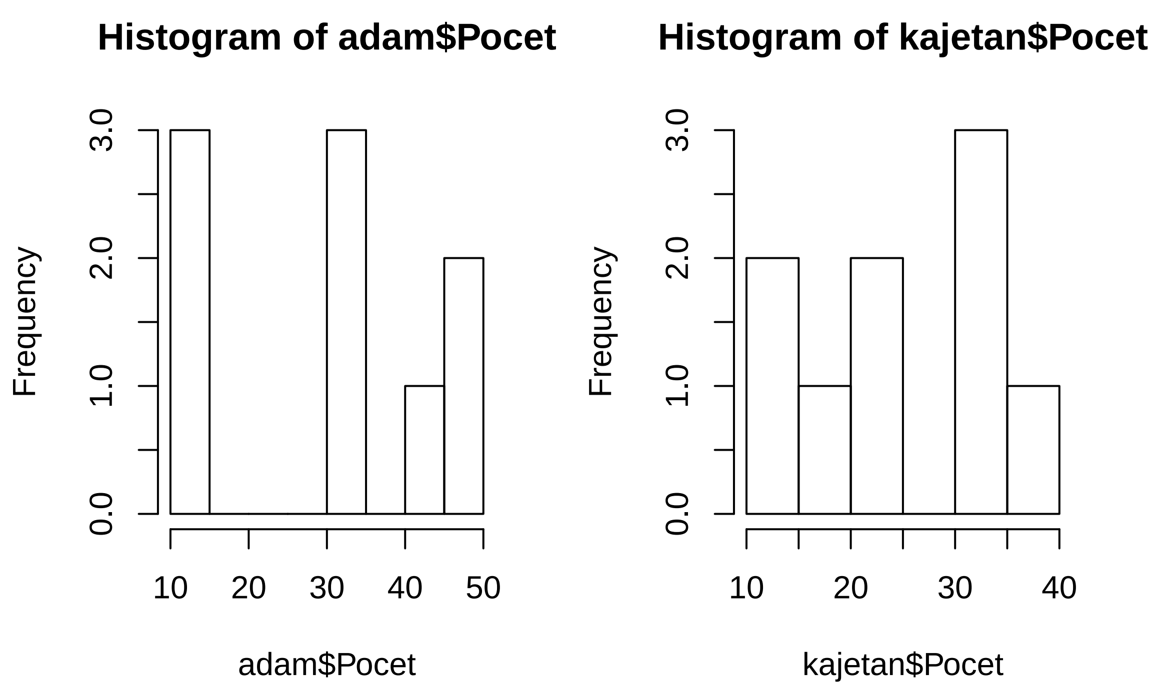


3. část

- Zamerte se zvlaste na faktory ruka a velikost kola. Otestujte hypotezu o schodnosti rozptylu pro jednotlive urovne a vhodnym testem overte stejnost strednich hodnot. Provedte Tukey HSD a Fisher LSD test pro parove porovnani stednich hodnot jednotlivych skupin s vybranou korekci p-hodnoty.

