

# Metody sociálních výzkumů

---

Velmi skromný úvod do statistiky.

Motto: "Jsou tři druhy lži-lež prostá, lež odsouzeníhodná a statistika."

# Statistika

---

- Význam slova-vychází ze slova stát, s jeho administrativou a zaznamenáváním údajů k vojenským a daňovým účelům.
  - Při statistické analýze výsledků zpracování používáme metody POPISNÉ (deskriptivní) a metody INDUKTIVNÍ statistiky.
  - **Deskriptivní (popisná statistika)** -zabývá se uspořádáním souborů, jejich popisem a účelnou sumarizací.
  - Př. popis výběrového souboru policistů v Praze.
  - **Induktivní statistika** -umožňuje ze získaných dat vytvářet obecné závěry s udáním stupně jejich spolehlivosti.
-

# Deskriptivní statistika

---

## POPISNÁ STATISTIKA

Umožňuje přehledné uspořádání dat

- Statistické třídění-rozdělení prvků ze souboru do skupin (tříd) podle předem určených znaků.
  - Např. třídění novorozenců podle pohlaví, třídění zemřelých podle pohlaví, věku a zaměstnání.
  - **Četnost**-počet prvků zařazených do určité třídy - **absolutní četnost** (např. počet novorozených holčiček je 159), **relativní četnost** je dána poměrem absolutní četnosti a rozsahu souboru (počet holčiček z celkového souboru novorozenců je  $159/300$ ), udává se nejčastěji v procentech. Kumulativní četnost vzniká postupným načítáním četností.
-

# Deskriptivní statistika

---

- **Deskriptivní (popisná statistika)** zabývá se uspořádáním souborů, jejich popisem a účelnou sumarizací.
  - Př. popis výběrového souboru policistů v Praze-cvičení.
  - Tam, kde je to možné, uveďte průměr, medián a modus souboru.
-

# Míry středu (polohy)

---

Míry středu (Measures of Central Tendency):

- ❑ Průměr - počítá se jako aritmetický průměr, určuje polohu rozdělení na číselné ose. Lze počítat pro intervalová a poměrová data.
  - ❑ Medián - rozděluje distribuci na 2 shodné poloviny (je prostřední hodnotou).
  - ❑ Modus - je hodnota veličiny, která se vyskytuje nejčastěji.
-

# Modus

---

- Je hodnota, která se vyskytuje nejčastěji.
  - Lze počítat u všech typů proměnných, u nominálních je jedinou mírou středu.
  - Má-li veličina normální (Gaussovo) rozdělení, pak medián, modus i průměr jsou totožné!
-

# Míry středu (polohy)

---

- Cvičení:
  - Vypočtete průměrnou výplatu sociálních pracovníků v jedné neziskové organizaci. Vypočtete medián z výplat.
  - 15000,-Kč-4 pracovníci, 16200-1 pracovník, 16300-4 pracovníci, 17500-5 pracovníků, 17000-2 pracovníci, 19000-6 pracovníků, 23000-2 pracovníci, 25000-2 pracovníci.
-

# Analýzy a interpretace výsledků

---

- TYPY PROMĚNNÝCH-při zaznamenávání informací (dat) využíváme různé škály. **Pokud data dále zpracováváme, musíme si vždy uvědomit, s jakým typem škály pracujeme!**
  - **Nominální proměnné** = jednoduchý klasifikační systém (např židle, stoly, tabule; nebo typ odpovědi-správně, nesprávně). Nominální dichotomická proměnná-žena, muž. Nominální polytomická proměnná-např. národnost.
  - **Ordinální proměnné (pořadové)** - určují pořadí, například pořadí běžců v závodě (ale nevíme, jaký čas je od sebe dělil, mohl být zcela různý)
  - **Kardinální proměnné** - můžeme s nimi provádět veškeré matematické operace (plus, minus, krát, děleno).
-



# Analýzy a interpretace výsledků- střední hodnoty

---

- ❑ STŘEDNÍ HODNOTY
- ❑ PRŮMĚR-používáme, když čísla můžeme opravdu sčítat, tj. pracujeme s kardinálními proměnnými. Znaky jsou kvantitativní, měření na číselné stupnici. Je velmi citlivý na **odlehle hodnoty!**
- ❑ Výpočet- $x_1+x_2+x_3+\dots+x_n/n$
- ❑ Pozor!-tzv. průměrný plat v ČR-proč ho dosahuje tak málo lidí?

