

# Statistika v sociálním lékařství

---

## 7. seminář



# STATISTIKA – definice

---

- ❑ věda studující **hromadné jevy**
- ❑ **jev** – výsledek pokusu n. pozorování
- ❑ věda, jejímž předmětem jsou **výsledky HROMADNÝCH POZOROVÁNÍ** → jejich sběr, analýza + využití pro rozhodování a předpovědi
- ❑ **Hromadné jevy**, kt. nelze před provedením pokusu n. pozorování zcela přesně předvídat → **HROMADNÉ NÁHODNÉ JEVY**

*Pozn.: NÁHODY nelze vyloučit, ale lze je studovat exaktními metodami teorie pravděpodobnosti*

---

# STATISTIKA

---

**Těžiště stat.metod** spočívá v

- ❑ racionálním přístupu k řešení problému
- ❑ plánování výzkumu
- ❑ správné interpretaci a objektivizaci závěrů

*Myšlení musí předcházet počítání !!!*

Základem statistiky je **teorie pravděpodobnosti.**

---

# Role STATISTIKY v medicíně

---

využití na úrovni:

- **populační** – úroveň a vývoj zdrav.situace;, odhad a hodnocení přínosu zdravotnických programů, důležitý nástroj při šetřeních zdrav.stavu populace a jeho determinant
  - **individuální** – stanovení správné diagnózy (mnoho kvalit.+kvantitat.údajů) + odhad prognózy
  - **medicínský výzkum** – ze stat. vychází moderní epidemiologické metody (příprava, realizace, interpretace i využití výsledků)
-

# Role STATISTIKY v medicíně

---

- Medicína má **pravděpodobnostní** povahu → lidé jsou odlišní, podkladem medicínských rozhodnutí jsou biologické **variabilní znaky** – proto výsledky nelze předpovědět, vždy provázeny nejistotou (variabilita dat je důsledkem působení velkého počtu drobných náhodných vlivů)
- Statistické metody – studium a měření variability a zvládat nejistotu- **odhad míry nejistoty pomocí teorie pravděpodobnosti** a tak **závěry**( tj. induktivní úsudky) **objektivizovat**

**Proto je výuka statistiky obsahem osnov všech lékařských fakult !!!**

---

# Výuka statistiky na LF

---

- ❑ Všude v Evropě, pregraduálně i postgraduálně (omezený časový prostor, stručná skripta)
  - ❑ Úvod do statistického myšlení, rozvoj logického a kritického myšlení, nejen vlastní technická stránka zpracování výsledků
  - ❑ **Myšlení musí předcházet počítání!**
  - ❑ Přínos pro všechny budoucí lékaře (nejen vědecké pracovníky) – schopnost pochopit a využít poznatky odborné literatury, jejich kritické zhodnocení, schopnost odlišit mezi domněnkou a validní, věrohodnou informací
-

# Pohled do minulosti....

---

- **A. Bradford Hill** – *Principles of Medical Statistics* (1937) – učebnice statistiky pro celé generace britských lékařů, základna komunikace mezi lékaři a statistiky
  - Problém: technická a numerická stránka výpočtu. Jak to vypočítat?
  - **A současnost ...výpočetní technika** → co vypočítat a jak získané výsledky interpretovat?
    - pozitiva: rychle a spolehlivě lze zpracovat objemné soubory dat, náročnější metody
    - negativa: neinformovaný uživatel → nesmysly
-

# Základní statistické pojmy

---

Od 30. let 20. století rozvoj teorie pravděpodobnosti a indukční statistiky

## Základní a výběrový soubor

- **ZS:** souhrn prvků (osob, případů nemoci, pokusů), jejichž vlastnosti chceme poznat – *konečný* n. *nekonečný rozsah*
- **VS:** vybraná část ZS, kterou budeme skutečně zkoumat (měření, dotazníky, testy)

## Vyčerpávající (úplné) vs. výběrové (neúplné) šetření

- Vyčerpávající – sledujeme všechny prvky ZS (př.sčítání lidu, PN..)
- Výběrové – vyšetříme pouze vybranou část ZS tzv. **výběr**

Výběr a ZS spojuje tzv. **statistická indukce** – **zobecnění** výsledků z výběru na celý ZS (usuzujeme-li induktivně z vlastností stat. jednotek VS na vlastnosti všech stat. jednotek ZS)

Objektivizujeme indukční závěry pomocí **teorie pravděpodobnosti**.(odhad míry nejistoty)

---



# Náhodný výběr - typy výběrů

*(náhodnost výběru je podmínkou správnosti statistické indukce).*

---

**podmínka reprezentativnosti** (správnost induktivních závěrů závisí na tom, jak dobře reprezentuje výběr celý základní soubor)

**Za náhodný považujeme výběr získaný postupem, který zaručuje, že každý prvek ZS má na začátku výběru stejnou pravděpodobnost, že bude vybrán.**

1. **Prostý náhodný výběr** – losováním nebo pomocí tabulek
2. **Náhodný výběr mechanický** (systematický) – např. počáteční písmeno příjmení
3. **Náhodný výběr oblastní** (stratifikovaný) – rozdělení do oblastí, strat
4. **Párový výběr (mačování)** – např. k osobám s jistou vlastností (nemocí) osoby se stejnou vlastností a bez nemoci

