

PSY 117/454 - Statistická analýza dat**Jaro 2011**

VYUČJÍCÍ

- Mgr. Stanislav Ježek, PhD. – garance kurzu, přednášky, semináře
jezek@fss.muni.cz, 549494616, konzultační hodiny: středa 13 – 15, FSS 2.53
- Mgr. Jan Širůček – přednášky, semináře
sirucek@fss.muni.cz, 549498263, konzultace po dohodě emailem, FSS 2.45
- Mgr. Jan Šerek – přednášky, semináře
serek@fss.muni.cz, 549498263, konzultace po dohodě emailem, FSS 2.51

CHARAKTERISTIKA KURZU

Cílem kurzu je seznámit studenty se základy statistiky používané v psychologické praxi a výzkumu. Studenti získají porozumění základním prvkům statistiky a dovednost je aktivně i pasivně používat. Studenti získají dovednost připravit data pro statistické zpracování, spočítat základní statistiky, otestovat běžné typy hypotéz. Kurz klade důraz i na komunikaci, tj. slovní popis výsledků i schopnost porozumět takto popsaným výsledkům v empirických kvantitativních studiích. V rámci kurzu budou studenti seznamováni paralelně s českou i anglickou terminologií, aby byli po skončení kurzu schopni dále studovat a používat internetové zdroje.

Návaznosti kurzu

Kurz úzce souvisí s výukou metodologie. Mnoho problémů v metodologii má statistický základ a naopak mnohé problémy či omezení statistiky je potřeba zohledňovat v metodologii. Studijním programem stanovené pořadí absolvování kurzů je statistika – metodologie a v souladu s tím je koncipován obsah těchto povinných kurzů. Toto propojení je nezbytné pro vypracování diplomové práce a je součástí požadavků při státní bakalářské zkoušce.

Na statistice stojí také značná část psychometrie a tvoří tak přirozený základ pro studium psychodiagnostiky na magisterském stupni.

Předpokladem pro studium základů statistiky je běžné středoškolské matematické vzdělání.

ORGANIZACE KURZU

Počet kreditů:	5
Ukončení kurzu:	zkouška
Přednášky:	1 hodina týdně
Semináře:	1 hodiny týdně

Do seminárních skupin se studenti zapisují prostřednictvím informačního systému MU.

INFORMAČNÍ SYSTÉM MU

V informačním systému MU části Studijní materiály k předmětu PSY117/PSY454 umístíme podklady k tématům uvedeným v sylabu především v podobě prezentací z přednášek a odkazů na další zdroje informací. Na stejném místě jsou též k dispozici pokyny nezbytné k plnění písemných úkolů, popř. doplňky k tomuto sylabu.

V informačním systému naleznou studenti také interaktivní osnovu k předmětu. Tato osnova obsahuje pouze informace uvedené také v tomto sylabu. V případě nekonzistencí jsou směrodatné informace uvedené v tomto sylabu.

KOMUNIKACE S VYUČJÍCÍMI

Preferovaným komunikačním kanálem pro osobní komunikaci týkající se kurzu je email. Použijte prosím „PSY117“ resp. „PSY454“ v předmětu svých emailů. Urychlíte tím jejich vyřízení.

Emaily adresujte dr. Ježkovi. Pouze v případě záležitostí, které se týkají specificky seminářů Mgr. Širůčka nebo Mg. Šerka, pište přímo jim.

V případě dotazů týkajících se látky kurzu, využívejte prosím v maximální možné míře předmětové diskuzní fórum. Je pravděpodobné, že odpověď na Váš dotaz by mohla zajímat i Vaše spolužáky, a byla by proto škoda uzavřít takovou komunikaci do soukromí emailů. Všichni vyučující kurzu diskuzní fóra pravidelně sledují a přispívají do nich.

K důležitým ohlášením používáme hromadný email studentům.

POŽADAVKY NA UKONČENÍ KURZU

Seminární práce

V průběhu semestru mají studenti za úkol pracovat na níže uvedených úkolech. Jejich 100% realizace je předpokladem k tomu, aby se student mohl přihlásit ke zkoušce. Úkoly jsou zadávány na seminářích a v případě potřeby doplňovány informacemi v IS MU.

