

Statistika

Mgr. Radmila Krkošková, Ph.D.

- **Rozsah předmětu:** 2 + 1, 6 kreditů, Zk
- **Vyučující z katedry Informatiky a Matematiky:**

Mgr. Radmila Krkošková, Ph.D.

Ing. Zuzana Neničková, Ph.D.

Podmínky ukončení předmětu a udělení kreditů:

- **70% účast** na seminářích
- **Průběžný test: max 30 bodů** – započítají se do známky ze zkoušky
(koná se v 11. týdnu, tj. od 4. 5. 2020, obsah: 1.- 6. kap. **nejsou povoleny žádné „pomůcky“**)
- **Zkouška:** je písemná, **max 70 bodů**,
(možnost používat: skripta, poznámky, kalkulačku apod. - **nikoliv počítač! nikoliv mobil!**)
- **Nepřecenit možnost použití skript !!!**

Celkové hodnocení: max 100 bodů,
k úspěšnému absolvování je zapotřebí
získat **alespoň 60 bodů**:

0 až 59 bodů – „**F** - 4“

60 až 64 bodů - „**E** - 3“

65 až 69 bodů - „**D** – 2,5“

70 až 79 bodů - „**C** - 2“

80 až 89 bodů - „**B** – 1,5“

90 až 100 bodů - „**A** - 1“

Povinná a doporučená literatura

Povinná:

- J. Ramík, R. Stoklasová: *Statistika*. Studijní opora SU OPF, Karviná, 2013.
- **Intranet:** L:/ramik/PUBLIC/Statistika/Statistika.pdf

Doporučená:

- J. Ramík, Š. Čemerková: *Statistika A*. Skriptum SU OPF, Karviná, 2000.
- T.H. Wonnacot, R.J. Wonnacot: *Statistika pro obchod a hospodářství*, Victoria Publishing, Praha, 1993.
- L. Cyhelským J. Kahounová, R. Hindls: *Elementární statistická analýza*. Management Press, Praha 1996.
- R. Hindls, S. Hronová I. Novák: *Analýza dat v manažérském rozhodování*. Grada, Praha 1999.
- <http://www.statsoft.com> – elektronická učebnice v angličtině

Co dále?

- Co je statistika a čím se zabývá
- Proč statistiky „lžou“
- Něco z historie statistiky v dávnověku
- Také něco z novodobých dějin statistiky
- K čemu je statistika v obchodě a podnikání
- Jak využít počítače PC ve statistice

Co je to statistika?

3 významy slova **statistika**:

1. **Statistika** je tabulka, graf, číselná informace

2. **Statistika** je soubor metod (postupů, činností) zabývajících se informacemi zejména v číselné podobě

3. **Statistika** je funkce náhodné veličiny („terminus technicus“)

Co je statistika pro nás?

Statistika je soubor (souhrn, systém) metod (postupů, činností) zabývajících se **informacemi** v **číselné** a **nečíselné podobě**, jmenovitě jejich:

- **sběrem** (statistické šetření, dotazníkový průzkum, dotazování, řízený rozhovor aj.)
- **popisem** (strukturalizací, uložením v počítači)
- **analýzou** (rozborem, statistickými metodami)
- **vyhodnocením** (interpretací - vysvětlením)...

Statistiky lžou...

