

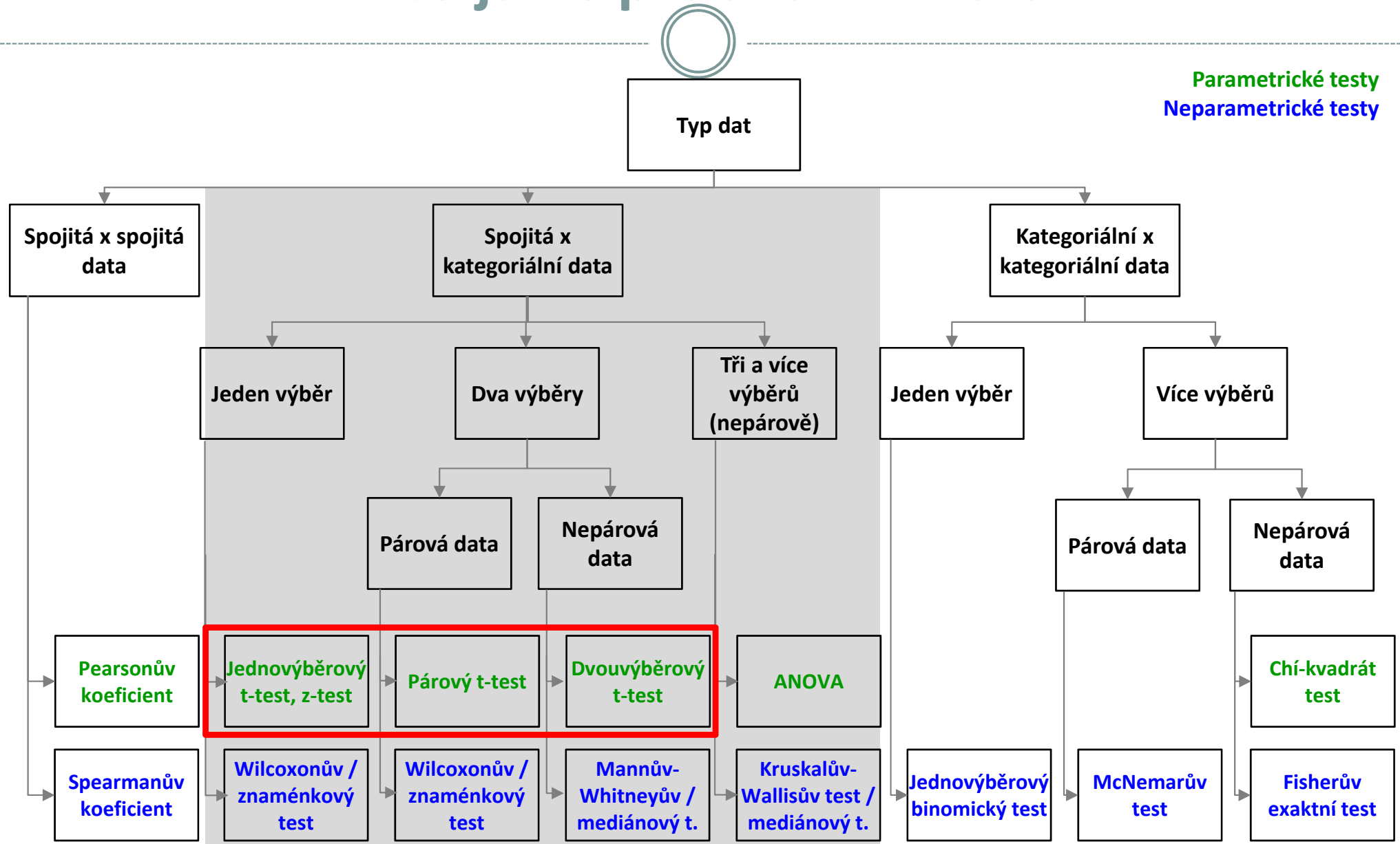
Biostatistika



Opakování Shrnutí statistických testů

Základní rozhodování o výběru statistických testů

- co jsme probírali minule



Parametrické statistické testy o parametrech tří a více výběrů

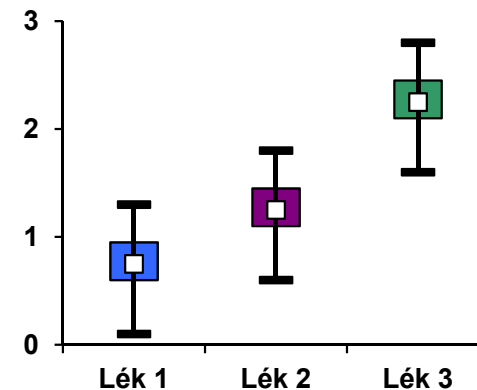
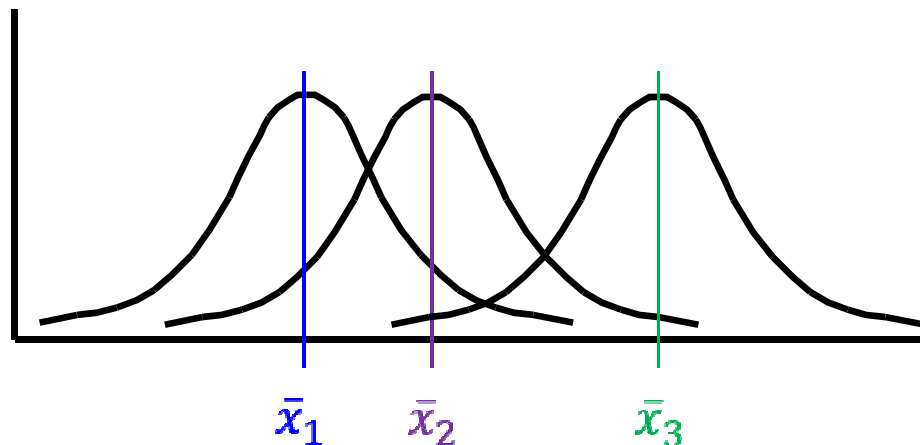


ANOVA

Analýza rozptylu (ANOVA) jednoduchého třídění



- Srovnáváme tři a více skupin dat, které jsou na sobě nezávislé (mezi objekty neexistuje vazba).
