# IV. Náhodná veličina

# Náhodná veličina…

***Náhodná veličina* (NV)** = Číselný výsledek náhodného pokusu.

Výsledky - obecně různé vlivem náhodných činitelů  mají různé pravděpodobnosti realizace

***Náhodná veličina* (NV)** = odpovídá kvantitativnímu znaku populačního souboru (je jeho **zobecněním**)

# Rozdělení (pravděpodobnosti) náhodné veličiny

**RNV** = Pravidlo (předpis), které každé číselné hodnotě nebo množině hodnot přiřazuje pravděpodobnost, že náhodná veličina nabude této hodnoty nebo hodnoty z tohoto intervalu

# Rozdělení náhodné veličiny

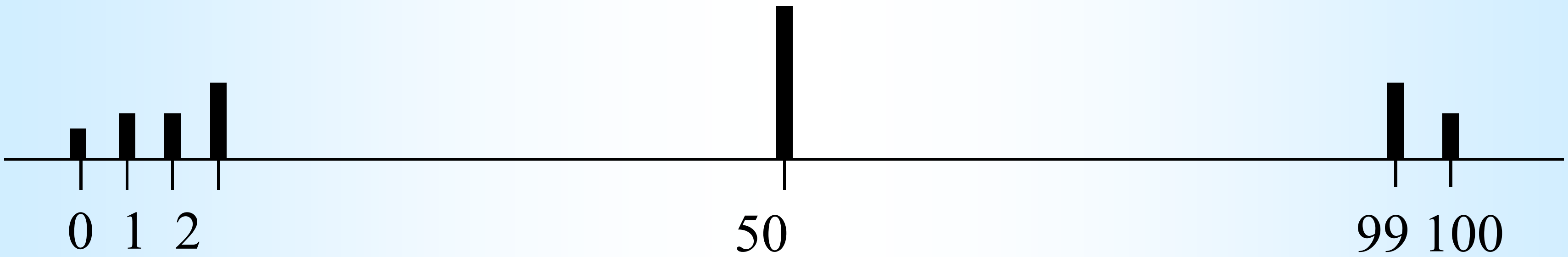
***Rozdělení pravděpodobnosti*** náhodné veličiny = úplné poznání NV:

- stanovení hodnot, jichž může NV nabývat
- znalost pravděpodobností, s nimiž NV nabývá určité hodnoty, nebo hodnoty z nějakého intervalu

# Náhodná veličina – Příklad 1

## NV diskrétního typu:

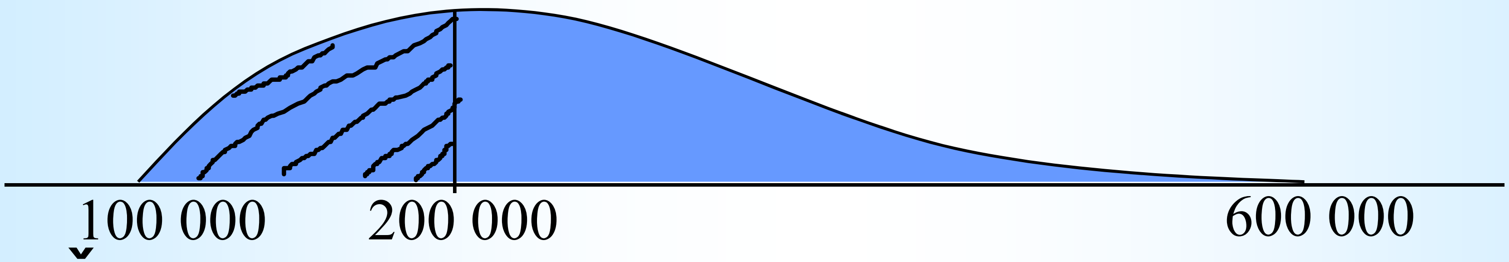
**1.** Jistý hotel má 100 pokojů, počet obsazených pokojů je náhodná veličina $X$ s možnými hodnotami $x = 0,1,2,\ldots,100$



