

# 4EK211 Základy ekonometrie

ZS 2016/17 Cvičení 1: Opakování ze statistiky



---

LENKA FIŘTOVÁ

KATEDRA EKONOMETRIE, FAKULTA INFORMATIKY A STATISTIKY

VYSOKÁ ŠKOLA EKONOMICKÁ V PRAZE

# Z čeho studovat

---

## 1) Z KNIHY

→ Krkošková, Š., Ráčková, A., Zouhar, J.: Základy ekonometrie v příkladech, 2. vydání, VŠE, 2010

→ Hušek, R.: Ekonometrická analýza, Oeconomica, 2007

→ Hušek, R.: Aplikovaná ekonometrie: teorie a praxe, Oeconomica, 2009

## 2) Z WEBU: <http://jakplavejak.cz/node/1>

→ prezentace ze cvičení

## 3) PŘIJÍT NA KH: úterý 18:00 – 19:30, pátek 9:15-10:45

**PRŮBĚŽNÝ TEST: 30 BODŮ – 3.11.**

# 1.1 Charakteristiky polohy

---

1. Střední hodnota

2. Medián

3. Modus

A) Nejčastěji se vyskytující hodnota v populaci.

B) Dlouhodobý průměr výsledků při opakovaném náhodném výběru z populace.

c) Hodnota oddělující horní polovinu populace od spodní.

# 1.1 Charakteristiky polohy

---

1. Střední hodnota

2. Medián

3. Modus

A) Nejčastěji se vyskytující hodnota v populaci.

B) Dlouhodobý průměr výsledků při opakovaném náhodném výběru z populace.

c) Hodnota oddělující horní polovinu populace od spodní.

# 1.1 Charakteristiky polohy

1. Ve 2. čtvrtletí 2014 činila \_\_\_\_\_  
mzda 25 500 Kč.

a) průměrná      b) mediánová

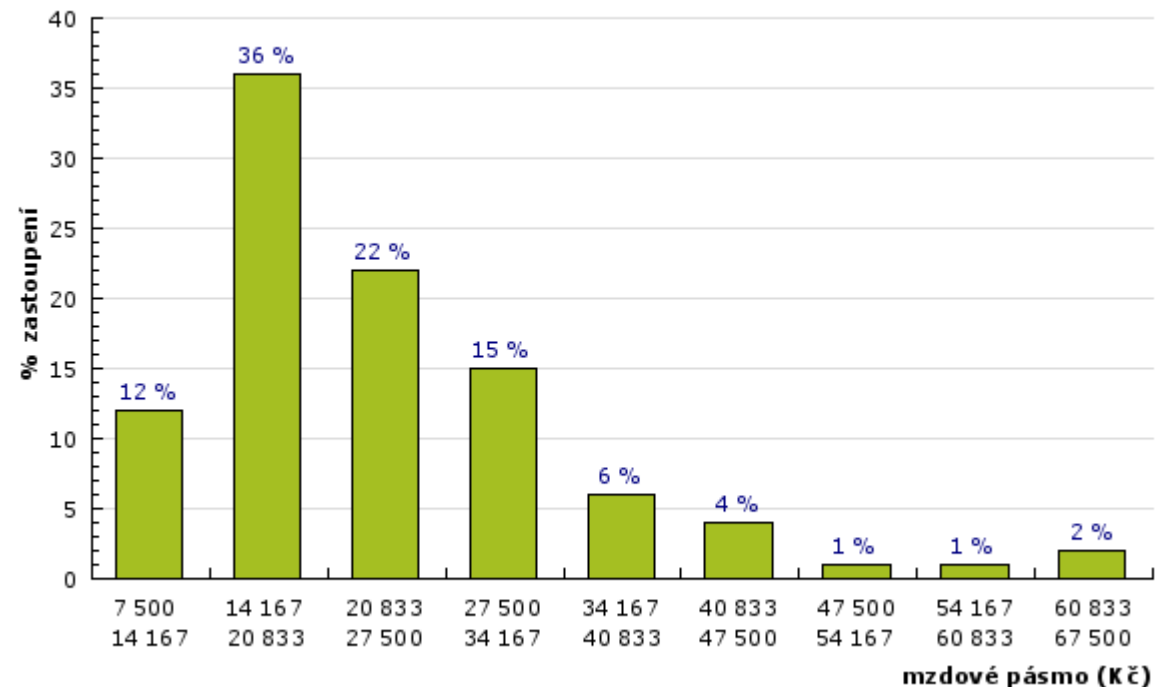
2. \_\_\_\_\_mzda činila 21 385 Kč.