- Příklady: srovnání krevního tlaku u třech skupin pacientů léčených léky A, B a C; srovnání kognitivního výkonu podle čtyř kategorií věku.

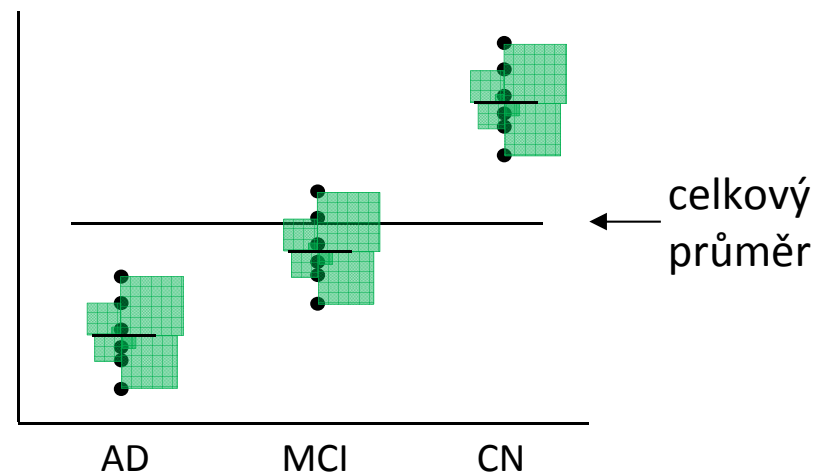
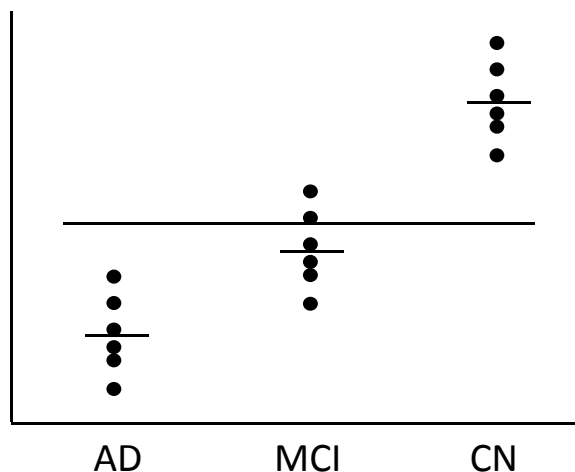


- Předpoklady: **normalita dat ve VŠECH skupinách, shodnost (homogenita) rozptylů VŠECH srovnávaných skupin**, nezávislost jednotlivých pozorování.
- Testová statistika:
$$F = \frac{S_A / df_A}{S_e / df_e}$$

ANOVA – princip



- Srovnání variability (rozptylu) mezi výběry s variabilitou uvnitř výběrů.

