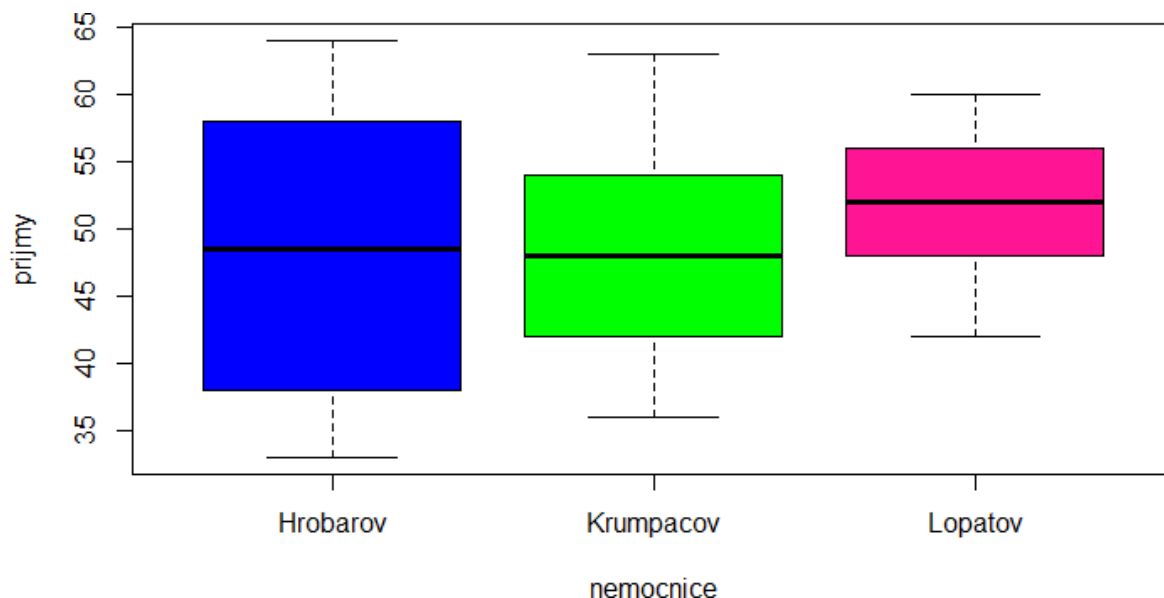


1. Nulová hypotéza: střední hodnoty se rovnají.
Alternativní: střední hodnoty se liší.

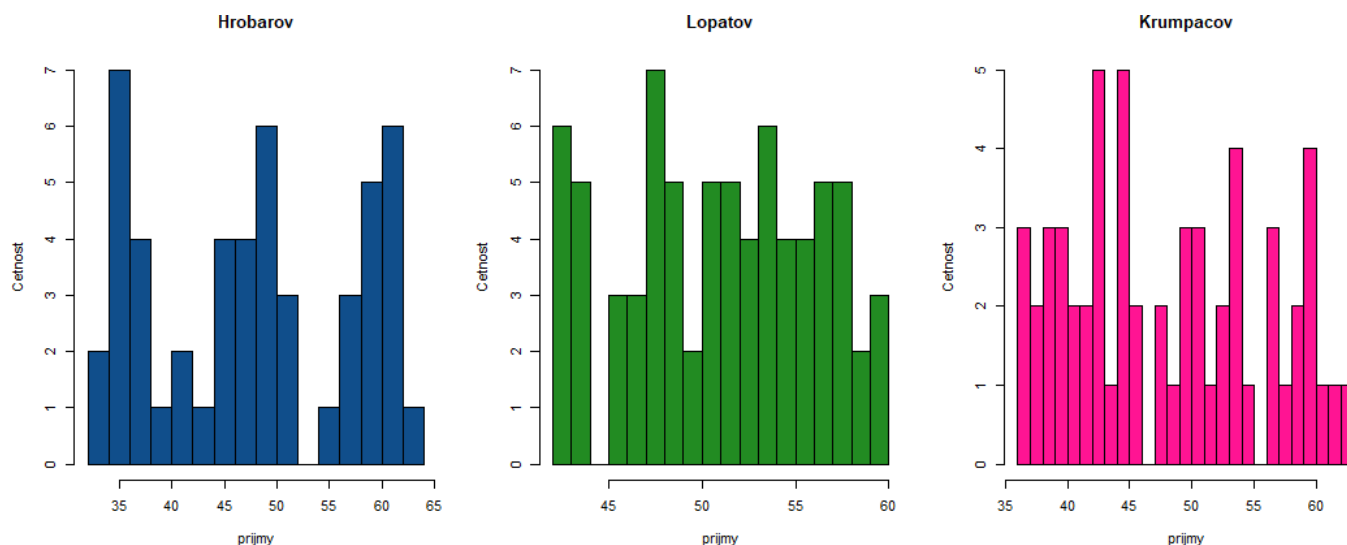
```
cesta<-"C:/Users/Serhii/Documents/Aplikovana_statistika+_R/rymicka.txt"
test<-read.table(cesta, header = TRUE)
boxplot(prijmy~nemocnice, data = test, col = c("blue","green", "deeppink"))
```



```
hist(test[grep("Hrobarov",test$nemocnice),2],breaks=21,col="dodgerblue4",main="Hrobarov",
xlab="prijmy",ylab = "Cetnost")
```

```
hist(test[grep("Lopatov",test$nemocnice),2],breaks=21,col="forestgreen",main="Lopatov",
xlab="prijmy",ylab = "Cetnost")
```

```
hist(test[grep("Krumpacov",test$nemocnice),2],breaks=21,col="deeppink",main="Krumpacov",
xlab="prijmy",ylab = "Cetnost")
```