*Pozn.: reprezentativnost může být porušena i při sběru dat, např.*

- *neúplné chybějící údaje*
- *nevhodně zvolené otázky*
- *nejednoznačnost odpovědi atd.*

# Statistické šetření

---

## Etapy statistického šetření

1. **Plán šetření** ( předurčuje všechny ostatní etapy a použitelnost výsledků v praxi)
2. **Sběr dat**
3. **Popis a technické zpracování** (deskriptivní statistika)
4. **Rozbor, závěry a interpretace** (induktivní statistika)

---

*J. W. Goethe: „Kdo splete první knoflík, ten se už pořádně nezapne“*

# Co nás čeká....

---

- *Deskriptivní statistika* – popis a technické zpracování dat ( třídění, tabulky a grafy, statistické ukazatele)
  - *Induktivní statistika*
    1. Odhady parametrů ZS z výběrových charakteristik
    2. Testování statistických hypotéz
    3. Hodnocení závislostí – kvantitativních a kvalitativních veličin
-

# Popis statistického souboru (deskriptivní statistika)

---

Základní metody deskriptivní statistiky (popis souboru):

Význam DS není v pouhém popisu, ale je to nezbytné východisko k usuzování z výběru na základní soubor (tj. indukci).

## 1. Třídění

**cíl:** uspořádat + zpřehlednit velký soubor dat

## 2. Prezentace dat

**tabulky + grafy**

**cíl:** znázornit rozložení četností sledovaných znaků

## 3. Statistické charakteristiky (ukazatele)

**cíl:** charakterizovat sledované znaky pomocí výstižných ukazatelů

---

# Deskriptivní statistika - typy veličin

(*způsob popisu i třídění závisí na typu veličin*)

---

**KVALITATIVNÍ** (kategoriální) – **slovní** určení, *nelze měřit číselně*, lze pouze klasifikovat do různých kategorií (pohlaví, věk, ...)

1. **Nominální** – lze vyjádřit pouze slovně, nelze seřadit
  - a) **alternativní** – existují pouze 2 varianty (kuřák x nekuřák, muž x žena, ...)
  - b) **množné** – existují  $> 2$  varianty (diagnózy, barva vlasů, ...)
2. **Ordinální** – lze je seřadit dle nějaké míry (ZŠ – SŠ – VŠ, silný – slabý kuřák – nekuřák)

**KVANTITATIVNÍ** – lze vyjádřit *pouze číselně* (jejich obměny charakterizovány polohou na číselné ose)

1. **Diskrétní** (nespojité) – nabývají oddělených hodnot, vyjádřeny celými čísly (počet cigaret, počet onemocnění)
  2. **Spojité** – jejich hodnoty na sebe plynule navazují, desetinná čísla (výška, hmotnost, ...)  
v praxi lze spojité znaky převést na diskrétní
-

# Statistické třídění

---

- **Rozdělení souboru dat do skupin** (tříd, kategorií, intervalů) **podle** předem určených **třídících znaků** → základem třídění je vymezení skupin
  - Podle počtu třídících znaků
    - tř. **jednostupňové** ( → jednoduchá tabulka)
    - tř. **vícestupňové** (kombinační) (→ kombinační tabulka)
  - Produktem třídění je tabulka **rozdělení (rozložení) četností**
-

# Statistické třídění

---

- **Kvalitativní** veličiny – třídy vytvořeny již před sběrem dat
    - např.: pohlaví: muž, žena
    - vzdělání: ZŠ, SŠ, VŠ
    - kuřák: ano, ne
-

# Statistické třídění

---

## □ Kvantitativní veličiny

a) diskrétní (malý počet obměn) viz kvalit. veličiny -  
př. počet dětí v rodině, počet vykouřených cigaret

b) spojité – nejjednodušší je seřazení hodnot dle velikosti (↑přehlednost, chyby)

- při větším počtu obměn: **redukce dat** (shrnutí do tříd, některé hodnoty spojujeme) – př. měření TK → mizí rozdíl mezi sp. a dis. veličinami, ztráta informací(citlivě!), ale usnadnění prezentace

Tříděním zjišťujeme zastoupení jednotlivých tříd (četnosti) –  
absolutní, relativní, kumulativní

---



# Třídění

a) Třídění **jednostupňové** – rozdělení souboru podle kuřáckých návyků

b) Třídění **dvoustupňové** – rozdělení souboru podle kuřáckých návyků a vzdělání

c) Třídění **třístupňové** – rozdělení souboru podle kuřáckých návyků, vzdělání a pohlaví

**a**

	CELKEM
Nekuřák	389
Slabý kuřák	274
Silný kuřák	261
CELKEM	924

**b**

	ZŠ	SŠ	VŠ	CELKEM
Nekuřák	269	74	46	389
Slabý kuřák	213	44	17	274
Silný kuřák	197	50	14	261
CELKEM	679	168	77	924

**c**

MUŽI	ZŠ	SŠ	VŠ	CELKEM
Nekuřák	130	39	27	196
Slabý kuřák	110	24	8	142
Silný kuřák	140	32	8	180
CELKEM	380	95	43	518

ŽENY	ZŠ	SŠ	VŠ	CELKEM
Nekuřák	139	35	19	193
Slabý kuřák	103	20	9	132
Silný kuřák	57	18	6	81
CELKEM	299	73	34	406