```
#hypoteza shodnosti rozptylu pro jednotlive urovne
alpha = 0.95

# podminky f-testu: normalita, tj. histogram a shapiro-wilk test
par(mfrow = c(1,2))
hist(adam$Pocet, breaks=6)
hist(kajetan$Pocet, breaks=6)
```




```
shapiro.test(adam$Pocet)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  adam$Pocet
## W = 0.88219, p-value = 0.1656
```

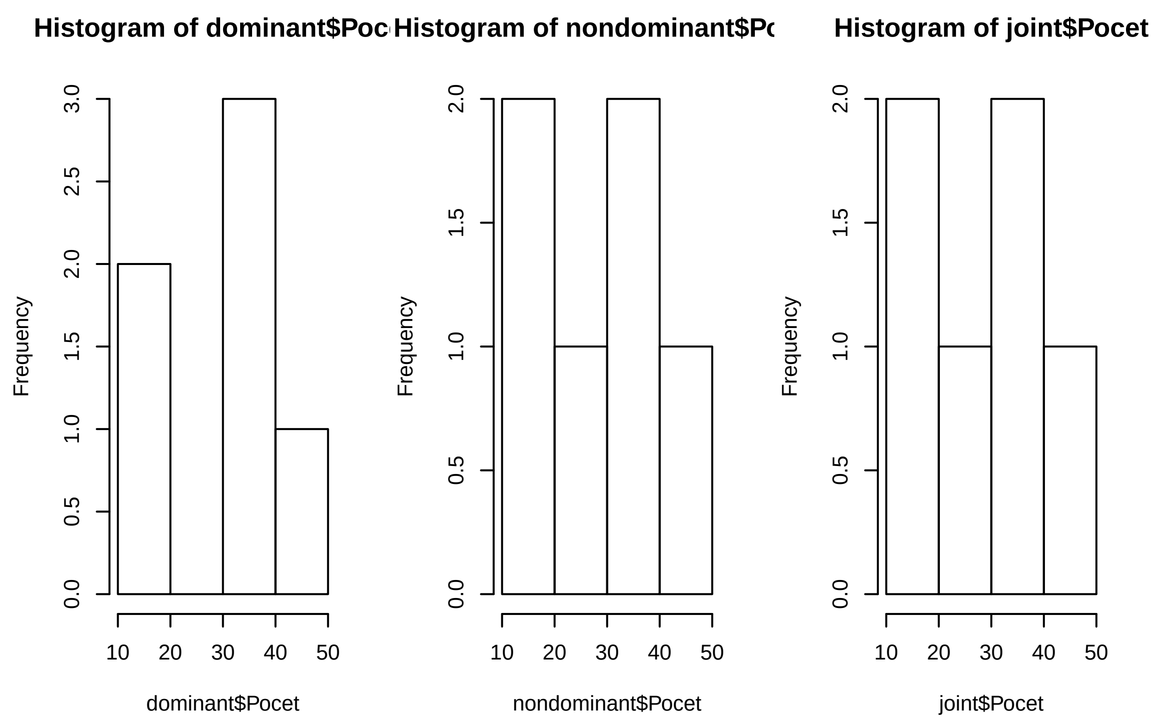
```
shapiro.test(kajetan$Pocet) # p-hodnoty > 0.05, tj. nezamitame normalitu
```

```
##
##  Shapiro-Wilk normality test
##
## data:  kajetan$Pocet
## W = 0.92229, p-value = 0.4116
```

```
var.test(adam$Pocet, kajetan$Pocet, ratio = 1, alternative = "two.sided", conf.level = alpha) # p-hodno
```

```
##
##  F test to compare two variances
##
## data:  adam$Pocet and kajetan$Pocet
## F = 2.693, num df = 8, denom df = 8, p-value = 0.1827
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##   0.607445 11.938599
## sample estimates:
## ratio of variances
##           2.692962
```

```
## RUKA
par(mfrow = c(1,3))
hist(dominant$Pocet)
hist(nondominant$Pocet)
hist(joint$Pocet)
```




```
shapiro.test(dominant$Pocet)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  dominant$Pocet
## W = 0.89281, p-value = 0.3332
```

```
shapiro.test(nondominant$Pocet)
```

```
##
##  Shapiro-Wilk normality test
##
```

```
## data:  nondominant$Pocet
## W = 0.91686, p-value = 0.4831
```

```r
shapiro.test(joint$Pocet) # p-hodnoty > 0.05, tj. nezamitame normalitu
```

```
##
##  Shapiro-Wilk normality test
##
## data:  joint$Pocet
## W = 0.9526, p-value = 0.7613
```

```r
var.test(dominant$Pocet, nondominant$Pocet, ratio = 1, alternative = "two.sided", conf.level = alpha) #
```

```
##
##  F test to compare two variances
##
## data:  dominant$Pocet and nondominant$Pocet
## F = 0.82808, num df = 5, denom df = 5, p-value = 0.8411
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1158741 5.9177775
## sample estimates:
## ratio of variances
##          0.8280802
```

```r
var.test(dominant$Pocet, joint$Pocet, ratio = 1, alternative = "two.sided", conf.level = alpha) # p-hod
```

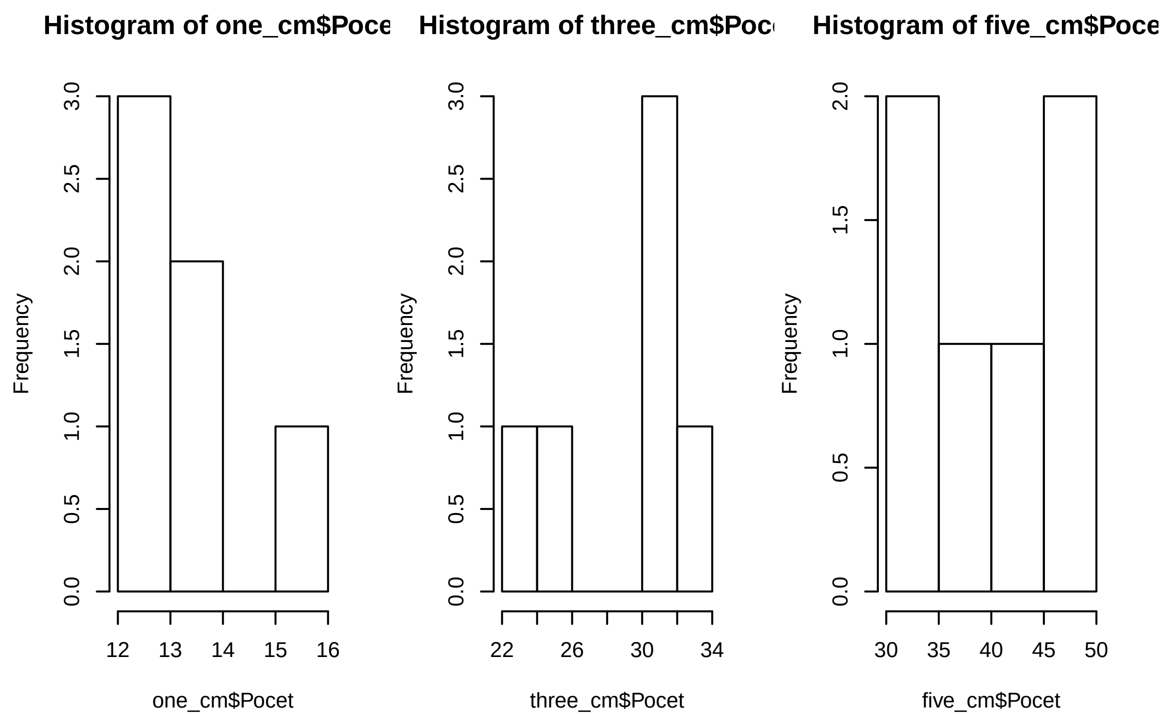
```
##
##  F test to compare two variances
##
## data:  dominant$Pocet and joint$Pocet
## F = 1.1023, num df = 5, denom df = 5, p-value = 0.9175
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1542474 7.8775351
## sample estimates:
## ratio of variances
##           1.102311
```

