

# Induktivní statistika

---

- normální rozdělení
  - rozdělení výběrových průměrů
-

# Normální rozdělení

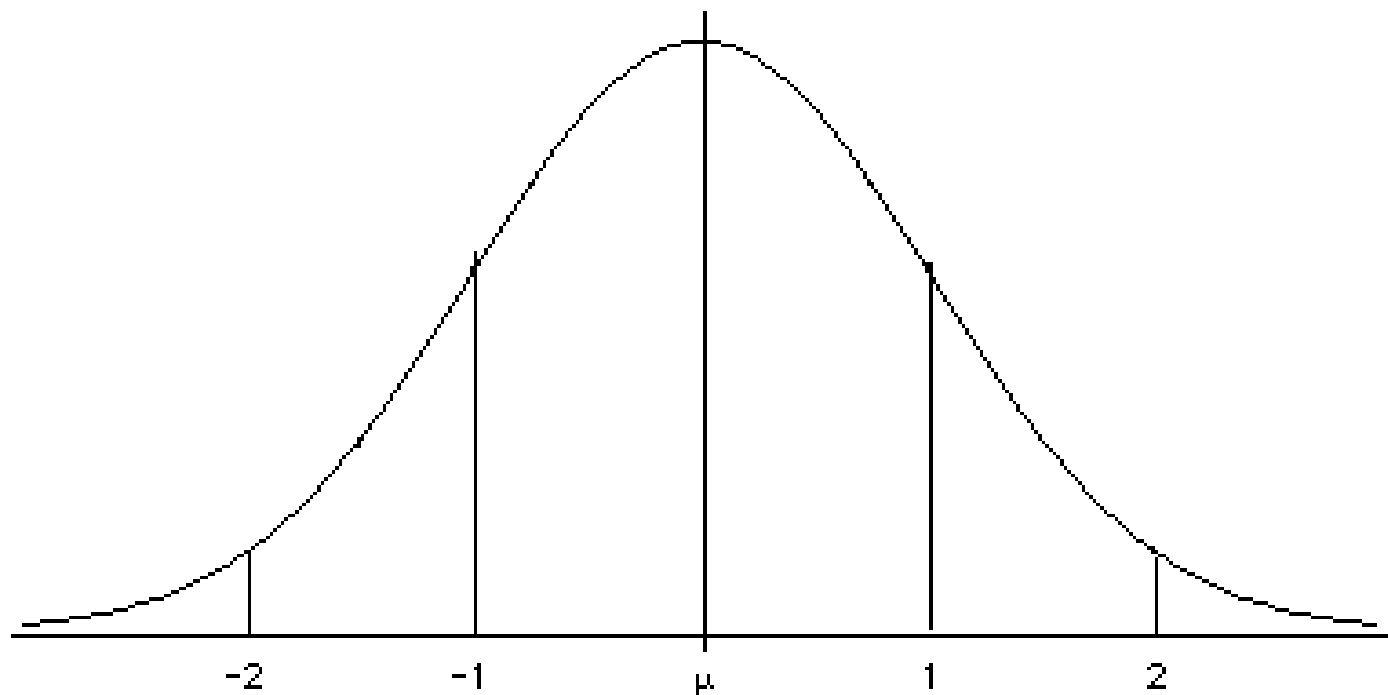
---

- normální rozdělení je symetrické, unimodální, zvonovitého tvaru
- označuje se i jako Gaussova křivka

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2}$$

# Normální rozdělení

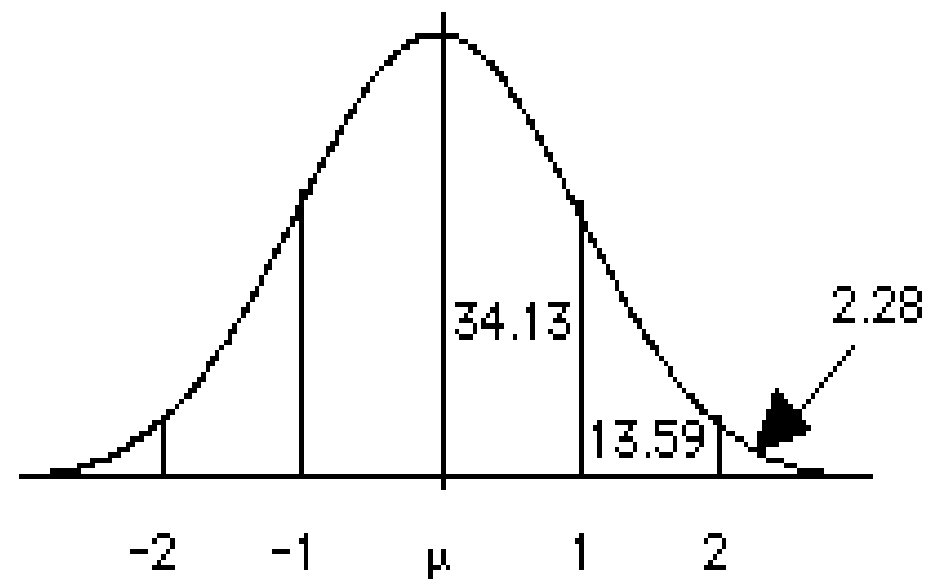
---



# Normální rozdělení

---

- 34.13% skóru spadá mezi průměr a 1 směr. odchylku
- 13.59% hodnot spadá mezi 1. a 2. směr. odchylku
- 2.28% hodnot spadá mezi 2. a 3. směr. odchylku



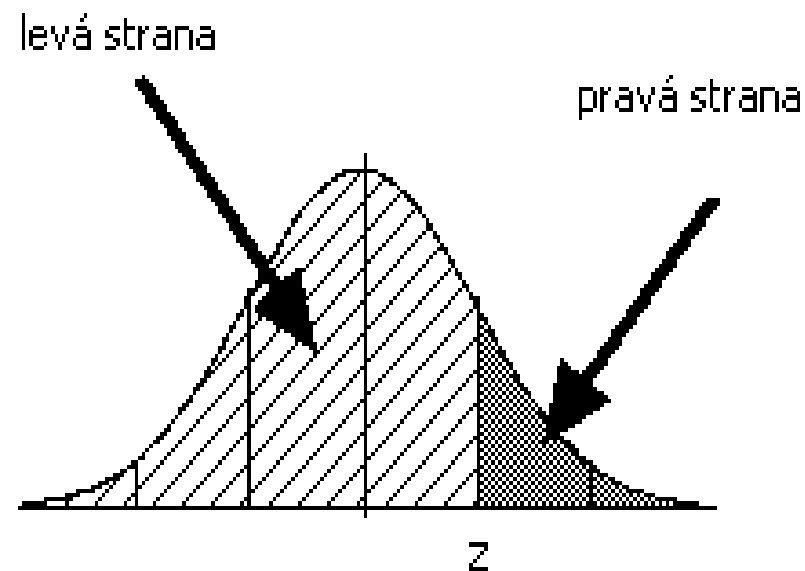
# Normální rozdělení

---

- tabulka normálního rozdělení (z rozdělení)
  - důležitý nástroj, obvykle jako apendix v učebnicích statistiky (spolu s dalšími tabulkami)
  - umožňuje zjistit hustotu oblasti pod křivkou (tj. pravděpodobnost) pro jednotlivé z-skóry
-

# Normální rozdělení

z	levá strana	pravá strana
0.00	0.5000	0.5000
0.01	0.5040	0.4960
...	...	...
0.30	0.6179	0.3821
0.31	0.6217	0.3783
...	...	...
1.00	0.8413	0.1587
...	...	...



