ANOVA & spol.

11. 11. 2015

Jan Šerek

PSY252 Statistická analýza dat II

Program dnešní přednášky

- jednofaktorová (one-way) ANOVA
- faktoriální (two…-way) ANOVA
- ANCOVA (ANOVA s kovariáty)
- MANOVA (ANOVA s více závislými)

ANOVA (analysis of variance)

- Liší se 2 skupiny v průměrné hodnotě nějaké proměnné? → **t-test**
- Liší se 3 (a více) skupiny v průměrné hodnotě nějaké proměnné? → **ANOVA**
  - „Liší se děti z úplných rodin, neúplných rodin a náhradní péče ve své partnerské spokojenosti?“
  - „Liší se průměrná tepová frekvence participantů, kteří byli vystavení podnětu A, podnětu B a žádnému podnětu (kontrolní skupina)?“

ANOVA (analysis of variance)

- Liší se 2 skupiny v průměrné hodnotě nějaké proměnné? → **t-test**
- Liší se 3 (a více) skupiny v průměrné hodnotě nějaké proměnné? → **ANOVA**
  - „Liší se děti z úplných rodin, neúplných rodin a náhradní péče ve své partnerské spokojenosti?“
  - „Liší se průměrná tepová frekvence participantů, kteří byli vystavení podnětu A, podnětu B a žádnému podnětu (kontrolní skupina)?“

1 nezávislá kategorická → 1 závislá intervalová

ANOVA (analysis of variance)

- Liší se 2 skupiny v průměrné hodnotě nějaké proměnné? → **t-test**
- Liší se 3 (a více) skupiny v průměrné hodnotě nějaké proměnné? → **ANOVA**
  - „Liší se děti z úplných rodin, neúplných rodin a náhradní péče ve své partnerské spokojenosti?“
  - …vá frekvence participantů, kteří … A, podnětu B a žádnému … pina)?“

v jazyku ANOVY se tato nezávislá kategorická proměnná nazývá **faktor**, který má určité **úrovně**

1 nezávislá kategorická → 1 závislá intervalová

ANOVA

ANOVU v zásadě provádíme ve 2 krocích:

KROK 1: **Existuje mezi skupinami nějaká odlišnost?**

– spočítáme statistiku F a testujeme, zda překračuje kritickou hodnotu

Pokud NE → konec

Pokud ANO → KROK 2: **Mezi jakými skupinami konkrétně tato odlišnost existuje?**

– máme o tom hypotézy → plánované kontrasty

– nemáme o tom hypotézy → post-hoc testy

ANOVA jako regrese

Hodnota závislé proměnné člověka *i*

$$Y_i = \mu + \varepsilon_i$$

Nevysvětlená individuální variabilita

Celkový průměr

ANOVA jako regrese

Hodnota závislé proměnné člověka *i*

$$Y_i = \mu + \varepsilon_i$$

Nevysvětlená individuální variabilita

Celkový průměr

$$Y_{ij} = \mu + \color{red}{\alpha_j} + \varepsilon_{ij}$$

Vliv toho, že je člověk členem skupiny *j*

ANOVA jako regrese

Hodnota závislé proměnné člověka *i*

$$Y_i = \mu + \varepsilon_i$$

Nevysvětlená individuální variabilita

Celkový průměr

$$Y_{ij} = \mu + {\color{red}\alpha_j} + \varepsilon_{ij}$$

Podstata ANOVY

Jak dobře je závislá proměnná vysvětlena modelem, který předpokládá odlišnost skupin ($\alpha \neq 0$)? Nepostačí nám stejně dobře model, který předpokládá, že se skupiny neliší?

ANOVA jako regrese

Souvisí socioekonomický status rodiny s tím, jak často dítě používá internet?

Nezávislá kategorická proměnná: **socioekonomický status**

3 hodnoty: nízký, střední, vysoký

Závislá intervalová proměnná: **frekvence používání internetu**

Liší se děti z rodin s nízkým, středním a vysokým SES v tom, jak často používají internet?

ANOVA jako regrese

$[\text{inter}]_i = [\text{průměrný inter}] + \varepsilon_i$

ANOVA jako regrese

$[\text{inter}]_i = [\text{průměrný inter}] + \varepsilon_i$

$[\text{inter}]_i = [\text{průměrný inter}] + \mathbf{b[ses]} + \varepsilon_i$

ANOVA jako regrese

$[\text{inter}]_i = [\text{průměrný inter}] + \varepsilon_i$

$[\text{inter}]_i = [\text{průměrný inter}] +$ **b[ses]** $+ \varepsilon_i$

Každá kategorická proměnná o *k* hodnotách (úrovních) může být vyjádřena souborem *k-1* binárních dummy proměnných.

