

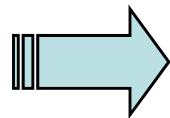
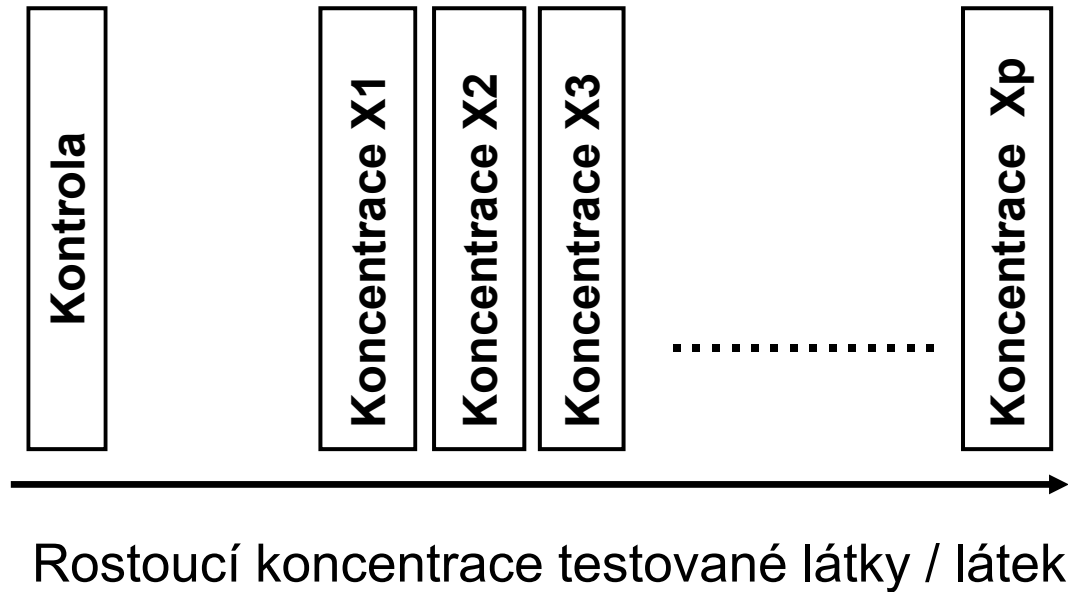


17. ANOVA – Analýza rozptylu

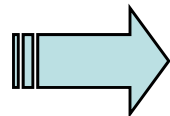


Analýza rozptylu - ANOVA

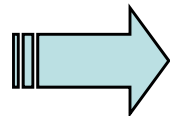
Základní technika
sloužící
k posouzení rozdílů
mezi více úrovněmi
pokusného zásahu



Celkově významné změny v reakci biologického systému



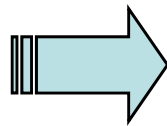
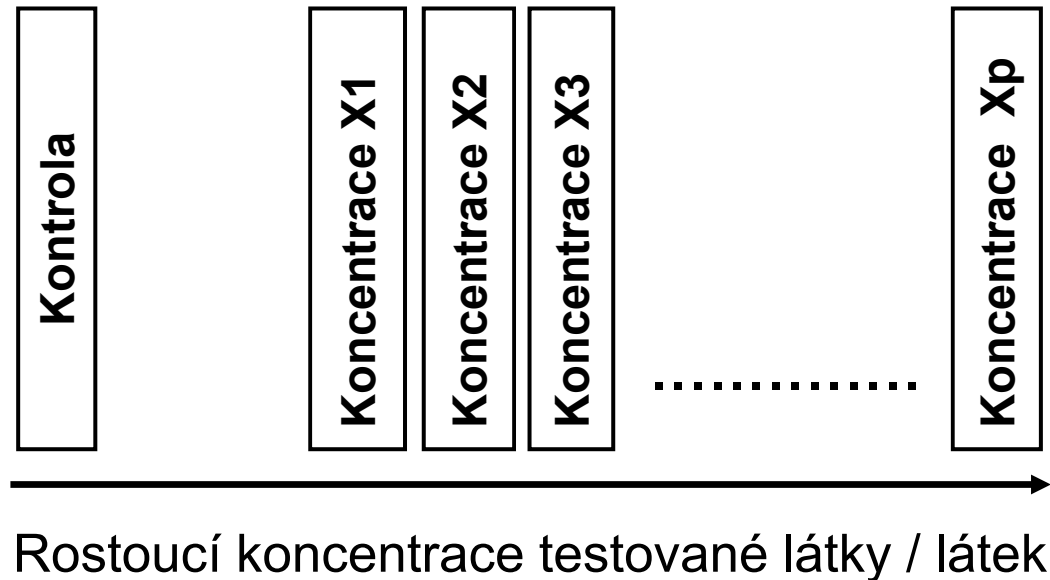
Vzájemné rozdíly účinku jednotlivých dávek



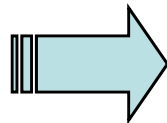
Rozdíly účinku dávek od kontroly

Analýza rozptylu - ANOVA

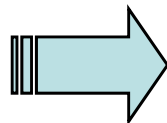
Významné kroky
analýzy, vedoucí k
efektivnímu srovnání
variant



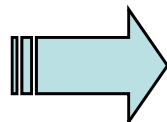
Splnění předpokladů analýzy
Transformace dat