a) Průměrná      b) Mediánová

3. Přibližně \_\_\_\_\_ zaměstnanců  
má podprůměrnou mzdu.

a) 33 %      b) 66 %

<http://www.czso.cz/csu/csu.nsf/informace/cpmz090514.docx>



<http://www.platy.cz/partner/region/ceska-republika>

# 1.1 Charakteristiky polohy

1. Ve 2. čtvrtletí 2014 činila \_\_\_\_\_  
mzda 25 500 Kč.

a) průměrná      b) mediánová

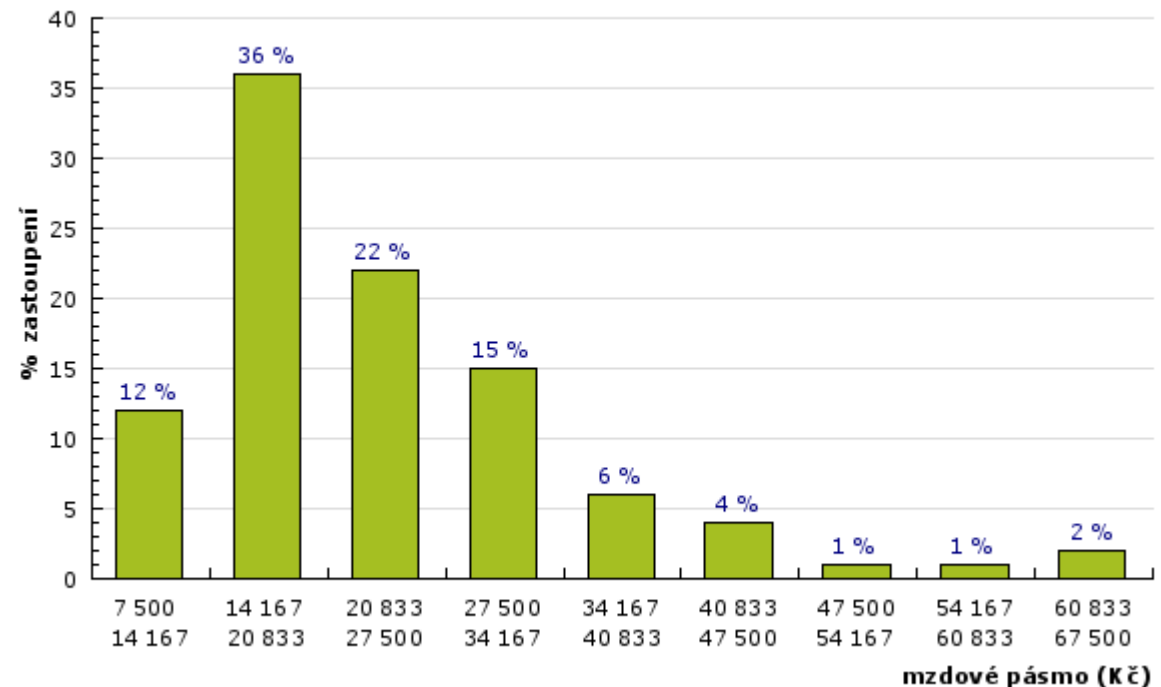
2. \_\_\_\_\_mzda činila 21 385 Kč.

a) Průměrná      b) Mediánová

3. Přibližně \_\_\_\_\_ zaměstnanců  
má podprůměrnou mzdu.

a) 33 %      b) 66 %

<http://www.czso.cz/csu/csu.nsf/informace/cpmz090514.docx>



<http://www.platy.cz/partner/region/ceska-republika>

# 1.1 Charakteristiky polohy

---

Následující data představují hodinovou mzdu deseti studentů.

60	90	90	90	100	110	110	120	130	600
----	----	----	----	-----	-----	-----	-----	-----	-----

Určete:

1. průměr
2. medián
3. modus

# 1.1 Charakteristiky polohy

---

Následující data představují hodinovou mzdu deseti studentů.