```r
var.test(nondominant$Pocet, joint$Pocet, ratio = 1, alternative = "two.sided", conf.level = alpha) # p-
```

```
##
##  F test to compare two variances
##
## data:  nondominant$Pocet and joint$Pocet
## F = 1.3312, num df = 5, denom df = 5, p-value = 0.7613
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1862711 9.5130095
## sample estimates:
## ratio of variances
##           1.331164
```

```r
## KRUH
par(mfrow = c(1,3))
hist(one_cm$Pocet)
hist(three_cm$Pocet)
```

```r
hist(five_cm$Pocet)
```




```r
shapiro.test(one_cm$Pocet)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  one_cm$Pocet
## W = 0.92664, p-value = 0.5544
```

```r
shapiro.test(three_cm$Pocet)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  three_cm$Pocet
## W = 0.8307, p-value = 0.109
```

```r
shapiro.test(five_cm$Pocet) # p-hodnoty > 0.05, tj. nezamitame normalitu
```

```
##
##  Shapiro-Wilk normality test
##
## data:  five_cm$Pocet
## W = 0.88926, p-value = 0.3143
```

```r
var.test(one_cm$Pocet, three_cm$Pocet, ratio = 1, alternative = "two.sided", conf.level = alpha) # ZAMI
```

```
##
##  F test to compare two variances
##
```

```
## data:  one_cm$Pocet and three_cm$Pocet
## F = 0.11002, num df = 5, denom df = 5, p-value = 0.0301
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.01539515 0.78624240
## sample estimates:
## ratio of variances
##          0.1100196
```

```
var.test(one_cm$Pocet, five_cm$Pocet, ratio = 1, alternative = "two.sided", conf.level = alpha) # ZAMIT
```

```
##
##  F test to compare two variances
##
## data:  one_cm$Pocet and five_cm$Pocet
## F = 0.028866, num df = 5, denom df = 5, p-value = 0.00139
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.004039244 0.206287311
## sample estimates:
## ratio of variances
##         0.02886598
```

```
var.test(three_cm$Pocet, five_cm$Pocet, ratio = 1, alternative = "two.sided", conf.level = alpha)
```

```
##
##  F test to compare two variances
##
## data:  three_cm$Pocet and five_cm$Pocet
## F = 0.26237, num df = 5, denom df = 5, p-value = 0.1683
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.03671384 1.87500430
## sample estimates:
## ratio of variances
##          0.2623711
```

```
# hypoteza shodnosti strednich hodnot
# podminky t-testu: normalita, tj. histogram a shapiro-wilk test; uz jsme provadeli
t.test(adam$Pocet, kajetan$Pocet, conf.level = 0.95, alternative = "two.sided", var.equal = TRUE) #neza