3 typy SES→ 2 binární proměnné **vys** a **str**

vys = 1 & **str** = 0 → vysoký SES

vys = 0 & **str** = 1 → střední SES

vys = 0 & **str** = 0 → nízký SES

ANOVA jako regrese

$[\text{inter}]_i$ = [průměrný inter] + ε_i

$[\text{inter}]_i$ = [průměrný inter] + **b[ses]** + ε_i

Každá kategorická proměnná o *k* hodnotách (úrovních) může být vyjádřena souborem *k-1* binárních dummy proměnných.

3 typy SES→ 2 binární proměnné **vys** a **str**

vys = 1 & **str** = 0 → vysoký SES
vys = 0 & **str** = 1 → střední SES
vys = 0 & **str** = 0 → nízký SES

$[\text{inter}]_i = \mathbf{b_0} + \mathbf{b_1[vys]} + \mathbf{b_2[str]} + \varepsilon_i$

ANOVA jako regrese

$[\text{inter}]_i$ = [průměrný inter] + ε_i

$[\text{inter}]_i$ = [průměrný inter] + **b[ses]** + ε_i

Každá kategorická proměnná o *k* hodnotách (úrovních) může být vyjádřena souborem *k-1* binárních dummy proměnných.

3 typy SES→ 2 binární proměnné **vys** a **str**

Průměrná frekvence dětí z rodin s nízkým SES

O kolik se liší průměrná frekvence dětí z rodin s vysokým SES

O kolik se liší průměrná frekvence dětí z rodin se středním SES

$[\text{inter}]_i = \mathbf{b_0} + \mathbf{b_1[vys]} + \mathbf{b_2[str]} + \varepsilon_i$

ANOVA jako regrese

Jestliže $\mathbf{b_1} = 0$ a $\mathbf{b_2} = 0$, znamená to, že SES nemá žádný vliv a všechny skupiny mají stejnou průměrnou frekvenci.

Potom by nám postačil základní model predikující frekvenci pouze z celkové průměrné frekvence a nevysvětlitelné individuální variability.

Vysvětlí nám model předpokládající nenulové $\mathbf{b_1}$ a/nebo $\mathbf{b_2}$ něco navíc?

Průměrná frekvence dětí z rodin s nízkým SES

O kolik se liší průměrná frekvence dětí z rodin s vysokým SES

O kolik se liší průměrná frekvence dětí z rodin se středním SES

$$[\text{inter}]_i = \mathbf{b_0} + \mathbf{b_1[vys]} + \mathbf{b_2[str]} + \varepsilon_i$$

ANOVA – statistika F

Model sum of squares

SS_M – kolik variability závislé proměnné dokáže vysvětlit model, který předpokládá odlišnost skupin (tj. že záleží na členství ve skupině)

SS_M = ∑ velikost skupiny *j* * (průměr skupiny *j* – celkový průměr)^2

Mean squares: **$MS_M = SS_M / df_M$**

df_M = (počet skupin – 1)

Resudial sum of squares

SS_R – kolik variability závislé proměnné zůstává nevysvětleno tímto modelem

SS_R = ∑ (hodnota člověka *i* ze skupiny *j* – průměr skupiny *j*)^2

Mean squares: **$MS_R = SS_R / df_R$**

df_R = (celkový počet lidí – počet skupin)

Pozn.: Total sum of squares: $SS_T = SS_M + SS_R$ = celkový rozptyl * (celkový počet lidí – 1)

ANOVA – statistika F

$F = MS_M / MS_R$

- poměr toho, co model vysvětlit dokáže, ku tomu, co vysvětlit nedokáže
- čím vyšší F, tím více záleží na rozdělení lidí do jednotlivých skupin, **tj. tím více se skupiny od sebe liší v závislé proměnné**
- jde o výběrovou statistiku, která má specifické rozložení, definované dvojicí stupňů volnosti (df_M, df_R)
- můžeme určit kritickou hodnotu (na určité hladině významnosti) a testovat, zda ji hodnota F v našem výzkumu překračuje, **tj. testovat statistickou významnost nalezených rozdílů mezi skupinami**