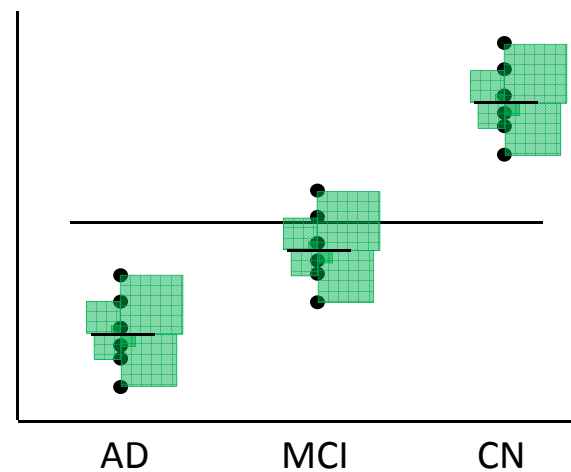
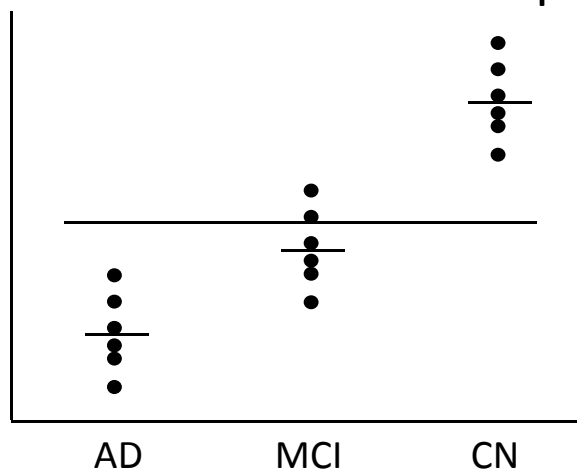


- Tabulka analýzy rozptylu jednoduchého třídění (One-Way ANOVA):

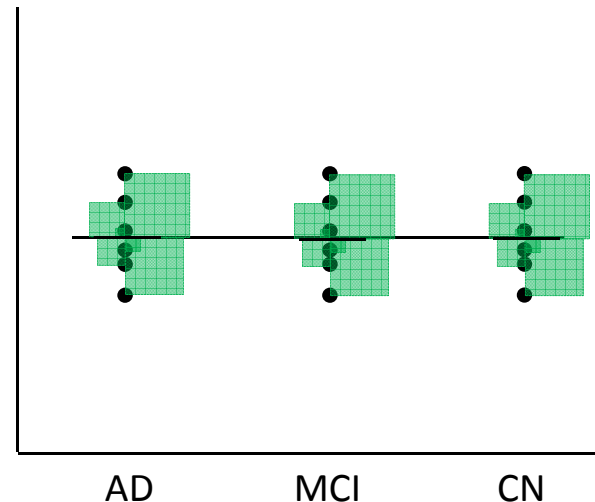
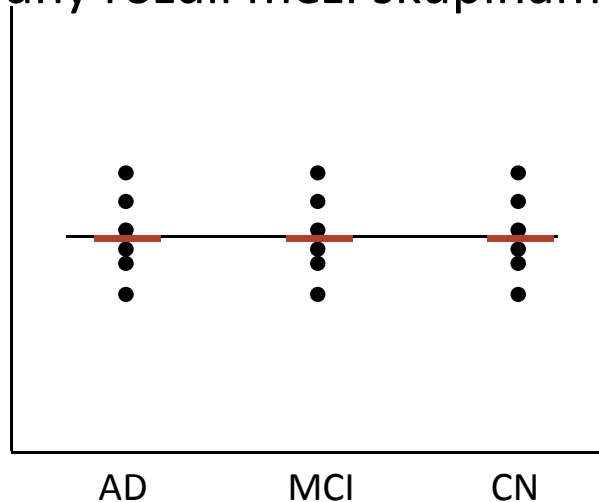
Variabilita	Součet čtverců	Počet stupňů volnosti	Průměrný čtverec	F statistika	p-hodnota
Mezi skupinami	S_A	$df_A = k - 1$	$MS_A = S_A / df_A$	$F = \frac{S_A / df_A}{S_e / df_e}$	p
Uvnitř skupin (reziduální var.)	S_e	$df_e = n - k$	$MS_e = S_e / df_e$		
Celkem	S_T	$df_T = n - 1$			

ANOVA – 2 ukázkové situace

- Rozdíl ve všech třech skupinách:



- Žádný rozdíl mezi skupinami:



ANOVA jednoduchého třídění



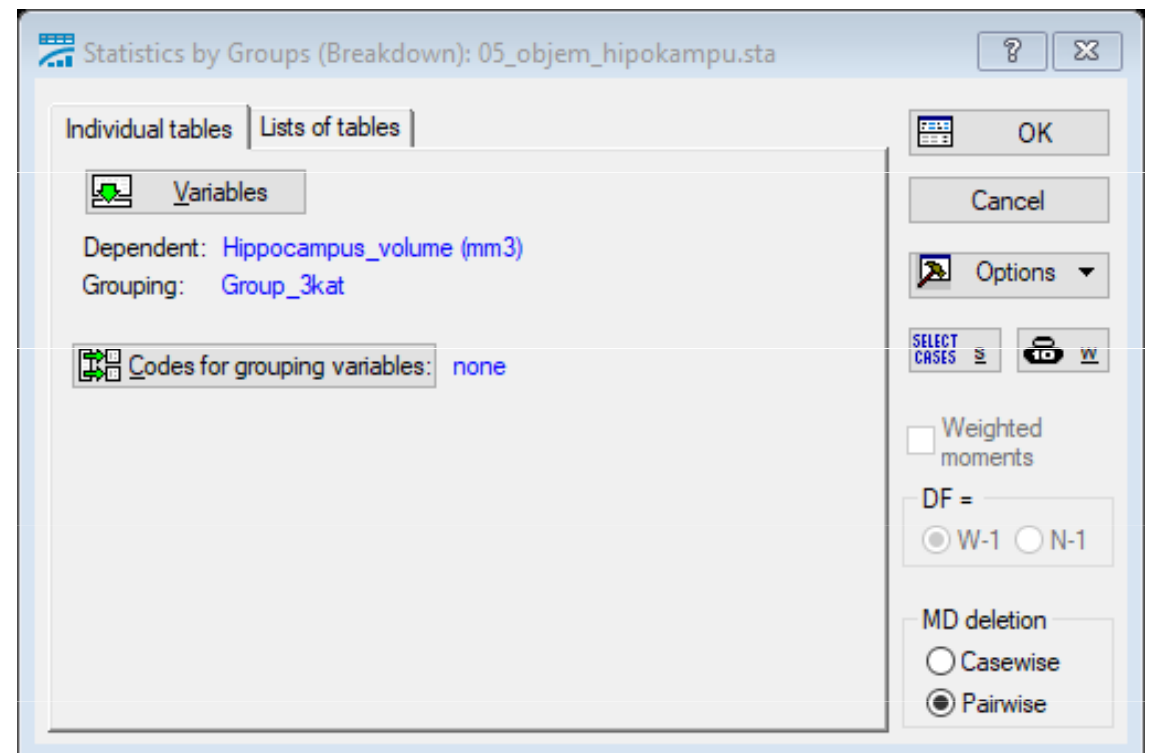
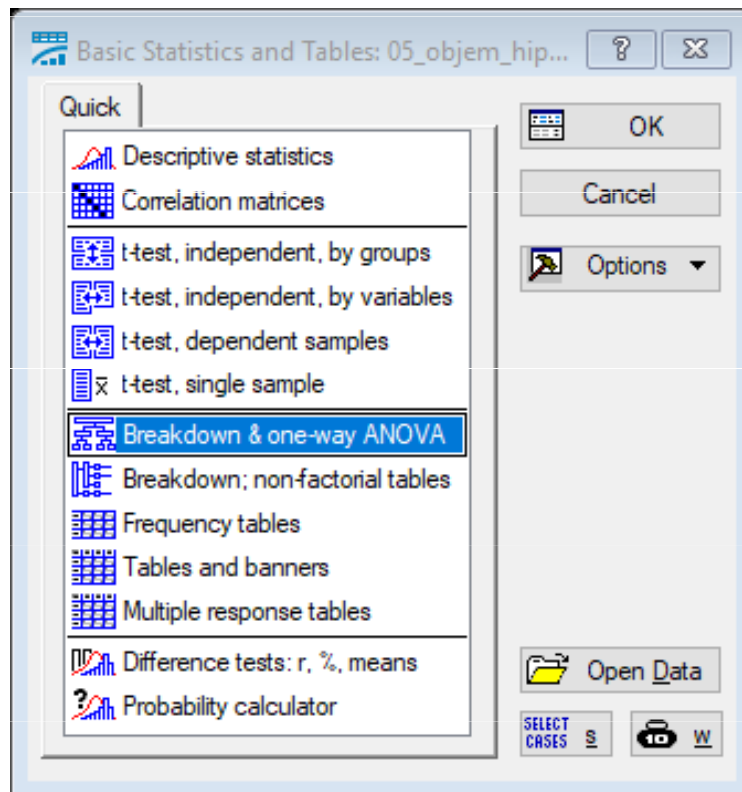
- **Příklad:** Chceme srovnat, zda se liší objem hipokampu podle typu onemocnění (3 - pacienti s AD; 2 - pacienti s MCI; 1 - zdravé kontroly).
- Tzn. hypotézy budou mít tvar: $H_0 : \mu_{AD} = \mu_{MCI} = \mu_{CN}$
 $H_1 : \text{nejméně jedno } \mu_i \text{ je odlišné od ostatních}$
- **Postup:**
 1. Popisná sumarizace objemu hipokampu podle typu onemocnění.
 2. Ověření normality hodnot ve VŠECH skupinách.
 3. Ověření shodnosti rozptylů skupin.
 4. Aplikujeme statistický test.
 5. Nulovou hypotézu zamítneme nebo nezamítneme:
P < 0,001 → zamítáme nulovou hypotézu → Rozdíl v objemu hipokampu podle typu onemocnění je statisticky významný (na hladině významnosti $\alpha = 0,05$.)