```

```
##
##  Two Sample t-test
##
## data:  adam$Pocet and kajetan$Pocet
## t = 1.0171, df = 16, p-value = 0.3242
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -6.384906 18.162684
## sample estimates:
## mean of x mean of y
##  30.66667  24.77778
```

```
### toto bude lepsi nahradit pairwise testem
t.test(dominant$Pocet, nondominant$Pocet, conf.level = 0.95, alternative = "two.sided", var.equal = TRU
```

```
##
```

```
##  Two Sample t-test
##
## data:  dominant$Pocet and nondominant$Pocet
## t = 0.25765, df = 10, p-value = 0.8019
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -15.29581  19.29581
## sample estimates:
## mean of x mean of y
##  29.16667  27.16667
```

```r
t.test(dominant$Pocet, joint$Pocet, conf.level = 0.95, alternative = "two.sided", var.equal = TRUE) #ne
```

```
##
##  Two Sample t-test
##
## data:  dominant$Pocet and joint$Pocet
## t = 0.3234, df = 10, p-value = 0.7531
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -13.74259  18.40925
## sample estimates:
## mean of x mean of y
##  29.16667  26.83333
```

```r
t.test(nondominant$Pocet, joint$Pocet, conf.level = 0.95, alternative = "two.sided", var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  nondominant$Pocet and joint$Pocet
## t = 0.043874, df = 10, p-value = 0.9659
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -16.59498  17.26165
## sample estimates:
## mean of x mean of y
##  27.16667  26.83333
```

```r
t.test(one_cm$Pocet, three_cm$Pocet, conf.level = 0.95, alternative = "two.sided", var.equal = FALSE) #
```

```
##
##  Welch Two Sample t-test
##
## data:  one_cm$Pocet and three_cm$Pocet
## t = -8.7487, df = 6.087, p-value = 0.000114
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -19.82019 -11.17981
## sample estimates:
## mean of x mean of y
##  13.66667  29.16667
```

```r
t.test(one_cm$Pocet, five_cm$Pocet, conf.level = 0.95, alternative = "two.sided", var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
```

```
##
## data:  one_cm$Pocet and five_cm$Pocet
## t = -8.008, df = 5.2884, p-value = 0.0003764
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -35.08817 -18.24516
## sample estimates:
## mean of x mean of y
##  13.66667  40.33333
```

```
t.test(three_cm$Pocet, five_cm$Pocet, conf.level = 0.95, alternative = "two.sided", var.equal = TRUE) #