# Náhodná veličina – Příklady

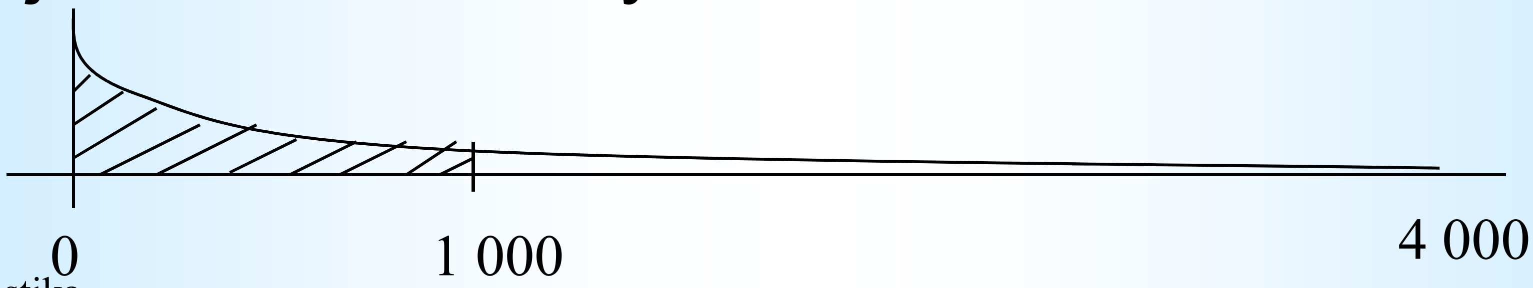
**2**. Roční mzda zaměstnanců podniku

- NV nabývá hodnoty 100 000 až 6000 000 Kč

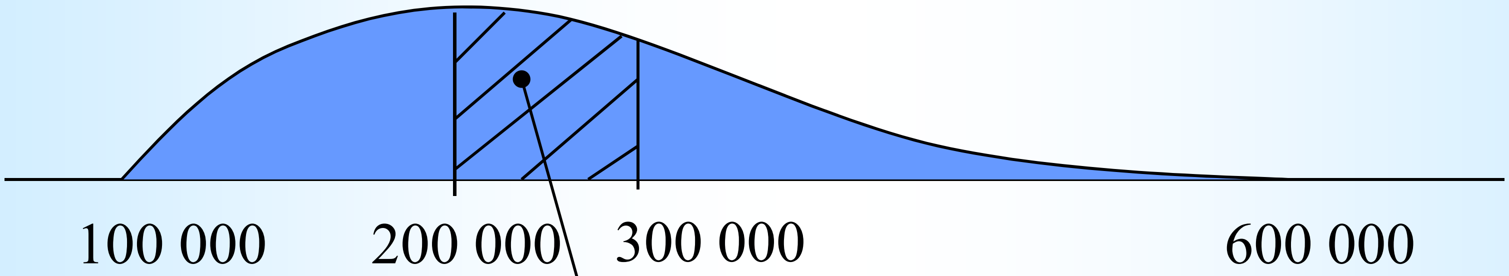



**3.** Životnost výrobku (žárovky) je náhodná veličina $X$, která může (teoreticky) nabýt jakékoliv hodnoty $x \geq 0$

P(X ≤ 200 tis.)



# Náhodná veličina – Příklad 2




PRST že roční mzda zaměstnance leží mezi 200 až 300 tis. Kč

# Způsoby vyjádření RNV

1. **Pravděpodobnostní funkce**
   (typ diskrétní NV)
2. **Hustota pravděpodobnosti**
   (typ spojité NV)
3. **Distribuční funkce**
   (oba typy NV: diskrétní / spojitý)

# Pravděpodobnostní funkce

*Pravděpodobnostní funkce* **(PF) -**
$p(x)$ – každé hodnotě $x \in \boldsymbol{D}$ přiřazuje odpovídající pravděpodobnost:

$$p(x) = P(X = x)$$

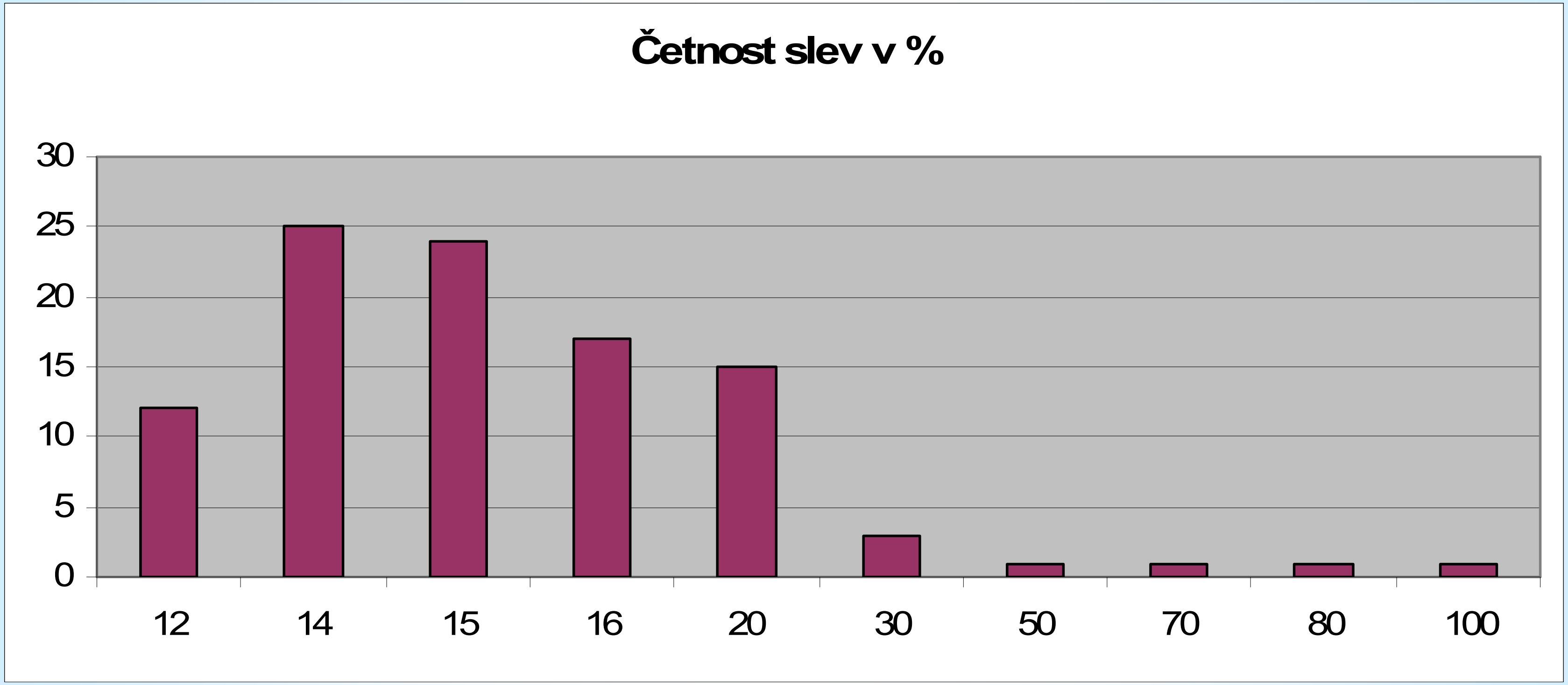
PF $p(x)$ splňuje vztahy:

$$\sum_{x \in D} p(x) = 1 \qquad P(a \le X \le b) = \sum_{x=a}^{b} p(x)$$

- pravděpodobnost, že náhodná veličina nabude hodnoty z intervalu $[a,b] \subset \boldsymbol{D}$, je rovna součtu pravděpodobností hodnot z tohoto intervalu
  ($\boldsymbol{D}$ je množina diskrétních hodnot)