Relevantnost kontroly
(vliv vlastní aplikace látek)



Vhodnost modelu ANOVA pro účely testu



Vlastní srovnání variant
Minimalizace chyb při ověřování hypotéz

Analýza rozptylu - ANOVA

***SPLNĚNÍ PŘEDPOKLADŮ ANOVA JE NEZBYTNOU PODMÍNKOU
POUŽITÍ TÉTO TECHNIKY***

ANOVA
= parametrická
analýza dat

1. Předpoklad nezávislosti
opakování experimentu

2. Homogenita
rozptylu v rámci
pokusných variant

3. Normalita rozložení
v rámci pokusných
variant

ALTERNATIVOU JSOU NEPARAMETRICKÉ METODY

Analýza rozptylu - ANOVA

Předpoklady analýzy rozptylu jsou nezbytné pro dosažení síly testu

• **Symetrické rozložení hodnot a normalita odchylek** od hodnoceného modelu ANOVA. Velkou část dat lze adekvátně normalizovat použitím logaritmické transformace. Předpoklad lognormální transformace může pochopitelně být teoreticky vyloučen u mnoha datových souborů obsahujících diskrétní parametry, kde je indikována vhodnost jiného typu transformace. U asymetricky rozložených a u diskrétních dat je nutné využít neparametrické alternativy analýzy rozptylu.

• **Statistická nezávislost reziduí** vyhodnocovaného modelu ANOVA. Pokud odhad a posouzení korelačních vztahů mezi pokusnými variantami není přímo předmětem výzkumu, lze jejich vliv na vyhodnocení odstranit znáhodněním dat v rámci pokusných variant - tedy změnou pořadí v náhodné. Rozsah vlivu těchto autokorelačních vztahů musí být ovšem primárně omezen správností experimentálního uspořádání.

• **Homogenita rozptylu** je nutným předpokladem pro smysluplnost vzájemných srovnání pokusných variant. U testů toxicity by splnění tohoto předpokladu mělo být ověřováno (Bartlettův test), neboť vážné rozdíly (až řádové) v jednotkách testovaného parametru mohou nastat v důsledku inhibice dávkami látky. Nehomogenita rozptylu je často ve vztahu k nenormalitě (asymetrii) dat a lze ji odstranit vhodnou normalizující transformací.

• **Aditivita** jako předpoklad týkající se složitějších experimentálních uspořádání. Exaktní otestování aditivity více pokusných faktorů je procedura poměrně náročná na experimentální design vyvážený co do počtu opakování. Je rovněž obtížné testovat interakci na nestandardních datech, neboť případná transformace může změnit charakter odchylek původních dat od hodnoceného modelu ANOVA.

Analýza rozptylu - ANOVA

Omezení aplikace ANOVA lze řešit

• **Chybějící data.** Vážným problémem jsou chybějící údaje o celé skupině kombinací testovaných látek, například u faktoriálních pokusů, kdy je znemožněno hodnocení experimentu jako celku.

• **Různé počty opakování** Jde o typický jev pro experimentální datové soubory. Při různých počtech opakování v experimentálních variantách jsou testy ANOVA citlivější na nenormalitu dat. Pokud jsou počty opakování zcela odlišné (až na řádové rozdíly), je nutno použít neparametrické techniky nebo analýzu rozptylu nevyvážených pokusů.

• **Odlehlé hodnoty.** Ojedinelé odlehlé hodnoty musí být před parametrickou analýzou rozptylu vyloučeny.

• **Nedostatek nezávislosti mezi rezidui modelu.** Jde o závažný nedostatek, zkreslující výsledek F-testu. Velmi často je tato skutečnost důsledkem špatného provedení nebo naplánování experimentu.

• **Nehomogenita rozptylu.** Velmi častý nedostatek experimentálních dat, často související s nenormalitou rozložení nebo s odlehlými hodnotami.

• **Nenormalita dat.** I v tomto případě lze situaci upravit vyloučením odlehlých hodnot nebo normalizující transformací.

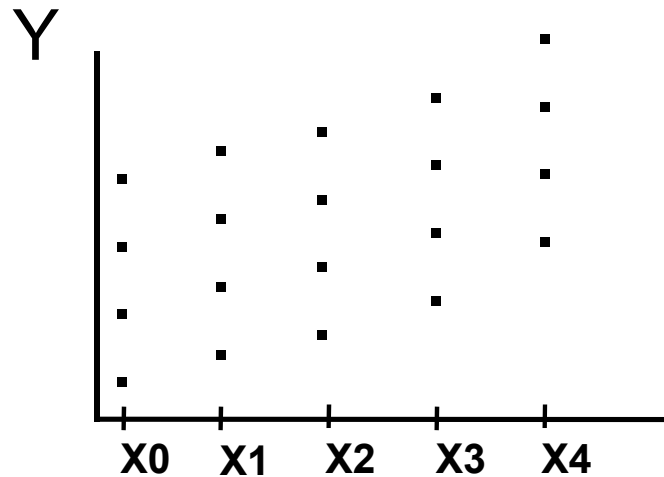
• **Neaditivita kombinovaného vlivu více pokusných zásahů.** Tuto situaci lze testovat jednak speciálními testy aditivity nebo přímo F testem kontrolujícím významnost vlivu interakce pokusných zásahů. Při významné interakci je nutné prozkoumat především její charakter ve vhodném experimentálním uspořádání.