Práce lze odevzdávat pouze elektronicky v odevzdávací ve studijních materiálech předmětu. Jméno vkládaného souboru s první seminární prací musí být **S1.doc** (resp. **S2.doc** pro případnou druhou)¹ a nic víc; IS k němu automaticky přidá jméno vkládajícího studenta. Komentář/popisek nechejte nevyplněný. Ohodnocené a okomentované práce budou do odevzdávacího navraceny do 14 dnů od termínu odevzdání práce.

Seminární práce jsou hodnoceny na následující škále: **přijata (10b), nepřijata (0b)**. „Nepřijetím“ je míněno vrácení k přepracování. Na přepracování má autor týden ode dne vrácení. Vynikajícím pracím může být uděleno ještě 5 bodů jako bonus.

Stanovené rozsahy prací jsou záměrné a závazné. Práce nedosahující minima i práce překračující maximum budou automaticky vráceny bez komentářů jako nepřijaté.

Opravené práce vkládejte do ISu do stejné odevzdávací jako práce původní. Jako jméno vkládaného souboru nyní použijte **S1o.doc**. Komentář/popisek opět nechejte nevyplněný.

Seminární práce I: Zamyšlení nad komunikováním statistiky v médiích

Zamyšlení nad způsoby, jimiž je statistika používána či často zneužívána v médiích (popř. v odborné komunikaci). Úkolem studenta je vyhledat v populárních médiích (noviny, časopisy, lépe serióznější) článek či zprávu, která se opírá o nějaké statistiky (výzkum, průzkumy apod.), dohledat originální zdroj těchto dat (výzkumná zpráva apod.) a zamyslet se nad způsobem prezentace těchto statistik a jeho důsledky. Práce by měla obsahovat krátké shrnutí vybrané zprávy (zejm. závěry), shrnutí statistických metod (ukazatele, zobrazení) použitých v originálním zdroji, úvaha nad tím, co bylo vybráno (a případně pozměněno) do článku a jaké důsledky to má pro význam sdělení. Jádrem úvahy by měly být otázky jako: Jak dobře, přesně shrnuje novinový článek původní výzkumnou zprávu? Nepřehání článek závěry původní studie (přílišné zobecnění, optimismus)? Jsou zmíněny i potenciální problémy originální studie?

Tato seminární práce se zaměřuje na tzv. statistickou gramotnost (statistical literacy), kterou označujeme schopnost aktivně i pasivně komunikovat statistické výsledky. Být statisticky gramotný znamená nenechat se opít každým rohlíkem s pěknými čísly a grafy; zároveň však neodmítat apriori každou prezentaci statistických sdělení. Věnujte pozornost také induktivním aspektům statistiky. Na internetu je mnoho informací k tomuto tématu; můžete vyhledávat též *misuse statistics* nebo *abuse statistics*. Viz též Utts (2003), Gigerenzer (2002), Wainer (2007), Savage Wainer (2008) ve studijních materiálech.

Lze využít i internetové noviny a časopisy. Lze použít i zpravodajství v anglickém jazyce.

Seminární práci lze zpracovat i na jiné téma, např. z oblasti historické provázanosti statistiky a psychologie, vizualizace dat či kontroverzí kolem statistického testování hypotéz. Pokud chcete zpracovávat nějaké takové téma, učiňte písemnou nabídku emailem svému vedoucímu seminární skupiny, a to do **31.3**. Nabídka by měla na několika řádcích obsahovat základní tezi, kterou chcete se své eseji sledovat. Zvolené téma lze zpracovat pouze po odsouhlasení vedoucím seminární skupiny (nejde o cenzuru, jen chceme předejít zbytečným zklamáním).

Závazný rozsah práce je **4000 - 5000** znaků vč. mezer.

Termín: **1. 5.**

Seminární práce II: Výběrové šetření – skupinový projekt

Záměrem tohoto úkolu je vyzkoušet si realizaci jednoduchého výběrového šetření. Úkolem pětičlenných týmů bude zvolit si jednoduchou výzkumnou otázku (př. výskyt nějakého jevu, názoru v nějaké populaci), pokusit se na ni získat odpověď jednoduchým výběrovým šetřením a celý projekt prezentovat v podobě posterů na semináři.