# Příklad další náhodné veličiny:
# „Kolo štěstí“

# Příklad další náhodné veličiny: „Kolo štěstí“



# Kolo štěstí – náhodná veličina

| $x_i$ - Sleva % | $n_i$ - Četnost | $p_i$ - Pr-st |
|---|---|---|
| 12 | 12 | 0,12 |
| 14 | 25 | 0,25 |
| 15 | 24 | 0,24 |
| 16 | 17 | 0,17 |
| 20 | 15 | 0,15 |
| 30 | 3 | 0,03 |
| 50 | 1 | 0,01 |
| 70 | 1 | 0,01 |
| 80 | 1 | 0,01 |
| 100 | 1 | 0,01 |
| Suma | 100 | 1,00 |

# Hustota pravděpodobnosti

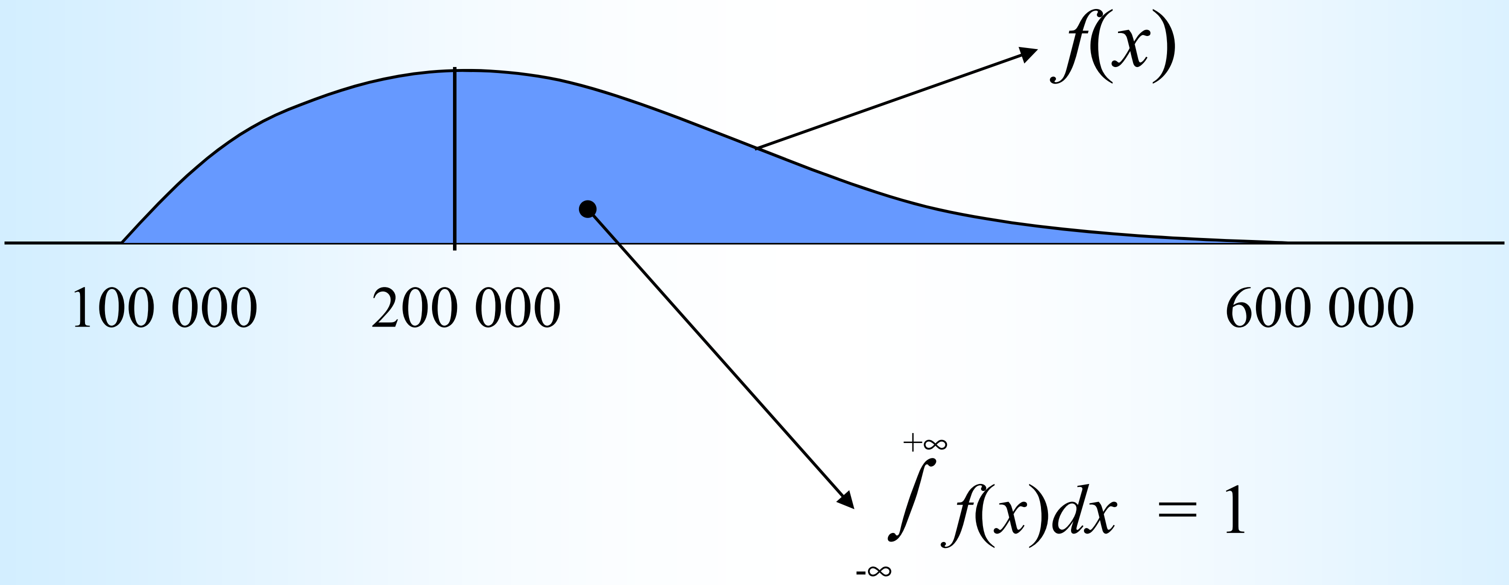
- **HP** - $f(x)$ - nezáporná funkce splňující podmínku:

$$\int_{-\infty}^{+\infty} f(t)dt = 1$$

- **Celá plocha pod grafem funkce** $f(x)$ – nad osou $x$ je rovna 1

# **Příklad funkce hustoty:**

Průměrná měsíční mzda zaměstnanců: NV nabývá hodnoty 100 až 600 tis. Kč



# Distribuční funkce

- **DF** - $F(x)$ – je definovaná na R vztahem:

$$F(x) = P(X \leq x)$$

- $F(x)$ – je neklesající funkce splňující:

