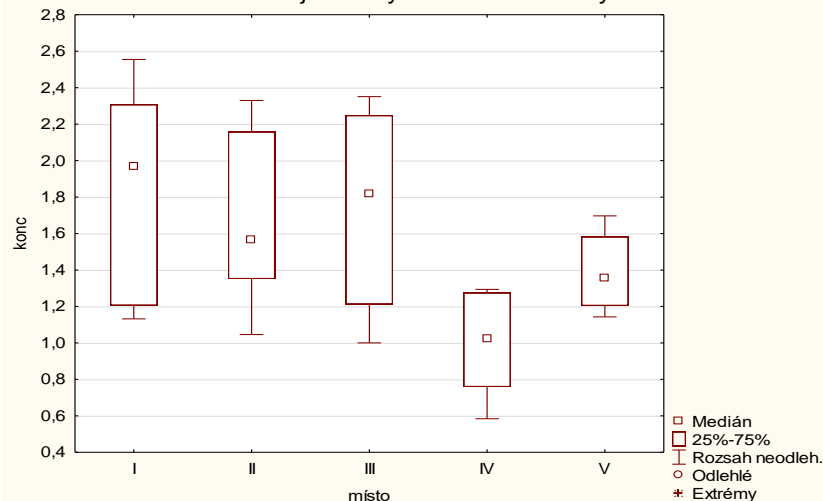
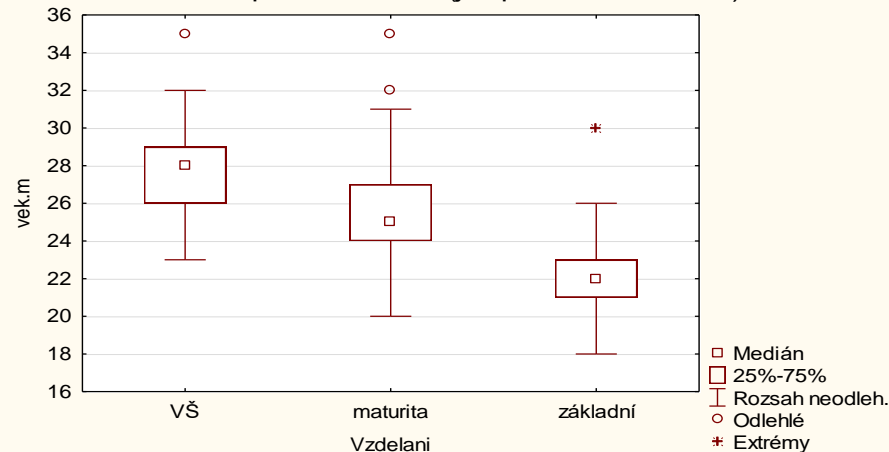


# Několik výběrů tříděných podle jednoho faktoru

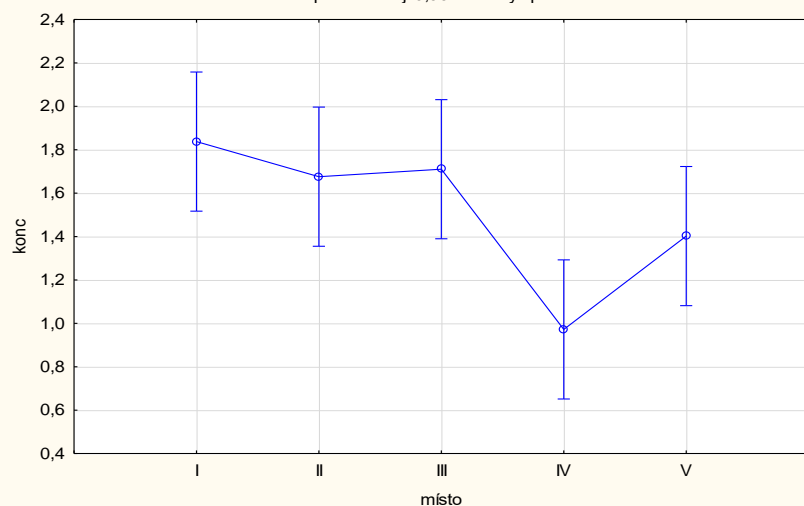
Koncentrace mědi v játrech ryb na 5 místech řeky.



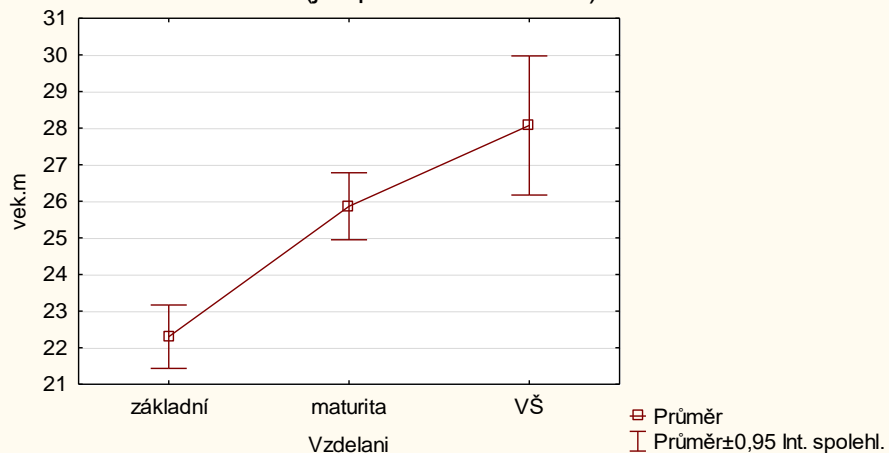
Věk rodiček podle vzdělání (jen první či druhé dítě).