# Normální rozdělení - příklady

---

- postup při zjišťování pravděpodobnosti z tabulky:
    - načrtnout si normální rozdělení, s hodnotou průměru a směr. odch.
    - zakreslit hledanou hodnotu (v přibližné vzdálenosti od průměru), vystínovat hledanou oblast
    - převést hodnotu  $X$  na  $z$ -skór
    - najít v tabulce pravděpodobnost
-

# Normální rozdělení - příklady

---

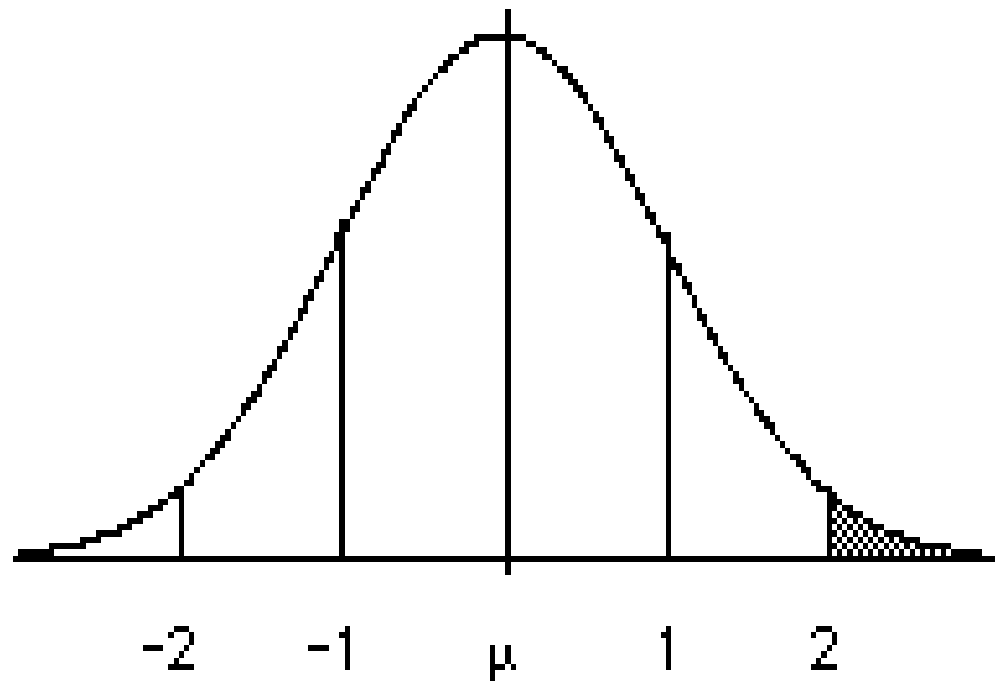
- Jaká je pravděpodobnost, že náhodně vybraná osoba z populace bude mít IQ 130 nebo vyšší? ( $\mu = 100$ ,  $\sigma = 15$ )
-



# Normální rozdělení - příklady

---

□  $z = 2$



# Normální rozdělení - příklady

---

- Jaká je pravděpodobnost, že náhodně vybraná osoba z populace bude mít IQ 130 nebo vyšší?
  - $z = 2$
  - $p = 0.0228$  tj. **2,3%**
-

# Normální rozdělení - příklady

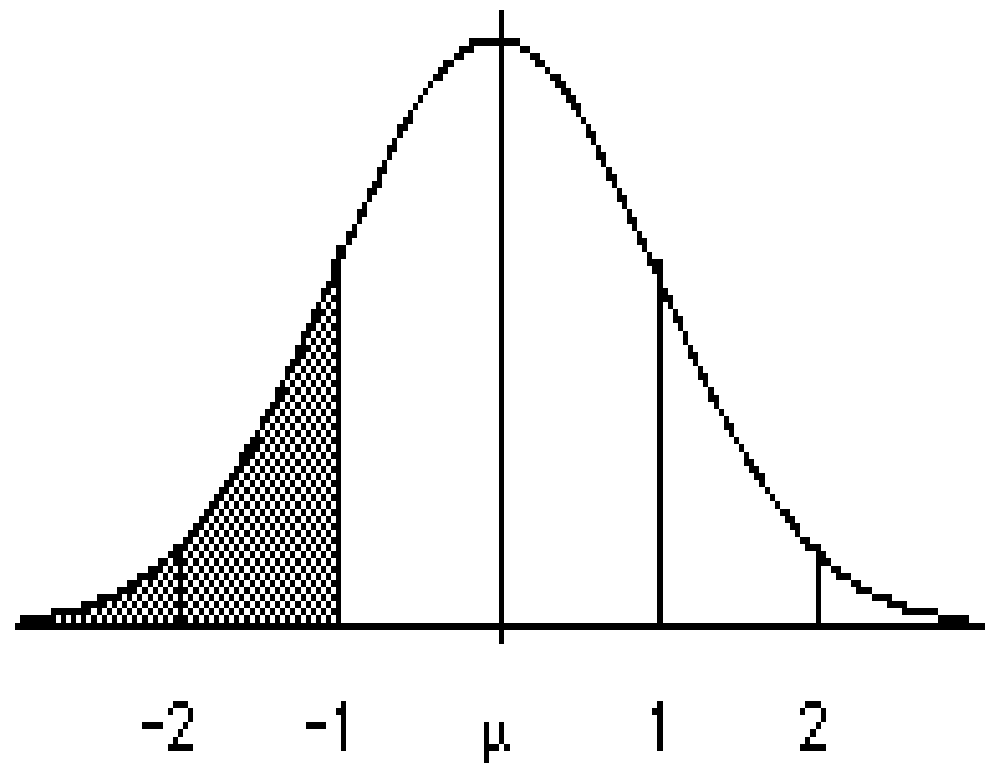
---

- Jaká je pravděpodobnost, že náhodně vybraná osoba z populace bude mít IQ 85 nebo nižší?
-

# Normální rozdělení - příklady

---

□  $z = -1$



# Normální rozdělení - příklady

---

- Jaká je pravděpodobnost, že náhodně vybraná osoba z populace bude mít IQ 85 nebo nižší?
  - $z = -1$
  - $p = 0.1587$  tj. **15,9%**
-