# Statistické třídění

---

## Obecná pravidla

- ❑ Zvolené intervaly musí pokrývat celou škálu naměřených hodnot
  - ❑ Středy intervalů – okrouhlá čísla
  - ❑ Horní hranice předcházejícího intervalu musí být  $<$  než dolní hranice intervalu následujícího
  - ❑ Počet intervalů - obvykle 10 - 15
-

# Statistické třídění – postup

---

- **Variační šíře**  $/R/ = X_{\max} - X_{\min}$
  - **R se vydělí požadovaným počtem intervalů, zaokrouhlení** → délka intervalu (stejně dlouhé, ale i nestejně délky)
  - Středy intervalů (okrouhlá čísla)
  - Hranice intervalů – dolní hranice 1.intervalu je menší než  $X_{\min}$ , horní hr. posledního intervalu je větší než  $X_{\max}$
  - Třídní /intervalová/ četnost – zastoupení jednotlivých tříd (četnosti) – absol., relat., kumulativní
-

# Jak zvolit intervaly?

---

1. Počet intervalů (tříd)
2. Délka intervalů
3. Hranice intervalů

V úvahu se bere:

- počet dat (velikost souboru)
  - přesnost měření
  - výzkumný cíl
-

# Statistické třídění

---

Tříděním zjišťujeme zastoupení jednotlivých tříd (četnosti)

**Absolutní četnost** – kolik z naměřených hodnot spadá do jednotlivých intervalů

**Kumulativní četnost** – součet všech předchozích intervalů  
*15 mužů (6 + 9) má  $VC < 3,8$*

**Relativní četnost** - % z celkového počtu měření  
*četnost 6 ..... 3% z 200*

**Kumulativní četnost** – obdoba kumulativní četnosti v %

**Grafy** – tvar rozložení

---

# Příklad: Navrhněte třídění VKP pro 200 zdravých mužů ve věku 40 – 50 let

---

1. **Variační šíře (rozpětí)** – od největší naměřené hodnoty odečteme nejmenší

$$6,59 - 3,08 = 3,51$$

2. **Stanovení počtu intervalů** – závisí na mnoha faktorech (velikost souboru, podrobnost,...)

3. **Délka intervalu** – rozpětí/ počet intervalů

$$3,51/10 = 0,351 \text{ délka 1 intervalu}$$

pravidlo: a) okrouhlé číslo

b) ne víc deset. míst než měřená veličina

zaokrouhlit na **0,40**

4. **Hranice intervalu** – počátek – od nejmenšího čísla 3,08 tj. 3,00  
(otevřené v.s. uzavřené intervaly)

$$1. \text{ interval } 3,00 - 3,39 \quad \text{nebo} \quad <3,00 - 3,4)$$

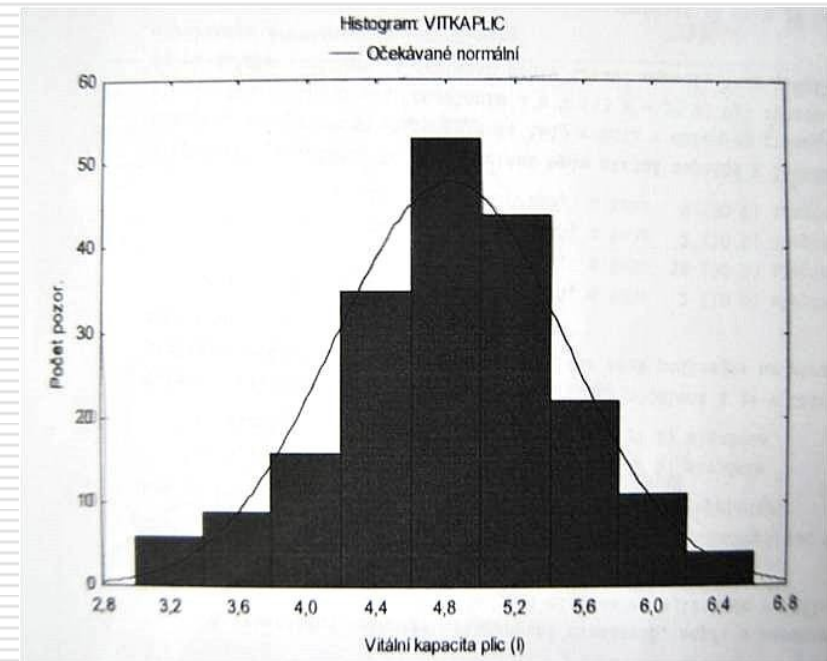
$$2. \text{ interval } 3,40 - 3,79 \quad <3,4 - 3,8)$$

---

# Tabulka: Vitální kapacita plic

		Tabulka četnosti: VITKAPLIC			
OD	DO	Četnost	Kumulativní četnost	Rel. četnost	Kumulativní rel. četnost
3,000000	<=x<3,400000	6	6	3,00000	3,0000
3,400000	<=x<3,800000	9	15	4,50000	7,5000
3,800000	<=x<4,200000	16	31	8,00000	15,5000
4,200000	<=x<4,600000	35	66	17,50000	33,0000
4,600000	<=x<5,000000	53	119	26,50000	59,5000
5,000000	<=x<5,400000	44	163	22,00000	81,5000
5,400000	<=x<5,800000	22	185	11,00000	92,5000
5,800000	<=x<6,200000	11	196	5,50000	98,0000
6,200000	<=x<6,600000	4	200	2,00000	100,0000
6,600000	<=x<7,000000	0	200	0,00000	100,0000
ChD		0	200	0,00000	100,0000