60	90	90	90	100	110	110	120	130	600
----	----	----	----	-----	-----	-----	-----	-----	-----

Určete:

1. průměr      150 Kč
2. medián      105 Kč
3. modus      90 Kč

Jak na to: <http://jakplavejak.cz/node/29>

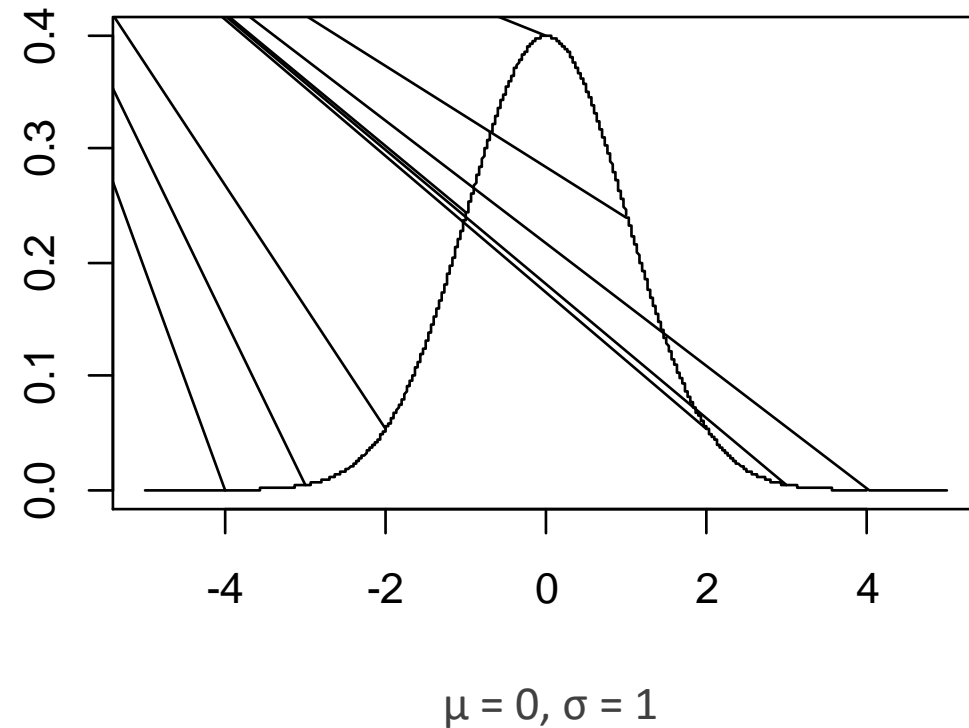


# 1.2 Normální rozdělení

---

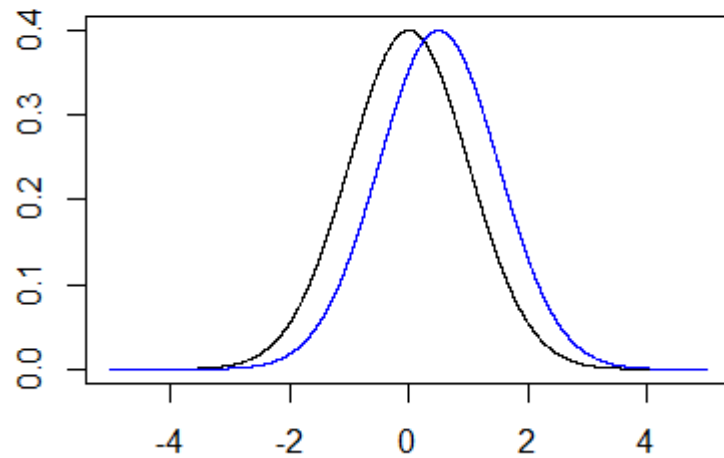
Co je to hustota pravděpodobnosti?  
Co je to distribuční funkce?

Co z toho je na obrázku?



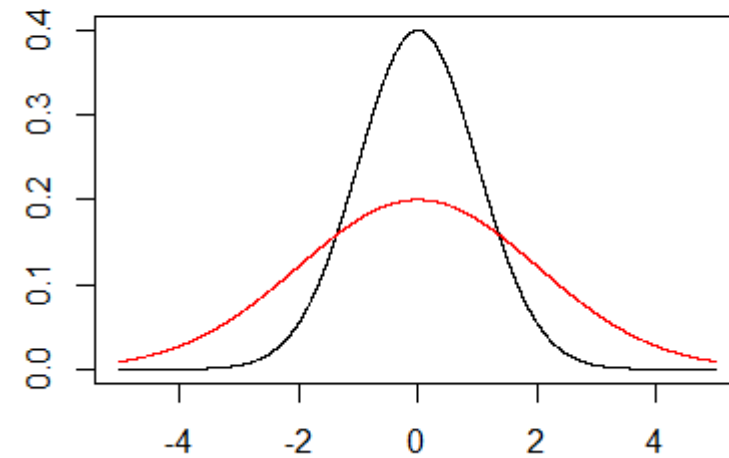
# 1.2 Normální rozdělení

Která veličina má vyšší střední hodnotu?



$\mu = 0, \sigma = 1$        $\mu = 0,5, \sigma = 1$

Která veličina má větší rozptyl?



