

Statistika a statistické myšlení

Základní pojmy a metody



**SLEZSKÁ
UNIVERZITA**

FAKULTA VEŘEJNÝCH
POLITIK V OPAVĚ

doc. Ing. Petr Sed'a, Ph.D.

Jak, kde a kdy mne najdete?

Spojení na přednášejícího:

- e-mail: petr.seda@math.slu.cz
- konzultační hodiny: ve čtvrtek dle domluvy, místnost č. 202 (Na Rybníčku 626/1)

Informace o předmětu:

<https://is.slu.cz/auth/el/fvp/leto2024/UVSSPUK051/index.qwarp> – interaktivní osnova (informační systém SLU)

Co Vás čeká v tomto předmětu?

- vstupní kurs do problematiky statistického zkoumání a analýzy

Co „**Základy statistiky**” přinesou budoucímu pracovníkovi v oblasti sociální práce?

- seznámí jej se základními metodami statistické analýzy,
- naučí správně interpretovat statistické výsledky,
- pomůže vybrat a správně použít statistické metody,
- ukáže mu, jak lze tyto metody využít v praxi a při zpracování závěrečných prací.

Obsah předmětu:

1. Statistika a statistické myšlení.
2. Základy teorie pravděpodobnosti, podmíněná pravděpodobnost.
3. Náhodná veličina a rozdělení pravděpodobnosti.
4. Vybraná rozdělení náhodné veličiny.
5. Základní statistické pojmy. Typy statistických znaků. Uspořádání dat.
6. Popisné statistiky kategoriálních dat,
7. Popisné statistiky číselných dat,
8. Úvod do teorie odhadu.
9. Principy testování hypotéz, základní jednovýběrové testy parametrických hypotéz.
10. Testování neparametrických hypotéz. Testy dobré shody.
11. Závislost kategoriálních dat, kontingenční tabulky.
12. Závislost číselných dat, korelační a regresní analýza.

Podmínky pro zápočet:

- 1 kontrolní test na konci semestru
 - test lze jednou opakovat
 - kombinace teoretických znalostí a praktických výpočtů

celkem je nutno dosáhnout úspěšnost více než 50%

Neučte se nazpaměť!

Smyslem předmětu není učit se nazpaměť definice, vzorce a postupy, ale rozumět jim a umět je použít.

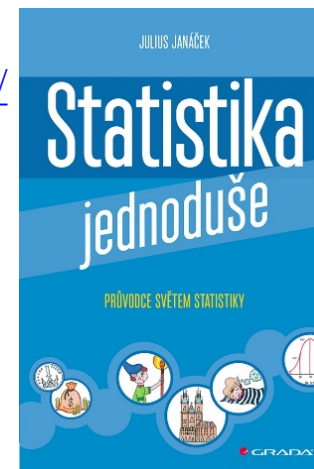
při testech lze používat:

- libovolnou kalkulačku
- vlastní poznámky z přednášek a cvičení
- libovolnou doporučenou literaturu

Z čeho studovat statistiku?

Zdroje informací

- Studijní materiály v IS SLU,
- Aktivní účast na přednáškách a cvičeníh,
- Ramík J. a Čemerková Š. *Statistika A*. Opava, Karviná: SLU, 2000.
- Janáček J. *Statistika jednoduše*. Praha: Grada, 2022.
- <https://www.bookport.cz/kniha/statistika-jednoduse-11215/>



Proč máte studovat statistiku?

Nemusíte být dobrým statistikem, ale musíte umět statisticky myslet

- statistika je všude kolem nás,
 - statistika je integrální součástí společenských věd,
 - statistika umožňuje provádět kvalifikovaná rozhodnutí na základě dostupných dat,
 - znalost statistických metod je důležitou konkurenční výhodou na trhu.
-
- statistika není složitá, ale je logická,
 - statistiku lze studovat i bez vyšší matematiky.

Sedm cílů předmětu:

1. Pochopíte, k čemu statistika slouží a jaké možnosti Vám nabízí.
2. Naučíte se vnímat procesy probíhající kolem nás jako zdroje statistických údajů.
3. Naučíte se statistiky číst, vyhodnocovat i vytvářet.
4. Naučíte se rozlišovat mezi fakty a lživými statistikami nebo povrchními závěry.
5. Osvojíte si základní statistické metody a nástroje.
6. Budete schopni dokládat svá tvrzení a rozhodnutí exaktními a nezpochybnitelnými fakty.
7. Získáte konkurenční výhodu nad těmi, kteří statistické nástroje nepoužívají.

Co je to statistika?

Různé definice pojmu statistika:

- a) Statistika je souhrn metod, zabývajících se sběrem, popisem, analýzou a interpretací údajů získaných z pozorování nebo experimentů.
- b) Statistika je vědecká disciplína, která se zabývá hromadnými jevy a procesy, studuje jejich variabilitu a zabývá se jejich analýzou.

Co říkají ekonomové?

Statistika je nástroj na podporu rozhodování.