# Normální rozdělení - příklady

---

- postup při zjišťování z-skóru z tabulky:
    - načrtnout si normální rozdělení
    - vystínovat oblast odpovídající zadané pravděpodobnosti
    - v tabulce vyhledat příslušný z-skór
    - vypočítat z něj hrubý skór
-

# Normální rozdělení - příklady

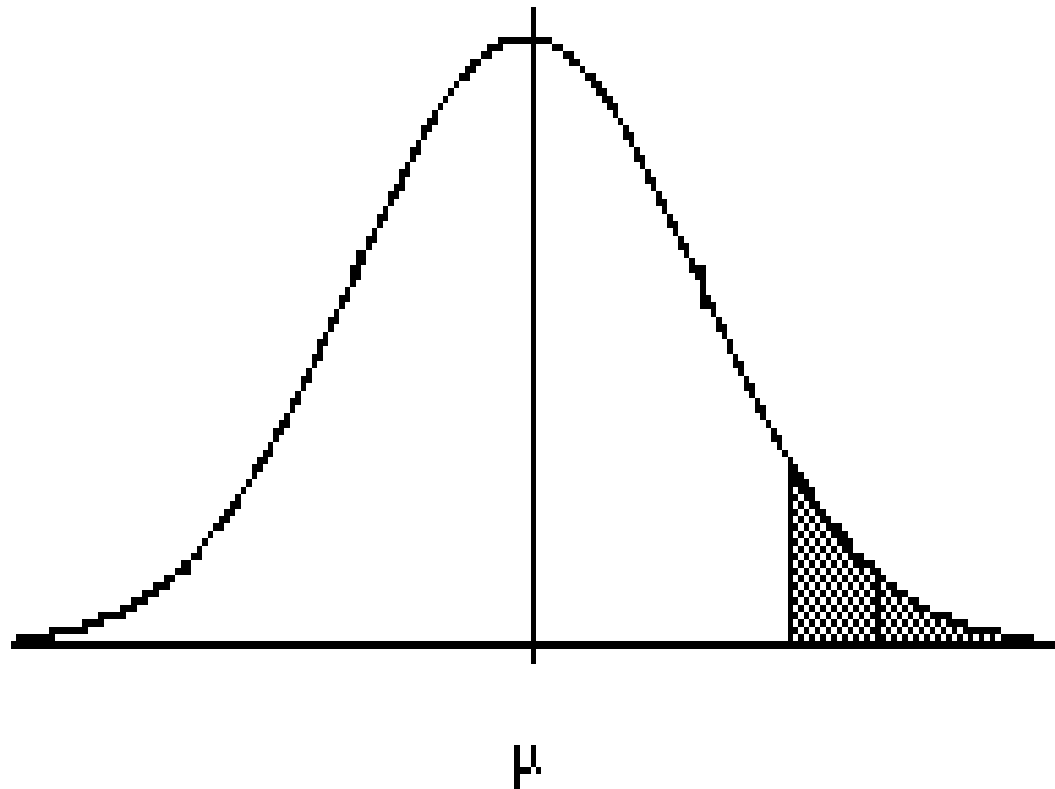
---

- Jakou minimální hodnotu IQ musí člověk mít, aby patřil mezi 5% osob s nejvyššími hodnotami IQ?
-

# Normální rozdělení - příklady

---

□  $p = 0.05$





# Normální rozdělení - příklady

---

- Jakou minimální hodnotu IQ musí člověk mít, aby patřil mezi 5% osob s nejvyššími hodnotami IQ?
  - $p = 0.05$
  - z tabulky:  **$z = 1.65$**
  - $X = (1.65) \cdot (15) + 100 = \mathbf{124.75}$
-

# Normální rozdělení - příklady

---

- někdy chceme zjistit pravděpodobnost, že skór bude spadat do určitého intervalu
  - postup:
    - načrtnout graf a vystínovat zadanou oblast
    - oba (ohraničující) skóry převést na z-skóry
    - vyhledat pravděpodobnosti  $<$  nebo  $>$  skóru
    - sečíst či odečíst pravděpodobnosti
-

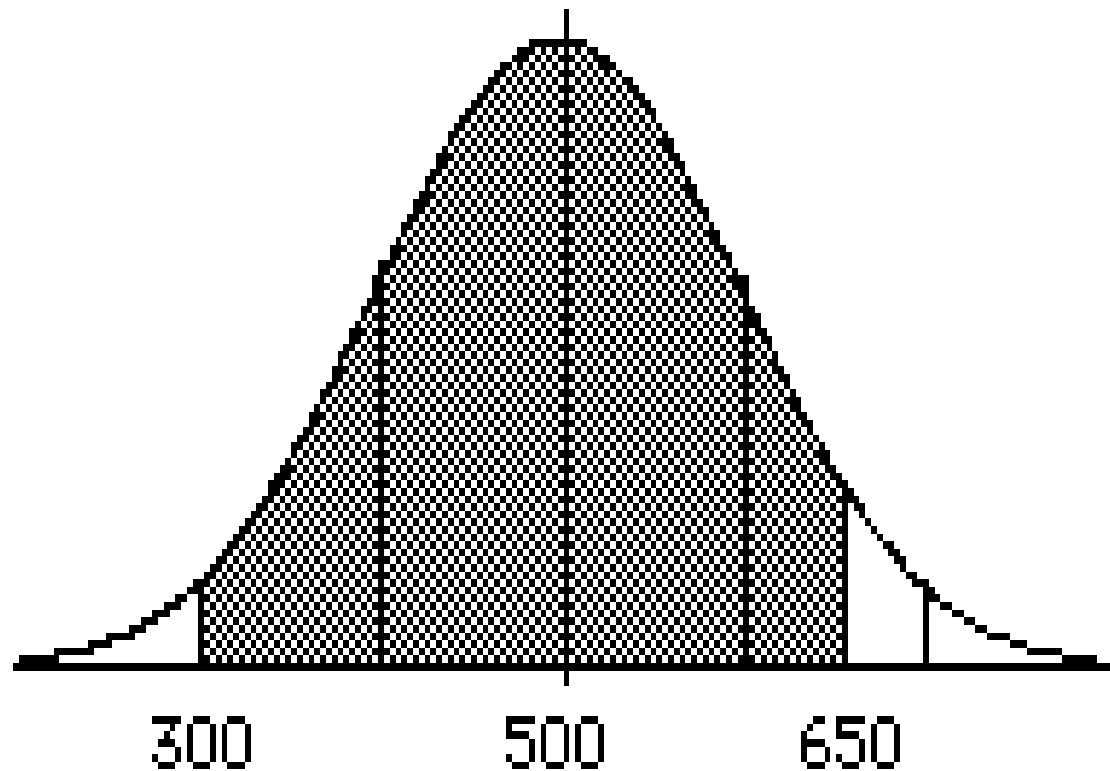
# Normální rozdělení - příklady

---

- Jaká je pravděpodobnost, že náhodně vybraný student bude v testu z psychologie skórovat mezi 300 a 650 body? ( $\mu = 500, \sigma = 100$ )
-