$\mu = 0, \sigma = 1$        $\mu = 0, \sigma = 2$

# 1.3 Normální rozdělení

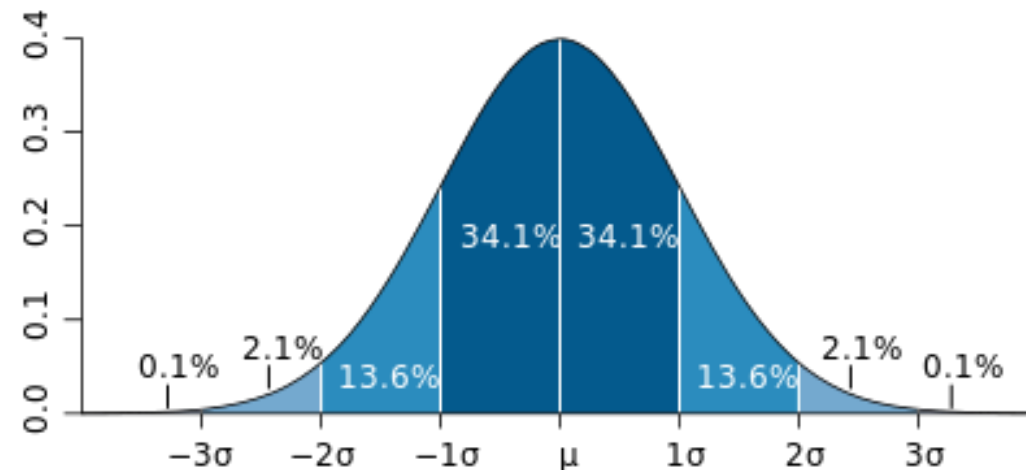
---

Co je to rozptyl?

Co je to směrodatná odchylka?

Jak byste tyto pojmy vysvětlili někomu, kdo o statistice nic neví?

Uvažujme náhodnou veličinu se střední hodnotou 5 a rozptylem 0. Co můžeme říct o této náhodné veličině?



[http://commons.wikimedia.org/wiki/File:Standard\\_deviation\\_diagram.svg](http://commons.wikimedia.org/wiki/File:Standard_deviation_diagram.svg)

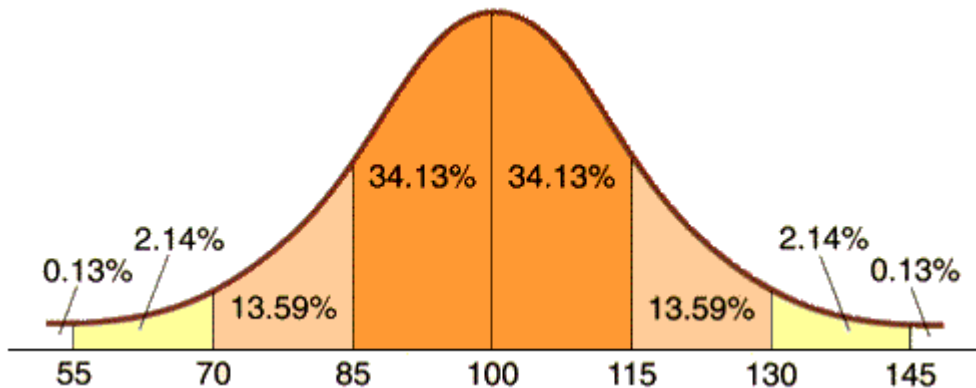
# 1.3 Normální rozdělení

---

Příklad:

Jaká je střední hodnota IQ v populaci?

Jaká je směrodatná odchylka?



[http://www.aceintelligence.com/iq\\_test/IQ\\_test\\_graph.gif](http://www.aceintelligence.com/iq_test/IQ_test_graph.gif)

# 1.4 Další důležitá rozdělení

---

- a) Chí-kvadrát rozdělení
- b) Studentovo t-rozdělení
- c) Fisherovo rozdělení

Jak tato rozdělení vypadají?

Platí:

Pokud mají nezávislé, stejně rozdělené veličiny  $X_1, X_2, \dots, X_n$  rozdělení  $N(0; 1)$ , má součet jejich druhých mocnin **chí-kvadrát** rozdělení s  $n$  stupni volnosti.

Jsou-li veličiny  $U \sim N(0; 1)$ ,  $Z \sim \chi^2(n)$  nezávislé, má veličina  $\frac{U}{\sqrt{Z}} \sqrt{n}$  **t-rozdělení** s  $n$  stupni volnosti.