ANOVA – předpoklady

- **nezávislost pozorování** (→ ANOVA pro opakovaná měření)
- **normalita rozložení** (v rámci každé skupiny)
  - narušení nevadí, pokud jsou skupiny stejně velké + mají velikost alespoň okolo 30
  - neparametrická alternativa – Kruskal-Wallisův test
- **homogenita rozptylů** (skupiny mají stejné rozptyly)
  - Levenův test – chceme, aby byl nesignifikantní
  - $s^2_{max} / s^2_{min} < 3$
  - narušení by nemělo vadit, pokud jsou skupiny stejně velké
  - při narušení lze použít Welchovo F

ANOVA – SPSS

Analyze → Compare Means → One-Way ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	SS_M	df_M	MS_M		
Within Groups	SS_R	df_R	MS_R		
Total	SS_T	$df_M + df_R$			

ANOVA

Máme hypotézy o konkrétních rozdílech mezi skupinami.

H1: Děti z rodin s nízkým SES používají internet méně často než ostatní děti.

H2: Děti z rodin se středním SES používají internet méně často než děti z rodin s vysokým SES.

ANOVA – plánované kontrasty

- umožňují porovnat jednotlivé skupiny v jednom kroku bez nutnosti korigovat hladinu významnosti (a snižovat tak sílu testu)
- jen když máme dopředu hypotézy
- kontrastů lze provést tolik, kolik je **počet skupin – 1**
- každý kontrast srovnává 2 průměry
  - průměr skupiny nebo průměr více skupin dohromady
  - např. **NÍZ** vs. **STŘ+VYS** nebo **STŘ** vs. **VYS**
- **ortogonální** (nezávislé) **kontrasty**
  - skupina použitá v jednom srovnání není použitá v dalším
- **neortogonální kontrasty**

ANOVA – plánované kontrasty

- zkoumáme, zda daný kontrast (rozdíl mezi dvěma průměry) signifikantně přispívá k variabilitě vysvětlené modelem (SS_M)
- abychom to zjistili, jakoby překódujeme hodnoty dummy proměnných, aby odhadnuté parametry (b_1, b_2 atd.) odrážely požadované kontrasty

$$[\text{inter}]_i = \mathbf{b_0} + \mathbf{b_1[vys]} + \mathbf{b_2[str]} + \varepsilon_i$$

$$[\text{inter}]_i = \mathbf{b_0} + \mathbf{b_1[kontrast1]} + \mathbf{b_2[kontrast2]} + \varepsilon_i$$

Kategorie	Kontrast 1 NÍZ vs. STŘ+VYS	Kontrast 2 STŘ vs. VYS
Vysoký SES	**1**	**-1**
Střední SES	**1**	**1**
Nízký SES	**-2**	**0**

ANOVA – plánované kontrasty

- zkoumáme, zda daný kontrast (rozdíl mezi dvěn gnifikantně přispívá k variabilitě vysvětlené modelem (SS_M)
- abychom to zjistili, jakoby překódujeme hodnot něnných, aby odhadnuté parametry (b_1, b_2 atd.) odrážely pož asty

Součet pro každý kontrast musí být 0

$[inter]_i$ = b_0 + tr] + ε_i

$[inter]_i$ = ntrast1] + b_2[kontrast2] + ε_i

Srovnávané skupiny musí mít odlišná znaménka

Skupina, kterou nechceme zahrnout → 0

Kategorie	Kontrast 1 NÍZ vs. STŘ+VYS	Kontrast 2 STŘ vs. VYS
Vysoký	1	-1
Střední	1	1
Nízký S	-2	0

Skupiny brané dohromady musí mít stejné číslo

ANOVA – post-hoc testy

- používáme, pokud nemáme dopředu jasné hypotézy
- srovnávají vše se vším – každou skupinu s každou
- mají v sobě mechanismy zohledňující zvýšené riziko chyby I. typu
- z principu jsou oboustranné
- je jich mnoho – liší se v několika parametrech:
  - konzervativní (ch. II. typu!) / liberální (ch. I. typu!)
  - ne/vhodné pro rozdílně velké skupiny
  - ne/vhodné pro rozdílné skupinové rozptyly

ANOVA – post-hoc testy

Doporučení podle A. Fielda:

- stejně velké skupiny a skupinové rozptyly (ideální situace): **REGWQ** nebo **Tukey**
- pokud si chceme být jistí, že neuděláme chybu I. typu: **Bonferroni**
- pokud jsou velikosti skupin trochu/hodně rozdílné: **Gabriel**/**Hochberg GT2**
- pokud pochybujeme o shodnosti skupinových rozptylů: **Games-Howell**