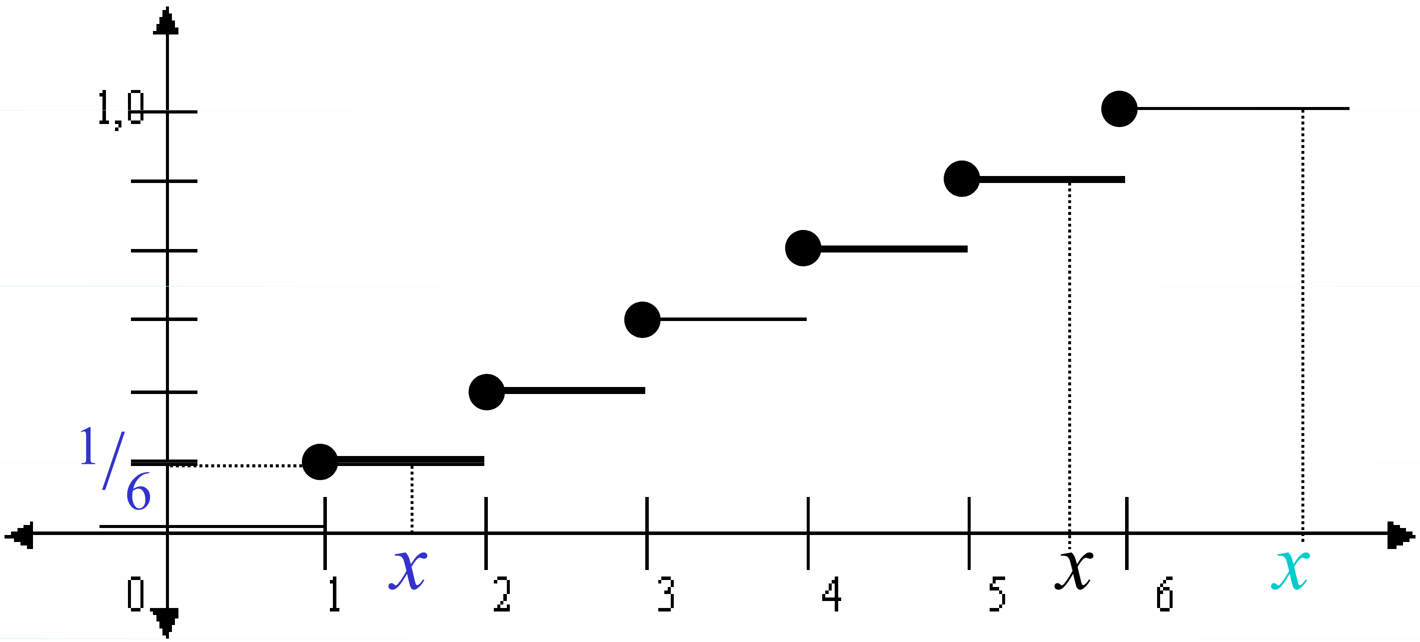
  $lim\ F(x)$ = 0 pro $x \rightarrow -\infty$