Současný efekt:  $F(4, 30)=4,8214, p=,00401$   
Vertikální sloupce označují 0,95 intervaly spolehlivosti



Průměry věku rodiček podle vzdělání (jen první či druhé dítě)



## Analýza rozptylu, analýza variance [analysis of variance, ANOVA]

Mám několik skupin (výběrů) a ptám se, **zda jsou jejich průměry srovnatelné?**

Hypotéza:  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$  střední hodnoty všech  $k$  skupin jsou stejné

Alternativa  $H_1$ : alespoň jedna střední hodnota se významně liší od ostatních

Myšlenka: jsou-li průměry skupin srovnatelné, potom bude *variabilita mezi skupinovými průměry* podstatně menší než *variabilita uvnitř skupin*.

Analýza tedy pracuje s odhadem rozptylů a testuje je pomocí F-testu.

Poznámka: Není správné testovat vztahy po dvojicích pomocí t-testů, protože velmi narůstá chyba prvního druhu:

Počet průměrů ( $k$ )	Hladina signifikance užívaná v $t$ -testech			
	0.05	0.02	0.01	0.001
2	0.05	0.02	0.01	0.001
3	0.13	0.05	0.03	0.003
4	0.21	0.09	0.05	0.006
5	0.23	0.13	0.07	0.009
10	0.63	0.37	0.23	0.034
20	0.92	0.71	0.52	0.109
$\infty$	1.00	1.00	1.00	1.00

Nejjednodušší situace:

## Jednocestná analýza rozptylu, analýza jednoduchého třídění

[one-way ANOVA, single faktor ANOVA]

Předpoklady testu:

- **Skupiny jsou náhodnými, vzájemně nezávislými výběry.** Tomu se říká **zcela náhodné uspořádání** [completely randomized experimental design]. Jiná uspořádání zahrnují například opakovaná měření na stejném jedinci, mláďata z jednoho vrhu, vegetační snímky z jedné lokality apod.
- **Každý výběr pochází z normálního rozdělení  $N(\mu_i, \sigma^2_i)$**
- **Homoskedasticita, tj. variance výběrů jsou stejné:  $\sigma^2_1 = \sigma^2_2 = \dots = \sigma^2_k$**

Skupiny jsou rozlišené **jedním faktorem**, např. příslušnost ke druhu. Jednotlivé druhy pak představují **úroveň faktoru** [levels of the factor].

Zároveň nás zajímají právě ty konkrétní druhy, jedná se o **model s pevnými efekty** [fixed effect model]. Odlišný model s *náhodnými efekty* popíšeme později.

## Jednocestná analýza rozptylu - značení

### Značení:

$$\begin{array}{lll}
 \text{1. skupina: } (X_{11}, X_{12}, X_{13}, \dots, X_{1n_1}) & \bar{X}_{1\cdot} = \frac{1}{n_1} \sum_{t=1}^{n_1} X_{1t} & s_1^2 = \frac{\sum_{t=1}^{n_1} (X_{1t} - \bar{X}_{1\cdot})^2}{n_1 - 1} \\
 \text{2. skupina: } (X_{21}, X_{22}, X_{23}, \dots, X_{2n_2}) & \bar{X}_{2\cdot} = \frac{1}{n_2} \sum_{t=1}^{n_2} X_{2t} & s_2^2 = \frac{\sum_{t=1}^{n_2} (X_{2t} - \bar{X}_{2\cdot})^2}{n_2 - 1} \\
 \vdots & & \\
 \text{k. skupina: } (X_{k1}, X_{k2}, X_{k3}, \dots, X_{kn_k}) & \bar{X}_{k\cdot} = \frac{1}{n_k} \sum_{t=1}^{n_k} X_{kt} & s_k^2 = \frac{\sum_{t=1}^{n_k} (X_{kt} - \bar{X}_{k\cdot})^2}{n_k - 1}
 \end{array}$$

$N = n_1 + n_2 + \dots + n_k$  ... celkový počet pozorování;

počty pozorování ve skupinách se mohou lišit

$$\bar{X}_{\cdot\cdot} = \frac{1}{N} \sum_{i=1}^k \sum_{t=1}^{n_i} X_{it} = \frac{1}{N} \sum_{i=1}^k n_i \cdot \bar{X}_{i\cdot} = \sum_{i=1}^k \left( \frac{n_i}{N} \right) \cdot \bar{X}_{i\cdot}$$

Poměrná část, součet vážených průměrů

**Reziduály** neboli odchylky od skupinových průměrů:  $X_{it} - \bar{X}_{i\cdot}$

## Jednocestná analýza rozptylu – rozklad variability

### Variabilita uvnitř skupin neboli reziduální rozptyl:

Suma čtverců reziduálů:

[error sum of squares,  
residual SS]

$$SS_E = \sum_{i=1}^k \sum_{t=1}^{n_i} (X_{it} - \bar{X}_{i\cdot})^2$$

$$DF_E = N - k$$

Průměrný čtverec odchylky od skupinového průměru:

[within group mean square,  
error mean square, residual MS]

$$MS_E = \frac{SS_E}{N - k} = \frac{\sum_{i=1}^k \sum_{t=1}^{n_i} (X_{it} - \bar{X}_{i\cdot})^2}{\sum_{i=1}^k (n_i - 1)}$$

Odhad společné variance  $\sigma^2 = \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$   
Předpoklad testu je, že variabilitu každé skupiny můžeme dobře odhadnout tímto výpočtem.