ANOVA – postup v softwaru STATISTICA



1. V menu **Statistics** zvolíme **Basic Statistics**,
vybereme **Breakdown & one-way ANOVA**

2. Zvolíme proměnné



ANOVA – postup v softwaru STATISTICA



3. Záložka *ANOVA & Tests*:

Statistics by Groups - Results: 05_objem_hipokampu.sta

DEPENDENT: 1 variable: Hippocampus_volume (mm3)

GROUPING: 1-Group_3k(3): 1 2 3

Quick | Descriptives | **ANOVA & tests** | Post-hoc

Analysis of Variance

Perform Welch's F-Test

Tests of homog. of variances

Levene tests

Brown-Forsythe tests

Categorized normal prob. plots

Categorized half-normal p-plots

Categorized detrended p-plots

Plot of means vs. std. devs

Interaction plots

Plot confidence intervals for means: 95.00 %

p-value for highlighting: .05

Summary

Cancel

Options

By Group...

ANOVA

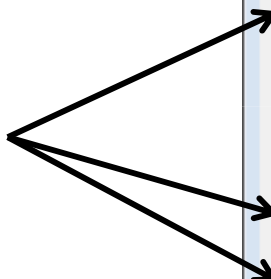
testy homogenity
rozptylů

ANOVA – postup v softwaru STATISTICA



4. Záložka *Post-hoc*:

post-hoc testy



Statistics by Groups - Results: 05_objem_hipokampu.sta

DEPENDENT: 1 variable: Hippocampus_volume (mm3)

GROUPING: 1-Group_3k(3): 1 2 3

Quick | Descriptives | ANOVA & tests | Post-hoc

Variables: Hippocampus_volume (mm3)

LSD test or planned comparison

Scheffé test

Newman-Keuls test & critical ranges

Duncan's multiple range test & critical ranges

Tukey honest significant difference (HSD)

Tukey HSD for unequal N (Spjotvoll/Stoline)

Alpha level for critical ranges: .050

p-value for highlighting: .05

Summary

Cancel

Options

By Group...