Popisné statistiky (Tabulka_vitální kapacita.sta)						
Proměnná	Průměr	Sm. odch.	Minimum	Maximum	N	Počet ChD
VITKAPLIC	4,825550	0,667008	3,080000	6,590000	200	0



# Deskriptivní statistika

## – **tabulky a grafy**

---

- TABULKY – výsledky třídění zapisujeme do tabulek rozdělení četností( vodorovné řádky- legenda, svislé sloupce- hlavička). *Zásady – skr. str. 10*
  - GRAFY – **jaké je rozložení** sledované veličiny? (*symetrické x asymetrické, jednovrcholové x vícevrcholové, U n. J rozložení..*)
    - **výběr vhodného ukazatele** pro popis souboru
-



Stupnici grafu volíme tak, aby celá  
plocha grafu byla rovnoměrně zaplněna.

! Vhodná volba stupnic na osách  $x, y$  !

22 / CHAPTER 2

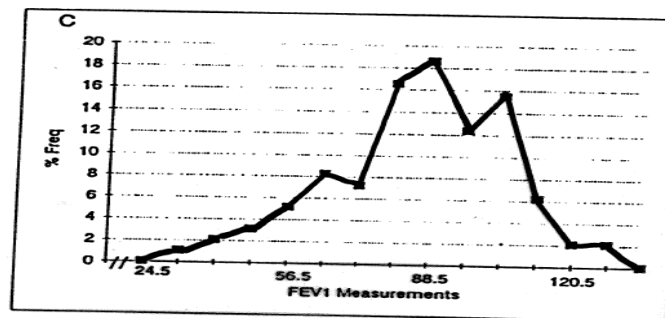
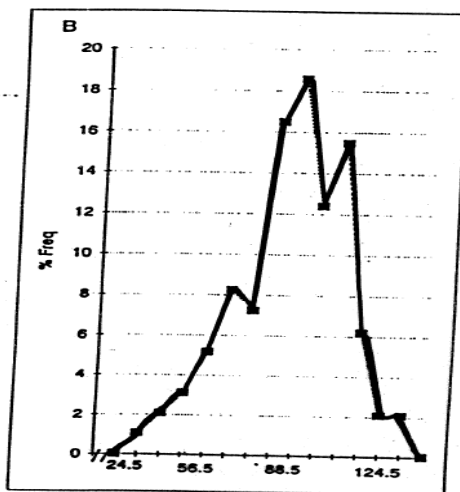
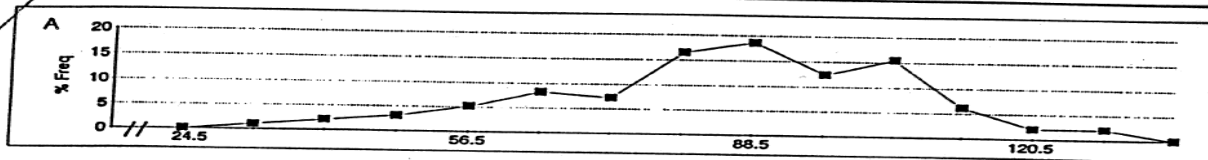
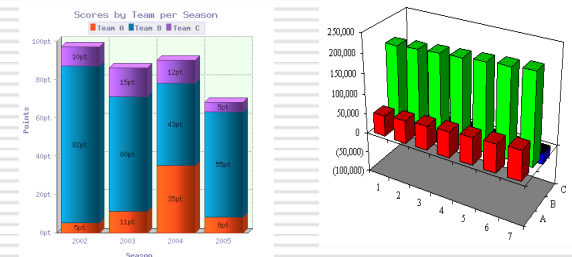


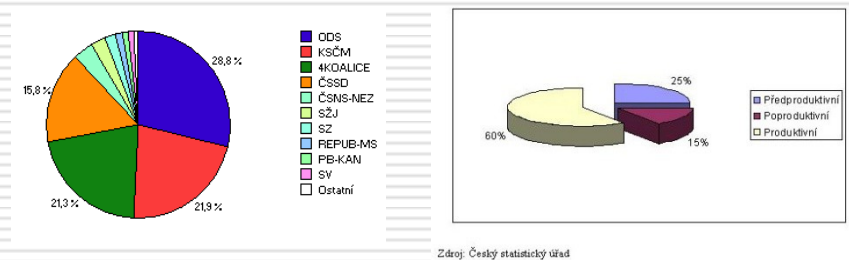
Figure 2-6. Effects of altering the relationship between the x- and y-axis.