Má-li veličina  $K_1$  chí-kvadrát rozdělení s  $n_1$  stupni volnosti a veličina  $K_2$  chí-kvadrát rozdělení s  $n_2$  stupni volnosti, pak má veličina  $\frac{K_1/n_1}{K_2/n_2}$  **Fisherovo rozdělení** s  $n_1, n_2$  stupni volnosti.

# 1.4 Počítání se stř. hodnotami a rozptyly

---

Jsou dány **nezávislé** náhodné veličiny  $X$  a  $Y$ , kde:

$$E_x = 10, \text{ var}_x = 1$$

$$E_y = 5, \text{ var}_y = 2$$

Spočítejte:

a)  $E[4x]$

b)  $E[x + 5]$

c)  $E[x + y]$

d)  $E[4x + 3y]$

e)  $\text{var}[4x]$

f)  $\text{var}[x + 5]$

g)  $\text{var}[x + y]$

h)  $\text{var}[4x + 3y]$

[http://nb.vse.cz/~zouharj/zek/repet\\_resene.pdf](http://nb.vse.cz/~zouharj/zek/repet_resene.pdf)

# 1.4 Počítání se stř. hodnotami a rozptyly

---

Jsou dány **nezávislé** náhodné veličiny X a Y, kde:

$$E_x = 10, \text{var}_x = 1$$

$$E_y = 5, \text{var}_y = 2$$

Spočítejte:

$$\text{a) } E[4x] = 40$$

$$\text{e) } \text{var}[4x] = 16$$

$$\text{b) } E[x + 5] = 15$$

$$\text{f) } \text{var}[x + 5] = 1$$

$$\text{c) } E[x + y] = 15$$

$$\text{g) } \text{var}[x + y] = 3$$

$$\text{d) } E[4x + 3y] = 55$$

$$\text{h) } \text{var}[4x + 3y] = 34$$

[http://nb.vse.cz/~zouharj/zek/repet\\_resene.pdf](http://nb.vse.cz/~zouharj/zek/repet_resene.pdf)

# 1.5 Náhodný výběr

---

Náhodné veličiny  $x_1, x_2, \dots, x_{10}$  jsou nezávislé a pochází z rozdělení  $N(\mu, \sigma^2)$ .

Výběrový průměr: 
$$\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i$$

Výběrový rozptyl: 
$$s^2 = \frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{x})^2$$

Jaké rozdělení má:

a)  $\bar{x}$

[http://nb.vse.cz/~zouharj/zek/repet\\_resene.pdf](http://nb.vse.cz/~zouharj/zek/repet_resene.pdf)



# 1.5 Náhodný výběr

---

Náhodné veličiny  $x_1, x_2, \dots, x_{10}$  jsou nezávislé a pochází z rozdělení  $N(\mu, \sigma^2)$ .

Výběrový průměr: 
$$\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i$$

Výběrový rozptyl: 
$$s^2 = \frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{x})^2$$

Jaké rozdělení má:

a)  $\bar{x} \sim N(\mu; \frac{\sigma^2}{10})$  protože...

$$E(\bar{x}) = \frac{1}{10} \sum_{i=1}^{10} E(x_i) = \frac{1}{10} \sum_{i=1}^{10} \mu = \mu$$

$$\text{var}(\bar{x}) = \left(\frac{1}{10}\right)^2 \cdot \sum_{i=1}^{10} \text{var}(x_i) = \left(\frac{1}{10}\right)^2 \cdot \sum_{i=1}^{10} \sigma^2 = \frac{\sigma^2}{10}$$

Co by se stalo se střední hodnotou  $\bar{x}$ , kdybychom měli nikoli 10, ale 1 000 náhodných veličin? A kdyby jich bylo nekonečno?

[http://nb.vse.cz/~zouharj/zek/repet\\_resene.pdf](http://nb.vse.cz/~zouharj/zek/repet_resene.pdf)

# 1.5 Náhodný výběr

---

Náhodné veličiny  $x_1, x_2, \dots, x_{10}$  jsou nezávislé a pochází z rozdělení  $N(\mu, \sigma^2)$ .

Výběrový průměr: 
$$\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i$$

Výběrový rozptyl: 
$$s^2 = \frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{x})^2$$

Jaké rozdělení má:

b) 
$$U = \frac{\bar{x} - \mu}{\sqrt{\frac{\sigma^2}{10}}}$$

[http://nb.vse.cz/~zouharj/zek/repet\\_resene.pdf](http://nb.vse.cz/~zouharj/zek/repet_resene.pdf)

# 1.5 Náhodný výběr

---

Náhodné veličiny  $x_1, x_2, \dots, x_{10}$  jsou nezávislé a pochází z rozdělení  $N(\mu, \sigma^2)$ .