Modely analýzy rozptylu

Model I. Pevný model

X_0	X_1	X_2	X_3	X_4
.
.
.
.
.
.
.
.
.

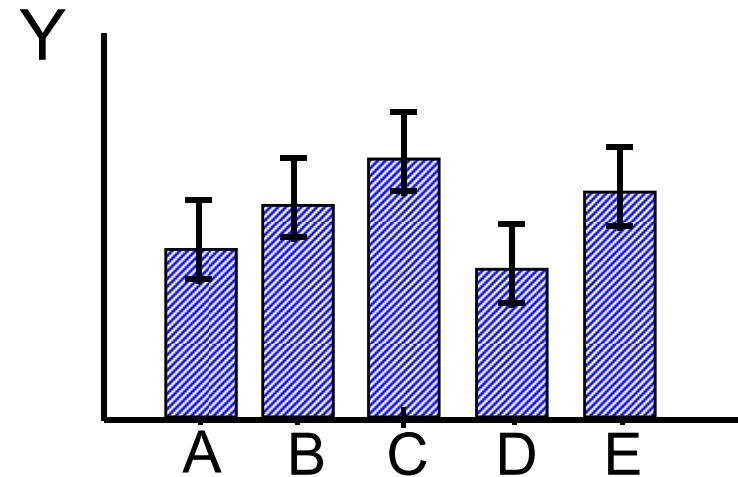
$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$



Model II. Náhodný model

A	B	C	D	E
.
.
.
.
.
.
.
.
.
.

$$y_{ij} = \mu + A_i + \varepsilon_{ij}$$



ANOVA – základní výpočet

- Základním principem ANOVY je porovnání rozptylu připadajícího na:
 - Rozdělení dat do skupin (tzv. effect, variance between groups)
 - Variabilitu objektů uvnitř skupin (tzv. error, variance within groups), předpokládá se, že jde o náhodnou variabilitu (=error)

1. Variabilita mezi skupinami

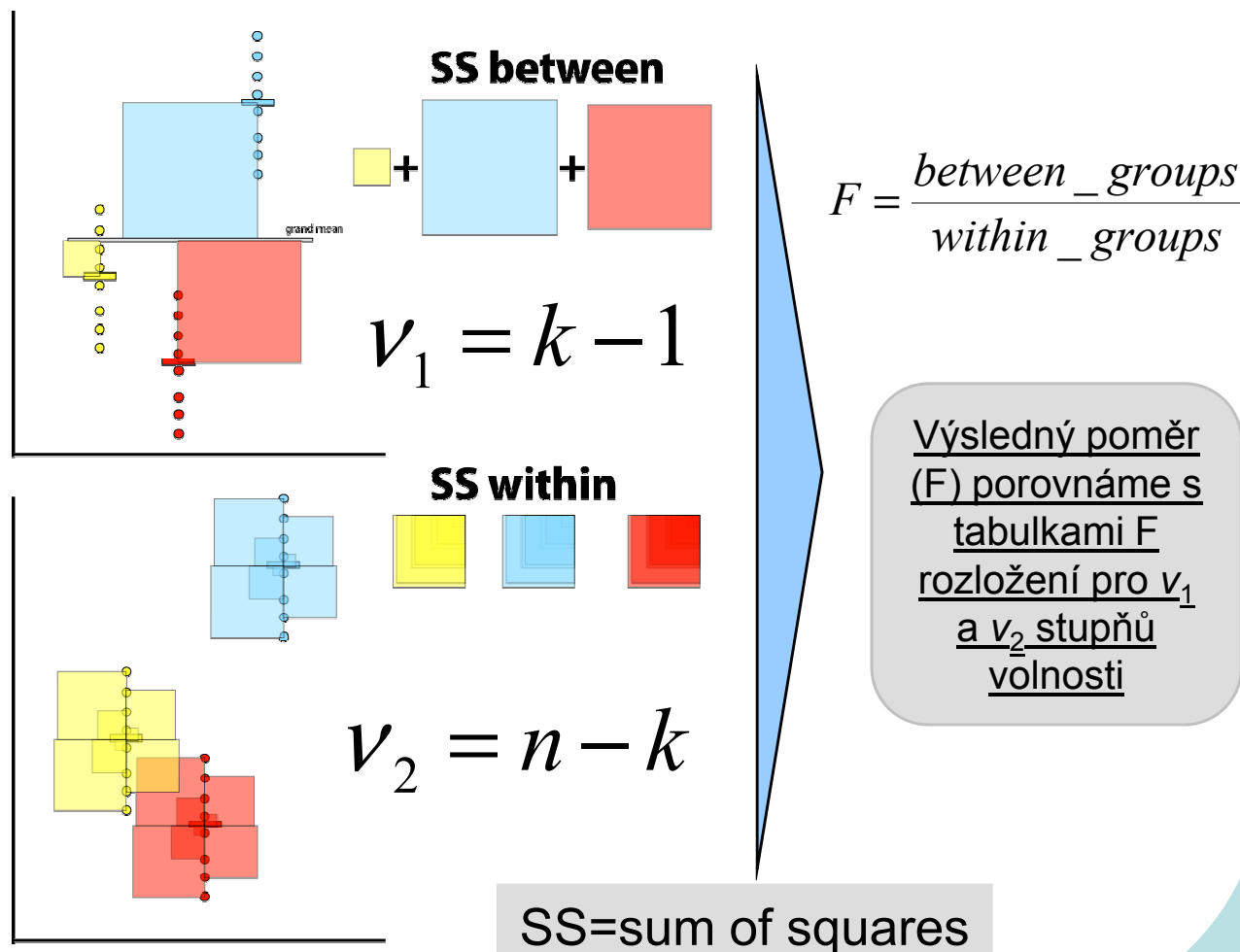
Rozptyl je počítán pro celkový průměr (tzv. grand mean) a průměry v jednotlivých skupinách dat