# Grafy - kvalitativní veličiny

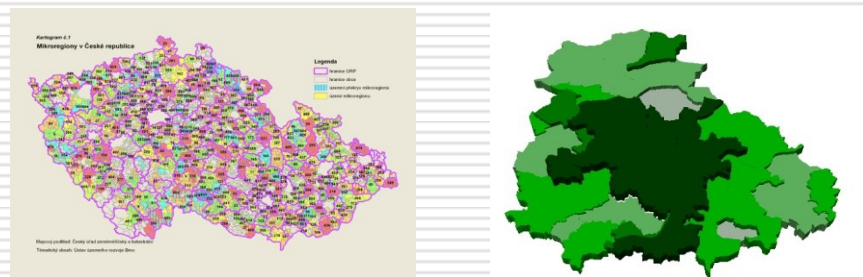
- Sloupcový graf (sloupce oddělené mezerou)



- Výšečový graf (struktura)



- Kartogram (regionální srovnání)



# Grafy

---

- **Bodový graf** - východisko studia závislosti kvantitativních veličin – informace o typu, směru a síle závislosti

K prezentaci výsledků statistického třídění (kvant.veličin) se obvykle používají –

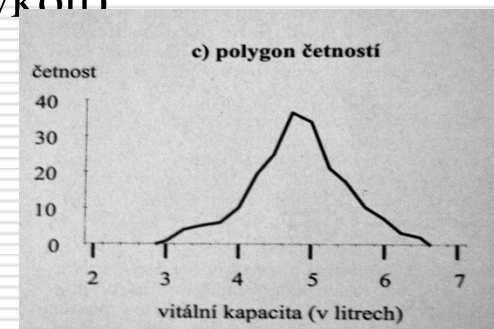
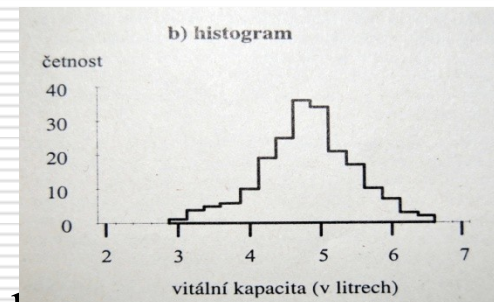
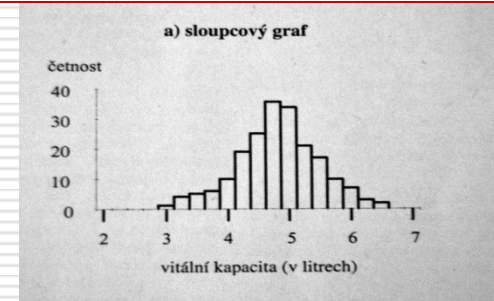
- **sloupcový graf**
- **histogram**
- **polygon četnosti**

*(tj. grafy znázorňující rozdělení četností –skr. str.10-11))*

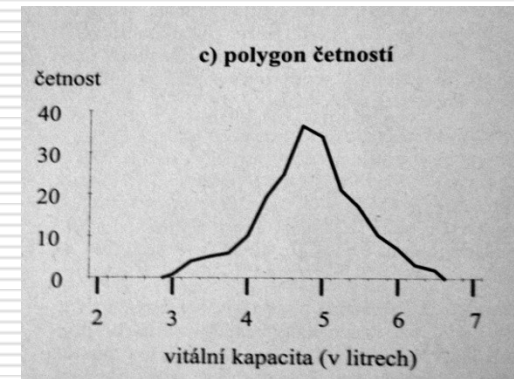
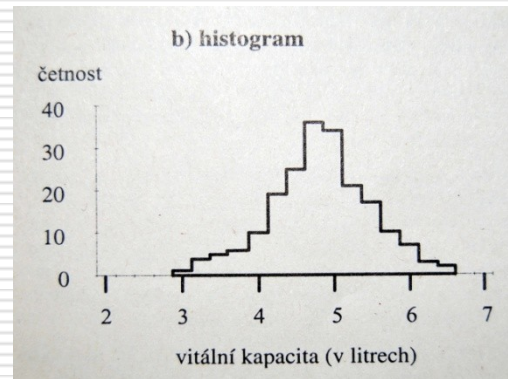
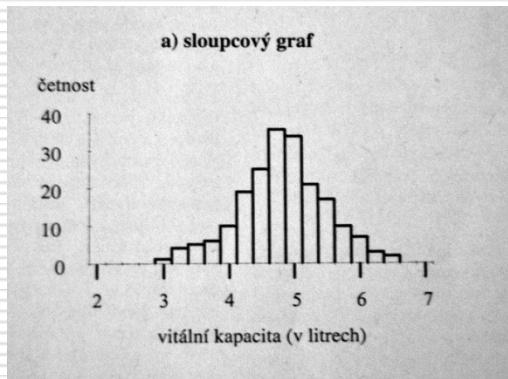
---

# Kvantitativní veličiny

- ❑ **Sloupcový graf** (plošný graf)
- ❑ **Histogram** (pouze obrysy sloupců)  
(spojnicový g.)
- ❑ **Polygon četností** (středky sloupců se spojí křivkou)  
(spojnicový g.)



# Grafy znázorňující frekvenci rozložení veličiny



osa **X** : naměřené hodnoty sledování veličiny

osa **Y** : četnost (abs. nebo v %) intervalů

## Tvar rozložení četností:

- Symetrické x asymetrické
- Jednovrcholové x vícevrcholové
- Podoba s teoretickými modely rozložení četností

# Statistické charakteristiky (ukazatele)

---

- **Kvalitativní veličiny** - relativní ukazatele – viz RS
- **Kvantitativní veličiny**
  - 1/ ukazatele polohy (střední hodnoty) - **aritmetický průměr**
    - **medián**
    - **modus**
    - **kvantil, percentil**
  
  - 2/ ukazatele variability - **rozpětí**
    - **rozptyl – směrodatná odchylka – variační koeficient**
    - **kvantily, percentily (nejméně dva)**  
(*šikmost, špičatost*)

**Volba ukazatele:**

1. **Tvar (typ)rozložení (symetrické X asymetrické)**
  2. **Typ sledovaného znaku**
-

# Statistické charakteristiky

---

Ukazatelé polohy i variability charakterizují rozdělení NV jak ve výběru (výběrové charakteristiky), tak v celém základním souboru (parametry).