Neznámý anglický lord

Motto 1: *„Statistika je zvlášť rafinovanou formou lži“*

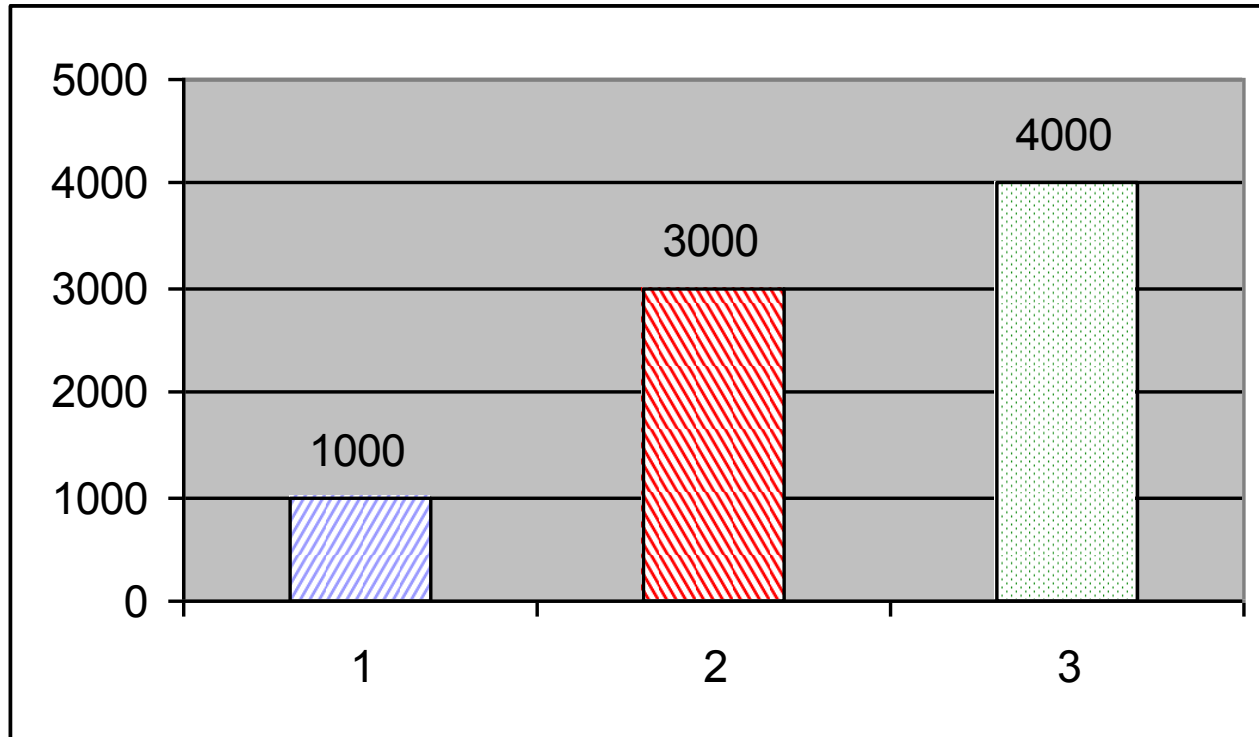
Motto 2: *„Věřím jen těm statistikám, které jsem sám zfalšoval“*

Churchill

Motto 3: *„Statistika nuda je, má však cenné údaje“*

Svěrák - Uhlíř

Příklad: Počet kriminálních činů ve městě M



Špatná zpráva: Kriminalita vzrostla (o 300 nebo dokonce 400 %)!

Dobrá zpráva: Rychlost růstu kriminality se snížila na polovinu (o 50 resp. 100 %)!

Mediánové platy v různých profesích (podle ČTV)

profese	mediánový plat
Lékaři	43 174
Právníci	41 725
Programátoři	41 164
Vědci	34 342
Učitelé	26 168

Víme, co je medián?!

Něco z historie statistiky v dávnověku

- „Statistika“ ve starém Egyptě, Mezopotámii, Číně...
- Nejstarší „statistika“ = popis státu = zobrazení daného zeměpisného, hospodářského a politického stavu

- "Status" znamená „stav“, ale také „stát“, neboť „stát sám je stav“
- Jedno z prvních státovědných děl: Francesco Sansovina: „**Del governo et administratione di diversi regni**“ („O vládě a správě v různých královstvích“, Itálie, 16. století)

Také něco z novodobých dějin statistiky

- Belgičan *Adolphe Quételet* (19. stol.):
homme moyen - průměrný člověk - ideální typ, o který se příroda snaží, ale který je **nereálný**