Stupně volnosti jsou odvozeny od počtu skupin (= počet skupin -1)

2. Variabilita uvnitř skupin

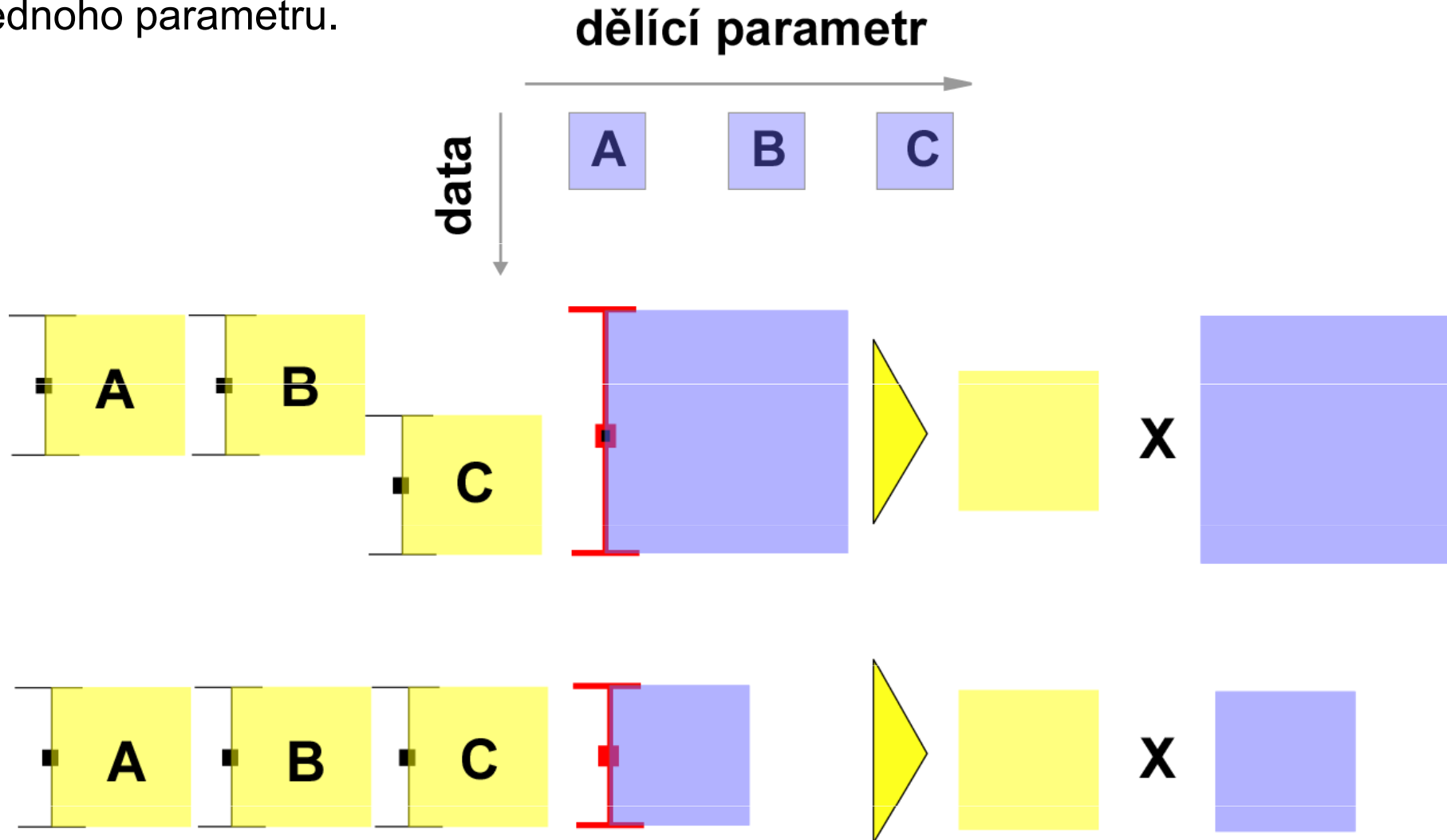
Rozptyl je počítán pro průměry jednotlivých skupin a objekty uvnitř příslušných, celková variabilita je pak sečtena pro všechny skupiny

Stupně volnosti jsou odvozeny od počtu hodnot (= počet hodnot - počet skupin)



Jednoduchý ANOVA design

Nejjednodušším případem ANOVA designu je rozdělení na skupiny podle jednoho parametru.