ANOVA – reportování

$F(df_M, df_R)$ = …, p = …, η^2 nebo ω^2 = …

- vždy uvádět deskriptivy pro každou skupinu – alespoň velikost, průměr, směrodatnou odch.
- vždy dopočítat velikost účinku (interpretujeme jako R^2 v lineární regresi)

  $\boldsymbol{\eta}^2 = SS_M / SS_T$

  $\boldsymbol{\omega}^2 = [SS_M – (df_M)MS_R] / [SS_T + MS_R]$
- df_M a df_R musejí být uváděny v tomto pořadí

U kontrastů uvádíme: $t(df)$ = …, p = …, d nebo r = …

- r = odmocnina$[t^2 / (t^2 + df)]$

„V modelu je pouze jeden faktor. Člověk je však ve skutečnosti obvykle členem více typů skupin najednou, což může mít vliv!“

„Provedeme více ANOV pro různé faktory (skupiny).“

„Tím se však vrátí známý problém s nárůstem rizika chyby I. typu. Navíc přijdeme o možnost posoudit vliv všech faktorů najednou v jednom modelu.“

„Můžeme přidat přímo do modelu další nezávislé kategorické proměnné – a spočítat tzv. **faktoriální ANOVU**.“

Faktoriální ANOVA

- ANOVA s více kategorickými nezávislými (faktory)
- uplatnění v **experimentálních** designech, kde pracujeme s několika druhy experimentální manipulace nebo kde chceme zohlednit kromě experimentální manipulace i další proměnné (např. pohlaví)
- uplatnění v **neexperimentálních** designech, kde chceme posoudit vliv více kategorických prediktorů najednou

Typy faktorů ve vícefaktoriálních designech

- Fixed factors
  - všechny úrovně faktoru, o které nám jde, jsou v našem výzkumu zahrnuty
  - *„Liší se užívání internetu mezi třemi typy SES?“*
  - *„Liší se užívání internetu podle pohlaví?“*
- Random factors
  - úrovně faktoru, zahrnuté v našem výzkumu, představují pouze náhodný vzorek z větší populace
  - *„Liší se užívání internetu mezi zeměmi?“*
  - *„Liší se užívání internetu podle školy, kterou adolescent navštěvuje?“*

One-way ANOVA

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

Faktoriální ANOVA

$$Y_{ijk} = \mu + \alpha_j + \beta_k + \gamma_{j \times k} + \varepsilon_{ijk}$$

Vliv toho, že je člověk členem skupiny *j*
MAIN EFFECT

Vliv toho, že je člověk členem skupiny *k*
MAIN EFFECT

One-way ANOVA

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

Faktoriální ANOVA

$$Y_{ijk} = \mu + \alpha_j + \beta_k + \gamma_{j\,x\,k} + \varepsilon_{ijk}$$

Vliv toho, že je člověk členem skupiny *j*
MAIN EFFECT

Vliv toho, že je člověk členem skupiny *k*
MAIN EFFECT

Vliv toho, že je člověk členem kombinace skupin *j* a *k*

One-way ANOVA

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

Faktoriální ANOVA

$$Y_{ijk} = \mu + \alpha_j + \beta_k + \gamma_{j \times k} + \varepsilon_{ijk}$$

Vliv toho, že je člověk členem skupiny *j*
MAIN EFFECT

Vliv toho, že je člověk členem skupiny *k*
MAIN EFFECT

Vliv toho, že je člověk členem kombinace skupin *j* a *k*

Interakce

One-way ANOVA

$$Y_{ij} = \mu + \color{red}{\alpha_j} + \varepsilon_{ij}$$

Faktoriální ANOVA

$$Y_{ijk} = \mu + \color{red}{\alpha_j} + \color{red}{\beta_k} + \color{red}{\gamma_{j \times k}} + \varepsilon_{ijk}$$

Vliv toho, že je člověk členem skupiny *j*
MAIN EFFECT

Vliv toho, že je člověk členem skupiny *k*
MAIN EFFECT

Vliv toho, že je člověk členem kombinace skupin *j* a *k*

Moderace

Interakce

Interakce (moderace)

- situace, kdy vliv dvou (či více) nezávislých proměnných na závislou proměnnou není pouhým součtem jejich jednotlivých vlivů
- s měnící se hladinou jedné nezávislé se mění vliv druhé nezávislé na závislou proměnnou
- nezávislá proměnná nemusí mít žádný přímý vliv (main effect) na závislou proměnnou, ale může ji ovlivňovat tím, že ovlivňuje vliv druhé nezávislé
- při interpretaci interakcí je obvykle velmi užitečné znázornění formou grafu