- **Výběrové charakteristiky** – náhodné veličiny, jejichž hodnotu počítáme z dat výběrového souboru. Jejich hodnota se mění náhodně výběr od výběru.
  - **Parametry** základního souboru- pro daný ZS pevná čísla(neměnné konstanty), jejichž hodnotu neznáme
-

# Ukazatele polohy (střední hodnoty)

---

Většina hodnot, kterých mohou NV nabývat , se **kupí kolem** nějakého **pevného bodu**, zpravidla kolem středu rozdělení četností. Tento bod charakterizuje polohu souboru na číselné ose a ukazatele vystihující tuto vlastnost se nazývají **ukazatele polohy** (střední hodnoty) .

---



# Ukazatele polohy (střední hodnoty)

---

- **Aritmetický průměr(m)** – součet pozorovaných hodnot dělený počtem sledovaných jednotek
- **Medián( $m_e$ )** – hodnota prostředního člena souboru, kt. je uspořádan podle velikosti (u sudého počtu – průměr ze dvou prostředních hodnot ), dělí soubor dat uspořádaný dle velikosti na dvě stejné části . Pořadová charakteristika
- **Modus ( $m_o$ )** – hodnota nejčastější, pro daný soubor nejvíce typická, leží v modálním intervalu, lze i grafickou metodou
- **Kvantily** (obecný název)- pořadové charakteristiky, dělí soubor dat na části o stanovené velikosti( tj. kvanta).**Percentily**( kvanta uvedena v %), Percentil  $P_x$  – x pořadí percentilu, percentily dělí soubor uspořádaný dle velikosti na části obsahující 1% celk. počtu stat. jednotek.**Decily**  
Medián = 5.decil a 50. percentil

---

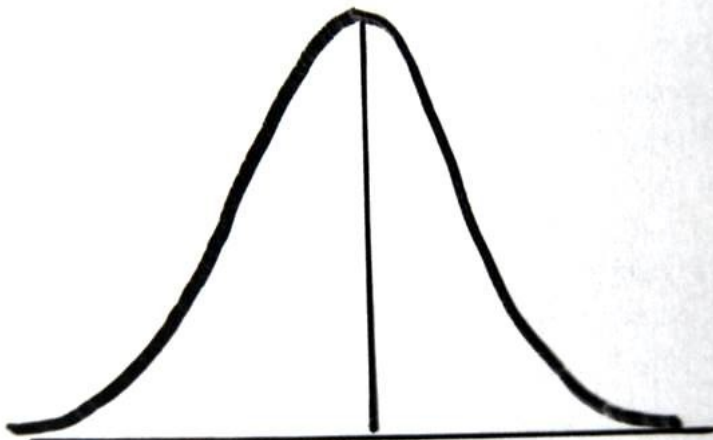
**aritmetický průměr nemá smysl počítat u asymetrických rozložení (náchylný k extrémním hodnotám)**

# Rozložení hodnot

---

Symetrické/asymetrické rozložení hodnot

SYMETRICKÉ ROZLOŽENÍ

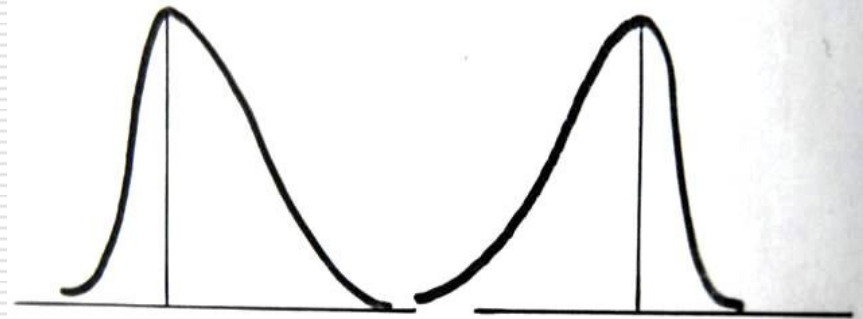


$$m = m_0 = m_2$$

ASYMETRICKÉ ROZLOŽENÍ

PRAVOSTR. ASYM.

LEVOSTR. ASYM.