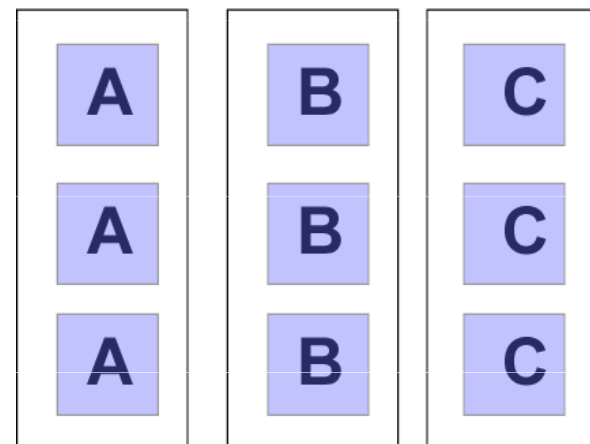
Nested ANOVA

- Rozdělení skupin na náhodné podskupiny (např. opakování experimentu)
- Cílem je zjistit, zda data v jedné skupině nejsou pouhou náhodou
- Nejprve je testována shoda podskupin v hlavních skupinách,
 - pokud jsou shodné, je vše v pořádku
 - pokud nejsou, stále lze zjišťovat, zda se variabilita uvnitř hlavních skupin liší od celkové variability

jednoduchá ANOVA



nested ANOVA

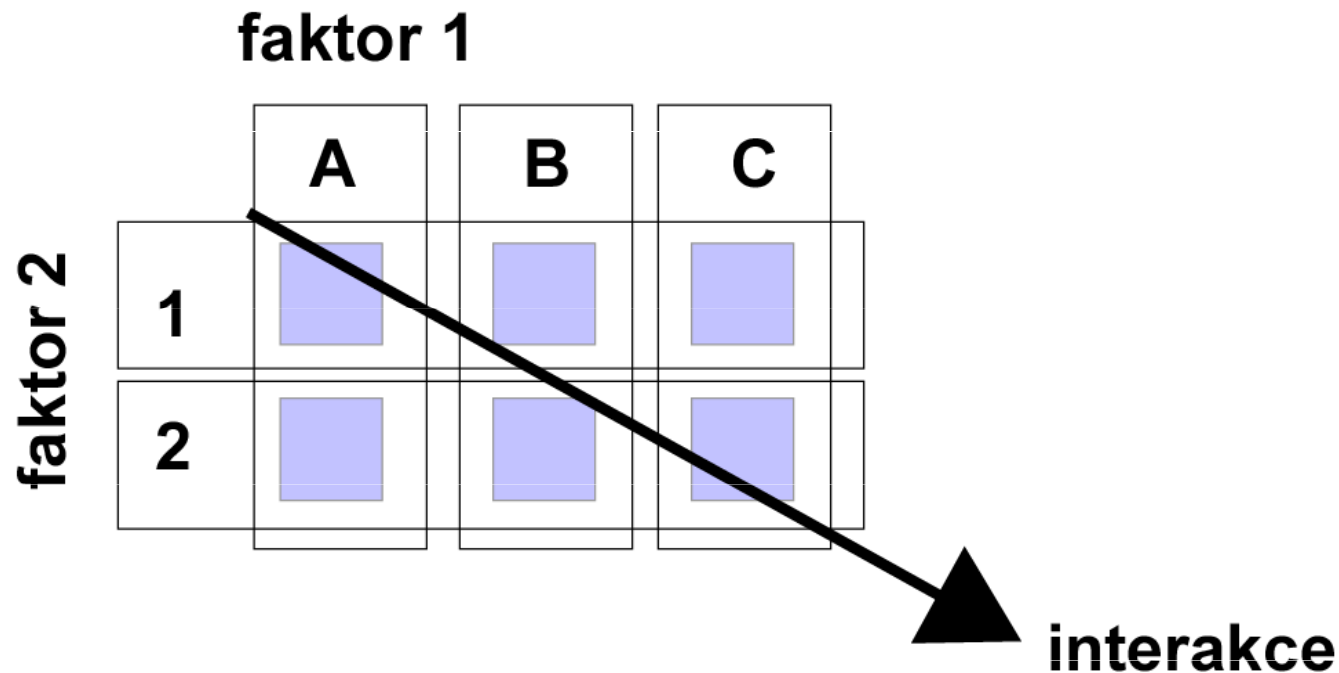


Two way ANOVA

Pro rozdělení do kategorií je zde více parametrů

Na rozdíl od nested ANOVY nejde o náhodná opakování experimentu, ale o řízené zásahy (např. vliv pH a koncentrace O_2)

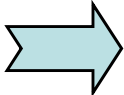
Kromě vlivu hlavních faktorů se uplatňuje i jejich interakce



Modely analýzy rozptylu - základní výstup

*Základním výstupem analýzy rozptylu je
Tabulka ANOVA - frakcionace komponent rozptylu*

Zdroj rozptylu	St. v.	SS	MS	F
Pok. zásah (mezi skupinami)	a - 1	SS_B	$SS_B/(a - 1)$	MS_B/MS_E
Uvnitř skupin	N - a	SS_E	$SS_E/(N - a)$	
Celkem	N - 1	SS_T		