### Variabilita mezi skupinami (definovanými faktorem A):

[among group mean square,  
group MS,  
MS<sub>G</sub>]

$$SS_A = \sum_{i=1}^k n_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$$

$$DF_A = k - 1$$

každé pozorování přispívá

$$MS_A = \frac{SS_A}{k - 1} = \frac{\sum_{i=1}^k n_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2}{k - 1}$$

Za platnosti  $H_0$  má  $MS_A$  hodnotu blízko nule.

Celková variabilita (spíše z tradice):

$$SS_{TOT} = \sum_{i=1}^k \sum_{t=1}^{n_i} (X_{it} - \bar{X}_{\cdot\cdot})^2$$

$$DF_{TOT} = N - 1$$

$$MS_{TOT} = \frac{SS_{TOT}}{DF_{TOT}} = \frac{SS_E + SS_A}{DF_E + DF_A}$$

**OPRAVA výroku v učebnici Lepš & Šmilauer, str. 155 dole:**

*„Lze ukázat, že pokud platí nulová hypotéza o rovnosti průměrů, potom i  $MS_G$  je odhadem společné variance  $\sigma^2$ .“*

Variabilita mezi skupinami (definovanými faktorem A,  $MS_G \approx MS_A$ ):

$$SS_A = \sum_{i=1}^k n_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2 \longrightarrow MS_A = \frac{SS_A}{k-1} = \frac{\sum_{i=1}^k n_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2}{k-1}$$

**Za platnosti  $H_0$ :**  $\mu_1 = \mu_2 = \dots = \mu_k$  budou  $\bar{X}_{1\cdot} \cong \bar{X}_{2\cdot} \cong \bar{X}_{3\cdot} \cong \bar{X}_{k\cdot} \cong \bar{X}_{\cdot\cdot}$ , a proto bude  $SS_A$  součet samých „malých“ čtverců a výsledek bude blízko nule. Čím menší součet, tím lépe data odpovídají nulové hypotéze.

Takový součet čtverců ale nemůže popisovat variabilitu původních dat  $X_{ij}$ .

## Jednocestná analýza rozptylu – test modelu

Říkáme, že celkovou variabilitu v datech rozkládáme na variabilitu vysvětlenou modelem ( $SS_A$ ) + variabilitu reziduální ( $SS_E$ ).

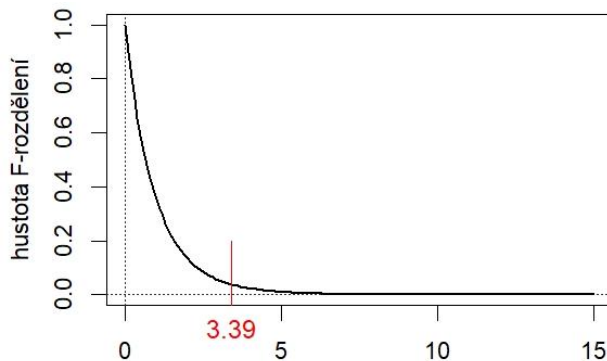
Tyto součty dále vyjadřujeme jako průměrný čtverec:

$MS_A \approx$  průměrný čtverec odchylky na jednu skupinu;

$MS_E \approx$  průměrný čtverec rezidua na jedno pozorování.

Testová statistika:

$$F = \frac{\frac{MS_A}{df_A}}{\frac{MS_E}{df_E}} \sim F_{df_A, df_E}$$



- Když  $H_0$  platí a průměry jsou srovnatelné, potom je  $MS_A$  malé a  $F$ -statistika je blízko nule.
- Když  $H_0$  neplatí a alespoň jeden průměr je jiný než ostatní, potom variabilita mezi průměry  $SS_A$  vzroste. Je-li průměrný čtverec  $MS_A$  alespoň třikrát až čtyřikrát větší než  $MS_E$  (záleží na stupních volnosti), zamítáme platnost nulové hypotézy.

## Jednocestná analýza rozptylu – zadání v R

Příklad: koncentrace mědi v játrech ryb

```
> iris1<-aov(Sepal.Width~Species, data=iris)
summary(iris1)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Species	2	11.35	5.672	49.16	<2e-16 ***
Residuals	147	16.96	0.115		

univerzální funkce pro analýzu rozptylu

```
> oneway.test(Sepal.Width~Species, data=iris, var.equal=F)
```

One-way analysis of means (not assuming equal variances)

data: Sepal.Width and Species

F = 45.012, num df = 2.000, denom df = 97.402, p-value = 1.433e-14

Jen analýza jednoho faktoru, možnost var.equal=F/T