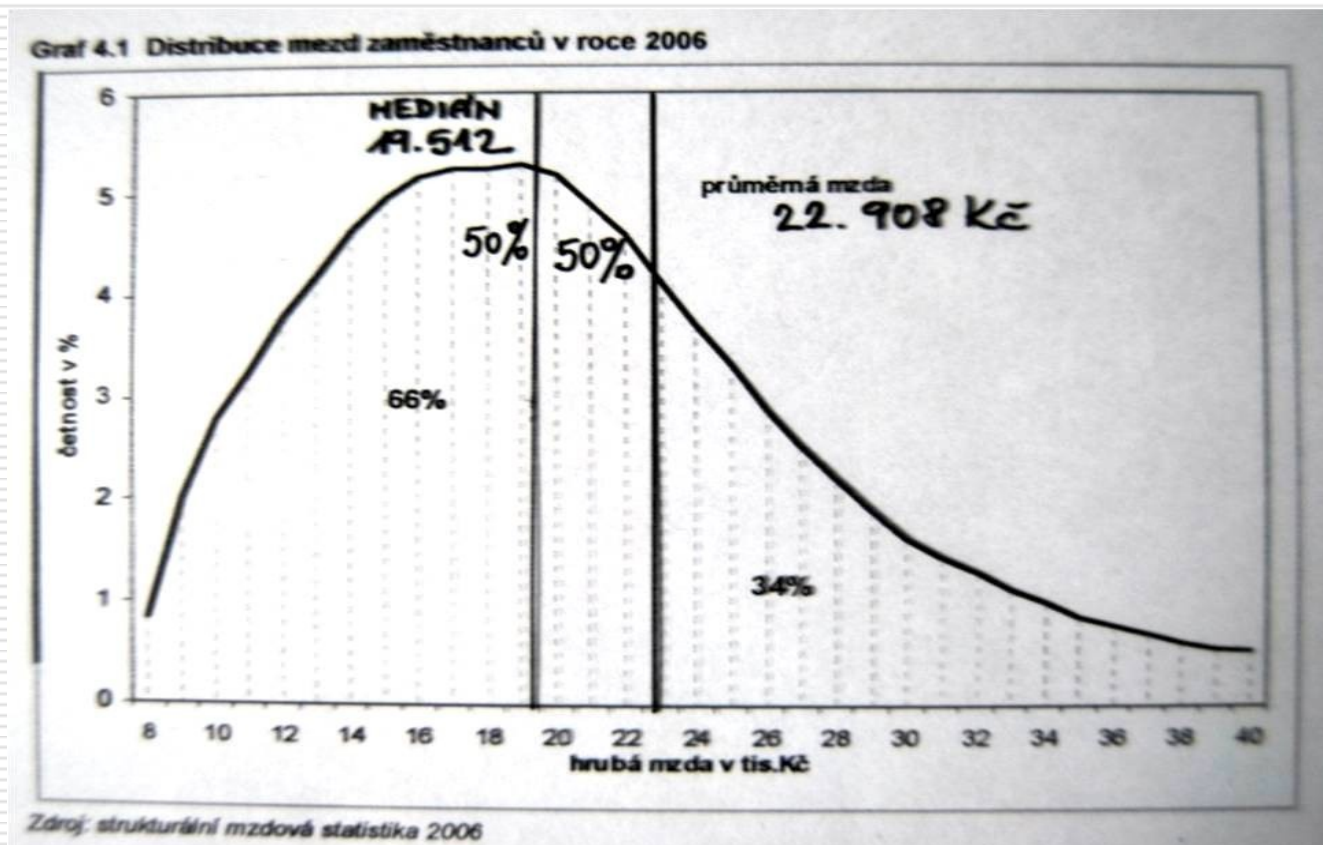


$$m_0 < m_2 < m$$

$$m < m_2 < m_0$$

# Př.: distribuce mezd zaměstnanců 2006

Obr.



# Ukazatele variability

---

- Další typická vlastnost NH je jejich **měnlivost** ( tj. **variabilita** )- hodnoty NV *kolísají v určitém rozmezí* kolem *středních hodnot*.

Ukazatelé kvantifikující míru tohoto kolísání ( rozptýlení ) se nazývají **ukazatelé variability**.

- Nejběžněji užívané ukazatele variability:
    1. Variační šíře(rozpětí)
    2. Rozptyl
    3. Směrodatná odchylka
    4. Variační koeficient
-

# Ukazatele variability

---

**Proč nestačí ukazatele střední hodnoty  
k výstižnému popisu souboru dat?**

Př.: Počet onemocnění u 2 skupin pěti kojenců v 1. roce života

1. skupina: 3, 4, 5, 6, 7     **m=5**

2. skupina: 0, 4, 5, 6, 10     **m=5**

obě skupiny mají stejný aritmetický průměr, ale liší se kolísáním hodnot  
– **VARIABILITOU**

**Spolu se střední hodnotou** by se měl uvádět **ukazatel variability**

**Průměr:** rozpětí, rozptyl, **směrodatná odchylka**, variační koeficient

**Modus, medián:** percentily

---

# Ukazatele variability

---

**Variční rozpětí (šíře):**  $R = X_{\max.} - X_{\min.}$  Pro  $N$  menší a rovno 10

**Rozptyl :** průměr čtverců odchylek jednotlivých pozorování od aritmetického průměru

$S^2$

$$S^2 = \frac{1}{n} \sum (x_i - m)^2$$

**Směrodatná odchylka:** odmocnina z rozptylu

$S$

$$S = \sqrt{S^2}$$

- ukazatel variability udávaný ve stejných jednotkách jako měřená veličina
- vypovídá o tom, o kolik se většina hodnot sledovaného znaku odchyluje od průměru

## Výpočet rozptylu

Porodní délka 5 novorozenců (cm)