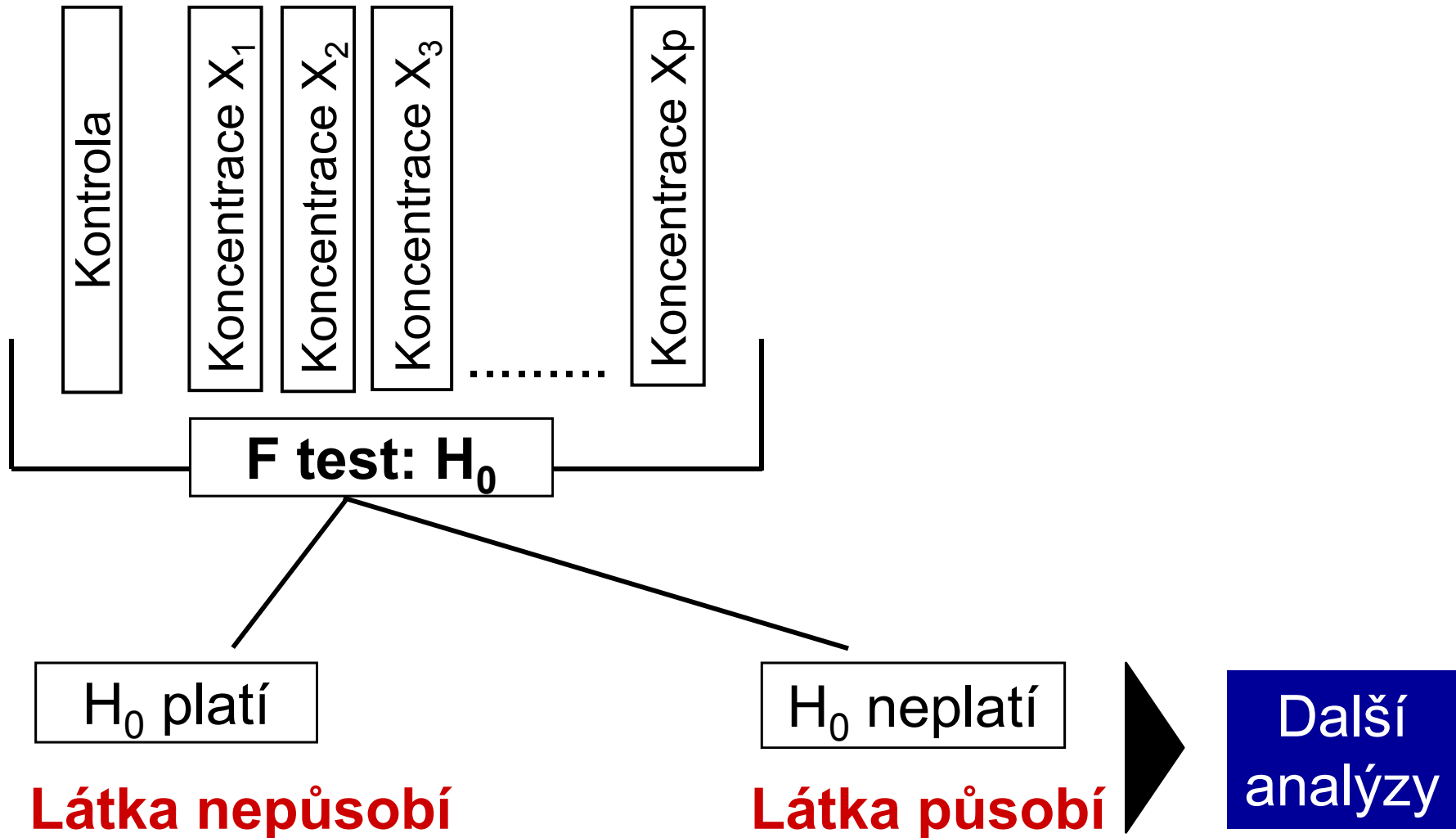
SS_B/SS_T  Kvantifikovaný podíl rozdílu mezi pokusnými zásahy na celkovém rozptylu