Výsledky ANOVA testu



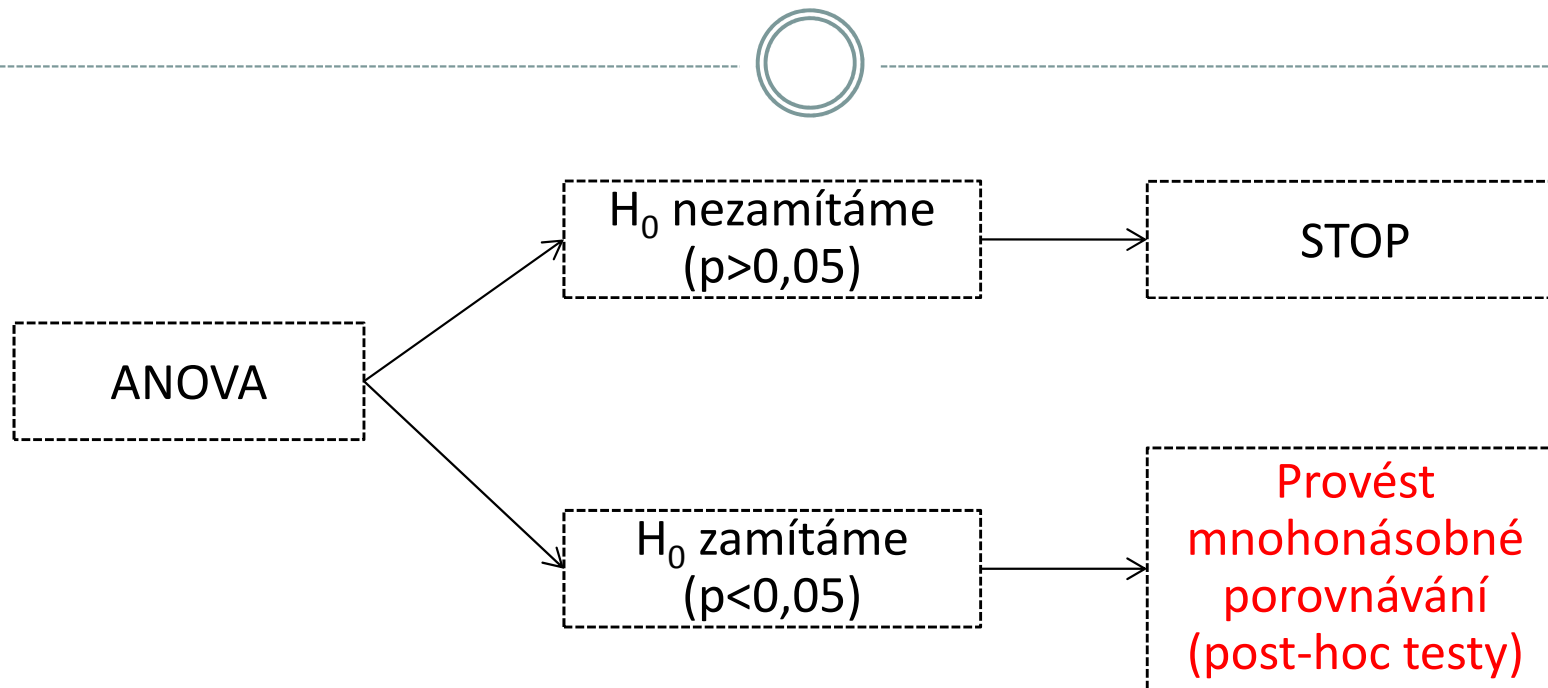
- Tabulka analýzy rozptylu jednoduchého třídění:

Variabilita	Součet čtverců	Počet stupňů volnosti	Průměrný čtverec	F statistika	p-hodnota
Mezi skupinami	$S_A =$ 71 422 222	$df_A = k - 1 =$ 2	$MS_A = S_A / df_A =$ 35 711 111	$F = \frac{S_A / df_A}{S_e / df_e} = 1103,6$	<0,001
Uvnitř skupin (reziduální var.)	$S_e =$ 26 857 142	$df_e = n - k =$ 830	$MS_e = S_e / df_e =$ 32 358		
Celkem	$S_T =$ 98 279 364	$df_T = n - 1 =$ 832			

- Výsledek ze softwaru STATISTICA:

Analysis of Variance (05_objem_hipokampu.sta)								
Marked effects are significant at p < .05000								
Variable	SS Effect	df Effect	MS Effect	SS Error	df Error	MS Error	F	p
Hippocampus_volume (mm3)	71422222	2	35711111	26857142	830	32358.00	1103.625	0.00

Další kroky analýzy



V našem příkladu $p < 0,05 \rightarrow$ provedeme post-hoc testy:

		Unequal N HSD; Variable: Hippocampus_volume Marked differences are significant at $p < .05000$		
Group_3kat		{1}	{2}	{3}
1	{1}	M=7054.3		
2	{2}		0.000022	0.000022
3	{3}		0.000022	0.000022

Poznámka



- Může nastat situace, kdy zamítneme H_0 u ANOVY, ale metodami mnohonásobného porovnávání nenajdeme významný rozdíl u žádné dvojice středních hodnot. K tomu dochází zvláště tehdy, když p-hodnota pro ANOVU je jen o málo nižší než zvolená hladina významnosti.
- Důvod: post-hoc testy (tzn. metody mnohonásobného porovnávání) mají obecně menší sílu než ANOVA, proto nemusí odhalit žádný rozdíl.