```
> anova(lm(Sepal.Width~Species, data=iris))
```

Analysis of Variance Table

Response: Sepal.Width

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Species	2	11.345	5.6725	49.16	< 2.2e-16 ***
Residuals	147	16.962	0.1154		

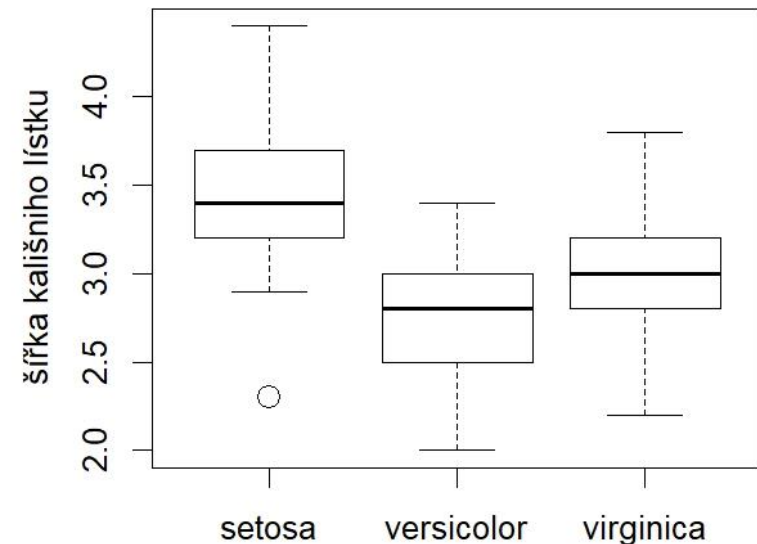
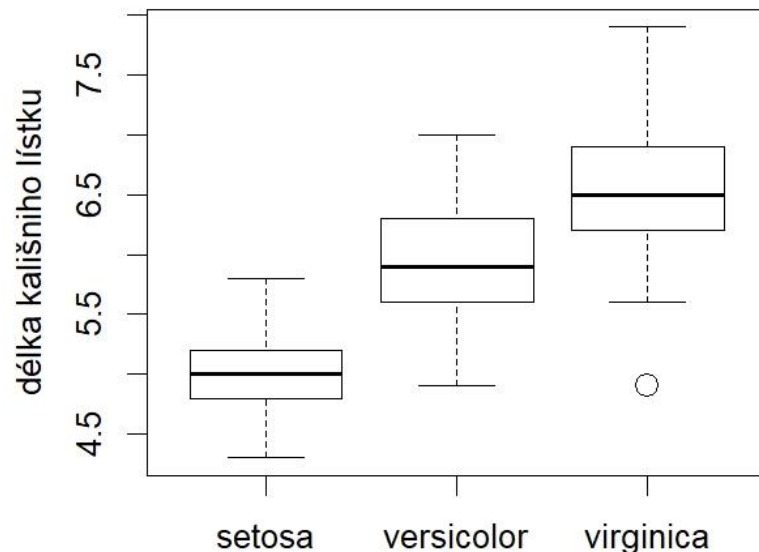
funkce porovnávající dva modely; model musím vytvořit předem



## Ověření předpokladů analýzy rozptylu

### 1) Vizuální ověření: boxploty, histogramy, q-q ploty po skupinách

- Normalita = hodnoty rozložené souměrně kolem průměru/mediánu, tedy mediány uprostřed krabice, oba vousy podobně dlouhé
- Shodnost rozptylů = krabice odpovídající mezikvartilovému rozpětí/ směrodatné odchylce jsou podobně vysoké u všech skupin
- Při velkém nesouladu zkusit transformaci nebo neparametrický test nebo zobecněné lineární modely (např. počty něčeho málo početného /0 až 3 kusy/ budou mít spíš diskrétní Poissonovo rozdělení)



## Ověření předpokladů analýzy rozptylu

### 2) Test shodnosti rozptylů ve skupinách = homoskedasticity

- **Leveneův test** – pracuje s absolutními odchylkami pozorování od skupinových průměrů  $|X_{it} - \bar{X}_{i\bullet}|$  a na tyto hodnoty aplikuje analýzu rozptylu.

$H_0$ : průměrné odchylky pozorování od skupinových průměrů jsou stejné ve všech skupinách.

Předpoklad: normální rozdělení hodnot v každé skupině.

R: `library(car); leveneTest(model)`

- **Brown & Forsythe test** – modifikace Leveneova testu na mediány  $|X_{it} - \tilde{X}_{i\bullet}|$ . Takto je test robustnější vůči odchylce od normálního rozdělení dat.

- **Bartlettův test** – klasický test, jehož nevýhodou je značná citlivost na porušení předpokladu o normálním rozdělení hodnot ve skupinách.

Je založený na porovnání logaritmu reziduálního rozptylu s váženým průměrem logaritmů rozptylů jednotlivých skupin:

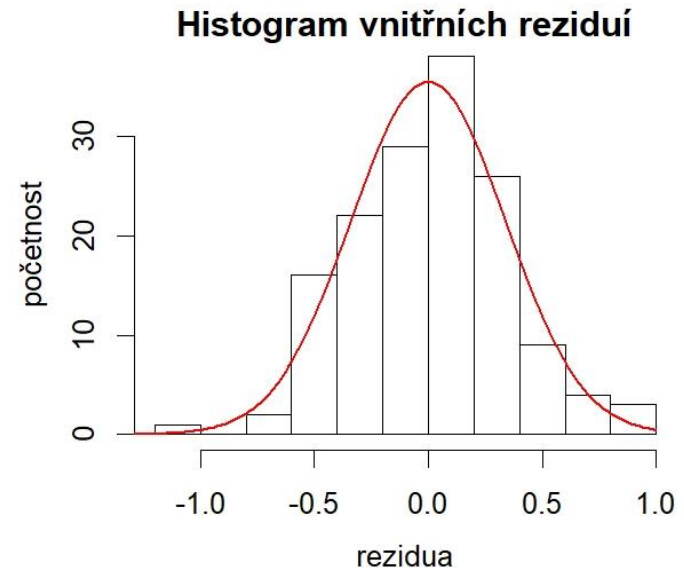
$$B = \frac{1}{C} \left( (N - k) \ln S^2 - \sum_{i=1}^k (n_i - 1) \ln S_i^2 \right)$$
, kde číslo **C** určují počty pozorování ve skupinách a vychází malinko větší než 1.

## Ověření předpokladů analýzy rozptylu

### 3) Test normality – Shapiro-Wilkův test na reziduály

Není třeba testovat každou skupinu zvlášť, přestože předpoklady jsou právě takto formulovány. Jestliže jsme nezamítnuli hypotézu o stejných variancích (rozptylech) ve skupinách, znamená to, že skupiny se odlišují ve středních hodnotách (průměrech), ale pozorování jsou kolem průměru rozptýleny podobně. Vezmu proto všechny reziduály ( $X_{it} - \bar{X}_{i\cdot}$ ) z  $k$  skupin dohromady a studuji rozdělení například pomocí Shapirova-Wilkova testu. **Tento postup je výhodný zejména v situaci, kdy jsou jednotlivé skupiny málo početné.**