Výběrový průměr: 
$$\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i$$

Výběrový rozptyl: 
$$s^2 = \frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{x})^2$$

Jaké rozdělení má:

b) 
$$U = \frac{\bar{x} - \mu}{\sqrt{\frac{\sigma^2}{10}}} \sim N(0; 1)$$

Jak se nazývá toto rozdělení?

Obecně, jak transformujeme libovolné normální rozdělení na toto rozdělení?

[http://nb.vse.cz/~zouharj/zek/repet\\_resene.pdf](http://nb.vse.cz/~zouharj/zek/repet_resene.pdf)

# 1.5 Testování hypotéz

Ve firmě je 100 zaměstnanců a my si myslíme, že jejich průměrná výška je 175 cm. Je to ale opravdu tak?

Mohli bychom všechny změřit a spočítat průměr. Nebo bychom mohli změřit jen některé, abychom sobě i jim ulehčili čas, a pomoci si svými poznatky ze statistiky.

Vybereme tedy **náhodně** 20 zaměstnanců (viz tabulka).

Předpokládáme, že výška zaměstnanců má normální rozdělení.

180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

---

Co je v tomto případě populací?

Co je v tomto případě výběrem?

Spočítejte výběrový průměr.

Spočítejte výběrový rozptyl a směrodatnou odchylku.

180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

---

Co je v tomto případě populací?

Co je v tomto případě výběrem?

Spočítejte výběrový průměr.

Spočítejte výběrový rozptyl a směrodatnou odchylku.

$$\bar{x} = 173,75$$

$$s^2 = 52,51$$

$$s = 7,2$$

180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

---

1. Zapište nulovou a alternativní hypotézu.
2. Spočítejte testovou statistiku.
3. Najděte kritickou hodnotu.
4. Učiňte nějaký závěr.

180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

1. Zapište nulovou a alternativní hypotézu.

2. Spočítejte testovou statistiku.

3. Najděte kritickou hodnotu.

4. Učiňte nějaký závěr.

$H_0: \mu = 175 \text{ cm}$

$H_1: \mu \neq 175 \text{ cm}$

Je to jednostranný nebo oboustranný test?

180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170



# 1.5 Testování hypotéz

1. Zapište nulovou a alternativní hypotézu.

2. Spočítejte testovou statistiku.

3. Najděte kritickou hodnotu.

3. Učiňte nějaký závěr.

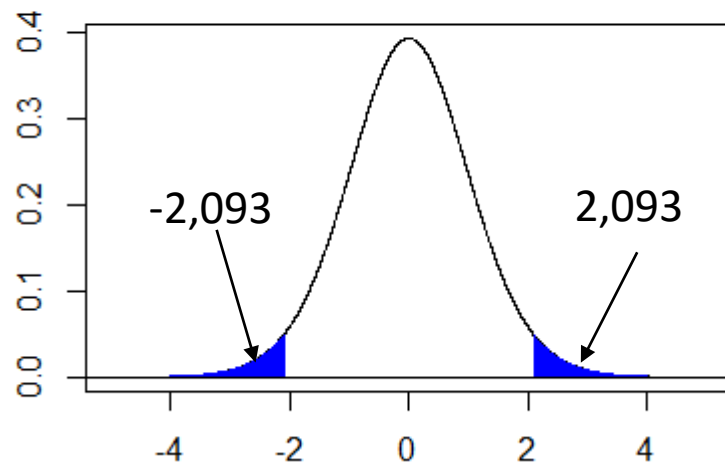
$$t = \frac{173,75 - 175}{\frac{7,2}{\sqrt{20}}} = -0,78$$

Jaké má tato statistika rozdělení?

180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

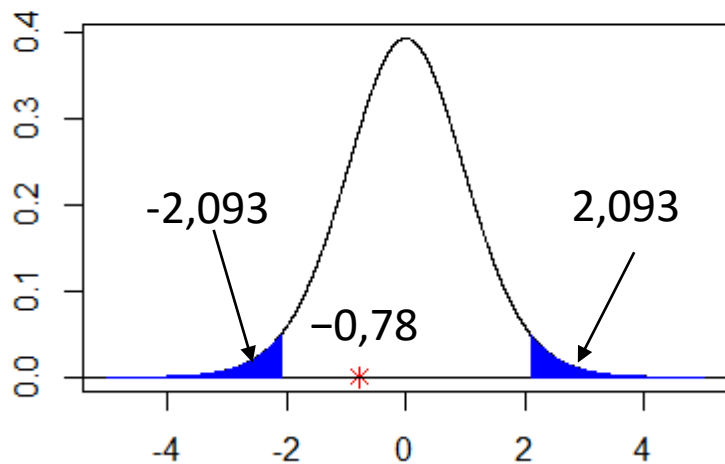
1. Zapište nulovou a alternativní hypotézu.
2. Spočítejte testovou statistiku.
3. Najděte kritickou hodnotu.
4. Učiňte nějaký závěr.