Základ budoucí statistiky:

koncept normálního rozdělení

(střední hodnoty a rozptylu)

18. a 19. století: - základy pro další rozvoj statistiky

- *Italové*: 3 Bernoulliiové (Jacob, Daniel, Nicolas)
- *Francouzi*: Lagrange, Laplace
- *Němci*: Euler a Gauss a jiní

Heslo statistiky:

POPULACE = „soubor všeho“

- **Počátek 20. stol.** - induktivní statistika: dříve popis každého detailu, nyní závěry o celku na základě

výběru = vzorku

- Heslo moderní statistiky: **VÝBĚR**

- **Zakladatelé moderní statistiky:**

Rusko: Čebyšev, Ljapunov a Markov

Anglie-Amerika: Fisher, Pearson a Neyman

Historický závěr

**Správné pochopení statistických
pojmu a metod je
předpokladem úspěšné práce
každého**

**(ekonomicky vzdělaného)
odborníka!**

Statistika a počítače

- **Počet PC/Mcintosh:** Svět > 1 000 000
v ČR asi 7 mil., na SU OPF 250 PC
(„každý má svůj NB“)

- **Internet:** Český statistický úřad – ČSÚ:
<http://www.czso.cz>

Statistický úřad EU – EUROSTAT (portál EUROPA):
<http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes>

- **Intranet** na SU OPF: <L:\RAMIK\PUBLIC\Statistika>

- **Excel:** Statistické funkce..., Analýza dat..., Řešitel...
- **Speciální statistické programy:**
SPSS, Statgraphics, Statistica aj. viz
<http://freestatistics.altervista.org>
- **Uplatnění PC ve výuce statistiky:**
eLearning, Moodle...
- **Elektronická učebnice statistiky:**
<http://www.statsoft.com>

1. Úkol statistiky: zpřehlednění dat

- **Cíl:** analýza informací a odhalení zákonitostí skrytých v datech

2 přístupy:

1. Popisná statistika

2. Induktivní statistika

- Ad 1: kategorizace, charakteristiky - nyní
- Ad 2: předpoklady o vzniku dat (pravděpodobnostní rozdělení) - později

Statistická jednotka

Objekty statistického zkoumání –

statistické jednotky

= to co zkoumá statistika

Příklady:

- obyvatelé v ČR, byty v ČR, domy v ČR,
- zákazníci supermarketu TESCO,
- zaměstnanci, firmy,
- organizace určitého typu: prodejny potravin, supermarkety (např. Hypernova),
- studenti SU OPF,
- voliči,
- výrobky,
- události (uzávěrky, ^{Statistika} úrahy, vrhy hrací kostkou) ²¹

Vymezení statistické jednotky

Statistická jednotka je vymezena alespoň ze 3 hledisek:

- **věcné hledisko**

(např. student VŠ mužského pohlaví)

- **prostorové hledisko**

(např. student SU OPF v Karviné)

- **časové hledisko** (např. v letošním akademickém roce student 1. ročníku)

Příklad ze sčítání lidu:

- **věcné hledisko**

všechny osoby

- **prostorové hledisko**

které se nacházejí na území České republiky

- **časové hledisko** v rozhodný okamžik, tj. o půlnoci z pátku 25. března na sobotu 26. března 2011

Statistický soubor

Statistický soubor = souhrn statistických jednotek **stejného** vymezení (věcného, prostorového, časového)

- Stat. soubor, který obsahuje **všechny** statistické jednotky daného vymezení - **základní soubor** (též **populační soubor** nebo krátce **populace**)
- Vybraná část základního souboru - **výběrový soubor**, též **vzorek**

Statistický znak

Statistické znaky = vlastnosti statistických jednotek statistických souborů

- znaky **kvalitativní** (někdy též slovní, textové nebo alfanumerické)

Příklady: pohlaví zákazníka, typ podniku, bydliště voliče, barva výrobku, chuť nápoje, spokojenost zákazníka apod.