```

```
##
##  Two Sample t-test
##
## data:  three_cm$Pocet and five_cm$Pocet
## t = -3.0274, df = 10, p-value = 0.01273
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -19.385315  -2.948018
## sample estimates:
## mean of x mean of y
##  29.16667  40.33333
```

```
### toto bude lepsi nahradit pairwise testem

pairwise.t.test(dat$Pocet,dat$Ruka,p.adjust.method="bonferroni")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  dat$Pocet and dat$Ruka
##
##             dominant joint
## joint       1        -
## nondominant 1        1
##
## P value adjustment method: bonferroni
```

```
pairwise.t.test(dat$Pocet,dat$Ruka,p.adjust.method="hochberg")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  dat$Pocet and dat$Ruka
##
##             dominant joint
## joint       0.97     -
## nondominant 0.97     0.97
##
## P value adjustment method: hochberg
```

```
pairwise.t.test(dat$Pocet,dat$Kruh,p.adjust.method="bonferroni")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
```

```
## data:  dat$Pocet and dat$Kruh
##
##   1       3
## 3 0.0004  -
## 5 8.4e-07 0.0069
##
## P value adjustment method: bonferroni
```

```
pairwise.t.test(dat$Pocet,dat$Kruh,p.adjust.method="hochberg")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  dat$Pocet and dat$Kruh
##
##   1       3
## 3 0.00027 -
## 5 8.4e-07 0.00229
##
## P value adjustment method: hochberg
```

Na začátku otestujeme normalitu jakožto podmínku F-testu. Pracujeme na 5% hladině významnosti. Normalitu testujeme pomocí Shapirova-Wilkova testu. Pro operátory nezamítáme nulovou hypotézu o normálním rozdělení. F-test pro operátory vrací p-hodnotu 0.1827, nezamítáme tedy nulovou hypotézu o rovnosti rozptylů. Přecházíme k faktoru Ruka. Pro všechny úrovně nezamítáme, dle Shapiro-wilkova testu, normalitu. F- test vrací pro páry dominant-nondominant, dominant-joint a nondominant-joint respektive p-hodnoty 0.84, 0.92 a 0.76 nezamítáme tedy rovnosti rozptylů. Jako poslední testujeme faktor Kruh. P-hodnoty S-W testu pro úrovně 1cm, 3cm a 5cm jsou postupně 0.55, 0.11 a 0.31, ani pro jednu z úrovní nezamítáme normalitu. F-test vrací pro dvojice 1cm-3cm, 1cm-5cm a 3cm-5cm p-hodnoty 0.030, 0.001 a 0.168 respektive. Pro dvojice 1cm-3cm a 1cm-5cm zamítáme nulovou hypotézu o rovnosti rozptylů. Dalším bodem je testování hypotézy shodnosti středních hodnot. Jedinou podmínkou použití t-testu je normalita, která je již otestována.