Statistika a statistické myšlení:

Cílem předmětu není učit statistiku, ale statisticky myslet.

Statistika – teoretická disciplína postavená na teorii pravděpodobnosti a vyšší matematice.

$$F(x|\mu; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{(u-\mu)^2}{2\sigma^2}\right] du$$

Statistické myšlení – schopnost používat data k pochopení variability procesů a hledání řešení problémů.

PROCES → NEURČITOST → DATA → POCHOPENÍ PROCESU

Nelze řídit proces, kterému nerozumíte.

Co je to statistické myšlení?

Statistické myšlení – „statistical thinking“:

- veškerá činnost se odehrává v rámci navzájem provázaných procesů,
- ve všech procesech existuje variabilita (různorodost),
- pochopení, resp. snížení variability je cesta k úspěchu.

Zamyslete se:

Porovnejte výrobní proces na automatizované lince a proces vývoje počasí.

Co mají společného? V čem se liší?

Jak může v obou případech pomoci statistika?

Statistika v každodenní praxi:

Statistiku děláme každý den – neustále, i když si to neuvědomujeme.

Neustále něco měříme, odhadujeme a experimentujeme. Ráno začínáme statisticky uvažovat ve chvíli, kdy se probudíme a přemýšlíme, jak dlouho ještě můžeme ležet v posteli. Stejným způsobem pokračujeme celý den – například když zvažujeme, zda má voda ve sprše správnou teplotu, kolik kávy si nasypat do šálku nebo v kolik hodin musíme vyrazit do školy nebo do práce.

*Neil R. Ullman (County College of Morris):
Statistical Thinking – Everyday Statistics*

Kořeny soudobé statistiky:

- **Předstatistické období**
 - starověk a středověk - sčítání lidu, majetku
 - zmínky prakticky u všech známých civilizací
- **Univerzitní statistika**
 - 16. - 18. století - Itálie, Německo
 - statistika jako popisná věda o státě
 - vznik termínu „statistika“ (die Statistik)
- **Politická aritmetika**
 - 17. století – Anglie
 - číselné údaje, výpočty z matrik (údajů o obyvatelstvu)
- **Teorie pravděpodobnosti**
 - od 16. století – velký nástup na přelomu 19. a 20. století
 - vyjádření nejistoty, možnosti odhadů a testů hypotéz

Sčítání lidu má historickou tradici:

zmínky ve Starém i Novém Zákoně

- 4. kniha Mojžíšova (Numeri = sčítání)
- evangelia – proč přišel Josef s Marií do Jeruzalému?

pravidelné sčítání lidu v Římě každých 5 let

- od 5. století před Kristem
- censoři – zodpovědní za statistiky

středověké statistiky

- městské knihy, venkovské urbáře
- církevní matriky

- Zamyslete se, proč vlastně vzniklo sčítání lidu a majetku.
- Kdy se konalo poslední sčítání lidu v České republice?
- Kde najdeme jeho výsledky?

Statistika jako vědecká disciplína:

- **Popisná (deskriptivní) statistika**
 - 18. století - statistika jako nástroj pro objasnění a hledání vztahů mezi hromadnými jevy (nejprve přírodní vědy, později i humanitní)
 - 19. století - statistika jako oficiální orgán státní správy
(Belgie – 1. státní statistický úřad na světě)
- **Matematická (induktivní) statistika**
 - konec 19. století - 20. století – využití teorie pravděpodobnosti, nástroje pro popis individuálních jevů
 - statistická analýza – odhady a testování hypotéz
 - statistika jako nástroj predikce a rozhodování

Hromadné a individuální jevy:

- **Hromadný jev**

- jev, který se opakuje na daném objektu za stejných podmínek nebo v dané množině objektů

- denní tržba v dané prodejně
 - počet dětí v českých domácnostech

Kolik procent českých domácností má 2 a více dětí?

- **Individuální jev**

- jev, který sledujeme u jednoho konkrétního objektu

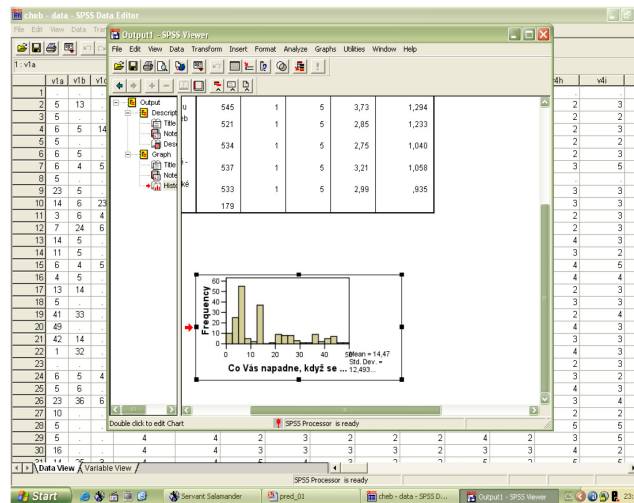
- dnešní tržba v dané prodejně
 - počet dětí v konkrétní domácnosti

S jakou pravděpodobností bude mít vybraná česká domácnost 2 a více dětí?