Termín: **poslední seminář v semestru**

Průběžné hodnocení

V průběhu semestru budou na semináře zařazeny tři krátké průběžné desetiminutové písemky. V každé bude možné získat 10b. Předběžné termíny průběžných testů jsou: **16.3.; 6.4.; 27.4.**

Množství bodů z průběžných písemek, které se bude počítat do celkového hodnocení, se počítá podle následujícího excelovského vzorce: $ZAOKROUHLIT((SUMA(P1;P2;P3)-MIN(P1;P2;P3))*1,5;0)$. Tento výpočet eliminuje jedno zakolísání směrem dolů (včetně případné absence).

Na průběžné písemky se náhradní termíny poskytují pouze výjimečně.

¹ Práce lze odevzdávat ve všech běžných editovatelných formátech textových dokumentů – doc, rtf, odt apod. Prosím, neodevzdávejte práce v obtížně editovatelných formátech, jako je např. pdf, djvu. Jde nám o možnost vkládání komentářů a čitelnost i na mobilních zařízeních.

Zkouška

Kurz je zakončen zkouškou. Zkouška má písemnou podobu a je možné za ni získat **50b**. K úspěšnému složení zkoušky je nutné získat minimálně **30b**.

Zkouší se v rozsahu látky, který je vymezen v tomto sylabu k předmětu PSY117/PSY454.

Celkové hodnocení

Celkem lze v průběhu kurzu získat **50b**; za závěrečnou zkoušku lze získat **50b**.

Celkové hodnocení bude používat následující stupnici

A: 100 - 92b B: 91 – 84b C: 83 – 77b D: 76 – 71b E: 70 – 65b F: 64 a méně bodů.

Uznávání dříve absolvovaných kurzů statistiky

1. **Kurzy dříve absolvované na katedře psychologie FSS** budou uznávány v plné míře². Svě žádosti o uznání v takovém případě směřujte sekretářce katedry psychologie.

2. **Kurzy absolvované jinde než na katedře psychologie FSS** budou uznávány tou formou, že jejich absolventi nebudou povinni plnit požadavky na práci v semestru (seminární práce, průběžné zkoušky) a bude jim za ně započítán plný počet bodů (40b). I nadále jsou však povinni absolvovat závěrečnou zkoušku. Žádosti o uznání v tomto případě zasílejte na mail jezek@fss.muni.cz. V předmětu zprávy použijte „uznani PSY117“ (bez diakritiky). V textu stačí uvést kód absolvovaného předmětu na FSS (typicky SOC108). V případě žádosti o uznání kurzů z jiné fakulty či univerzity, uveďte též odkaz na sylaby absolvovaných kurzů, popř. sylaby přímo přiložte. **Žádosti zasílejte do konce 2. týdne semestru.** Později zasláným žádostem nebude vyhověno.

3. **Specificky, absolventi kurzu SOC108** mohou žádat o úplné uznání PSY117, pokud ukončili SOC108 s hodnocením „A“.

Omluvy

Omluvy jsou přijímány pouze předem. Pozdější omluvy budou akceptovány, pouze pokud šlo o nepředvídatelné případy. Víte-li o tom, že budete mít ze závažných důvodů problémy s dodržáním některého z termínů, informujte nás o tom co nejdříve.

ÚČAST NA VÝZKUMU V RÁMCI KURZU

V rámci účasti kurzu mohou být studenti požádáni o účast na výzkumech jak pro zlepšení kvality kurzu samotného, tak i pro další účely spojené s výzkumným zaměřením katedry. Účast v těchto výzkumech je ryze dobrovolná a nemá vliv na úspěšné ukončení kurzu či získané hodnocení.