MS_B/MS_T  Statistická významnost rozdílu

Analýza rozptylu - obecný F test

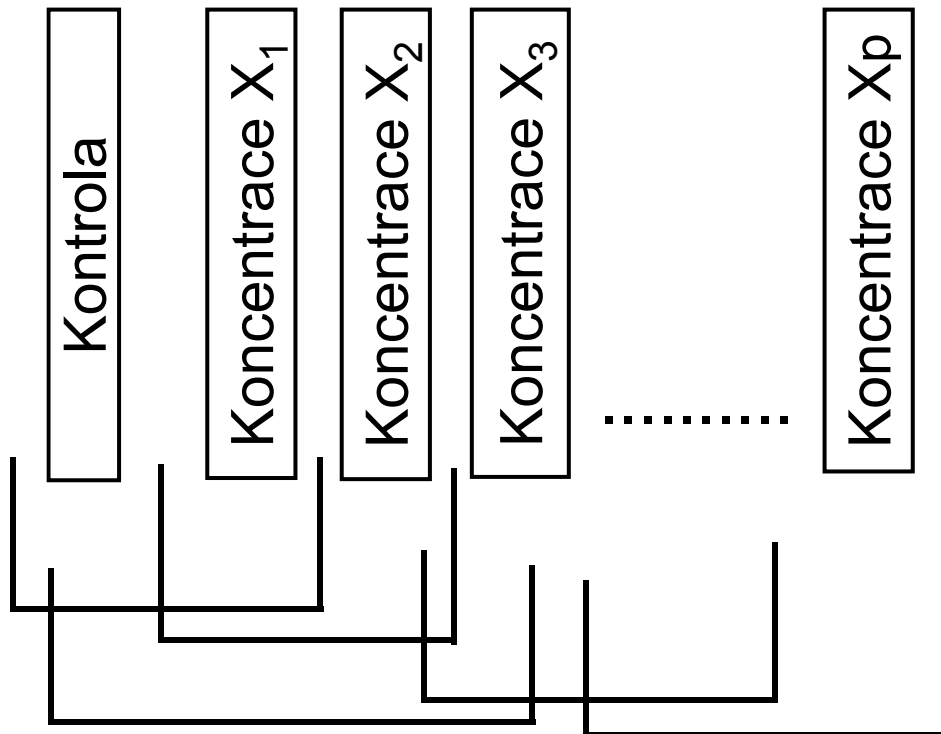
obecný F test

$$H_0: m_1 = m_2 = m_3 = \dots = m_p$$



Analýza rozptylu - Testy kontrastů

ANOVA: H_0 zamítnuta
Testy kontrastů



Rozdíly v smysluplných kombinacích ?

Plánované

Neplánované

Pro srovnání variant s kontrolou

Testování kontrastů
"Multiple range testy"

Parametrické

Neparametrické

Příklad: Anova - One way

Dávka rostlinného stimulantu (0, 4, 8, 12 mg/l)

$A = 4 ; n = 8$

I. ANOVA

Bartlett's test: $P = 0,9847$

K-S test: $P = 0,482 - 0,6525$ pro jednotlivé kategorie

Source	D. f.	SS	MS	F
Between Groups	3	305,8	101,9	8,56
Within Groups	28	322,2	11,9	
Total (corr.)	31	638,0		

II. Multiple Range Test

NKS -test

Level	Average	Homogenous Groups
0	34,8	x
4	41,4	x
12	41,8	x
8	52,6	x

Příklad: Anova - One way

I. Zásah: 4 klinická stadia virové choroby (napadá kr. buňky)

Sledovaná veličina: aktivita enzymu v těchto krevních buňkách

$$H_o : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

n = 3
MODEL = ?

	I	II	III	IV
	22,8	16,4	11,2	14,2
	19,4	17,8	18,2	10,1
	12,5	19,1	15,8	12,8
Σ	65,7	53,3	45,2	37,1
průměr	21,9	17,8	15,1	12,4

II.

Source	D.f.	MS	F	P
Between groups	3	49,6	8,39	0,0075
Within groups	8	5,9		
Total (corr.)	11	-		

III. Komponenta rozptylu:

$$\sigma_A^2 \sim S_A^2 = \frac{MS_A - MS_e}{n} = \frac{49,6 - 5,9}{3} = 14,57$$

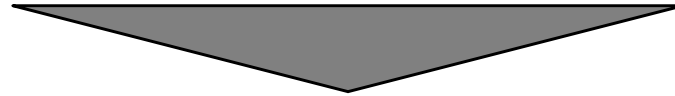
$$S_A^2 = 2,5 \cdot S_e^2$$

IV.

$$\rho_I \sim r_I = \frac{S_A^2}{S_A^2 + S_e^2} = 0,7142$$

Srovnání variant v testech

Srovnávání variant po celkovém testu ANOVA



Mnoho existujících algoritmů není vhodných pro konkrétní případ

Day and Quin
Ecological Monographs, 1989

Test	Využití	Poznámka
Dunnett Williams	Srovnání s kontrolou	Ex. i modifikace pro různá n.
ANOVA testy (F)	Orthogonální kontrasty	Plánovaná srovnání
Ryan Q test	Jednoduché kontrasty	Vyhodnocen jako nejlepší test