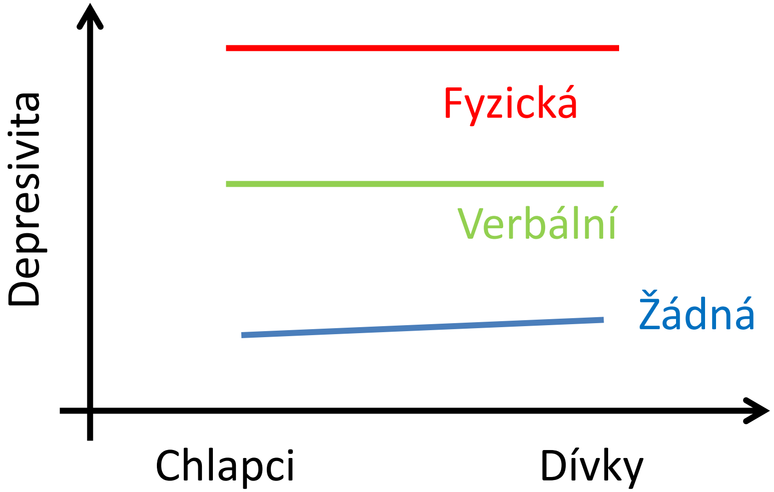
Interakce (moderace)

- dvě kategorické proměnné (případ faktoriální ANOVY)
  - Zážitek s různými typy školní šikany má jiný vliv na depresivitu u dívek a u chlapců.

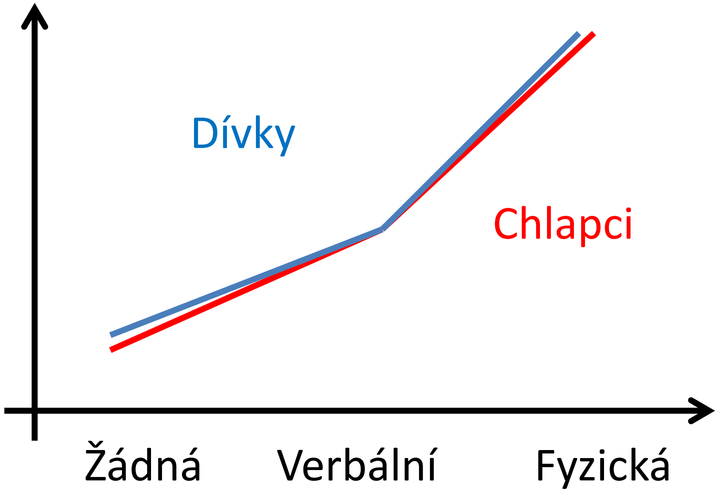
Interakce (moderace)

- dvě kategorické proměnné (případ faktoriální ANOVY)
  - Zážitek s různými typy školní šikany má jiný vliv na depresivitu u dívek a u chlapců.