180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

1. Zapište nulovou a alternativní hypotézu.
2. Spočítejte testovou statistiku.
3. Najděte kritickou hodnotu.
4. Učiňte nějaký závěr.



180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

1. Zapište nulovou a alternativní hypotézu.
2. Spočítejte testovou statistiku.
3. Najděte kritickou hodnotu.
4. Učiňte nějaký závěr.

Nezamítli jsme tedy nulovou hypotézu, a to na pětiprocentní hladině významnosti.

Co to znamená?

180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

Nyní se podívejme, jak vypadá skutečnost (většinou ji ale neznáme!)

Průměrná výška je 175 cm.

Směrodatná odchylka je 10,5 cm.

Učinili jsme při testu hypotézy správný závěr?

180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

Uvažujme následující situaci. Místo prvních dvou řádků jsme náhodně vybrali tyto zaměstnance.

$$\text{Testová statistika je } t = \frac{165-175}{\frac{9,6}{\sqrt{20}}} = -4,2$$

Může taková situace nastat?

Zamítneme v tomto případě nulovou hypotézu?

Pokud ano, jaké chyby se přitom dopustíme?

180	180	175	165	172	172	179	167	182	185
170	175	175	169	183	167	172	183	166	158
181	176	180	156	164	186	193	179	161	150
183	179	162	186	175	183	187	176	158	189
190	163	165	177	192	174	185	196	183	161
190	179	190	173	168	175	183	167	177	184
198	164	165	167	183	188	178	174	187	156
173	184	177	174	179	171	176	167	180	191
188	187	180	169	177	177	190	159	172	168
154	175	160	164	163	167	176	151	158	170

# 1.5 Testování hypotéz

---

Chyba I. druhu: zamítneme nulovou hypotézu, přestože platí

Chyba II. druhu: nezamítneme nulovou hypotézu, přestože platí alternativní hypotéza

p-hodnota: nejnižší hladina významnosti, při které můžeme zamítnout nulovou hypotézu. V předchozím případě je (při oboustranném testu) p-hodnota rovna 0,000485. [Co to znamená?](#)