  $lim\ F(x)$ = 1 pro $x \rightarrow +\infty$

# Příklad distribuční funkce pro diskrétní NV: Hrací kostka

$F(x)=P(X\leq x)$ – Pr-st,

že NV $X$ nabude hodnoty $\leq x$

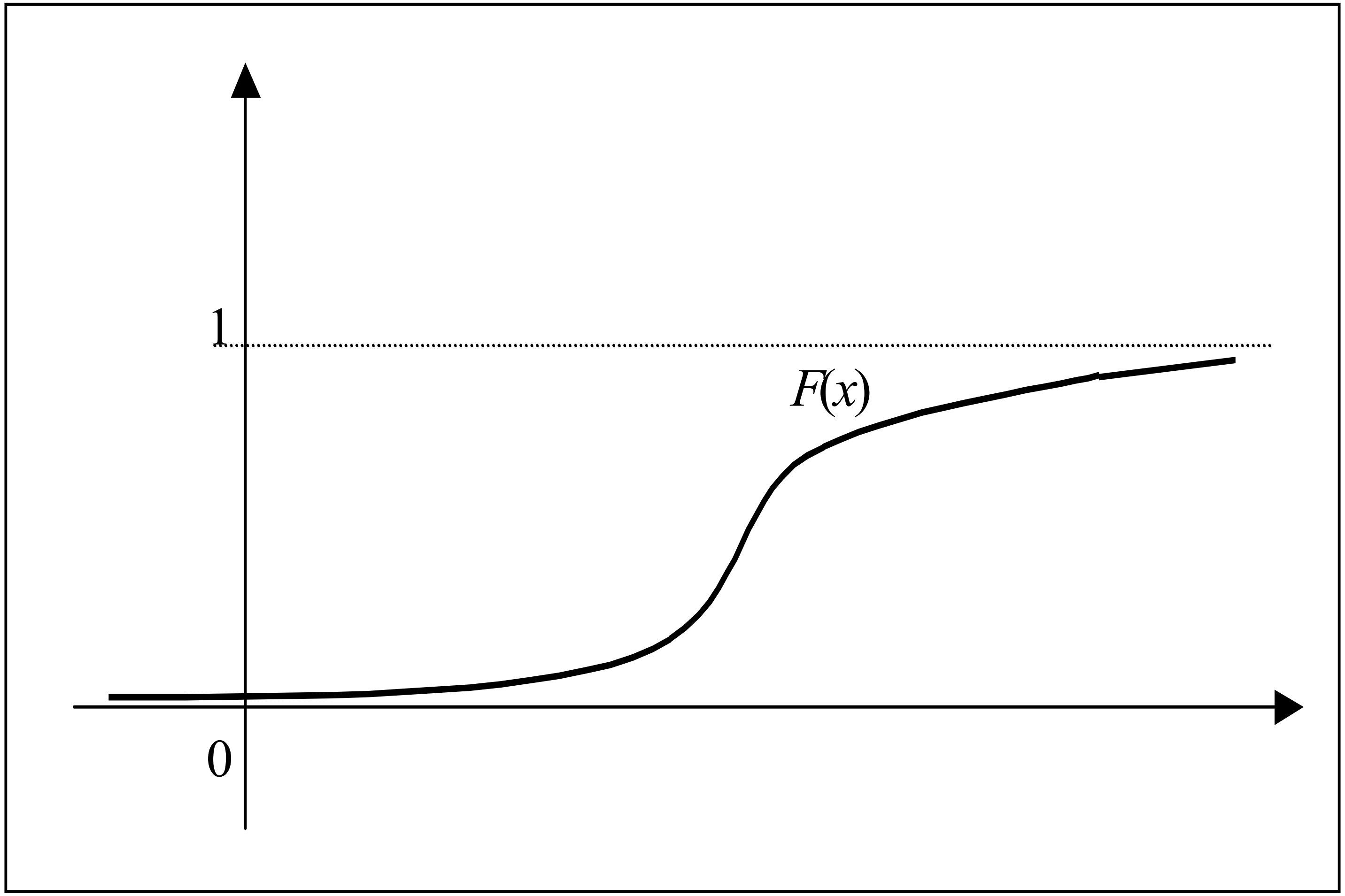



Za $x$ lze dosadit ***jakékoliv*** číslo!

# Distribuční funkce spojité NV

1. Neklesající spojitá funkce

2. Limity 0 a 1 pro $x \to \pm\infty$



# Vztahy mezi Hustotou a Distribuční funkcí

- Mezi HP a DF platí následující vztahy: HP je ***derivací*** DF:

$$f(x) = \frac{dF(x)}{dx}$$

pro všechna $x$ kde má DF derivaci

- Naopak: DF náhodné veličiny je ***neurčitým integrálem*** (primitivní funkcí) k HP, tj.

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

# Charakteristiky náhodné veličiny (NV)

- **Charakteristiky polohy:**
  - ♦ modus
  - ♦ medián
  - ♦ střední hodnota
  - ♦ kvantily
- **Charakteristiky variability:**
  - ♦ rozptyl
  - ♦ směrodatná odchylka

# Charakteristiky polohy

- Rozdělení pravděpodobnosti náhodné veličiny (NV) = úplné poznání NV:

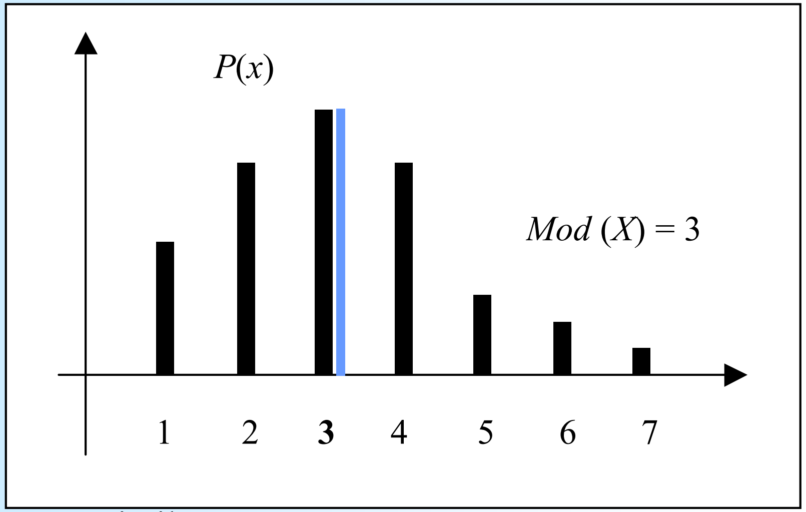
  - stanovení hodnot, jichž může náhodná veličina nabývat,

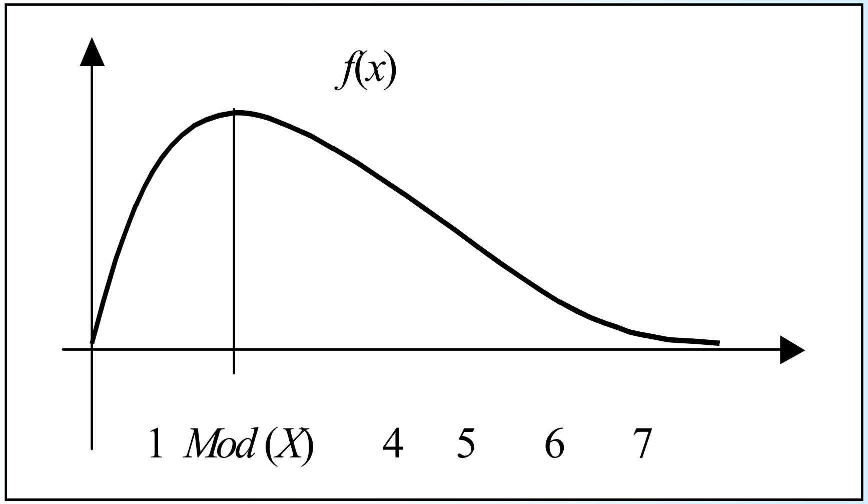
  - znalost pravděpodobností, s nimiž náhodná veličina nabývá určité hodnoty, nebo hodnoty z nějakého intervalu
- Charakteristiky polohy (ChP) = informace o náhodné veličině koncentrovaná do *jediného čísla*, které veličinu „dobře“ charakterizuje (způsob výpočtu je jednoznačně definován)
- Analogie s ChP statistického souboru!