LITERATURA

Základní zdroje

- Hendl, J.: *Přehled statistických metod zpracování dat. Analýza a metaanalýza dat*. Brno: Portál 2004.
- Osecký, L., Osecký, P.: *Receptář jednoduchých metod statistické indukce*. Brno, AV ČR 1996.³
- Urbánek, T.: K prezentaci výsledků statistických analýz, 1. část. *ČsPsych*, 2007 (51), 6, 601 – 609.
- Urbánek, T.: K prezentaci výsledků statistických analýz, 2. část. *ČsPsych*, 2008 (52), 1, 70 – 79.
- Cohen, J.: The Earth is round ($p < .05$). *American psychologist*, 1994 (49), 12, 997 – 1003.³
- Cohen, J.: A Power primer. *Psychological Bulletin*, 1992 (112), 1, 155 – 159.³
- Utts, J.: What educated citizens should know about statistics and probability. *American Statistician*, 2003 (57), 2, 74 – 79.
- Good, P. I., Hardin, J. W.: *Common errors in statistics (and how to avoid them)*. Wiley-Interscience 2003.
- Morgan, S. E., Reichert, T., Harrison, T. R.: *From numbers to words. Reporting statistical results for the social sciences*. Allyn & Bacon, 2002.
- Abelson, R. P.: *Statistics as principled argument*. Lawrence Erlbaum Associates, 1995.
- Wainer, H.: The most dangerous equation: Ignorance of how sample size affects statistical variation has created havoc for nearly a millenium. *American Scientist*, 2007, vol. 95, no. 3, s. 249 – 256.

Další užitečné publikace v češtině

Základy statistiky jsou staré, takže i poměrně letité knihy obsahují totéž, co ty nové s barevnými obrázky. Základy jsou ve všech knihách stejné, takže si student může vybrat takovou, která je napsána jemu srozumitelným jazykem. Starší

² Platí zejména pro studenty, kteří si zapsali PSY454 a již z bakalářského studia mají absolvován kurz PSY117.

³ Dostupné ve studijních materiálech

učebnice se od novějších liší především podstatně menším důrazem na velikosti účinku, statistickou sílu a použití neparametrických metod⁴. Počítejte s tím, prosím. Například:

Swoboda, H.: *Moderní statistika*. Praha: Svoboda, 1977.⁵

Lamsers, V., Růžička, L.: *Základy statistiky pro sociologii*. Praha: Svoboda, 1970.

O měření v psychologii pojednává přístupná kniha: Urbánek, T.: *Základy psychometrie*. PsÚ AV ČR a PsÚ FF MU, 2002.

Cizojazyčná literatura

S výjimkou Swobodova textu jsou všechny učebnice v češtině považovány studenty psychologie za obtížně srozumitelné (nikoli neprávem). Vřele doporučujeme studovat přednostně či paralelně z lety ověřených anglicky psaných učebnic.

Hlavní zdroje:

Glass, G. V., Hopkins, K. D.: *Statistical methods in education and psychology*, 3rd Ed. Allyn and Bacon, 1996.⁶

Field, A.: *Discovering statistics using SPSS*, 3rd Ed. Sage, 2009. (nebo 2. vydání z r. 2005)⁷

Alternativní zdroje:

Moore, D. S.: *The basic practice of statistics*, 4th Ed., Freeman, 2006⁸.

Moore, D. S., McCabe, G. P., Craig, B.: *Introduction to the Practice of Statistics*. 6th Ed., Freeman, 2007.

Coolican, H.: *Research methods and statistics in psychology*, 4th Ed. Hodder & Stoughton, 2004.

Welkowitz, J., Ewen, R. B., Cohen, J.: *Introductory statistics for the behavioral sciences*, 5th. Harcourt Brace, 2000⁹.

Blaikie, N.: *Analyzing quantitative data*. Sage, 2003¹⁰.

de Vaus, D.: *Analyzing social science data: 50 key problems in data analysis*. Sage, 2002¹¹.

Pro mnohé studenty je výhodné studovat paralelně i některou z učebnic používání SPSS. Příkladem vhodných jsou:

Morgan, G. A. et al: *SPSS for introductory statistics. Use and interpretation*, 2nd ed. LEA, 2004.

Pozor na velmi jednoduché příručky k SPSS; ty často nenabízejí žádný vhled do statistiky. Pro většinu uživatelů Windows znalých základů statistiky je používání SPSS snadné a intuitivní, takže pořizování takové příručky není nutné.

Velmi užitečným a přístupným slovníkem statistických a příbuzných termínů v anglickém jazyce je Everitt, B. E., Wykes, T.: *A dictionary of statistics for psychologists*. Arnold, 1999.

Příručka k tomu, jak dělat statistiku v Excelu: Gupta, V.: *Statistical analysis with Excel*. VJ Books, 2002.