Vypadá zhruba vpořádku(zhruba protože máme málo dat)

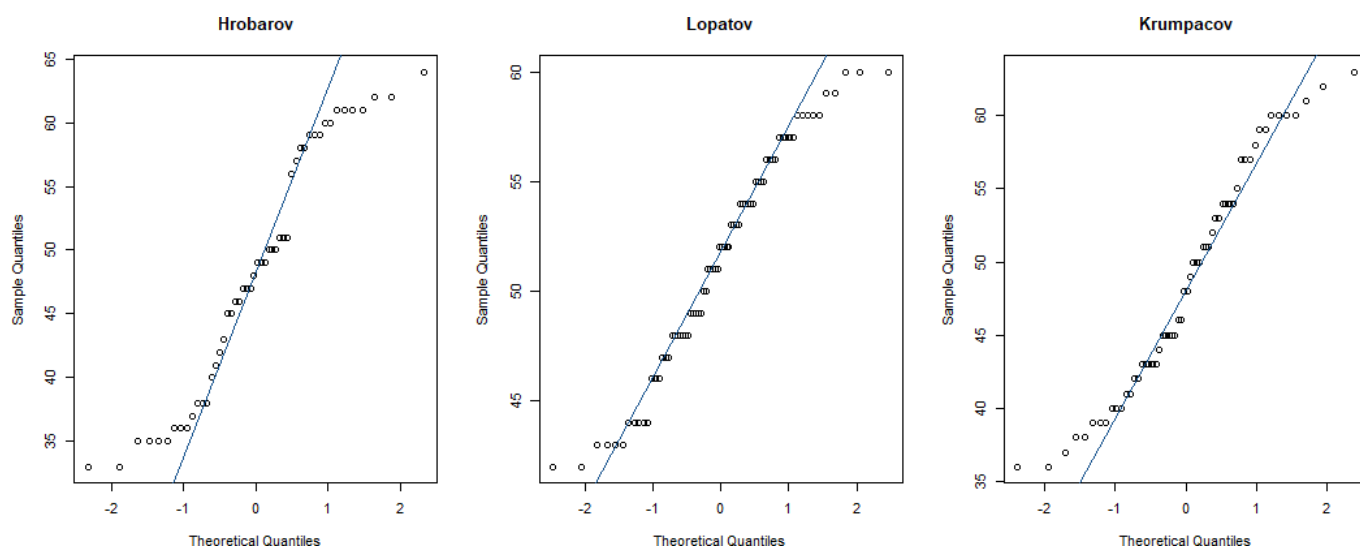
```
table(test$nemocnice)
Hrobarov Krumpacov Lopatov
  50     58     74
```

Metoda ANOVA není vyvážená

```
qqnorm(test[grep("Hrobarov",test$nemocnice),2], main="Hrobarov")
qqline(test[grep("Hrobarov",test$nemocnice),2],col="dodgerblue4")
```

```
qqnorm(test[grep("Lopatov",test$nemocnice),2], main="Lopatov")
qqline(test[grep("Lopatov",test$nemocnice),2],col="dodgerblue4")
```

```
qqnorm(test[grep("Krumpacov",test$nemocnice),2], main="Krumpacov")
qqline(test[grep("Krumpacov",test$nemocnice),2],col="dodgerblue4")
```