```
R: shapiro.test(model$residuals)  
hist(model$residuals)
```



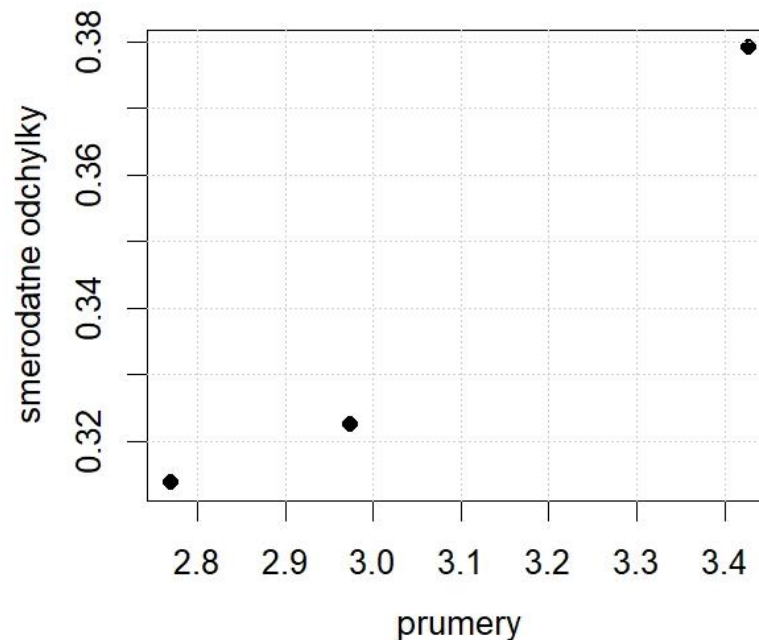
## Ověření předpokladů analýzy rozptylu

### 4) Kontrola vztahu rozptylu a střední hodnoty

Upozorní na možné problémy s nehomogenitou rozptylů. Zajímá nás situace, kdy s vyšším průměrem se zvětšuje také rozptyl. Toto je charakteristické zejména pro logaritmicko-normální rozdělení. Logaritmováním původních pozorování můžeme přiblížit rozptyly ve skupinách i celý výběr normálnímu rozdělení.

Častá vlastnost dat o koncentracích, hmotnosti a jiných objemových ukazatelů.

Vztah rozptylů a středních hodnot lze vypořádat také z boxplotů.



## Ověření předpokladů analýzy rozptylu

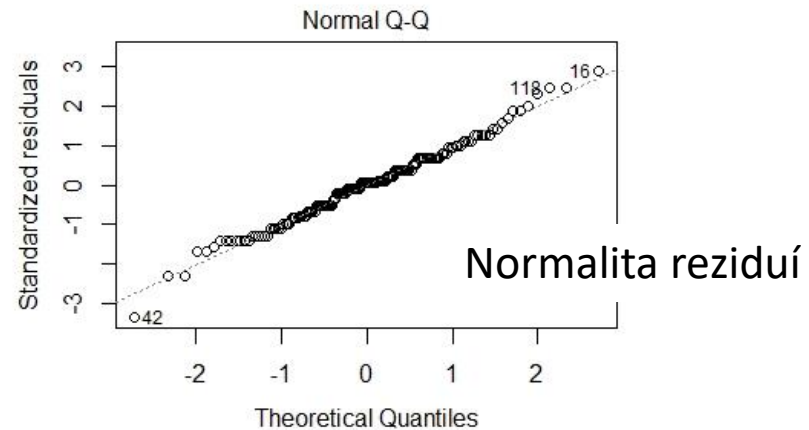
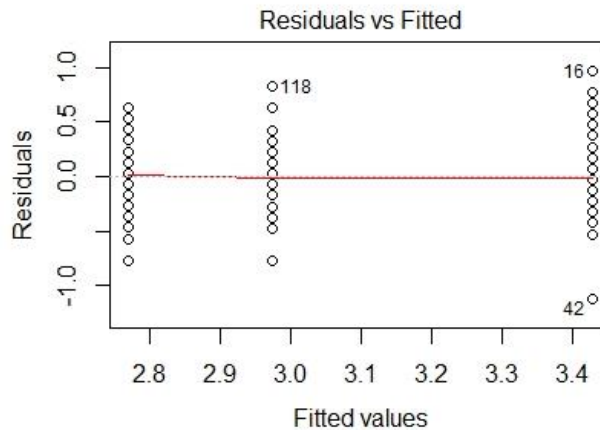
ANOVA je celkem robustní vůči narušení předpokladů, proto je jejich ověřování zlehčováno. Ověřování je však důležité například v těchto situacích:

- Mám jen málo pozorování v jednotlivých skupinách → dívám se hlavně na normalitu reziduálů a na vztah rozptylů a středních hodnot.
- Mám nevyvážený design [unbalanced design], tj. některé skupiny početné a jiné velmi nepočetné.
- Data mají výrazně šikmé rozdělení. Pomůže transformace?
- Leveneův F-test mohu posuzovat na hladině chyby 1. druhu 1 %.

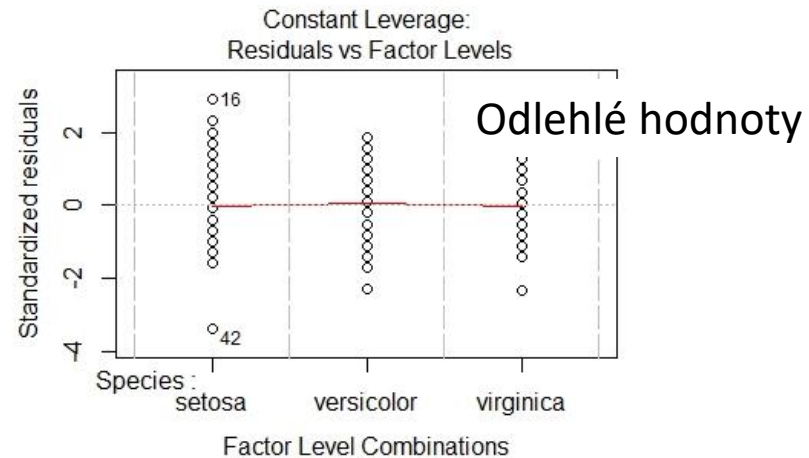
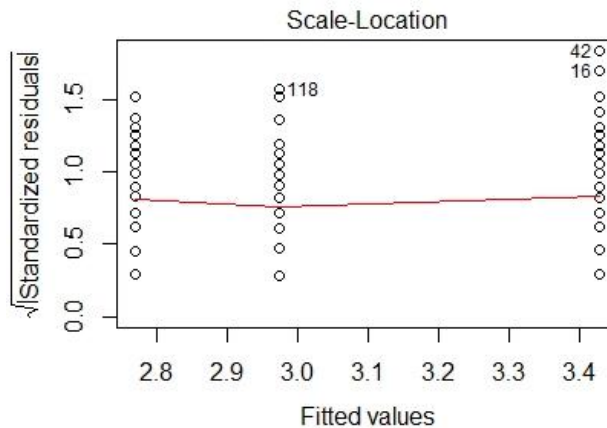
## Ověření předpokladů – možnosti softwaru

- `library(car); leveneTest(model)`
- `shapiro.test(model$residuals); hist(model$residuals)`
- `plot(model)`

Shodnost rozptylů,  
souměrnost reziduí  
kolem nuly.



Vztah průměrů  
a směrodatné  
odchylky  
– je tam  
závislost?



## Mnohonásobná porovnání [multiple comparisons]

Pokud jsme v předchozím „celkovém“ testu zamítli nulovou hypotézu o shodnosti středních hodnot, zajímá nás, který z průměrů se liší od kterého.

Na tuto otázku použijeme takové testy, které mají ošetřeno nebezpečí vzrůstající chyby 1. druhu, tedy že označím za průkazný rozdíl mezi průměry i tam, kde ve skutečnosti rozdíl není (jen moje náhodné výběry náhodou pokrývají pravou a levou část spektra možných hodnot a proto se průměry zdají odlišné...)

Kolika testů se úvahy týkají?

Mám-li  $k$  skupin, potom

porovnávám  $\frac{k \cdot (k-1)}{2}$  dvojic výběrů.

Příklad:  $k = 3 \rightarrow 3 \cdot 2 / 2 = 3$

	Skup. 1	Skup. 2	Skup. 3
Skup. 1		1 vs. 2	1 vs. 3
Skup. 2			2 vs. 3
Skup. 3			

## Mnohonásobná porovnání

### Tukeyho metoda HSD

*HSD = honest significance difference, tj. poctivě průkazný rozdíl*

$$|\bar{X}_{i\cdot} - \bar{X}_{j\cdot}| \geq q_{k, N-k}(1 - \alpha) \cdot \sqrt{\frac{S_E^2}{2} \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}$$

kde  $q_{k, N-k}$  je kvantil studentizovaného rozpětí

- Odhad rozptylu je spočítán na základě variability všech skupin, nejenom těch dvou, které právě srovnáváme.
- Test byl původně navržen pro vyvážený design (všechny  $n_i$  stejná). Používáme také pro nestejná  $n_i$ , jen výsledek je potom spíše konzervativní, tzn. že skutečná pravděpodobnost chyby 1. druhu může být menší než zvolená hladina 5 %.