Rozšiřující texty

Rozšiřující články a kapitoly jsou k dispozici ve studijních materiálech v ISu.

Abelson, R.: *Statistics as a principled argument*. Hillsdale: LEA, 1995.

Arbuthnot, J.: An argument for Divine Providence taken from the constant regularity observ'd in the birth of both sexes. *Philosophical Transactions*, 1710, vol. 27, s. 186 – 190.

Cowles, M.: *Statistics in psychology. An historical perspective*, 2nd Ed. Mahwah: LEA, 2001.

Emerson, J. D., Hoaglin, D. C.: Stem-and-leaf displays. In D. C. Hoaglin, F. Mosteller, J. W. Tukey (Eds.), *Understanding Robust and Exploratory Data Analysis*, 7 – 32. New York: Wiley, 1983.

Freeman, J., Walters, S.J., Campbell, M.J.: *How to display data*. Blackwell, 2008.¹²

Gaito, J.: Measurement Scales and Statistics: Resurgence of an Old Misconception. *Psychological Bulletin*, 1986, vol. 87, no. 3, s. 564 – 567.

Gigerenzer, G.: In the year 2054: Innumeracy defeated. In P. Sedlmeier, T. Betsch (Eds.), etc: *Frequency processing and cognition*. Oxford: Oxford University Press, 2002.

Lord, F. M.: On the statistical treatment of football numbers. *American Psychologist*, 1953, vol. 8, s. 750-751.

Nickerson, R. S.: Null hypothesis significance testing: a review of an old and continuing controversy. *Psychological methods*, 2000, vol. 5, no. 2, 241 – 301.

⁴ Totéž platí pro rozdíl mezi špatnými a dobrými učebnicemi.

⁵ Studenti si chválí srozumitelnost textu. Je v několika výtiscích v knihovně FSS.

⁶ Můj hlavní zdroj, mohu zapůjčit.

⁷ Velmi ceněná učebnice. Trochu přeskakuje základy. V knihovně je třetí vydání nejméně v 5 výtiscích prezenčně.

⁸ Profesor David Moore je velmi oceňovaným autorem statistických učebnic. Jako jeden z mála dokáže zjednodušovat a zároveň se vyhnout zavádějícím zjednodušením. Díky tomu ho oceňují nejen studenti, ale i statistikové.

⁹ Stručná, klasika s moderními prvky, dobrá. Nemaže med kolem huby, jde přímo k cíli. Dobré zaměření na sílu testu.

¹⁰ Učebnice pro odváznější. Učebnice doplňuje statistiku o kousky metodologie/epistemologie (co jsou to data, co je kauzalita apod.). Myslím, že není nutné si i tak obtížný úkol pochopení základů statistiky komplikovat epistemologickými problémy. Pokud ale máte tendenci se hodně zamýšlet nad tím, co jsou data, proměnné a proč "ti statistiku děláme", je to dobrý zdroj.

¹¹ Výtečný zdroj pro lidi, co absolvovali kurz statistiky, právě pracují na diplomce a potřebují poradit s některými praktickými problémy analýzy. Vhodné i jako doplňující zdroj při studiu. Zahrnuje všechny aspekty kvantitativního výzkumu. V každé kapitole popisuje jeden problém (vč. vysvětlení, v čem je vlastně problém) a jeho řešení.

¹² Vynikající, jednoduchý zdroj o grafickém zobrazování dat.

- Savage, S., Wainer, H.: Until Proven Guilty: False Positives and the War on Terror. *Chance*, 2008, vol. 21, no. 1, s. 59 – 62.
- Scholten, A. Z., Borsboom, D.: A reanalysis of Lord's statistical treatment of football numbers. *Journal of Mathematical Psychology*, 2009, vol. 53, s. 69 – 75.
- Tversky, A., Kahneman, D.: Belief in the law of small numbers. In D. Kahneman, P. Slovic, A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Oxford: Oxford University Press, 1982. Více též v přednášce při přebírání Nobelovy ceny na http://nobelprize.org/nobel_prizes/economics/laureates/2002/kahneman-lecture.html.