# Normální rozdělení - příklady

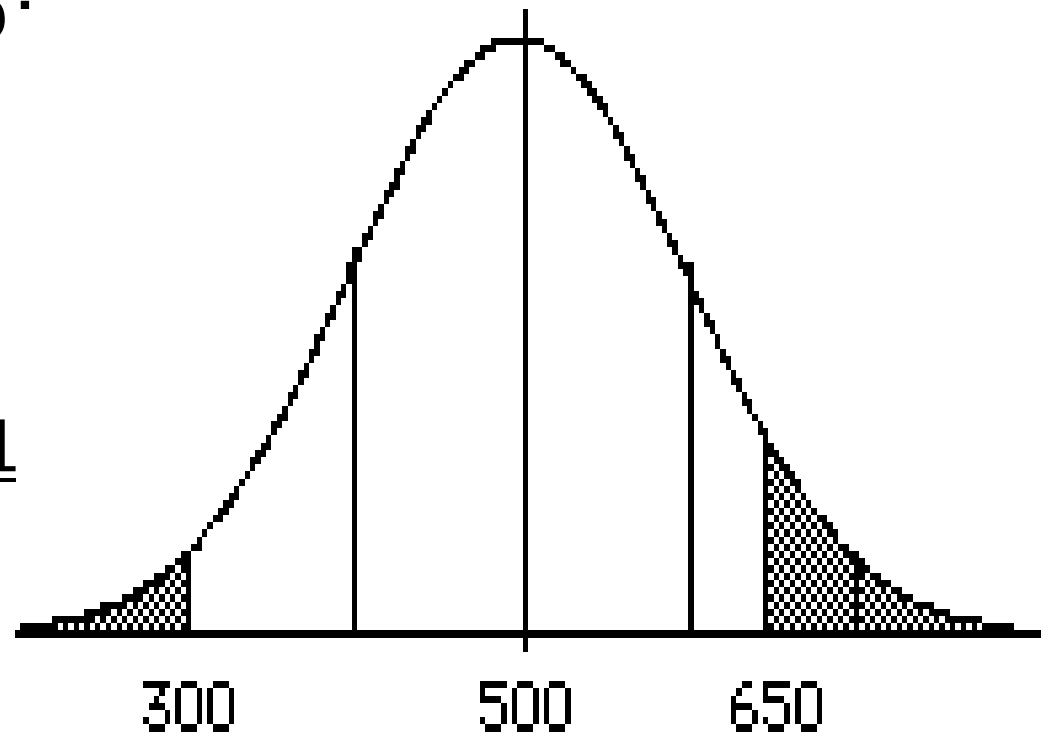
---



# Normální rozdělení - příklady

---

- alternativní postup  
spočítat  
pravděpodobnost  
skórování mimo  
zadaný interval a  
poté ji odečíst od 1



# Normální rozdělení - příklady

---

- Jaká je pravděpodobnost, že náhodně vybraný student bude v testu z psychologie skórovat mezi 300 a 650 body? ( $\mu = 500$ ,  $\sigma = 100$ )
  - $p(x \geq \frac{650 - 500}{100}) = p(z \geq 1.5) = 0.0668$
  - $p(x < \frac{300 - 500}{100}) = p(z \leq -2.0) = 0.0228$
  - $0.0668 + 0.0228 = .0896$
  - $p(300 < x < 650) = 1 - 0.0896 = \mathbf{0.9104}$
-

# Normální rozdělení - příklady

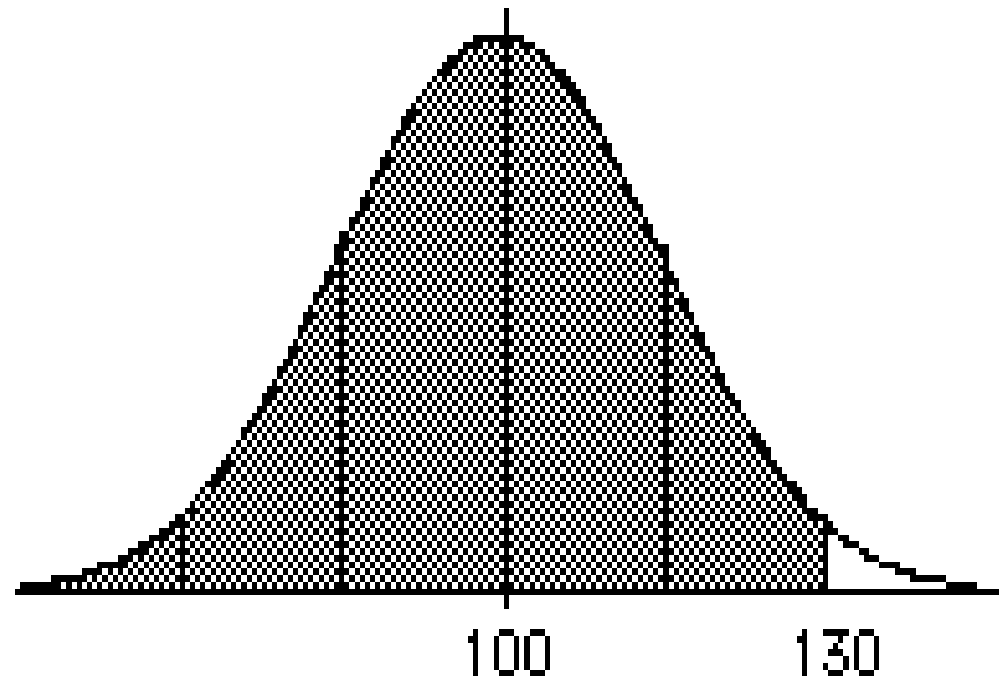
---

- pomocí tabulky normálního rozdělení je možno nalézt také hodnotu percentilu
  - příklad: kolik procent osob má nižší hodnoty IQ než člověk s IQ 130?
-

# Normální rozdělení - příklady

---

□  $z = 2$





# Normální rozdělení - příklady

---

□ Kolik procent osob má nižší hodnoty IQ než člověk s IQ 130?