žádná interakce



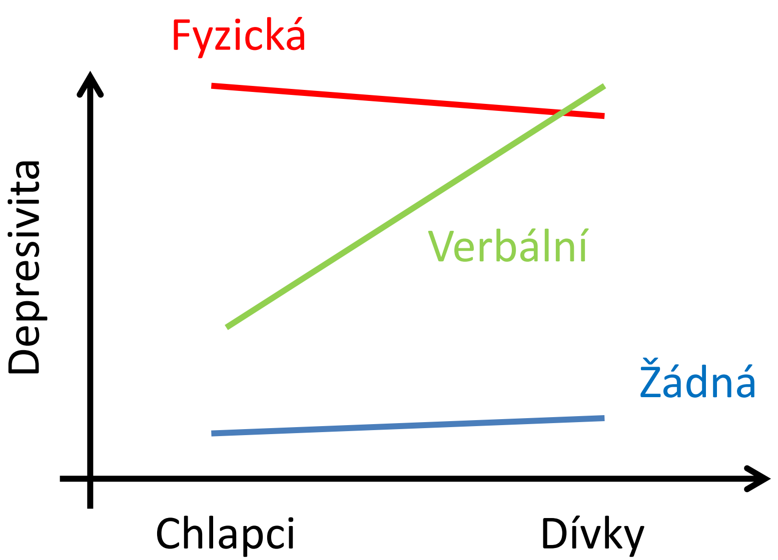
=



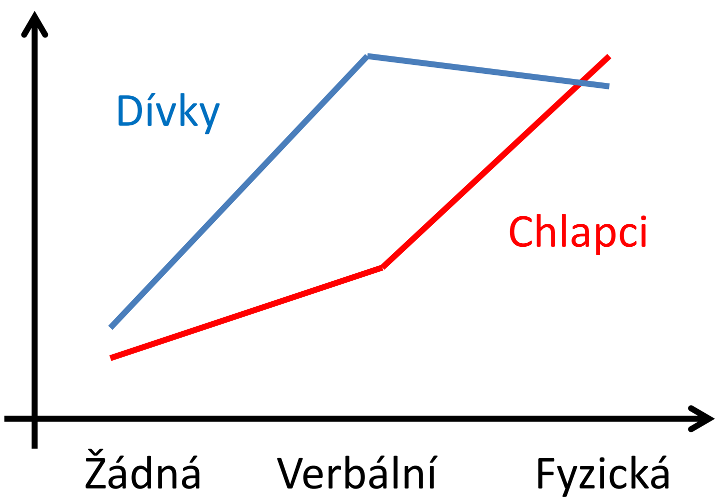
Interakce (moderace)

- dvě kategorické proměnné (případ faktoriální ANOVY)
  - Zážitek s různými typy školní šikany má jiný vliv na depresivitu u dívek a u chlapců.

interakce



=

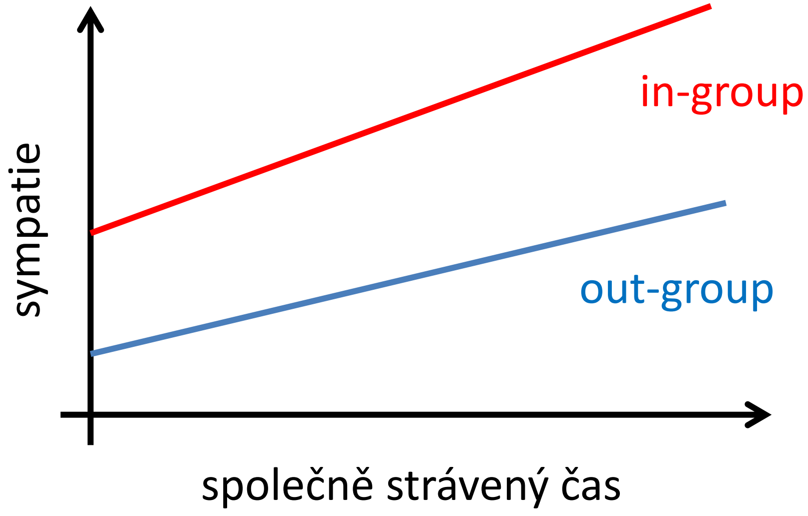


Interakce (moderace)

- dvě kategorické proměnné (případ faktoriální ANOVY)
  - Zážitek s různými typy školní šikany má jiný vliv na depresivitu u dívek a u chlapců.
- kategorická a intervalová proměnná
  - Společně strávený čas posiluje naše sympatie pouze k členům in-group, nikoli out-group.

Interakce (moderace)

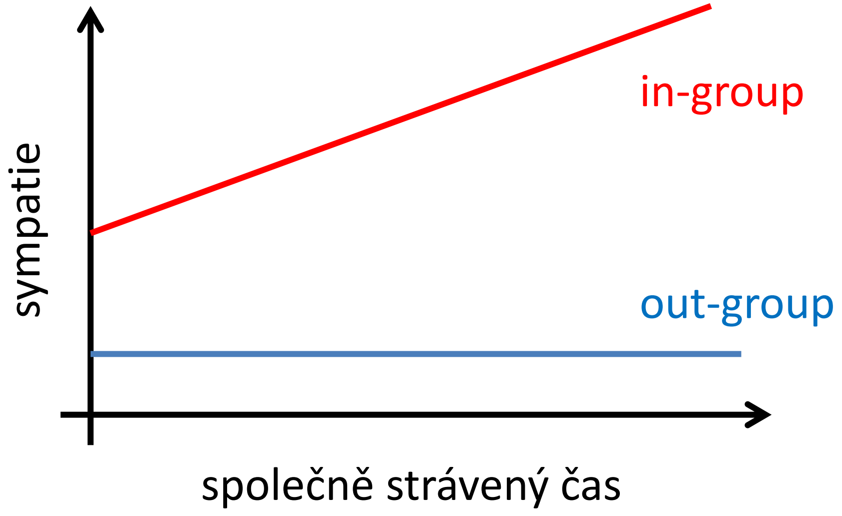
- dvě kategorické proměnné (případ faktoriální ANOVY)
  - Zážitek s různými typy školní šikany má jiný vliv na depresivitu u dívek a u chlapců.
- kategorická a intervalová proměnná
  - Společně strávený čas posiluje naše sympatie pouze k členům in-group, nikoli out-group.