Statistický populár pro rozšíření obzorů

- Best, J.: *Damn lies and statistics: untangling numbers from the media, politicians and activists*. Berkeley and Los Angeles: University of California Press, 2001.
- Best, J.: *More damn lies and statistics: how numbers confuse public issues*. Berkeley and Los Angeles: University of California Press, 2004.
- Huff, D.: *How to lie with statistics*. New York: W.W. Norton & Company, Inc., 1954.
- Woolfson, M. M.: *Everyday probability and statistics. Health, elections, gambling and war*. London: Imperial College Press, 2008.

Internetové zdroje

- Vynikající zdroj o statistice, spíše pro pokročilé. <http://www.statsoft.com/textbook/stathome.html>
- Zdroj o zdrojích: <http://en.wikipedia.org/wiki/Statistics>
- Online učebnice: <http://en.wikibooks.org/wiki/Statistics>

Software

Pro procvičování statistiky je nutná zejména zručnost v používání tabulkového kalkulátoru, např. MS Office Excel či OpenOffice.org Calc. Většinu základních statistik je možné spočítat v tabulkovém kalkulátoru. Stejně tak správa dat je v malých výzkumných projektech podstatně pohodlnější v tabulkových kalkulátorech než specializovaném statistickém software.

Na univerzitách jsou běžně využívány obecné statistické softwarové balíky, u nás SPSS a Statistica. Ty jsou však velmi drahé a pro běžnou praxi obsahují spoustu nepotřebných funkcí. Univerzální a drahý software lze dnes snadno nahradit použitím malých specializovaných aplikací na internetu. Jejich přehled naleznete např. na <http://statpages.org/>.

I když práce se statistickým software není těžištěm tohoto kurzu, doporučujeme studentům, aby se statistickými programy v průběhu kurzu zkoušeli pracovat a naučili se pracovat s jejich výstupy. SPSS budeme využívat i na seminářích.

PROBLÉMOVÉ OKRUHY

Níže uvedené problémové okruhy jsou součástí přednášek předmětu PSY117/454. Jedná se o relativně rozsáhlé okruhy, takže některé pojmy a přístupy uvedené v sylabu se na přednáškách objeví pouze v podobě odkazu na literaturu k samostudiu.

1. Proměnné, výzkumný kontext

Data, proměnné, úrovně měření, kvalita měření, organizace dat, kontrola dat. Tvorba datové matice (v Excelu a SPSS), kódování proměnných.

Literatura: Hendl 43–50; Urbánek 12–24

2. Zobrazování dat, četnosti, distribuce

Tabelace dat, šíře intervalů, minimum, maximum, odlehlá hodnota (outlier), absolutní a relativní četnosti (frekvence), kumulativní absolutní a relativní četnosti, rozložení (rozdělení) četností (dat), tvary rozložení (normální, bimodální, uniformní, pozitivně zešikmené, negativně zešikmené), normální (Gaussovo) rozložení, velikosti oblastí pod křivkou normálního rozložení, Poissonovo rozložení, graf absolutních a relativních četností, sloupcový graf, histogram

Literatura: Hendl 85–92; 100; 139; Glass, Hopkins, kap. 6; Good, Hardin (2003) 107–125; stonkolist: Emerson, Hoaglin (2003), Gaussova křivka: Swoboda 73–87.

3. Míry centrální tendence a variability, transformace

Modus, medián, průměr, vážený průměr, vhodnost použití různých měr centrální tendence, (variační) rozpětí, kvartilové rozpětí, směrodatná odchylka (populační, výběrová), rozptyl, vliv přičítání konstanty a násobení konstantou na m a s , z-skóry a další standardní skóry (T , IQ), normalizované skóry, percentily, šikmost, špičatost, krabicový graf s anténami

Literatura: Hendl 93–99; Urbánek 84 – 94; Glass, Hopkins 94

4. Pravděpodobnost

Pojetí pravděpodobnosti, počítání s pravděpodobnostmi, náhodné jevy, podmíněné pravděpodobnosti, Bayesův teorém, normální pravděpodobnostní rozložení a další běžná rozložení.

Literatura: Hendl 115–127; 134; Swoboda 29 – 33

5. Vztahy mezi proměnnými, korelace

Korelace – Pearsonův, Spearmanův, Kendallův koeficient a jejich vlastnosti. Koeficient determinace, kovariance. Kontingenční tabulka, marginální četnosti. Lineární vztah, monotónní vztah, pozitivní a negativní vztah. Těsnost vztahu. Bodový graf. Parciální a semiparciální korelace. Korelace mezi položkami, Cronbachovo α .