□ z tabulky: pro  $z = 2$

$p = 0.9772$

**97.72%** osob má nižší skór

---

# Rozdělení výběrových průměrů

---

- cílem indukční statistiky je odhadnout parametry populace z charakteristik vzorku (výběrového souboru)
  - např. odhadem průměru populace bude průměr vzorku
  - odhad je vždy zatížen určitou **výběrovou chybou**
-

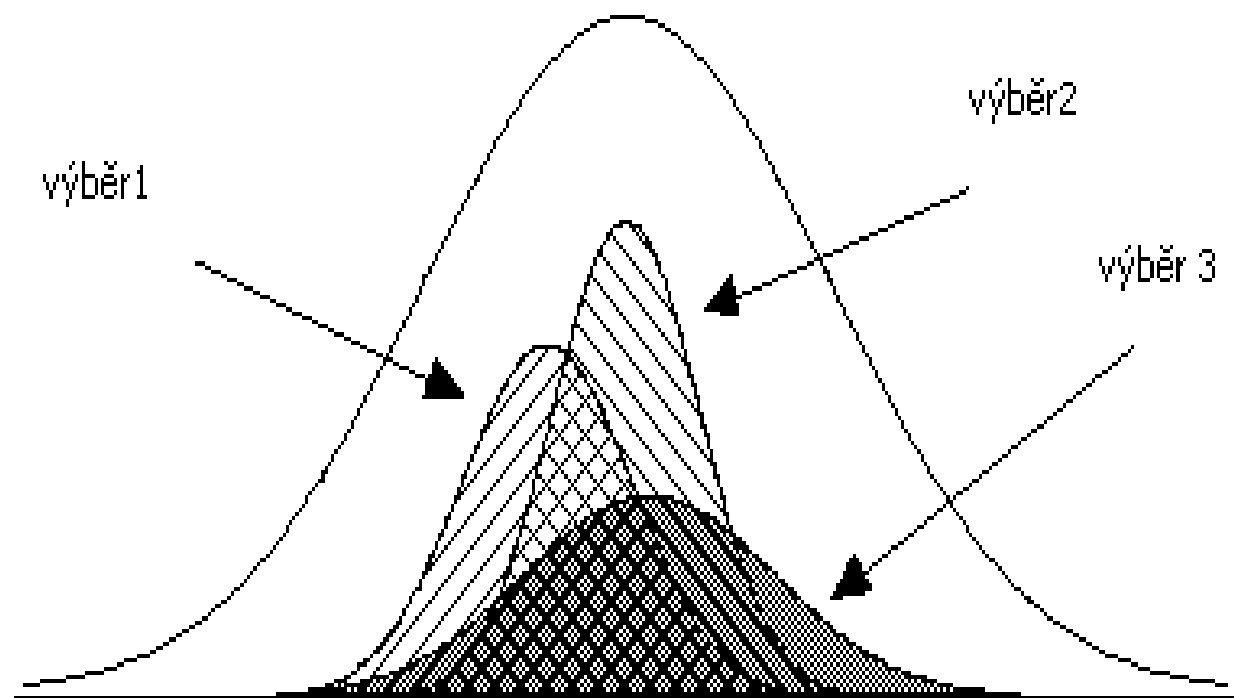
# Rozdělení výběrových průměrů

---

- předpokládejme, že z jedné populace vybereme 3 různé vzorky
  - budou se nejspíš navzájem lišit ve tvaru rozdělení hodnot, průměru i variabilitě
  - jak se rozhodneme, který z nich zvolit pro odhad průměru populace ??
-

# Rozdělení výběrových průměrů

---



# Rozdělení výběrových průměrů

---

- pokud bychom spočítali průměry ze všech možných výběrů o určité velikosti  $n$ , budou tvořit tzv. **rozdělení výběrových průměrů** (sampling distribution)
-

# Rozdělení výběrových průměrů

---

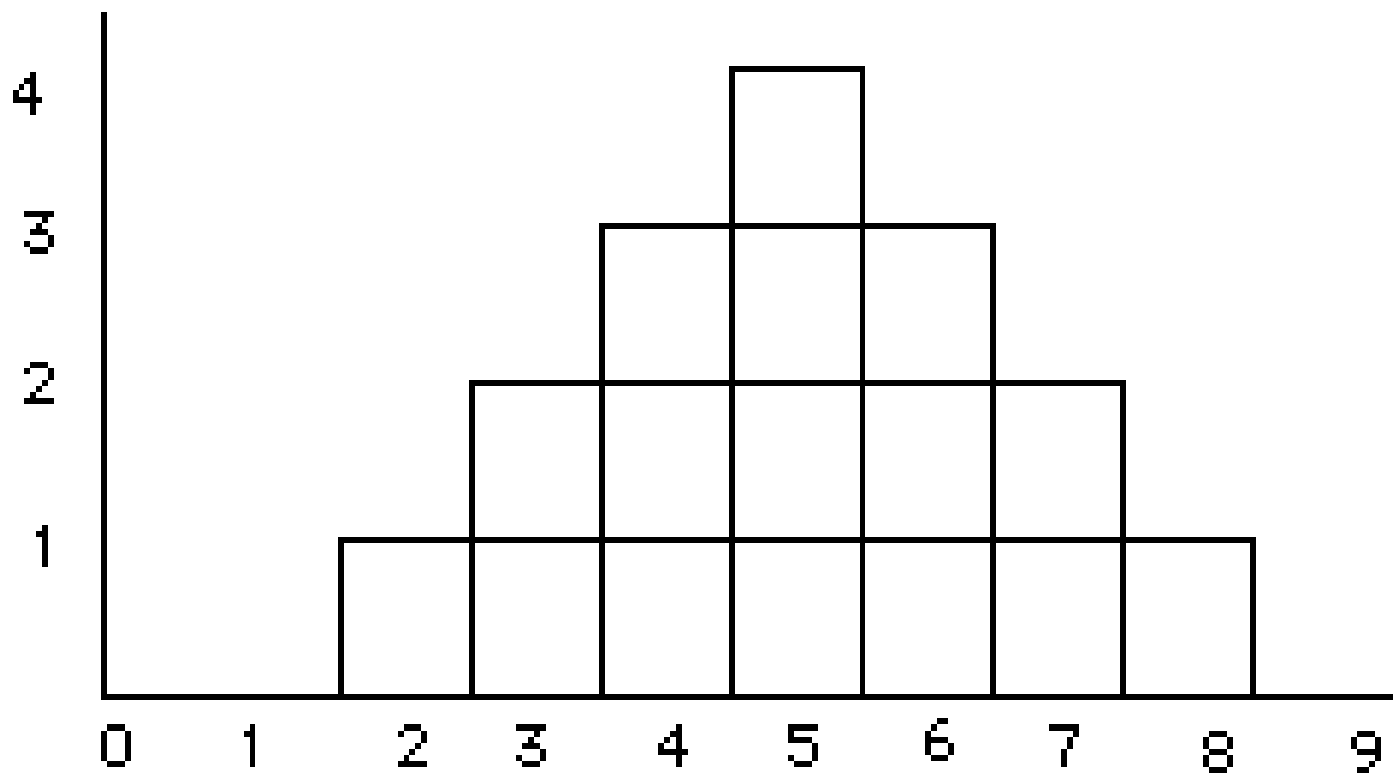
- **příklad:** populace hodnot 2, 4, 6, 8
  - průměr  $\mu = 5$
  - předpokládejme, že průměr neznáme a pokoušíme se ho odhadnout ze vzorku  $n=2$
  - v tabulce jsou uvedeny všechny možné výběrové soubory
-