# Charakteristiky polohy - Modus

***Modus*** = $Mod(X)$ NV $X$ s rozdělením pravděpodobnosti dané FP $p(x)$ (resp. FH $f(x)$ )

= hodnota $Mod(X)$, pro kterou je $p(x)$, resp. $f(x)$ maximální



# Charakteristiky polohy - Kvantily

- ***Kvantil*** - hodnota rozdělující obor hodnot NV $X$ v určitém poměru
- **100.*p-procentní kvantil*** $x_p$ - nejmenší takové číslo, které splňuje současně nerovnosti:

$$P\left(X \leq x_p\right) \geq p$$

$$P\left(X \geq x_p\right) \geq 1 - p$$

v případě spojité náhodné veličiny lze psát

$$F(x_p) = P(X \leq x_p) = p$$

# Charakteristiky polohy - Medián, Kvartily, Decily

Některé kvantily pro vybraná rozdělení pravděpodobnosti bývají tabelovány:

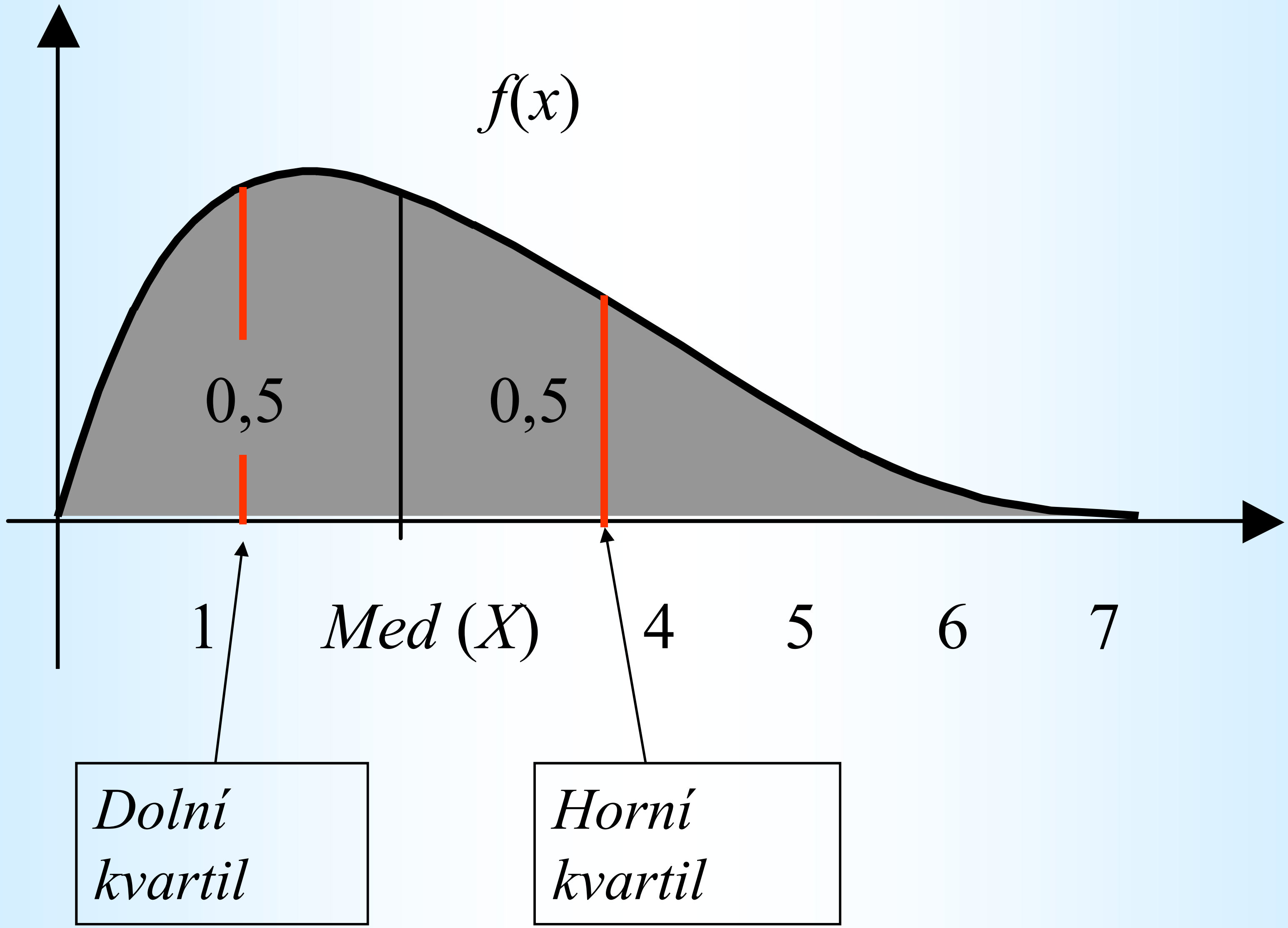
- *Medián* $Med(X)$ NV $X$ je 50-procentní kvantil:

$$Med(X) = x_{0,5}$$

- *Kvartil* NV je 25-procentní kvantil (jsou 3: horní, dolní, prostřední je medián)

- *Decil* NV je 10-procentní kvantil (je jich 9, prostřední je medián)