Literatura: Hendl 237–265

6. Lineární regrese

Statistická predikce, lineární vs. nelineární regrese, lineární a kvadratická funkce, odhad, modelování, regrese, reziduum, prediktor, závislá a nezávislá proměnná, zdroje variability, stanovení regresní přímky metodou nejmenších čtverců, regresní rozptyl a reziduální rozptyl, koeficient determinace jako ukazatel úspěšnosti regrese, homoskedascita, mnohočetná (mnohonásobná) regrese, logistická regrese

Literatura: Hendl 266 - 290

7. Statistická indukce, intervalové odhady

Vzorek(výběr), statistiky vs. parametry, estimační vlastnosti popisných statistik, výběrová rozložení, centrální limitní teorém, směrodatná chyba (průměru), výběrové rozložení průměru, relativní četnosti, rozptylu, bodové vs. intervalové odhady.

Literatura: Hendl 146–147, 150–174

8. Testování hypotéz

Statistická (nulová) hypotéza, výzkumná (alternativní) hypotéza, jednostranná vs. oboustranná hypotéza (test); Bayesovský přístup k testování hypotéz vs. Fisher-Pearson-Neymanovský (tradiční) přístup, úroveň (hladina) statistické významnosti, chyba I. a II. typu a jejich pravděpodobnost, (statistická) síla testu, jednovýběrový t-test, dvouvýběrový t-test (pro nezávislé výběry), párový t-test (z-test), Levenův test, testování korelačního koeficientu, velikost efektu, Cohenovo d , transformace z d na r a naopak.

Literatura: Hendl 175–220; 234; 401–412; 252–253; 413–419; Osečtí 1-36, Cohen 1992, 1994.

9. Testy pro nominální a ordinální proměnné

Parametrické vs. neparametrické testy, znaménkový test, test relativních četností, test dobré shody (χ^2), závislost kategoriálních proměnných (χ^2 , Cramerovo V , koeficient shody pozorovatelů λ), Wilcoxonovy testy (jednovýběrový, dvouvýběrový), Mann-Whitney U .

Literatura: Hendl 191–197; 220–232, 297–302, 304–313, 322–323

10. Analýza rozptylu

Problém s prováděním většího počtu testů, rybaření v datech, Bonferroniho korekce, princip analýzy rozptylu, rozptyl mezi skupinami (SS, MS), rozptyl uvnitř skupin (SS, MS), statistika F (F -test), analýza rozptylu s jedním faktorem (one-way), předpoklady analýzy rozptylu, post-hoc testy (S-N-K, Scheffe, LSD), velikost účinku (η^2 , ω^2), interakce faktorů

Literatura: Hendl 198–200, 337–346, 354–356

SEMINÁŘE

Semináře jsou zaměřeny na praktickou aplikaci znalostí prezentovaných na přednáškách a doplněných samostudiem. Jejich účelem je také prohloubit porozumění základním pojmům a myšlenkám. Účast na seminářích je v souladu se studijním řádem povinná.

ČASOVÁ OSNOVA

Datum	Téma přednášky	Téma semináře	Důležité termíny
23. 2.	Úvodní setkání	tvorba datové matice	
2. 3.	Proměnné, zobrazování dat, četnosti, distribuce		Žádosti o uznání
9. 3.	Míry centrální tendence a variability	statistická deskripce	
16. 3.	Pravděpodobnost	PP1	
23. 3.	Korelace	korelace	
30. 3.	Lineární regrese	regrese	
6. 4.	Pořadové korelace, shrnutí	PP2, pořadové korelace	
13. 4.	Statistická indukce, intervalové odhady	indukce	
20. 4.	Testování hypotéz	intervalové odhady	
27. 4.	Přehledy testů, síla testu	PP3, testování hypotéz	
4. 5.	Testy pro nominální a ordinální proměnné	neparametrické metody	1.5. seminární práce
11. 5.	Analýza rozptylu	analýza rozptylu	
18. 5.	Volný program	Prezentace posterů	