```
#Tukey HSD + Fisher LSD
#Celkovy aov
#Kruh bychom ani testovat, protoze jsou ruzne variance ve skupinach 1cm, 3cm, 5cm
aov_celk =  aov(Pocet ~ Jmeno + Ruka + Kruh, data = dat) #ruka neni stat. vyznamna
summary(aov_celk)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Jmeno        1  156.1   156.1   7.728   0.0167 *
## Ruka         2   19.1     9.6   0.473   0.6342
## Kruh         2 2152.1  1076.1  53.285 1.07e-06 ***
## Residuals   12  242.3    20.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
LSD1 <- LSD.test(aov_celk, "Jmeno"); #LSD1
#LSD2 <- LSD.test(aov_celk, "Ruka"); LSD2
#LSD3 <- LSD.test(aov_celk, "Kruh"); LSD3
TukeyHSD(aov_celk, "Jmeno", ordered = FALSE, conf.level = alpha)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Pocet ~ Jmeno + Ruka + Kruh, data = dat)
##
```

```
## $Jmeno
##                   diff      lwr       upr     p adj
## kajetan-adam -5.888889 -10.5045 -1.273273 0.0166546
```

```
#TukeyHSD(aov_celk, "Ruka", ordered = FALSE, conf.level = alpha)
#TukeyHSD(aov_celk, "Kruh", ordered = FALSE, conf.level = alpha)

#par(mfrow = c(2,3))
#plot(LSD1)
#plot(LSD2)
#plot(LSD3)
#plot(TukeyHSD(aov_celk, "Jmeno", ordered = FALSE,las=1))
#plot(TukeyHSD(aov_celk, "Ruka", ordered = FALSE,las=1))
#plot(TukeyHSD(aov_celk, "Kruh", ordered = FALSE,las=1))
```

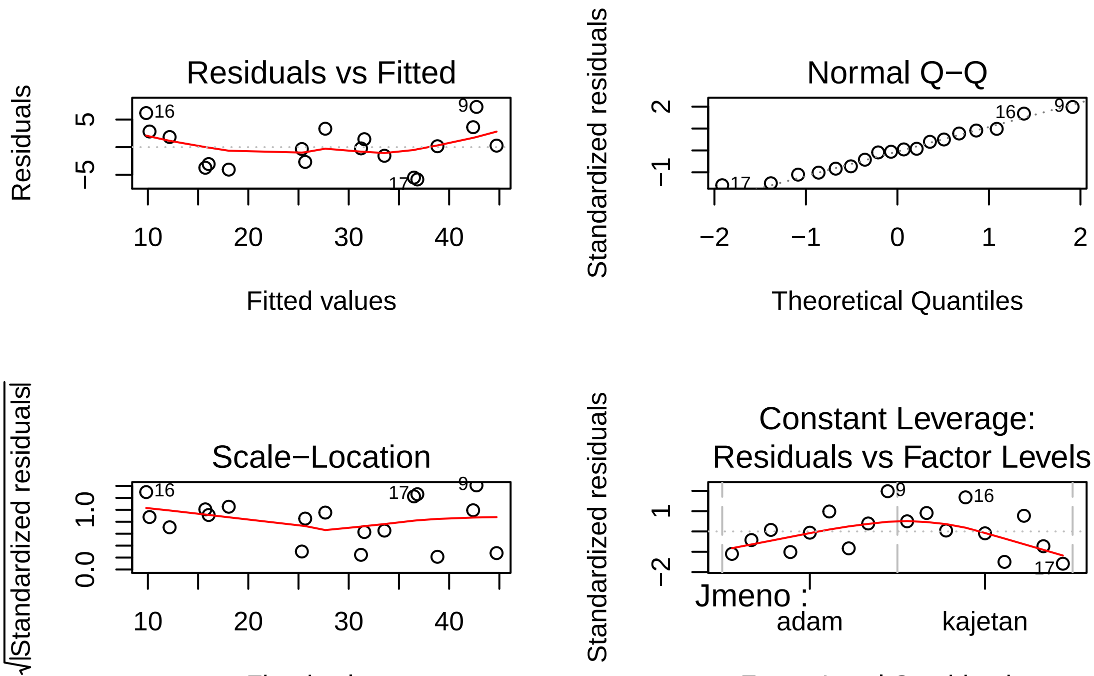
4.část

- Analyzute data pomoci ANOVA (vsechny promenne berte jako faktor) - s predpokladem, ze kazdy z ucastniku experimentu tvori jeden blok. - diskutujte vysledky, overte predpoklady, vykreslete QQ-plot, residua x fitted values, resida x cas, … - diskutujte vliv znahodneni experimentu a vyvoj rezidui v case (cislo mereni)

```
aov_celk
```

```
## Call:
##    aov(formula = Pocet ~ Jmeno + Ruka + Kruh, data = dat)
##
## Terms:
##                    Jmeno      Ruka      Kruh Residuals
## Sum of Squares   156.0556   19.1111 2152.1111  242.3333
## Deg. of Freedom         1         2         2        12
##
## Residual standard error: 4.493823
## Estimated effects may be unbalanced
```

```
par(mfrow = c(2,2))
plot(aov_celk)
```


Residuals vs Fitted
Residuals
Fitted values
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Scale−Location
√|Standardized residuals|
Fitted values
Constant Leverage:
Residuals vs Factor Levels
Standardized residuals
Jmeno :
adam
kajetan
Factor Level Combinations