# Rozdělení výběrových průměrů

<u>výběr</u>	<u>první skór</u>	<u>druhý skór</u>	<u>průměr vzorku</u>
1	2	2	2
2	2	4	3
3	2	6	4
4	2	8	5
5	4	2	3
6	4	4	4
7	4	6	5
8	4	8	6
9	6	2	4
10	6	4	5
11	6	6	6
12	6	8	7
13	8	2	5
14	8	4	6
15	8	6	7
16	8	8	8

# Rozdělení výběrových průměrů

---





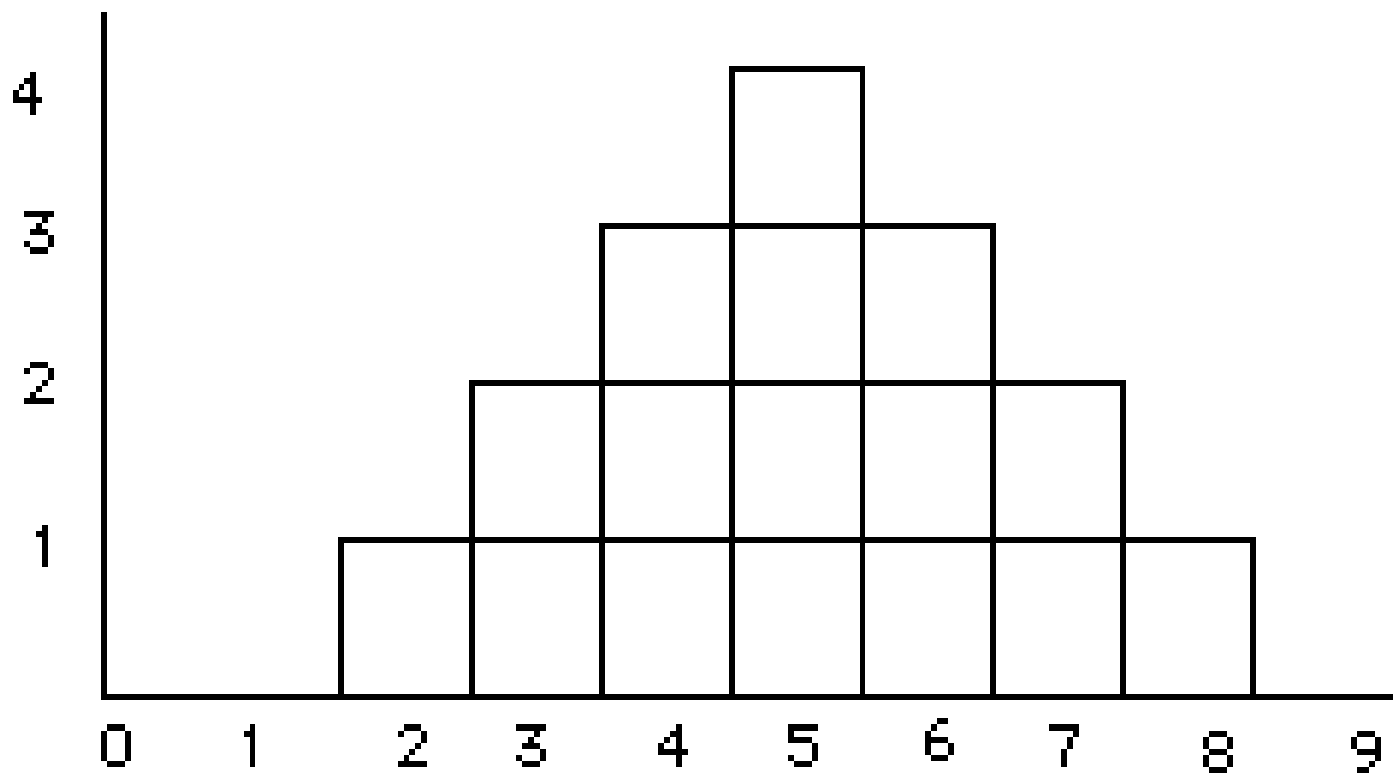
# Rozdělení výběrových průměrů

---

- jaká je pravděpodobnost, že z této populace vybereme vzorek s průměrem vyšším než 7?
-

# Rozdělení výběrových průměrů

---



# Rozdělení výběrových průměrů

---

- jaká je pravděpodobnost, že z této populace vybereme vzorek s průměrem vyšším než 7?
  - v rozdělení výběrových průměrů je takový vzorek jen 1 ze 16 – tj. pravděpodobnost takového průměru vzorku je  $1/16 = 0.0625$
-

# Rozdělení výběrových průměrů

---

- většina populací i vzorků je mnohem větší
  - ale existují určité základní vlastnosti rozdělení výběrových průměrů (RVP)
  - **tvar** – RVP se při dostatečně velkém vzorku (30 a více) blíží **normálnímu rozdělení**
-

# Rozdělení výběrových průměrů

---

- **průměr** – průměr průměrů všech teoretických výběrů je roven průměru populace
  - označuje se také jako očekávaná hodnota průměru vzorku
-

# Rozdělení výběrových průměrů

---

- **variabilita** – směrodatná odchylka RVP se označuje jako výběrová chyba (standard error) průměru
  - jde o směrodatnou odchylku výběrových průměrů od průměru populace
  - ukazuje, jak spolehlivý je odhad populačního průměru z průměru vzorku – tj. jak velkou chybou je odhad zatížen
-