# Analýzy a interpretace výsledků- střední hodnoty

- ❑ **MODUS**=hodnota, která se v souboru dat vyskytuje nejčastěji.
- ❑ 1,1,1,2,2,5,8,8,9,9,9,9,10,11,12,12,12
- ❑ Důležitá je zejména pro **nominální proměnné**
- ❑ **MEDIÁN**=hodnota, která rozdělí pozorování na dvě stejně velké skupiny (**důležitá u ordinálních proměnných**)
- ❑ Př. 61, 49, 35, 74, 53, 82 (medián bude průměr hodnot 53+61)

barva očí	četnost výskytu	
modrá	5	
zelená	12	
hnědá	10	
smíšená	4	
černá	2	

# Analýzy a interpretace výsledků

---

- MEDIÁN-př. Co je mediánem hodnocení závažnosti průběhu onemocnění? A=nejlehčí, F nejtěžší průběh.
  - C,E,B,D,A,A,B,F,C,C,D
-

# Analýzy a interpretace výsledků

---

- Řešení-seřadíme vzestupně
  - A,A,B,B,C,C,C,D,D,E,F
  - Mediánem je stupeň onemocnění C
  - Příklad. Co je mediánem u výplat:
  - 15 tis, 20 tis, 22 tis, 14 tis, 45 tis, 12 tis, 14 tis, 10 tis, 13 tis?
  - 10, 12, 13, 14, 14, 15, 20, 22, 45
-

# Analýzy a interpretace výsledků

---

- Řešení 10, 12, 14, 14, 15, 20, 22
  - Př. Vypočtete aritmetický průměr u předchozího zadání výplat: 15 tis, 20 tis, 22 tis, 14 tis, 30 tis, 12 tis, 14 tis, 10 tis?
  - Př. Určete modus u předchozího zadání.
-

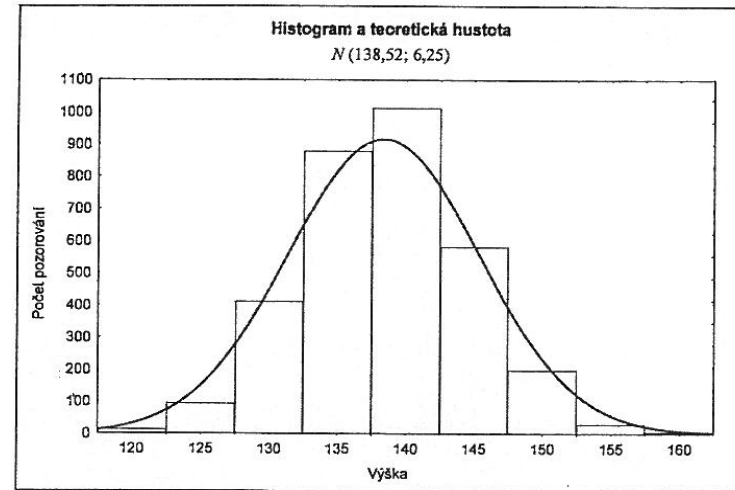
# Gausova křivka

---

- Unimodální symetrická křivka distribuce tvaru zvonu, s polovinou skóru nad průměrem a s polovinou skóru pod průměrem.
  - Tzv. Gausova křivka nebo-li normální rozdělení-tj. rozdělení, řídicí se určitým modelem (pro spojité veličiny)
  - Průměr  $\mu$
  - Rozptyl  $\sigma^2$  (variance);  $\sigma$  směrodatná odchylka
-

# Gausova křivka

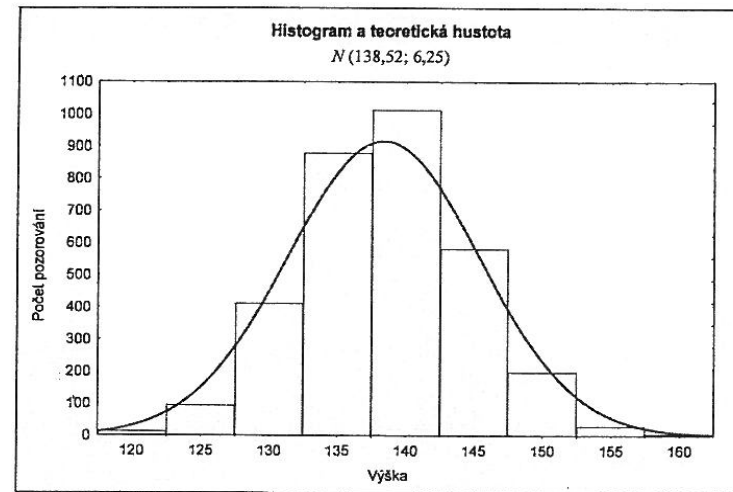
- V praxi se setkáváme s normálním rozložením u řady veličin (př. tělesná výška, tlak....)
- **Je-li veličina normálně rozdělena, pak průměr, medián a modus jsou shodné, mají stejnou hodnotu.**
- Směrodatná odchylka-parametr  $\sigma$ , určuje tvar hustoty normálního rozložení. Směrodatná odchylka je MIRA VARIABILITY.