- znaky **kvantitativní** (též číselné, metrické, měřitelné)

Příklady: tržby firmy za měsíc, cena výrobku, počet zákazníků za den, HDP státu v USD, výsledky vrhu hrací kostkou apod.

Kvalitativní znaky členíme na:

- **nominální znaky** (též jmenovité)
- **ordinální znaky** (též pořadové)

Hodnoty kvalitativních znaků = **kategorie**

Příklad 1: kategoriemi znaku „pohlaví zákazníka“ jsou „Muž“ a „Žena“ (nebo M a Ž, M – „Male“ a F – „Female“) – **kategorie jsou rovnocenné**

Příklad 2: kategoriemi znaku „spokojenost zákazníka“ mohou být 3 výrazy „nízká“, „průměrná“ a „vysoká“, nebo též 3 **kódy** „1“, „2“ a „3“ -

nejedná se o kvantitativní (číselný) znak !!!

Kategorie nejsou rovnocenné, lze je uspořádat

Kvantitativní znaky členíme na:

- **diskrétní znaky** (mají konečný nebo nekonečný počet hodnot)
- **spojité znaky** (mají **vždy** nekonečný počet hodnot)

Diskrétní znaky nabývají izolovaných číselných hodnot (celá čísla)

Příklady:

1. Počet zákazníků v prodejně za den - nabývá hodnot 0, 1, 2, 3,... atd., není shora omezen (alespoň teoreticky) - **nekonečný** diskrétní znak
2. Počet ok na hrací kostce je omezený, konkrétně nabývá hodnot 1, 2,..., 6, - **konečný** diskrétní znak

Kvantitativní znaky:

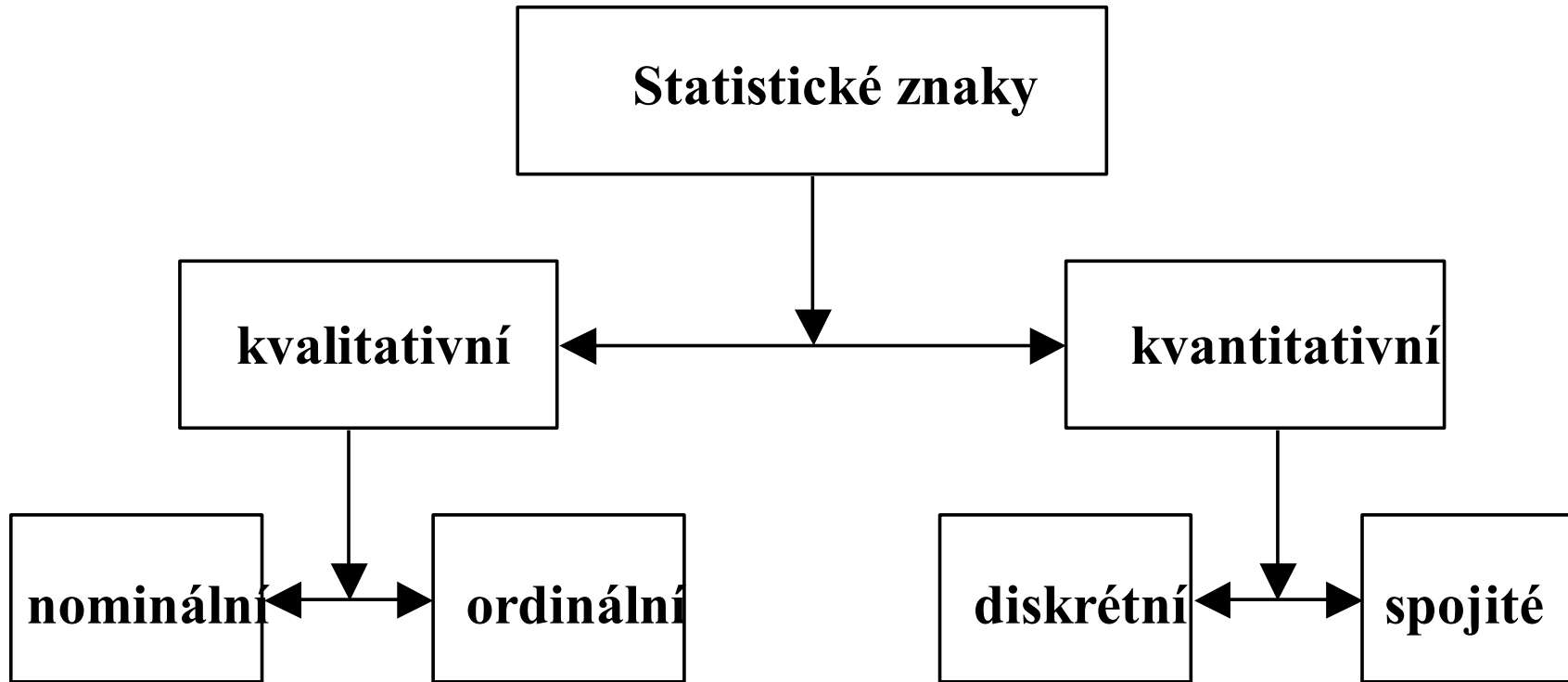
- **diskrétní znaky** (konečné nebo nekonečné)
- **spojité znaky** (vždy nekonečné)