Hrobarov nevypadá moc dobře, ale obecně nejde o kritickou situaci

Navíc mohu normalitu otestovat, např. pomocí Shapiro-Wilkova testu

```
shapiro.test(test[grep("Hrobarov",test$nemocnice),2])
```

Shapiro-Wilk normality test

data: test[grep("Hrobarov", test\$nemocnice), 2]

W = 0.92658, p-value = 0.004111

```
shapiro.test(test[grep("Lopatov",test$nemocnice),2])
```

Shapiro-Wilk normality test

data: test[grep("Lopatov", test\$nemocnice), 2]

W = 0.96021, p-value = 0.02005

```
shapiro.test(test[grep("Krumpacov",test$nemocnice),2])
```

Shapiro-Wilk normality test

data: test[grep("Krumpacov", test\$nemocnice), 2]

W = 0.9495, p-value = 0.01725

Hodnoty P-value jsou poměrně nízké, v nemocnici Hrobarov již s chybou 0,4 % zamítli bychom hypotézu, že data pocházejí z normálního rozdělení, ale ve skutečnosti nám poskytnuté údaje pocházejí z normálního rozdělení, takže toto je pouze důkazem toho, že ne vždy realizace náhodného výběru musí vypadat ideálně

```
sd(test[grep("Hrobarov",test$nemocnice),2])
```

```
[1] 9.627766
```

```
sd(test[grep("Lopatov ",test$nemocnice),2])
```

```
[1] 5.093426
```

```
sd(test[grep("Lopatov ",test$nemocnice),2])
```

```
[1] 7.771386
```

Směrodatné odchylky výrazně neliší

```
var(test[grep("Krumpacov",test$nemocnice),2])
```

```
[1] 60.39443
```

```
var(test[grep("Lopatov ",test$nemocnice),2])
```

```
[1] 25.94298
```

```
var(test[grep("Hrobarov ",test$nemocnice),2])
```

```
[1] 92.69388
```

```
var.test(test[grep("Hrobarov",test$nemocnice),2],test[grep("Lopatov",test$nemocnice),2])
```

F test to compare two variances

data: test[grep("Hrobarov", test\$nemocnice), 2] and test[grep("Lopatov", test\$nemocnice), 2]

F = 3.573, num df = 49, denom df = 73, p-value = 8.869e-07

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

2.158977 6.072282

sample estimates:

ratio of variances

3.572984

```
var.test(test[grep("Hrobarov",test$nemocnice),2],test[grep("Krumpacov",test$nemocnice),2])
```

F test to compare two variances

```
data: test[grepl("Hrobarov", test$nemocnice), 2] and test[grepl("Krumpacov", test$nemocnice), 2]
```

F = 1.5348, num df = 49, denom df = 57, p-value = 0.1195

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

0.8942177 2.6644019

sample estimates:

ratio of variances

1.534808

```
var.test(test[grepl("Lopatov", test$nemocnice), 2], test[grepl("Krumpacov", test$nemocnice), 2])
```

F test to compare two variances

```
data: test[grepl("Lopatov", test$nemocnice), 2] and test[grepl("Krumpacov", test$nemocnice), 2]
```

F = 0.42956, num df = 73, denom df = 57, p-value = 0.000708

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

0.2600274 0.6989141

sample estimates:

ratio of variances

0.4295592

Velmi malé hodnoty p-value ukazují, že můžeme bezpečně zamítnout schodnost rozptylu

Zkusme porovnat mediány

Základním rysem mediánu při popisu dat ve srovnání s průměrem (často jednoduše označovaným jako „průměr“) je, že není zkreslen malým podílem extrémně velkých nebo malých hodnot, a proto poskytuje lepší reprezentaci „typické“ hodnoty.

Použijeme neparametrickou ANOVU kruskal test

```
kruskal.test(prijmy~nemocnice, data = test)
```

Kruskal-Wallis rank sum test

data: prijmy by nemocnice

Kruskal-Wallis chi-squared = 5.8257, df = 2, p-value = 0.05432

p-value = 0.05432

Na 5 % by již zamítnuli Nullovou hypotézu, že mediány jsou schodné