Statistika v současnosti:

statistika v 21. století

- využití všech postupů, které si statistika během svého vývoje vytvořila nebo osvojila
- hledání dalších nástrojů a metod
- sblížování statistiky a informatiky, nástup osobních počítačů a Internetu
- využití statistiky v nových oborech (historie, umění, ...)



program SPSS
na počítači

Statistická jednotka a soubor:

statistická jednotka

- **objekt statistického sledování, nositel sledovaných vlastností**
- osoby (studenti 2. ročníku FVP)
- organizace (fakulty v ČR)
- věci (počítače v budově FVP)
- události (termín zápočtu ze statistiky)

statistický soubor

- **množina sledovaných vymezených statistických jednotek**
- vymezení věcné (studenti sociální práce)
- vymezení prostorové (studenti FVP)
- vymezení časové (studenti 2. ročníku v akademickém roce 2023/24)

Vyčerpávající a výběrové šetření:

soubor základní (populace) - všechny vymezené jednotky

např. všichni studenti FVP v akademickém roce 2023/24

vyčerpávající šetření - pouze výjimečně (sčítání lidu)

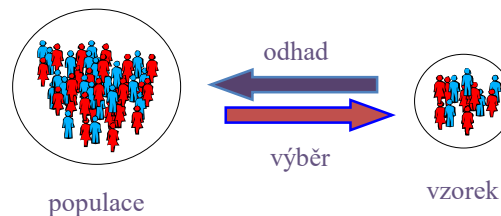
soubor výběrový (vzorek) - vybraná podmnožina základního souboru

např. studenti, kteří se zúčastnili anonymní ankety

výběrové šetření - většina statistických šetření (průzkumy, ankety, ...)

proč výběrové šetření?

- šetří čas a peníze
- řeší problém s nedostupnými jednotkami
- umožňuje provádět destruktivní testy
- výběrovým šetřením získáváme odhady vlastností základního souboru



Etapy statistického zjišťování:

1. **Formulace problému, stanovení cílů**
 - identifikace analyzovaných procesů a zdrojů jejich neurčitosti
2. **Sběr údajů**
 - získání dat popisujících dané procesy
 - primární data (pozorování, dokazování, experiment), sekundární
3. **Zpracování údajů**
 - transformace dat, odstranění chybných dat, třídění dat
 - výstupy ve formě tabulek, grafů, číselných charakteristik (ukazatelů)
4. **Analýza údajů**
 - odhady, testování hypotéz, extrapolace
 - výstupy ve formě doporučení na podporu rozhodování
5. **Využití výsledků**
 - pochopení sledovaných procesů, využití v praxi

Formy statistického zjišťování:

- **pozorování**
 - sledování pomocí smyslového vnímání (zrak, sluch, chuť)
např. vizuální kontrola jakosti výrobků
- **dotazování**
 - zjišťování pomocí výkazů a anket
např. dotazník zjišťující oblibu nového výrobku
- **experiment**
 - sledování chování za určitých podmínek
např. měření výkonu procesoru při různých frekvencích
- **sekundární výzkum**
 - využití již sebraných údajů – z výkazů, z internetu, ...
např. vyhodnocení inflace podle údajů ze Statistické ročenky

Státní statistická služba:

- 1841 – první statistický úřad v Evropě (Belgie)

Státní statistická služba - podle zákona PČR č. 89/95 Sb.

- Český statistický úřad (ČSÚ)
- ministerstva a orgány veřejné správy (např. FÚ, ŽÚ)

Zahraníční statistické orgány a organizace

- Statistická komise OSN
- EUROSTAT (Evropa)

Nestátní statistické zdroje

- nezávislé organizace (Median, SC&C, PAQ Research, CVVM, SANEP, STEM)
- podnikové statistiky

Český statistický úřad: www.czso.cz

úkoly Českého statistického úřadu:

- zabezpečuje sběr a zpracování údajů pro státní statistiku, včetně sčítání lidu
- vytváří a poskytuje statistické informace o sociálním, ekonomickém a ekologickém vývoji ČR
- vede národní účetnictví

struktura Českého statistického úřadu

- centrální orgán
- krajské reprezentace
- oblastní zastoupení – bývalé okresy a větší města

Statistická zjišťování v praxi:

sčítání lidu (osob, domácností a bytů)

- úplné šetření - census
- koná se cca každých 10 let (naposledy v roce 2021)
- je harmonizováno s podobnými akcemi v Evropě a ve světě

průzkumy veřejného mínění

- výběrová šetření formou dotazníků
- počet respondentů řádově několik tisíc

marketingové průzkumy

- výběrová šetření obvykle formou anket
- celostátní, regionální, resortní, firemní, ...

1. Ramík J. a Čemerková Š. *Statistika A*. Karviná: SLU, 2000. **(kapitola 1)**.





Děkuji za pozornost.