Spojité znaky nabývají všech možných číselných hodnot z určitého číselného intervalu, omezeného i neomezeného (reálná čísla)

Příklady:

1. Cena výrobku, doba životnosti výrobku - nabývá hodnot z intervalu $(0, +\infty)$ není shora omezen (alespoň teoreticky)
2. Hmotnost výrobku, rozměry výrobku apod. - nabývá hodnot z intervalu $[a, b]$, je omezený

Statistický znak



Příklad 1. Firma

Výběrový soubor, statistické jednotky: zaměstnanci

ID	Pohlaví	Věk	Stav	Vzdělání	Funkce	MzdaRok	KvalitaStravy
5060	M	65	rozvedený	SŠ	dělník	258800	4
1030	M	60	rozvedený	VŠ	manažér	630000	2
3049	M	60	ženatý	ZŠ	technik	436600	5
5047	M	60	vdovec	ZŠOU	dělník	240600	3
5061	M	60	vdovec	ZŠOU	dělník	241800	1
5087	M	60	vdovec	SŠ	dělník	239500	
5133	Z	60	vdaná	SŠ	dělník	241100	4
5177	Z	60	vdova	SŠ	dělník	239600	4
3030	Z	58	vdova	ZŠ	technik	422600	1
3014	Z	56	vdova	VŠ	technik	303600	3
5012	Z	56	vdova	ZŠOU	dělník	223100	4
5056	M	56	rozvedený	ZŠ	dělník	225200	5
5101	M	56	svobodný	ZŠOU	dělník	224600	4
5106	M	56	ženatý	ZŠOU	dělník	226100	7
5146	Z	56	vdaná	ZŠOU	dělník	224900	3
5153	M	56	rozvedený	SŠ	dělník	224500	4
5189	M	56	ženatý	ZŠOU	dělník	224600	1
5196	M	56	vdovec	ZŠOU	dělník	222800	3
1031	M	55	ženatý	VŠ	manažér	429000	
5016	M	55	rozvedený	SŠ	admin. pr.	259000	5
5021	Z	55	vdaná	ZŠOU	dělník	220200	
5062	Z	55	vdova	ZŠOU	dělník	221400	5
5107	M	55	rozvedený	ZŠOU	dělník	220500	4
5154	Z	55	vdova	ZŠOU	dělník	219200	5
5195	M	55	ženatý	ZŠOU	dělník	219400	6