**Obecně: je-li p-hodnota  $\leq 0,05$ , pak na 5% hladině významnosti zamítáme nulovou hypotézu.**

# 1.6 Kovariance a korelace

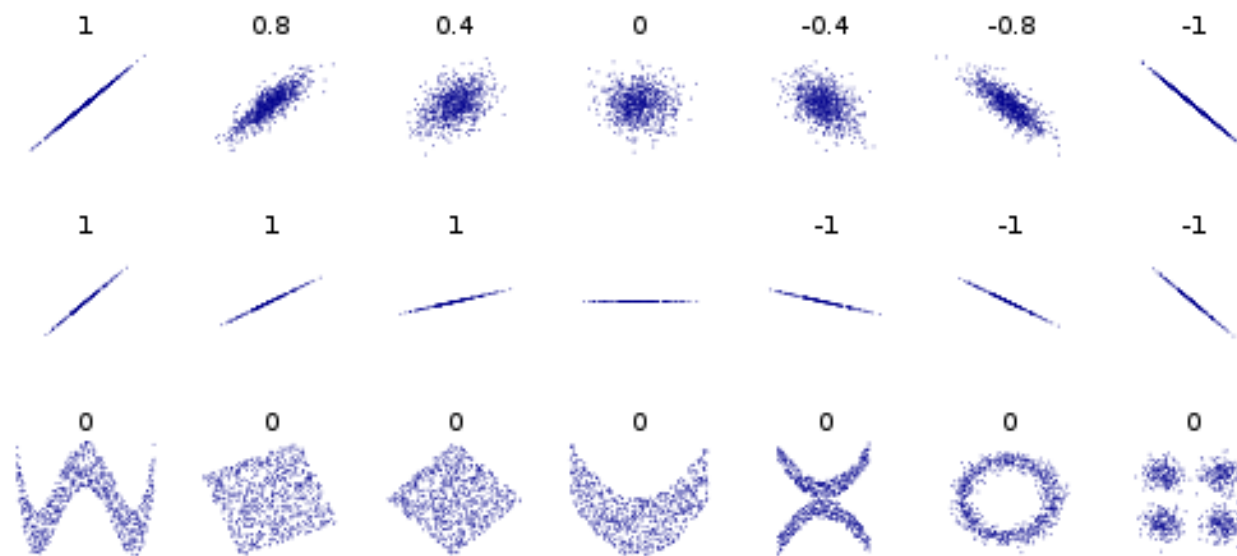
---

1. Vymyslete nějaký příklad pozitivně korelovaných, negativně korelovaných a nekorelovaných veličin.
2. Jaký je vztah mezi kovariancí a korelací?
3. Jakých hodnot může nabývat kovariance a jakých korelační koeficient?
4. Pokud je korelační koeficient roven nule, jsou veličiny určitě nezávislé?
5. Pokud jsou veličiny korelované, znamená to, že jedna je příčinou druhé?



# 1.6 Kovariance a korelace

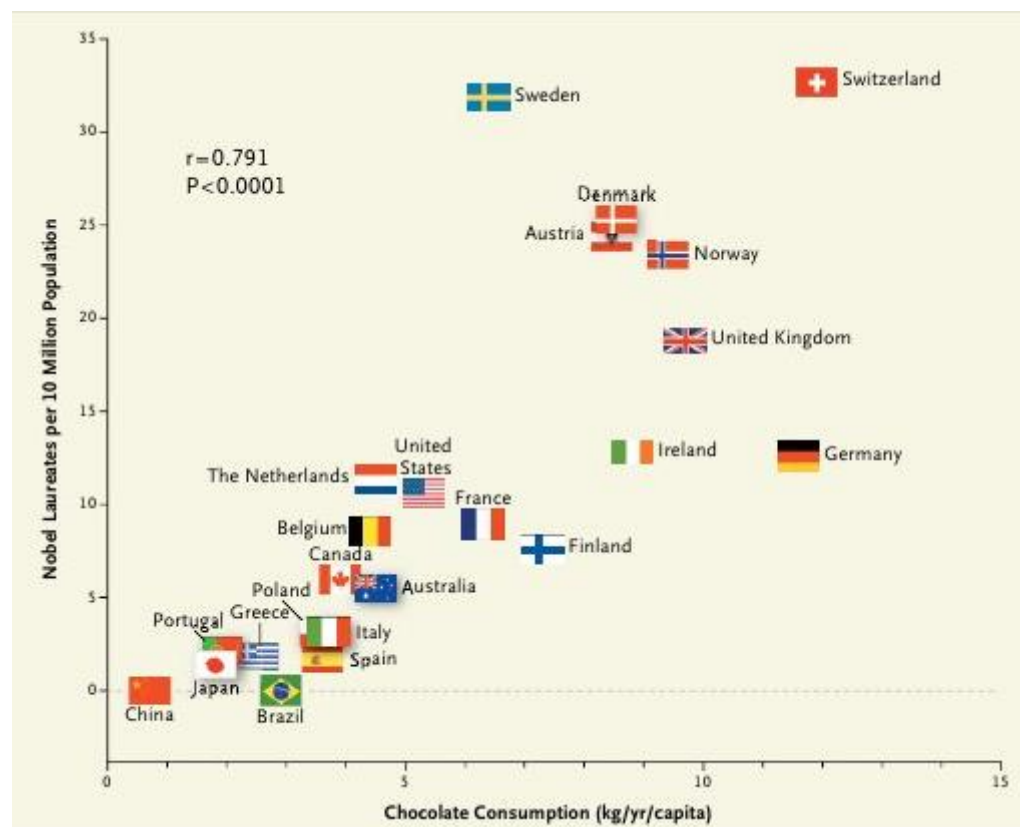
korelační koeficient



[https://en.wikipedia.org/wiki/Correlation\\_and\\_dependence#mediaviewer/File:Correlation\\_examples2.svg](https://en.wikipedia.org/wiki/Correlation_and_dependence#mediaviewer/File:Correlation_examples2.svg)

# 1.6 Kovariance a korelace

korelace a kauzalita: roční spotřeba čokolády versus počet nositelů Nobelovy ceny na 10 milion obyvatel



<http://i.stack.imgur.com/TWQaB.jpg>

# Otázky?

---

Další opakování ze statistiky např. zde:

[http://nb.vse.cz/~zouharj/zek/repet\\_resene.pdf](http://nb.vse.cz/~zouharj/zek/repet_resene.pdf)