$x_i$	$x_i - m$	$(x_i - m)^2$
49	- 1,6	2,56
50	- 0,6	0,36
50	- 0,6	0,36
51	+ 0,4	0,16
53	+ 2,4	5,76
253	0,0	9,20

$$m = \frac{253}{5} = 50,6 \text{ cm}$$

$$s^2 = \frac{9,20}{5} = 1,84 \text{ (cm}^2\text{)}$$

$$s = \sqrt{1,84} = 1,36 \text{ cm}$$

$$v. k. = \frac{1,36}{50,6} \cdot 100 = 2,69 \%$$

↓  
tzv. relativní ukazatel variability

# Ukazatele variability

---

**Variační koeficient** - relativní ukazatel variability , bezrozměrná veličina

$$v.k. = \frac{s}{m} \cdot 100 \quad (\%)$$

- podíl směrodatné odchyly a průměru

**Využití v.k.** 1/ srovnání variability dvou či více souborů se značně odlišnými průměry

2/ srovnání variability znaků uváděných v různých měrných jednotkách

---



# Využití variačního koeficientu

---

- 1) Je-li v. k.  $\geq 50\%$ , pak je soubor natolik nesourodý, že nemá smysl charakterizovat ho aritmetickým průměrem.
  - 2) Slouží ke srovnání variability 2 souborů, jejichž průměry se značně liší.  
Př.: tělesná výška 3 letých a 18 letých chlapců  
31 Ch **95,3  $\pm$  4,4 cm**      **v.k. = 4,6 %**  
181 Ch **173,4  $\pm$  6,3 cm**      **v.k. = 3,6 %**
  - 3) Slouží ke srovnání variability znaků uváděných v různých jednotkách.  
Př.: soubor 3 letých chlapců, tělesná výška (cm), hmotnost (kg)  
výška            95,3  $\pm$  4,4 cm      **v.k. = 4,6 %**  
hmotnost        14,9  $\pm$  1,7 kg      **v.k. = 11,4 %**
-

# Směrodatná odchylka (standardní odchylka)

---

- významná v indukční statistice, jako výběrová charakteristika je východiskem pro odhad směr.odch. ZS
  - Cennou vlastností sm.odch.je její **schopnost vymezit hranice**, ve kterých se nachází **určité množství stat. jednotek** → cca platí (za předpokladu normálního rozdělení výběrového souboru), že
-

# Směrodatná odchylka

---

- v intervalu  $m \pm s$  leží **68 %** údajů
- v intervalu  $m \pm 2s$  leží **95 %** údajů
- v intervalu  $m \pm 3s$  leží **99,7%** údajů

Lze určit interval pro libovolně zvolený násobek směrodatné odchylky.

---

## Ukazatele variability pro asymetrická rozložení četností

---

- **Variabilitu vyjadřujeme pomocí dvou kvantilů** – percentilů, decilů, **Kvantily** – pořadové ukazatele, které dělí soubor uspořádaný dle velikosti na části o stanovené velikosti (kvanta).
- Ve zdravotnické statistice – **percentily** (kvanta v procentech)
- Variabilita = určení intervalu, ve kterém se pohybuje 80% (P<sub>10</sub> – P<sub>90</sub>), případně 50% (P<sub>25</sub> - P<sub>75</sub>) pozorování.

Postup:

1. Určíme hodnotu pozorování, které představuje 10. percentil = dolní hranice intervalu
2. Určíme hodnotu pozorování, které představuje 90. percentil = horní hranice intervalu

$$P_x = n \cdot \frac{x}{100} + 1$$

# Percentilové růstové grafy

---

**Auxologie** – obor, který se komplexně zabývá růstem a vývojem člověka.

- umožňují pediatrům a rodičům, aby podle návodu připojeného ke grafům průběžně hodnotili všechna základní růstová data dítěte od narození až do jeho osmnácti let (tělesná výška, tělesná hmotnost, obvod hlavy, obvod paže, ...)
  - Zároveň je grafy seznamují s **variabilitou** těchto základních antropometrických rozměrů pro každou věkovou skupinu chlapců a dívek současné české populace
  - Zcela snadno tak lze zjistit, kolik např. měří nejmenší děti (3. -10. percentil), jak vysoké jsou největší děti (90. – 97. percentil) a kolik měří dítě zcela průměrné (50. percentil).
-

# Jaké charakteristiky použít? /dle typu rozložení/

---

## □ **Symetrická rozložení**

- průměr

- směrodatná odchylka

## □ **Asymetrická rozdělení**

- **modus, medián, 2 percentily**

(např.  $P_{10}$   $P_{25}$   $m_e$   $P_{75}$   $P_{90}$ )

*Transformace (např. logaritmická) – převede nesymetr. rozdělení na symetr., pak lze použít  $m$  a  $s$*

---

# Úkol:

---

Máme soubor 200 hodnot VCP, které jsme naměřili ve výběru 200 mužů (40-50 let)...

$$\underline{n = 200, m = 4,824, s = 0,668}$$

Stanovte (pomocí směr.odchylky a průměru) hranice - tj. intervaly, ve kterých se nachází 68 %, 95% a téměř 100 % naměřených hodnot VCP. (na 2 desetinná místa)

---

# Úkol:

---

Které percentily odpovídají  
jednonásobku a dvojnásobku  
směrodatné odchylky? (tj. intervaly  
 $m \pm s$  a  $m \pm 2s$  vyjádřete pomocí  
percentilů.)

---