

PSY117/454

Statistická analýza dat v psychologii

Přednáška 13

Smysluplné užití statistiky

VIII. Thou shalt not worship the 0.05 significance level.

Michael Driscoll, The American Mathematical Monthly, The Ten Commandments of Statistical Inference, Volume 84, Number 8, 1977 (p. 628)

Dichotomizace výsledků výzkumu

- Výsledek výzkumu je testováním zredukován na ano-ne

	H_0 přijata	H_0 zamítnuta
H_0 pravdivá (žádný efekt)	OK	chyba 1. typu α (její pravděpodobnost)
H_0 nepravdivá (efekt)	chyba 2. typu β	OK Síla ($1-\beta$)

Čím nižší je α , tím vyšší je β . Přesná podoba vztahu závisí na použitém testu. α i β mohou být nízké pouze při vysokých n .

Síla testu viz Hendl 401-411.

Síla testu

- ... pravděpodobnost, že výsledek našeho statistického testu bude znít „na hladině α je rozdíl(korelace) statisticky významný“, pokud určitý rozdíl skutečně existuje
 - ... pravděpodobnost nezamítnutí H_0 , je-li H_0 nepravdivá
 - $P(p < \alpha | H_1) = 1 - \beta$
 - Závisí na
 - Velikosti účinku – skutečném stavu věcí – čím větší je účinek, tím vyšší je síla testu.
 - Velikost účinku obvykle souvisí s variabilitou
 - Zvolené α – čím nižší α , tím nižší síla testu
 - Velikosti vzorku – čím větší vzorek, tím vyšší síla testu
 - Použitím testu – obecně neparametrické mají nižší sílu testu
 - Žádoucí hodnoty pro sílu testu jsou **> 0,8**
-


K čemu jsou úvahy o síle testu?

- K tomu abychom nedělali zbytečné výzkumy.
 - Cohen 1960 – J of Abnormal & Social Psychol – průměrná síla 0,48
 - Cohen 1992 – nic moc se za 30 let nezměnilo
 - Existuje-li rozdíl/vztah a my ho hledáme v tak malém vzorku, že síla je < 50%, pak to nemá smysl
 - Abychom nedělali manipulativní závěry
 - Chceme-li potvrdit hypotézu o **ne**existenci rozdílu/vztahu, **nelze** to udělat prostým **ne**vyvrácením nulové hypotézy.
 - K tomu, abychom si dokázali spočítat, **jak velký vzorek** potřebujeme pro svůj výzkum – náklady.
-

Jak spočítáme potřebnou velikost vzorku?

- ...těžko... pro každý statistický test se počítá jinak
 - Totéž platí i pro sílu testu
 - Obecně: Jak velký vzorek potřebuji na to aby mi *pro mě zajímavý účinek* vyšel s 80% pravděpodobností statisticky významný na hladině α ?
 - receptář Oseckých
 - **G*Power** <http://www.psych.uni-duesseldorf.de/aap/projects/gpower/>
 - Online např.
 - <http://www.stat.uiowa.edu/~rlenth/Power/index.html>
 - <http://www.dssresearch.com/toolkit/spcalc/power.asp>
-

Základní postup zpracování dat

- ❑ Pečlivá tvorba datové matice
 - ❑ Tvorba a transformace proměnných
 - ❑ Seznámení se s daty a explorace
 - Běžné popisné statistiky, zejm. četnosti
 - Všemožná zobrazení, grafy a kontingenční tabulky
 - Outlieři, chyby, chybějící data, nečekané souvislosti
pomůže lépe pochopit získaná data (a důvěřit jim)
 - ❑ Upřesnění a testování hypotéz
 - Nezapomenout na ověření předpokladů testů
 - Umíme-li několika způsoby, je dobré je vyzkoušet, zvláště když si nejsme jistí splněním předpokladů
- 
- dukující
o nám

Základní principy komunikování výsledků

Ve výsledcích uvádíme

- Způsob výpočtu/kódování proměnných
- Popisné statistiky odpovídající úrovni měření a následujícím testům
 - Často včetně intervalů spolehlivosti na M nebo r
- Ověření předpokladů testů, je-li nutné (dle konvence)
- Testy hypotéz
 - Testová statistika (t , z , F , r apod.) + df
 - $p = \dots$, nebo $p < (>) \alpha$
 - Velikost účinku (d , r^2 , OR)

Statistická gramotnost

- ❑ Korelace neimplikuje kauzalitu
- ❑ Je rozdíl mezi statistickou a praktickou významností (zvláště u velkých vzorků)
- ❑ Je rozdíl mezi zjištěním nulového účinku/rozdílu a konstatováním, že rozdíly nejsou statisticky významné (zvláště u malých vzorků)
- ❑ Data jsou „omylným“ zachycením jevů a vždy je třeba zohledňovat jejich vznik – otázky a postupy ... metodologie
- ❑ Zdánlivě velmi nepravděpodobné jevy a koincidence se vyskytují často, protože mají nesmírně mnoho možností se vyskytnout
- ❑ Nezaměňovat podmíněné pravděpodobnosti $P(A|B)$ a $P(B|A)$.
Nezaměňovat podmíněné pravděpodobnosti s nepodmíněnými.
- ❑ Variabilita je přirozená a „normální“ není totéž co „průměrný“

Bezpečné řízení

- ❑ Automobile Association Foundation for Traffic Safety zjistila, že pouze 1,5% řidičů telefonovala, když měli havárii, zatímco 10,9% řidičů bylo v okamžiku havárie vyrušováno další osobou v autě nebo rádiem.
 - ❑ Řada médií z toho činila závěr, že telefonování při řízení ohrožuje řidiče méně než další pasažéři nebo poslech rádia.
 - ❑ $P(\text{Telefonování}|\text{Havárie}) \neq P(\text{Havárie}|\text{Telefonování})$
-

M.A.G.I.C.

V čem se obecně liší dobrý statistický argument od špatného?

MAGNITUDE – velikost účinku

ARTICULATION – konkrétnost, jistota: t -test vs. ANOVA, χ^2

GENERALITY – aplikovatelnost na co nejširší třídu jevů

INTERESTINGNESS – pro odborníka překvapivý výsledek?

CREDIBILITY – věrohodnost: čím je výsledek překvapivější, tím lépe musí být doložen

- To jsou aspekty, které se ve výzkumné zprávě snažíme zdůraznit (a při provádění výzkumu v rámci možností zajistit)
-

Výzkumně-kriticky orientovaná Áčka

PSY117 – statistika a PSY252 – statistika II (pouze na jednooborová)

PSYXXX – psychometrika (pouze na jednooborové psychologii)

PSY112 – metodologie – úvod, klasika a PSY118 – metodologie –
kvalitativa

- Kritické myšlení při příjmu poznatků
 - Konzumace výzkumu
 - Realizace vlastních výzkumů
 - Psychodiagnostika

 - Chybí: filozofie, logika sociologie vědy
-

2 knihy na závěr

Mlodinow, L. (2009). **Život je jen náhoda. Jak náhoda ovlivňuje naše životy.** Praha: Slovart.

Porter, T. M. (1995). **Trust in numbers. The pursuit of objectivity in science and public life.** Princeton: Princeton University Press.