žádná interakce

Interakce (moderace)

- dvě kategorické proměnné (případ faktoriální ANOVY)
  - Zážitek s různými typy školní šikany má jiný vliv na depresivitu u dívek a u chlapců.
- kategorická a intervalová proměnná
  - Společně strávený čas posiluje naše sympatie pouze k členům in-group, nikoli out-group.



interakce

Interakce (moderace)

- dvě kategorické proměnné (případ faktoriální ANOVY)
  - Zážitek s různými typy školní šikany má jiný vliv na depresivitu u dívek a u chlapců.
- kategorická a intervalová proměnná
  - Společně strávený čas posiluje naše sympatie pouze k členům in-group, nikoli out-group.
- dvě intervalové proměnné
  - S rostoucím příjmem se oslabuje vztah mezi spokojeností v práci a celkovou životní spokojeností.

Faktoriální ANOVA

SES: Souvisí SES s frekvencí používání internetu?

pohlaví: Souvisí pohlaví s frekvencí používání internetu?

interakce: Má SES jinou souvislost s používáním internetu u chlapců než u dívek?

	Nízký SES	Střední SES	Vysoký SES
Chlapci			
Dívky			

Faktoriální ANOVA - předpoklady

Vše, co v případě one-way ANOVY

Pro každou kombinaci faktorů by měl být zastoupený dostatečný počet případů.

Lze posoudit na základě jednoduché kontingenční tabulky.

Případnou nevyváženost lze částečně zohlednit zvoleným typem analýzy.

Počet případů	Nízký SES	Střední SES	Vysoký SES
Kluci	60	8	34
Holky	2	72	35

Faktoriální ANOVA v SPSS

Analyze → Generalized Linear Model → Univariate…

celková vysvětlená variabilita (SS_M) je rozsekána zvlášť pro jednotlivé faktory

Source	Type X Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	SS_M	df_M	MS_M		
Intercept					
Faktor1	$SS_{Faktor1}$	$df_{Faktor1}$	$MS_{Faktor1}$		
Faktor2	$SS_{Faktor2}$	$df_{Faktor2}$	$MS_{Faktor2}$		
Faktor1*Faktor2	$SS_{Interakce\ F1*F2}$	$df_{Int.\ F1*F2}$	$MS_{Int.\ F1*F2}$		
Error	SS_R	df_R	MS_R		
Total					
Corrected Total	SS_T	df_M+df_R			

Každý faktor a interakce má vlastní statistiku F, proto lze posoudit, zda je signifikantním prediktorem závislé proměnné

Faktoriální ANOVA – reportování

Uvádíme zvlášť, jaký efekt měl každý faktor (main effect) nebo interakce faktorů:

$F(df_{Faktor}, df_R)$ = …, p = …, parciální η^2 …

- parciální $\eta^2 = SS_{Faktor} / (SS_{Faktor} + SS_R)$

+ případné kontrasty a post-hoc testy jako u ANOVY

V některých situacích má smysl předpokládat, že je závislá proměnná ovlivňována nejen faktory, ale i intervalovými nezávislými proměnnými. Potřebujeme tedy model, který bude **kombinovat kategorické a intervalové nezávislé proměnné**.

Proč zavádět intervalové nezávislé do ANOVY:

- snížíme množství nevysvětlené variability v modelu
- kontrolujeme, zda není vliv faktorů zkreslen nějakou související intervalovou proměnnou

→ přesnější posouzení vlivu faktorů

Příklad: Používání internetu může souviset s věkem člověka. Pokud budeme tuto proměnnou kontrolovat, získáme představu o vlivu SES na frekvenci používání internetu, který je „očištěný“ od možného vlivu věku.