# Medián, kvartily



# Charakteristiky polohy - Střední hodnota

- $X$ - diskrétní náhodná veličina

  $p(x)$ - pravděpodobnostní funkce

  *střední hodnota* $E(X)$ - součet hodnot vynásobených jejich pravděpodobností:

$$E(X) = \sum_x x p(x)$$

# Charakteristiky polohy - Střední hodnota

- $X$ - spojitá náhodná veličina

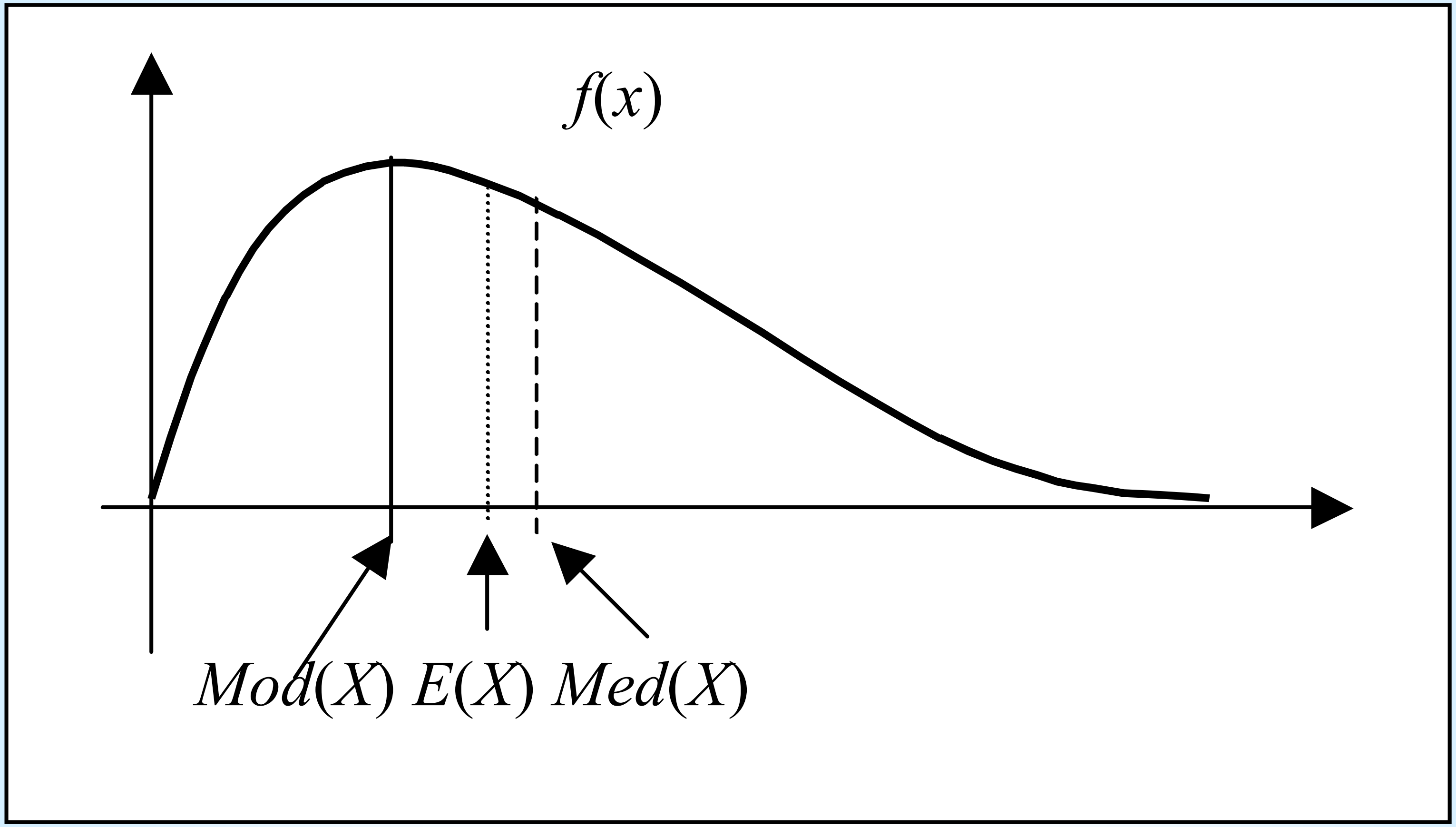
  $f(x)$ - funkce HP

***Střední hodnota* - $E(X)$**

*je* určitý integrál z hodnot $X$ násobených funkcí hustoty pravděpodobnosti $f(x)$ :

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

# Charakteristiky polohy NV mohou být různé!



# Vlastnosti střední hodnoty:

1. $k$ - konstantní NV: $E(k) = k$

2. $X$ - NV: $E(k.X) = k.E(X)$

3. $Y$ - NV: $E(X + Y) = E(X) + E(Y)$

4. $X$ a $Y$ **nezávislé** NV:

$$E(X.Y) = E(X).E(Y)$$

Pro závislé NV **neplatí**!!!

# Charakteristiky variability - Rozptyl a Směrodatná odchylka

- $X$ - diskrétní náhodná veličina
- $p(x)$ - pravděpodobnostní funkce
- ***rozptyl*** $Var(X)$ - součet kvadrátů odchylek od SH vynásobených jejich pravděpodobností:

$$Var(X) = \sum_x [x - E(X)]^2 p(x)$$

$$= \sum_x x^2 p(x) - [E(X)]^2$$

# Charakteristiky variability - Rozptyl a Směrodatná odchylka

- $X$ - spojitá náhodná veličina
- $f(x)$ - funkce hustoty
- *rozptyl* $Var(X)$ - integrál kvadrátů odchylek od SH vynásobených jejich hustotou:

$$Var(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx$$

- ***Směrodatná odchylka***:

$$\sigma(X) = \sqrt{Var(X)}$$

## Příklad 4:

Počet různých druhů zboží, které zákazník nakoupí při jedné návštěvě obchodního domu, je náhodná veličina $X$.