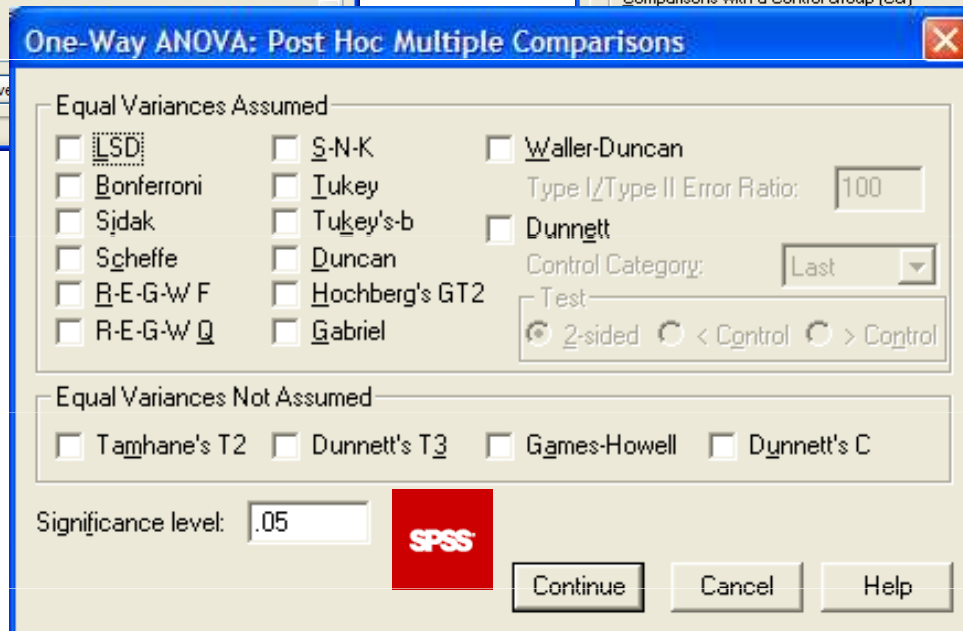
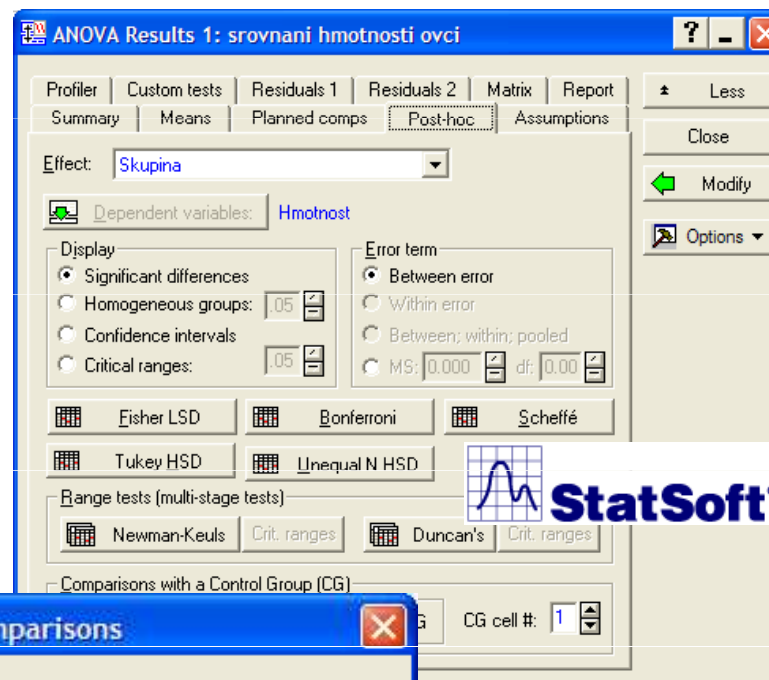
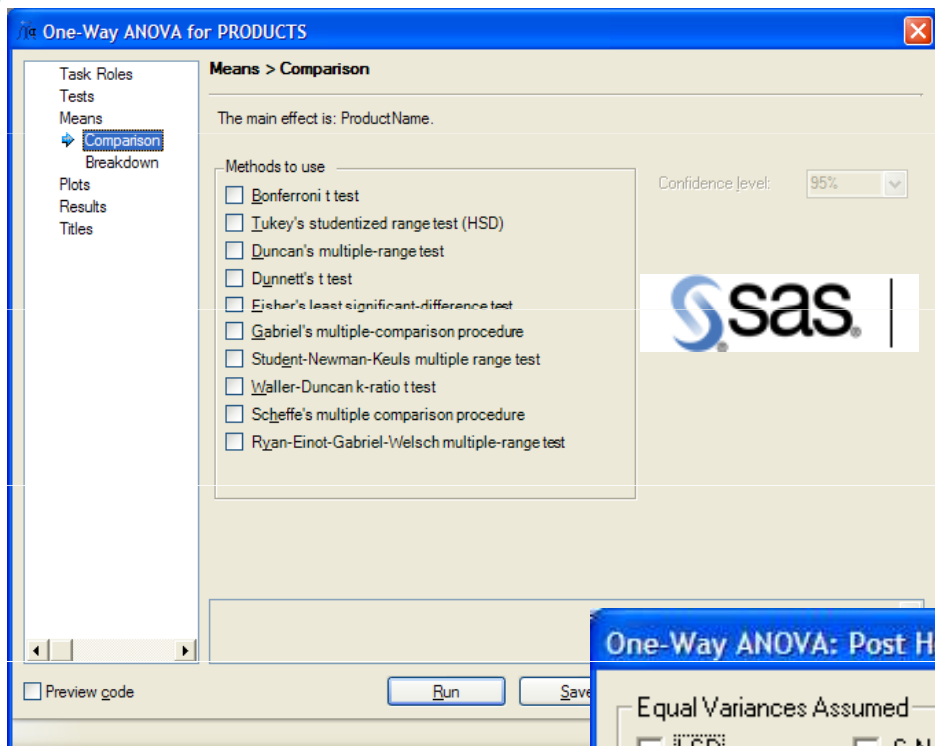
Testy pro jednoduché kontrasty

Scheffe	Tukey	LSD
Bonferroni	Dunn-Sidák	Kramer

Testy nevhodné

Duncan	Student - Newmann-Keuls	Waller-Duncan k ratio
--------	-------------------------	-----------------------

Řada post-hoc testů v různých SW



Hypotetické příklady - Multiple Range Tests

	<u>Level</u>	<u>Homogenous Group</u>
15	1	x
18	2	xx
22	3	xx
26	4	x
38	5	x

	<u>Level</u>	<u>Homogenous Group</u>
15	1	x
22	2	x
24	3	xx
29	4	x
30	5	x

	<u>Level</u>	<u>Homogenous Group</u>
15	1	x
18	2	xx
22	3	x
29	4	x
36	5	x