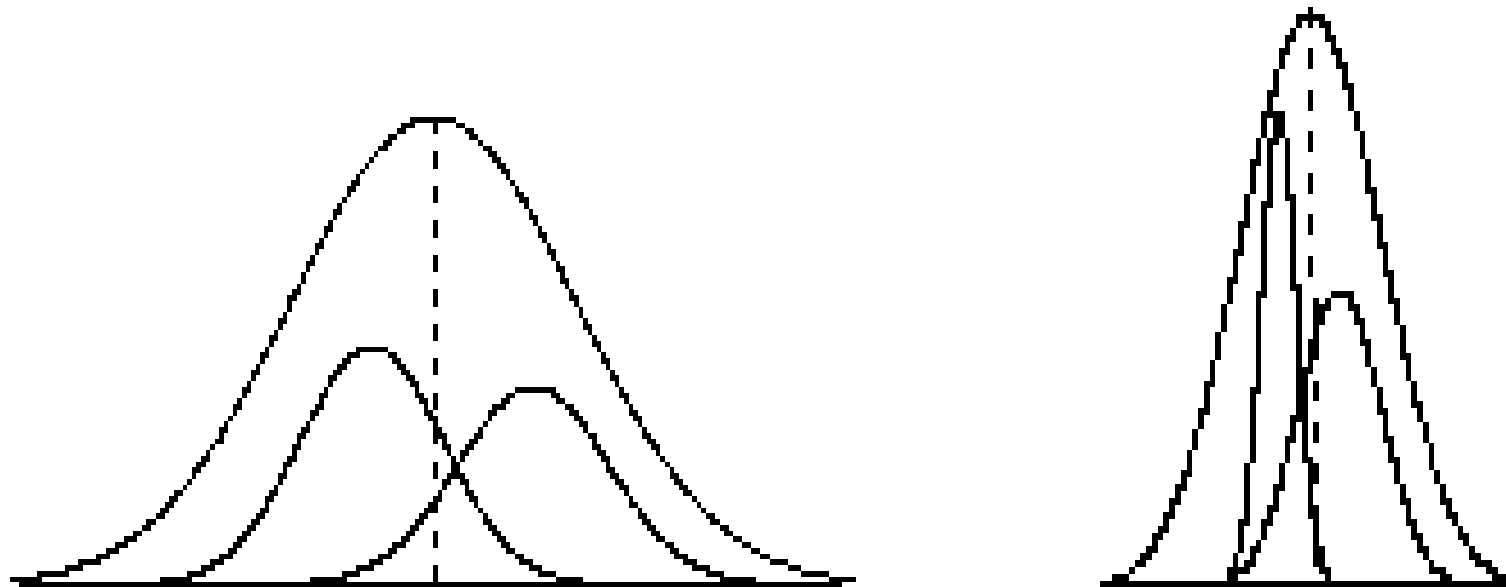
# Rozdělení výběrových průměrů

---

- velikost výběrové chyby je dána dvěma charakteristikami: variabilitou v populaci a velikostí výběru
  - **variabilita znaku v populaci:** čím je vyšší, tím je vyšší i variabilita výběrových průměrů
-

# Rozdělení výběrových průměrů

---



---



# Rozdělení výběrových průměrů

---

- velikost výběru – čím větší výběr ( $n$ ), tím lépe jeho průměr reprezentuje průměr populace (zákon velkých čísel)
-

# Rozdělení výběrových průměrů

---

□ vzorec pro výpočet výběrové chyby:

$$\sigma_x = \sigma / \sqrt{n}$$

# Rozdělení výběrových průměrů

---

- **centrální limitní věta** – pro každou populaci o průměru  $\mu$  a směrodatné odchylce  $\sigma$  se bude rozdělení výběrových průměrů výběrů blížit normálnímu rozdělení s průměrem  $\mu$  a směrodatnou odchylkou  $\sigma_x = \sigma/\sqrt{n}$
-

# Rozdělení výběrových průměrů

---

- **příklad:** když vybereme z populace náhodně vzorek 9 osob, jaká je pravděpodobnost, že jejich průměrné IQ bude větší nebo rovno 112?
-

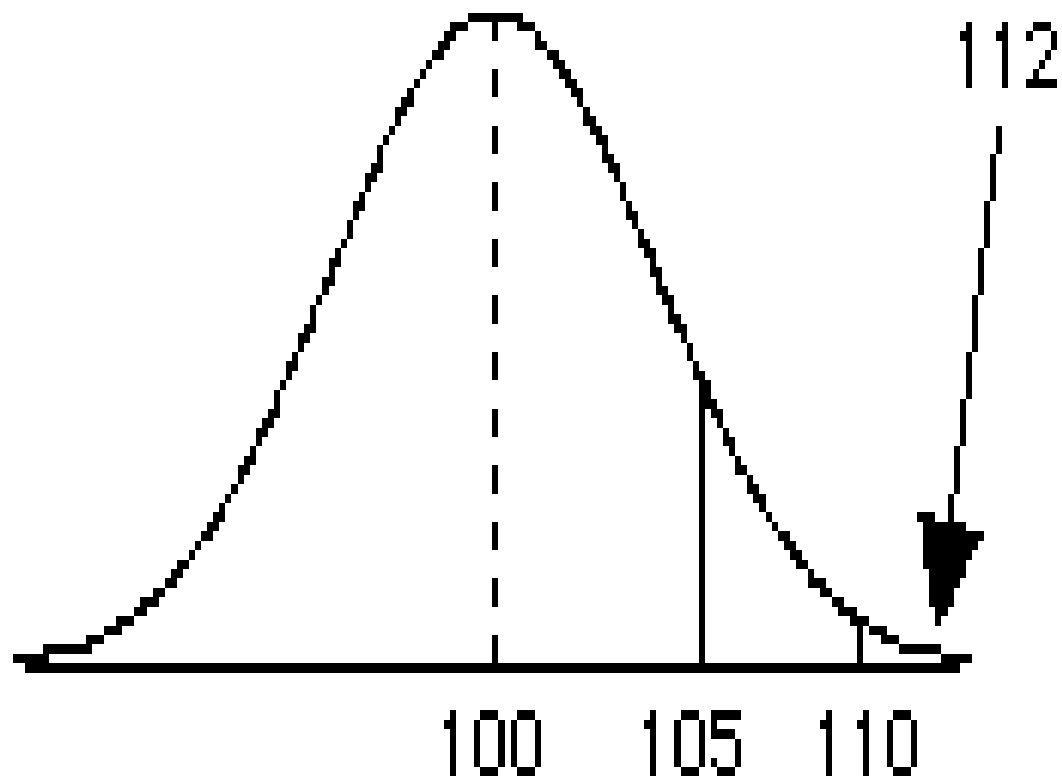
# Rozdělení výběrových průměrů

---

- **příklad:** když vybereme z populace náhodně vzorek 9 osob, jaká je pravděpodobnost, že jejich průměrné IQ bude větší nebo rovno 112?
  - $\mu = 100, \sigma_x = \sigma/\sqrt{n} = 15/3 = 5$
  - $z = (112-100)/\sigma_x = 12/5 = \mathbf{2.4}$
  - z tabulky  $P(Z \geq 2.4) = \mathbf{0.0082}$
-

# Rozdělení výběrových průměrů

---



# Příklad

---

- IQ ( $\mu=100$ ,  $\sigma=15$ )
  - jaká je **pravděpodobnost, že průměr výběru o velikosti  $n=25$  bude mezi hodnotami 94 a 106?**
-