Bylo zjištěno, že tato veličina nabývá hodnot:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ = | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| $p(x)$ = | 0,2 | 0,4 | 0,25 | 0,1 | 0,03 | 0,01 | 0,01 |

- Střední hodnota počtu druhů zboží zakoupeného jedním zákazníkem

  $\boldsymbol{E(X)}$ **=** 0*0,2+1*0,4+2*0,25+3*0,1+4*0,03 +5*0,01+6*0,01 = 1,43

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ = | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| $p(x)$ = | 0,2 | 0,4 | 0,25 | 0,1 | 0,03 | 0,01 | 0,01 |

# Příklad 4 pokračování:

- Pravděpodobnostní funkce $p(x)$ nabývá maximální hodnotu 0,4 pro $x = 1$ :

$$\boldsymbol{Mod(X) = 1}$$

- Medián:

$p(X \leq 1) = p(X=0) + p(X=1)$ = 0,2+0,4 = 0,6 ≥ 0,5

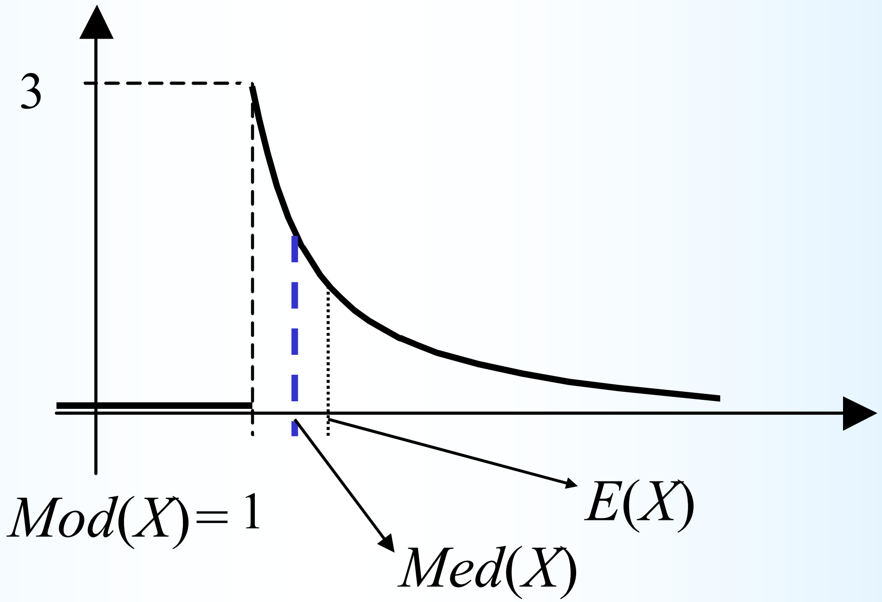
$p(X \geq 1) = p(X=1) + p(X=2) + \ldots + p(X=6)$ =
0,4+0,25+0,1+0,03+0,01+0,01 = 0,8 ≥ 1 - 0,5 = 0,5

Podle definice je medián:

$$\boldsymbol{Med(X) = 1}$$

- ***Var (X) =***
$0^2.0,2+1^2.0,4+2^2.0,25+3^2.0,1+4^2.0,03+5^2.0,01+6^2.0,01 - 1,43^2$ = 3,39 - 2,04 = **1,35**
- $\sigma(x)$ = √1,35 = **1,16**

# Příklad 5:




Hustota pravděpodobnosti náhodné veličiny $X$ je dána funkcí

$$f(x) = \frac{3}{x^4} \quad \text{pro } x \geq 1$$

$$= \quad 0 \quad \text{pro } x < 1$$

Snadno se ověří, že platí : $\int_{-\infty}^{+\infty} f(x)dx = 1$

# Příklad 5:

Svého maxima nabývá funkce hustoty zřejmě pro $x = 1 \Rightarrow \quad \boldsymbol{Mod(X) = 1}$

Všechny tři charakteristiky jsou rozdílné!

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \left[ -\frac{3}{2x^2} \right]_1^{\infty} = \frac{3}{2}$$

$$Med(X) = \sqrt[3]{2}$$

# Funkce náhodné veličiny

- $X_1, X_2,\ldots,X_n$ - NV
- ***„Statistika"*** - funkce náhodné veličiny: $Y = f(X_1, X_2,\ldots,X_n)$ - opět NV!!!
- K nejdůležitějším statistikám patří:

- ***Výběrový průměr***: $\overline{X} = \dfrac{X_1 + \ldots + X_n}{n}$

- ***Výběrový rozptyl***: $S^2 = \dfrac{(X_1 - \overline{X})^2 + \ldots + (X_n - \overline{X})^2}{n-1}$

- ***Výběrová směrodatná odchylka***: $S = \sqrt{S^2}$

# Charakteristiky statistik

- **Střední hodnoty**: $E(X_i) = \mu$ , $Var(X_i) = \sigma^2$

  výběrového průměru $\bar{X}$ : $E(\bar{X}) = \mu$

  výběrového rozptylu $S^2$ : $E(S^2) = \sigma^2$

  výběrové směrodatné odchylky $S$ :

  $$E(S) = \sigma$$

# Charakteristiky statistik

- **Rozptyly**:
  výběrového průměru $\overline{X}$: $$Var(\overline{X}) = \frac{\sigma^2}{n}$$

  výběrového rozptylu - neuvádíme (nebudeme využívat)