- Předpoklad homogenity variance je stále vyžadován a doporučení stejné velikosti skupin zde platí naléhavěji než pro hlavní test ANOVA modelu.

**R:** `TukeyHSD(x, which, ordered=FALSE, conf.level=0.95, ...)`

**x** není proměnná, ale objekt typu `model`, ve kterém jsou výsledky analýzy ANOVA (`aov`)



## Mnohonásobná porovnání – další čtení

Existuje samozřejmě více metod na porovnávání. Některé už jsou zastaralé, jiné vyžadují modernější přístup. Často také musíte rozumět pojmu „kontrasty v modelu“.

K dalšímu čtení vezměte Lepš & Šmilauer, str. 160 + řešení v R od str. 175.

## Mnohonásobná porovnání – prezentace výsledků

Lepš & Šmilauer, str. 163

- Situace, kdy skupina s „prostředním“ průměrem se neliší ani od menšího ani od většího průměru, zatímco „malý“ a „velký“ průměr se odlišují průkazně. Není ale možné, aby „prostřední“ skupina byla totožná s „menší“ i „větší“ skupinou. Došlo zde zřejmě k chybě druhého druhu (nezamítám hypotézu, která neplatí). Je to dáno nejistotou způsobenou nedostatkem pozorování.
- Stává se, že hlavní F-test ANOVA modelu zamítá hypotézu o rovnosti průměrů, ale při mnohonásobných porovnáních nedostanu žádný průkazný rozdíl. To je dáno faktem, že mnohonásobné testování nemá takovou sílu jako hlavní F-test. Dá se ale očekávat, že data s větším počtem pozorování by pravděpodobně odhalila průkazné rozdíly i v mnohonásobných porovnáních.

## Mnohonásobné porovnávání – poznámky k chybě 1. druhu

Komentář k problematice porovnávání *a posteriori* (skupiny porovnávám až po provedení celkového F-testu) a *a priori* (předem vyberu skupiny, které mě zajímají) najdete v učebnici Biostatistika, Lepš & Šmilauer (2016) na straně 160. Autoři odkazují dále na Biometry, Sokal & Rohlf (2010, 4. vydání) str. 228-237.

Dále tam najdete komentář ke skutečné pravděpodobnosti chyby 1. druhu v situaci, kdy provádím vícero testů na jedněch datech (stále strana 160). *Comparisonwise Type I error rate, experimentwise Type I error rate, familywise Type I error rate.*

## Neparametricky: Kruskalův-Wallisův test

Test založený na pořadí [\[analysis of variance by ranks\]](#).

- Předpoklady: výběry jsou vzájemně i „uvnitř“ nezávislé, tedy zcela náhodné experimentální uspořádání. Data nejlépe spojitá, ale lze testovat i pro data na ordinální stupnici.
- Pokud porovnávám jen 2 výběry, budou se čísla lišit od výsledků Mann-Whitneyova testu.
- Nulová hypotéza: distribuční funkce všech skupin jsou stejné, tedy speciálně také všechny mediány jsou stejné.
- Testová statistika:

$$H = (N - 1) \frac{\sum_{i=1}^k n_i (R_{i\cdot} - R_{\cdot\cdot})^2}{\sum_{i=1}^k \sum_{t=1}^{n_i} (R_{it} - R_{\cdot\cdot})^2} \sim \chi^2_{(k-1)}$$

## Neparametricky: Kruskalův-Wallisův test

R: `kruskal.test(formula, data, subset, na.action, ...)`

## Neparametricky: mnohonásobná porovnání

R: v knihovně `pgirmess` funkce