Obrázek 5.1: Histogram výběrového rozdělení tělesné výšky 3231 chlapců ve věku 9,5–10 let (délka třídního intervalu 5 cm) a teoretická hustota normálního rozdělení tělesné výšky

# Gausova křivka

- V NORMÁLNÍM ROZDĚLENÍ:
- 1. téměř 70 % hodnot leží méně než  $1 \sigma$  od průměru
- 2. 95% hodnot leží ve vzdálenosti menší než  $2 \sigma$  od průměru.
- 99 % hodnot leží ve vzdálenosti menší než  $2 \sigma$  od průměru.

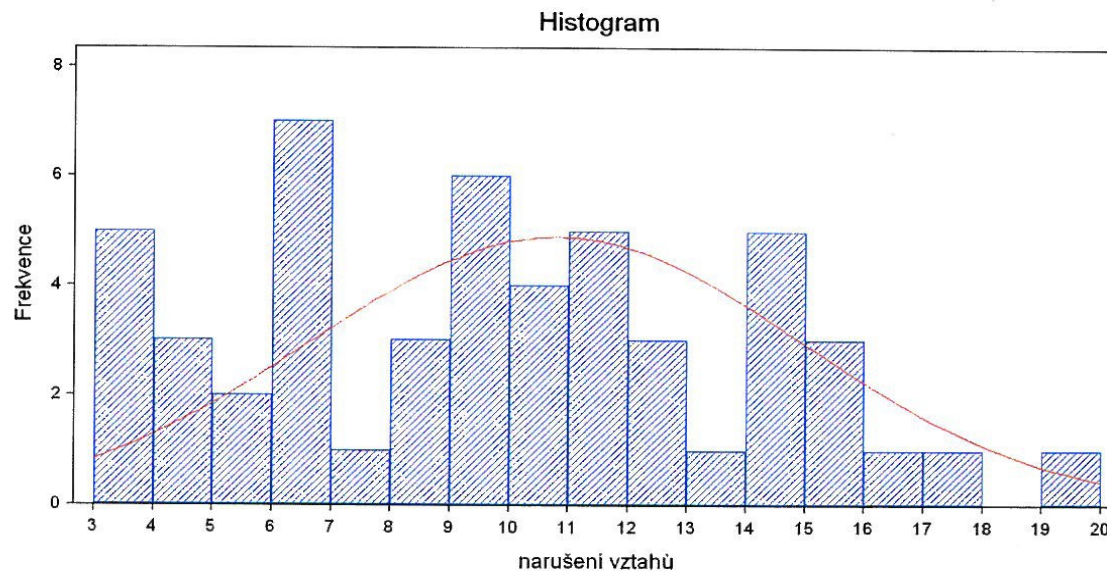


Obrázek 5.1: Histogram výběrového rozdělení tělesné výšky 3231 chlapců ve věku 9,5–10 let (délka třídního intervalu 5 cm) a teoretická hustota normálního rozdělení tělesné výšky



# Analýzy a interpretace výsledků

Obr. 14 Frekvence hrubých skóre Neuro-Qol oblast „narušení vztahů v důsl. SDN“



V oblasti „emoční potíže“ bylo minimum pro hrubé skóre 7, maximum 35. Pacienti se SDN dosahli min. 7, max. 35, průměru 21,3 (STD=7,00). Data splňují kritérium normálního rozložení.