# Příklad 1

---

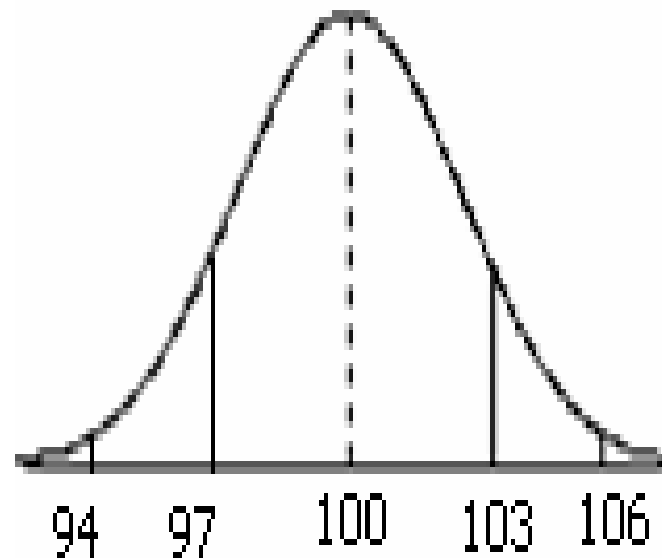
- IQ ( $\mu=100$ ,  $\sigma=15$ )
  - jaká je pravděpodobnost, že průměr výběru o velikosti  $n=25$  bude mezi hodnotami 94 a 106?
  - $z = (94-100) / (15/ \sqrt{25}) = -6/3 = -2$
  - $z = (106-100) / (15/ \sqrt{25}) = 6/3 = 2$
-



# Příklad 1

---

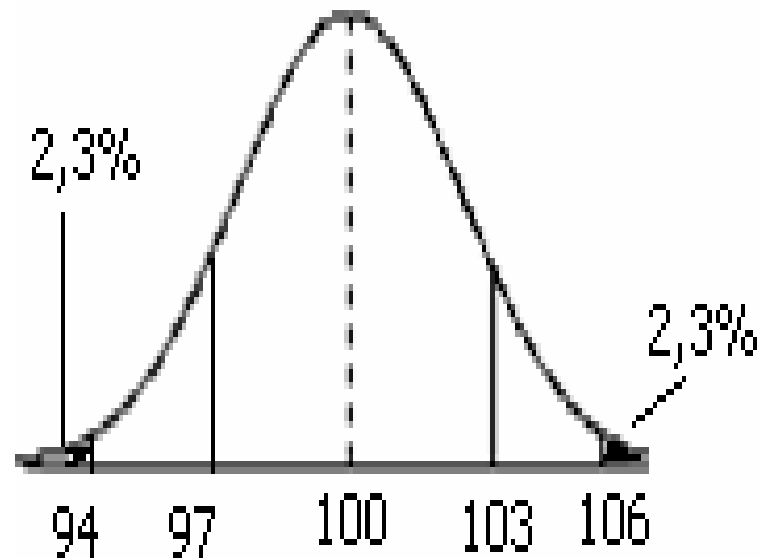
- najdeme v tabulce normovaného normálního rozdělení hodnotu pravděpodobnosti pro  $z=2$  a  $z=-2$



# Příklad 1

---

- hodnota je 0,023
- dohromady 0,046
- odečteme od 1,00
- výsledek **0,954**



# Příklad 1

---

- IQ ( $\mu=100$ ,  $\sigma=15$ )
  - jaká je pravděpodobnost, že průměr výběru o velikosti  $n=25$  bude mezi hodnotami 94 a 106?
  - **pravděpodobnost takového průměru je 95,4%**
-

# Příklad 2

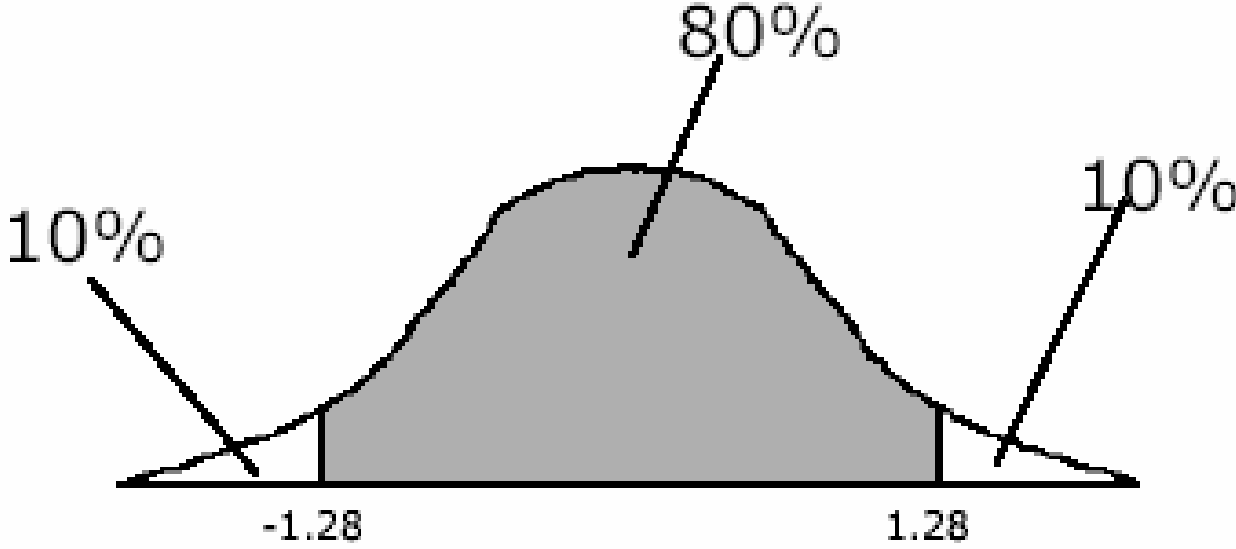
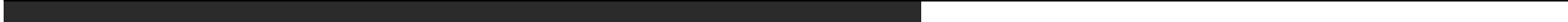
---

- IQ ( $\mu=100$ ,  $\sigma=15$ )
  - v jakém **rozsahu hodnot** bude pravděpodobně **80% všech průměrů výběrů** o velikosti  $n=25$ ?
-

# Příklad 2

---

- IQ ( $\mu=100$ ,  $\sigma=15$ )
  - v jakém **rozsahu hodnot** bude pravděpodobně **středních 80% všech průměrů výběrů** o velikosti  $n=25$ ?
  - potřebujeme zjistit hodnotu  $z$ , která odděluje pravděpodobnost 10% na obou stranách rozdělení
-



## Příklad 2

---

- IQ ( $\mu=100$ ,  $\sigma=15$ )
  - v jakém **rozsahu hodnot** bude pravděpodobně **středních 80% všech průměrů výběrů** o velikosti  $n=25$ ?
  - $z = 1,28$  a  $-1,28$  převedeme na hodnoty IQ
-

# Příklad 2

---

$$\bar{x} = \mu + z^*(\sigma/\sqrt{n})$$

$$\bar{x} = 100 + 1,28*(15/\sqrt{25}) = 100+1,28(3) = \mathbf{103,84}$$

$$\bar{x} = 100 + (-1,28)*(15/\sqrt{25}) = 100-1,28(3) = \mathbf{96,16}$$

---



## Příklad 2

---

- IQ ( $\mu=100$ ,  $\sigma=15$ )
  - v jakém rozsahu hodnot bude pravděpodobně středních 80% všech průměrů výběrů o velikosti  $n=25$ ?
  - **80% všech výběrových průměrů bude v rozsahu hodnot 96,16 - 103,84**
-