```
kruskalmc(x ~ groups, data= , probs=0.05)
```

## Analýza rozptylu

### Náhodný vs. pevný efekt faktoru [random vs. fixed effect]

#### Pevný efekt

- zajímají nás konkrétně zvolené úrovně faktoru;
- úrovně mohu zopakovat i v příštím experimentu;
- pro faktor s pevným efektem potřebuji reparametrizační podmínku;
- Př.: přidavek hnojiva, typ diety, rozdíl mezi druhy, mezi pohlavími

#### Náhodný efekt

- úroveň faktoru tvoří např. mláďata z jednoho vrhu, snímky z jedné lokality, paraziti z jedné ryby, apod.;
- celý faktor tedy rozlišuje různé vrhy mláďat, různé lokality, různé ryby, které ovšem chápeme jako náhodný výběr ze všech vrhů/lokalit/ryb daného typu.
- Primárně nás nezajímají rozdíly mezi měřenými vrhy/lokalitami/rybami, ale zajímá nás, zda se obecně vrhy/lokality/ryby mezi sebou liší.
- Úrovním faktoru s náhodným efektem pak říkáme **náhodné bloky**.

(Připouštíme ale i situaci, kdy nás zajímají právě ty zvolené vrhy/lokality/ryby. Potom je to faktor s pevnými efekty a řídí se pravidly pro pevné efekty.)

## Analýza rozptylu s více parametry

### Uspořádání experimentu [design of the experiment]

Faktoriálové uspořádání = design zahrnuje všechny kombinace úrovní faktorů  
 (například data „npk“ přímo v Rku. )

### Vyvážené uspořádání [balanced design]

- **Ideálně stejný počet pozorování** pro všechny kombinace úrovní faktorů (pro všechny buňky)
- S nevyvážeností počtu pozorování klesá síla testů.
- Proporční uspořádání = varianta, když není možné dodržet vyvážené uspořádání

$$n_{ij} = \frac{(\text{počet pozorov. s úrovní } i \text{ faktoru } A)(\text{počet pozorov. s úrovní } j \text{ faktoru } B)}{N} =$$

$$= \frac{\sum_{j=1}^b n_{ij} \cdot \sum_{i=1}^a n_{ij}}{N}$$

	B1	B2	...	Bj	...	Bb
A1				$n_{1j}$		
Ai	$n_{i1}$	$n_{i2}$	...	$n_{ij}$	...	$n_{ib}$
...				⋮		
Aa				$n_{aj}$		

## Analýza rozptylu

# Uspořádání experimentu

### Nevyvážený design:

- přestává platit vztah  $SS_{TOT} = SS_A + SS_B + SS_{AB} + SS_{Err}$
- klesají síly testů
- existují různé výpočtové metody rozkladu vysvětlené variability mezi jednotlivé efekty (faktory). Výsledky analýz z různých softwarů se proto mohou hodně lišit.



## Analýza rozptylu

### Uspořádání experimentu

Kde studovat více o typech rozkladu variability na součty čtverců:

- začněte v Lepš & Šmilauer, str. 219. Dále anglické učebnice.
- Základní R-ková funkce **aov ( )** používá první typ rozkladu, tzv. sekvenciální rozklad. V knihovně **car** je dostupná funkce **Anova ( )**, která nabízí druhý a třetí typ rozkladu (parciální a ortogonální rozklad). Například pro model se 2 pevnými efekty, který má nesignifikantní interakce doporučuje Lepš a Šmilauer použít právě typ rozkladu II.