# Analýzy a interpretace výsledků

## 7.3.3 Depresivní symptomatika měřená pomocí Zungovy škály depresivní výsledky

Minimální dosažená hodnota hrubého skóre u pacientů se SDN v Zungově testu byla 26, maximální hodnota 58, průměr 38,53, SD 7,46. Z celkového počtu 54 pacientů se SDN vykazovalo zvýšenou depresivní symptomatiku (cut-off 40) 24 pacientů, tj. 44%. Z počtu 24 pacientů s depresivní symptomatikou bylo 13 žen (tj. 70% z celkového počtu žen) a 11 mužů (tj. 35% z celkového počtu mužů). Žádný z pacientů nedosáhl cut-off 60 pro těžkou depresi. Při srovnání depresivní symptomatiky souboru pacientů se SDN a souboru zdravé populace vykazali signifikantně vyšší míru depresivní symptomatiky pacienti se SDN (viz tabulka 7.).

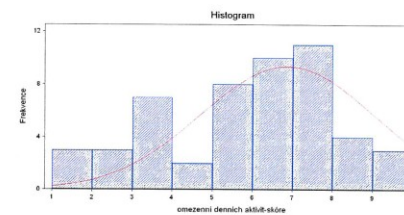
Tab.7 Porovnání míry depresivity u souboru pacientů se SDN a souboru zdravých jedinců

Zungova škála depresivní průměr pacientů	Zungova škála depresivity průměr zdraví	t-test	p	průměry jsou rozdílné
38,53 (STD7,46)	31,58 (STD 9,19)	4,50	0,0001	

## 7.3.4 Výsledky-dopad syndromu diabetické nohy na kvalitu života v oblasti psychosociální (omezení denních aktivit, narušení vztahů, emoční obtíže)

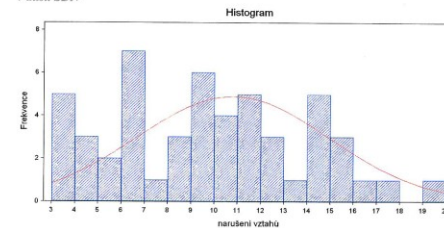
Z celkového počtu 54 pacientů se SDN vyplnilo dotazník Neuro-Qol, část oblast psychosociální 51 pacientů. Tato oblast obsahuje (po úpravách původní verze) celkem 14 položek, které pokrývají 3podoblasti. Minimální možné celkové hrubé skóre bylo celkem 14, maximální 70. Pacienti dosáhli min.14, maxima 62, průměr celkového skóre byl 38,84 (STD=11,96). V podoblasti „omezení denních aktivit“ bylo minimum 2, maximum 10. Pacienti se SDN dosáhli min.2, max.10, průměr byl 6, 01 (STD= 2,55). Data splňují kritérium normálního rozložení.

Obr.13 Frekvence hrubých skóre Neuro-Qol oblast „omezení denních aktivit v důsl. SDN“



V oblasti „narušení vztahů“ bylo minimum pro hrubé skóre 4, maximum pro hrubé skóre 20. Pacienti se SDN dosáhli min. 4, max. 20, průměru 10,13 (STD= 4,15). Data splňují kritérium normálního rozložení.

Obr. 14 Frekvence hrubých skóre Neuro-Qol oblast „narušení vztahů v důsl. SDN“



V oblasti „emoční potíže“ bylo minimum pro hrubé skóre 7, maximum 35. Pacienti se SDN dosáhli min. 7, max. 35, průměru 21,3 (STD=7,00). Data splňují kritérium normálního rozložení.

# Frekvenční distribuce

---

=tabulka vyjadřující frekvenci jednotlivých skóre (výsledků).

Cvičení-v rámci výzkumu emocí nás zajímá, jakou emoci považují studenti za „nejnakažlivější“: štěstí, vzrušení, smutek, vztek, strach, zájem.

Provedte a výsledky přehledně zaznamenejte do tabulky.

---

# Frekvenční distribuce

---

Kategorie      Frekv.(četnost)      Frekv.  
(procent)

- Vztek
  - Vzrušení
  - Strach
  - Štěstí
  - Zájem
  - Smutek
-

# Frekvenční distribuce

---

- ❑ Groupování frekvenční distribuce-
  - ❑ Redukujeme množství kategorií, tím zvýšíme přehlednost dat.
  - ❑ Viz předchozí cvičení-groupování na kategorie:
  - ❑ Pozitivní emoce (štěstí, zájem, vzrušení)
  - ❑ Negativní emoce (vztek, strach, smutek)
  - ❑ (kapitola 4 in Davis and Smith: An Introduction in Statistics nad Research Methods)
-