ANCOVA (analysis of covariance)

- ANOVA s jednou či více nezávislými intervalovými proměnnými (tzv. kovariáty)
- zavádět jen kovariáty, pro které existují **dobré důvody** (nenacpat tam vše, co jsme měřili)
- dobře zvolené kovariáty → zvýšení síly testu
  - kovariát odebere část nevysvětlené variability závislé proměnné, čímž se lépe projeví případný vliv faktorů
- špatně zvolené kovariáty → snížení síly testu
  - za každý přidaný kovariát ztrácíme jeden stupeň volnosti
- uplatnění v **experimentálních designech**, kde chceme statisticky kontrolovat nežádoucí rozdíly mezi skupinami
- uplatnění v **neexperimentálních designech**, kde chceme statisticky kontrolovat intervalové prediktory a posoudit tak nezkreslený vliv kategorických prediktorů

One-way ANOVA

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

ANCOVA

$$Y_{ijk} = \mu + \alpha_j + \beta X_{ij} + \varepsilon_{ijk}$$

Vliv toho, že je člověk členem skupiny *j*

Vliv intervalové proměnné x, tj. kovariátu

ANCOVA - předpoklady

- předpoklady ANOVY + předpoklady lineární regrese
- kovariát a skupinová příslušnost musí být nezávislé
  - pokud nejsou, je obtížné interpretovat výsledky
  - obdoba požadavku nekorelovaných prediktorů u vícenásobné lineární regrese
  - lze testovat dopředu jednoduchou ANOVOU (přičemž chceme, aby vyšla nesignifikantní)
- kovariát musí mít ve všech skupinách stejně silný vliv na závislou proměnnou
  - lze testovat zavedením interakce do modelu (přičemž chceme, aby vyšla nesignifikantní)

ANCOVA v SPSS

Analyze → Generalized Linear Model→ Univariate…

celková vysvětlená variabilita (SS_M) je rozsekána zvlášť pro kovariát(y) a faktor(y)

Source	Type X Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	**SS_M**	**df_M**	**MS_M**		
Intercept					
Kovariát1	**$SS_{Kovariát1}$**	**$df_{Kovariát1}$**	**$MS_{Kovariát1}$**		
Faktor1	**$SS_{Faktor1}$**	**$df_{Faktor1}$**	**$MS_{Faktor1}$**		
Error	**SS_R**	**df_R**	**MS_R**		
Total					
Corrected Total	**SS_T**	**df_M+df_R**			

můžeme si nechat zobrazit tzv. „**marginal means**“ (= jaké by byly skupinové průměry, kdyby se úroveň kovariátu nelišila napříč skupinami)

ANCOVA – reportování

Uvádíme, jaký efekt měl každý kovariát:

$F(df_{Kovariát}, df_R)$ = …, p = …, r = …

- pro jednotlivé kovariáty vždy $df_{Kovariát} = 1$
- $r = \text{odmocnina}[t^2 / (t^2 + df)]$

A uvádíme, jaký efekt měl každý faktor:

$F(df_{Faktor}, df_R)$ = …, p = …, parciální η^2 = …

- parciální $\eta^2 = SS_{Faktor} / (SS_{Faktor} + SS_R)$

+ případné kontrasty a post-hoc testy jako u ANOVY

MANOVA (multivariační ANOVA)

- ANOVA s více **závislými** intervalovými proměnnými
- posuzujeme vliv nezávislých proměnných na lineární kombinaci závislých proměnných
- pracujeme s multivariační obdobou F
- bereme v úvahu nejen (ne)vysvětlený rozptyl, ale i (ne)vysvětlenou kovarianci mezi závislými proměnnými

- výhody oproti sérii více ANOV
  - kontrolujeme nárůst rizika chyby I. typu
  - lze odhalit vztah ke <u>kombinaci</u> závislých proměnných
- nevýhody
  - obtížná interpretace výsledků
  - málokdy přinese nové informace oproti ANOVĚ
  - vyžaduje splnění dalších předpokladů, které nelze jednoduše otestovat v SPSS (multivariační normalita)

Úkol na seminář

- Otestujte hypotézy předpokládající nějaké kontrasty mezi 3 či více skupinami (one-way ANOVA s následnými kontrasty)

  +

- Otestujte hypotézu předpokládající interakci mezi 2 faktory (faktoriální ANOVA)
  - zde nemusíte dopočítávat kontrasty
  - můžete (ale nemusíte) do modelu zahrnout i nějaký kovariát
  - při interpretaci se neomezujte pouze na konstatování, že ne/byl nalezen signifikantní interakční efekt, ale rovněž popište (nejlépe na základě grafů), v čem konkrétně tato interakce spočívá