## Analýza rozptylu - Uspořádání experimentu

### Rozmístění pokusných ploch v terénu

Biotické i abiotické charakteristiky se v ploše kontinuálně mění => blízké plochy jsou si podobnější než plochy vzdálené. Proto byly vytvořeny standardní postupy pro uspořádání pokusů. Cílem je co nejvíce omezit vliv heterogenity pokusné plochy při zachování nezávislosti opakování.

#### Latinský čtverec

Každý řádek a každý sloupec obsahuje právě jednu plochu od každého typu zásahu.

Analyzujeme 3cestným modelem ANOVA bez opakování.

Pro každou plochu zaznamenám řádek (faktor A), sloupec (B) a typ zásahu (C).

columns

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

ROWS

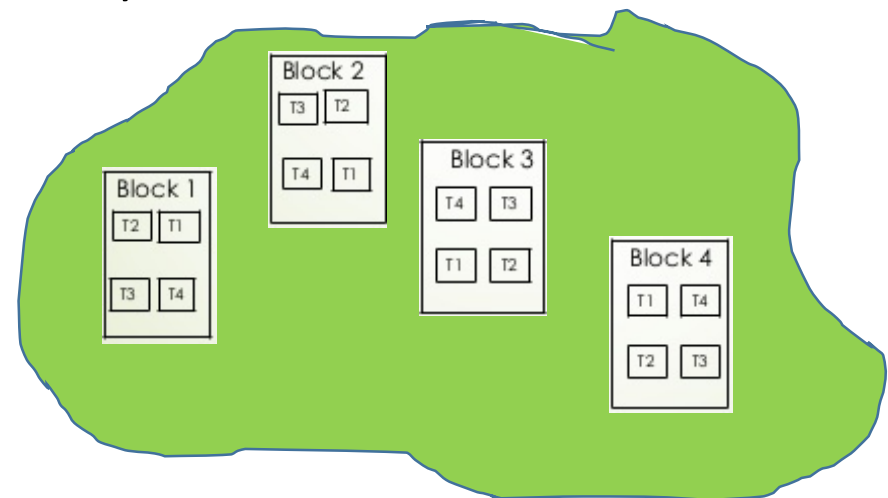
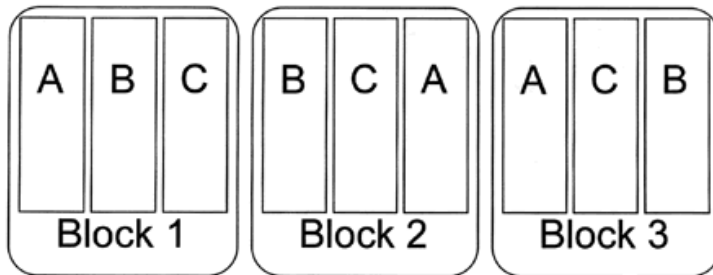
Variantou je také úplně znáhodněné uspořádání: typ zásahu na plochu určujeme náhodně (generátorem náhodných čísel) tak, aby každá plocha měla stejnou pravděpodobnost dostat kterýkoli druh zásahu.

## Analýza rozptylu - Uspořádání experimentu

### Rozmístění pokusných ploch v terénu

#### Úplné znáhodněné bloky [randomized completely blocks]

- Výhodné pro velkou plochu. Potřebný počet opakování určuje počet bloků. Umístění bloků je vybíráno tak, aby byly vnitřně maximálně homogenní a zároveň aby odlišnosti mezi bloky pokrývaly variabilitu podmínek, přes které chceme naše závěry zobecňovat.
- Toto uspořádání umožňuje dobře odlišit prostorovou variabilitu od vlivu experimentálního zásahu.
- Pro homogenní plochu (rozdíly mezi bloky minimální) se naopak snižuje síla testu, zvláště, když je počet opakování malý.



## Analýza rozptylu - Uspořádání experimentu

### Další typ uspořádání

#### Opakovaná měření [repeated measures]

- Situace, kdy na jednom subjektu opakovaně měřím/hodnotím stejnou charakteristiku.
- např. přírůstky na rostlině či zvířeti; reakce jedince na různé podněty: typ stravy, chuť vína, alergie;
- Subjekt je náhodný blok.
- Pozor na situaci, kdy jeden zásah může ovlivnit výsledek dalšího zásahu (např. při posuzování chuti vína není jedno, jestli jde suché víno po sladkém nebo naopak). Toto je třeba ošetřit např. dostatečným časovým odstupem. Nežádoucí vliv se dá snížit také větším počtem pokusných subjektů a náhodným pořadím provedených zásahů.

#### Vlastní uspořádání

Plánujete-li experiment s nestandardním uspořádáním, vždy nejdřív plán konzultujte se školitelem. Může se totiž stát, že data nepůjdou vyhodnotit žádnou statistickou metodou, nebo místo opakování naměříte tzv. *pseudoreplikace*.