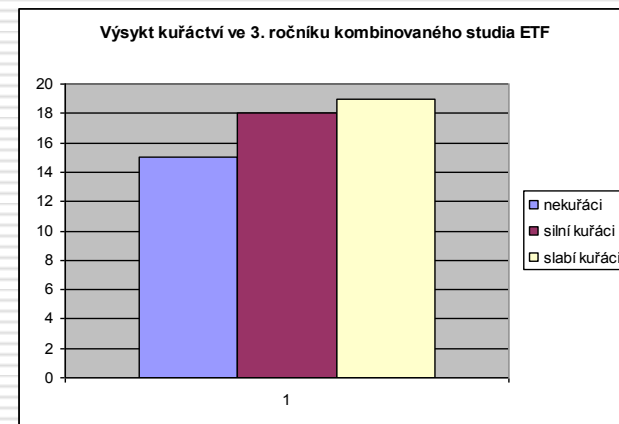
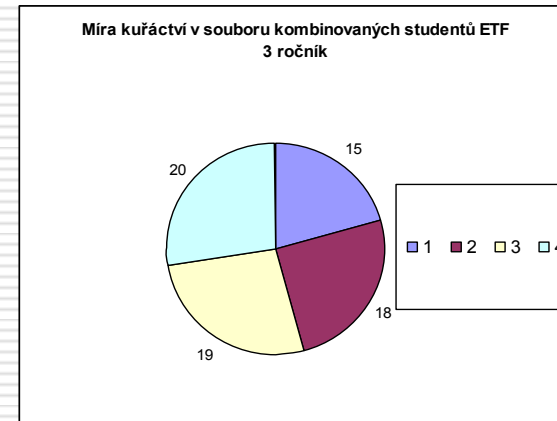
# Frekvenční distribuce

---

- ❑ Groupování frekvenční distribuce-
  - ❑ Redukujeme množství kategorií, tím zvýšíme přehlednost dat.
  - ❑ Viz předchozí cvičení-groupování na kategorie:
  - ❑ Pozitivní emoce (štěstí, zájem, vzrušení)
  - ❑ Negativní emoce (vztek, strach, smutek)
  - ❑ (kapitola 4 in Davis and Smith: An Introduction in Statistics nad Research Methods)
-

# Analýzy a interpretace výsledků

- ❑ GRAFICKÉ ZNÁZORŇOVÁNÍ
- ❑ Je velmi účinný způsob, jak prezentovat statistické údaje.
- ❑ Graf je přepsání číselných údajů do soustavy geometrických obrazců.
- ❑ Druhy grafů-bodový, sloupcový, spojnicový, kruhový



# Obraz řekne více než tisíc slov

---

- ❑ Aneb grafické znázornění
  - ❑ Příklad viz předchozí výzkum s emocemi (nominální data)
  - ❑ Koláčový graf (Pie Chart)-vhodný například pro procentuální frekvenční distribuci (celek = 100%).
  - ❑ Sloupcový graf (Bar Graph) – znázorňuje frekvenci dat v jednotlivých kategoriích. Vhodný pro nominální data.
-



# Obraz řekne více než tisíc slov

---

- Histogram-pro kvantitativní kategorie, které lze řadit od nejnižší po nejvyšší
  - Frekvenční polygon-od histogramu se liší graficky-křivka bodů
  - Čárový graf-souřadnice  $x, y$ . Na  $x$  nezávisle proměnnou, na  $y$  proměnnou, kterou měříme.
-

# Induktivní statistika

---

- Induktivní statistika-umožňuje ze získaných dat vytvářet obecné závěry s udáním stupně jejich spolehlivosti.
  - Stanovení  $H_0$  a  $H_1$
  - Potvrzení či vyvrácení  $H_1$  matematickým výpočtem na zvolené hladině pravděpodobnosti (nejčastěji  $p=0,05$ )
-

# Induktivní statistika - výpočet korelace

---

- ❑ KORELACE-Pearsonův korelační koeficient měří sílu závislosti mezi dvěma kardinálními proměnnými
  - ❑ Hodnota korelačního koeficientu se pohybuje v intervalu  $-1;1$
  - ❑ 0 je roven tehdy, pokud veličiny jsou nezávislé.
  - ❑ Korelace neznamena příčinnost!
  - ❑ Pro velký soubor dat může vyjít korelace i při slabém vztahu!
  - ❑ Síla korelace:  $r=1$  naprostá závislost
  - ❑  $r=1-0,9$  velmi vysoká závislost atd.
-