# Náhodný výběr

$X$ n - násobné nezávislé opakování náhodného pokusu, při kterém se realizuje $X$

dobrý výrobek $\mapsto 1$

vadný $\mapsto 0$

**Def.** Nechť $X$ je náhodná proměnná s distr. funkcí $F$

**Náhodný výběr** rozsah $n$ z daného rozdělení je $\mathbb{X} = (X_1, \ldots, X_n)$, kde $X_1, \ldots, X_n$ nezávislé n. p. se stejnou distr. funkcí $F$

$\mathbb{x} = (x_1, \dots, x_n)$

realizace náhodného výběru

Zvolit vhodný model

Odhadnout $\underline{E(X)}$ na základě hodnot získaných náhodným výběrem...

$$E(X),\ D(X),\ \sqrt{D(X)}$$

Odhady: <u>bodové</u> jedním číslem

pomocí tzv. <u>statistik</u>

výběrové charakteristiky

výběrový průměr

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

odhad střední hodnoty

# intervalové odhady


0,95

# Bodové odhady

## střední hodnoty a rozptylu

# Rozdělení pravděpodobnosti výběrových průměrů z normálního rozdělení

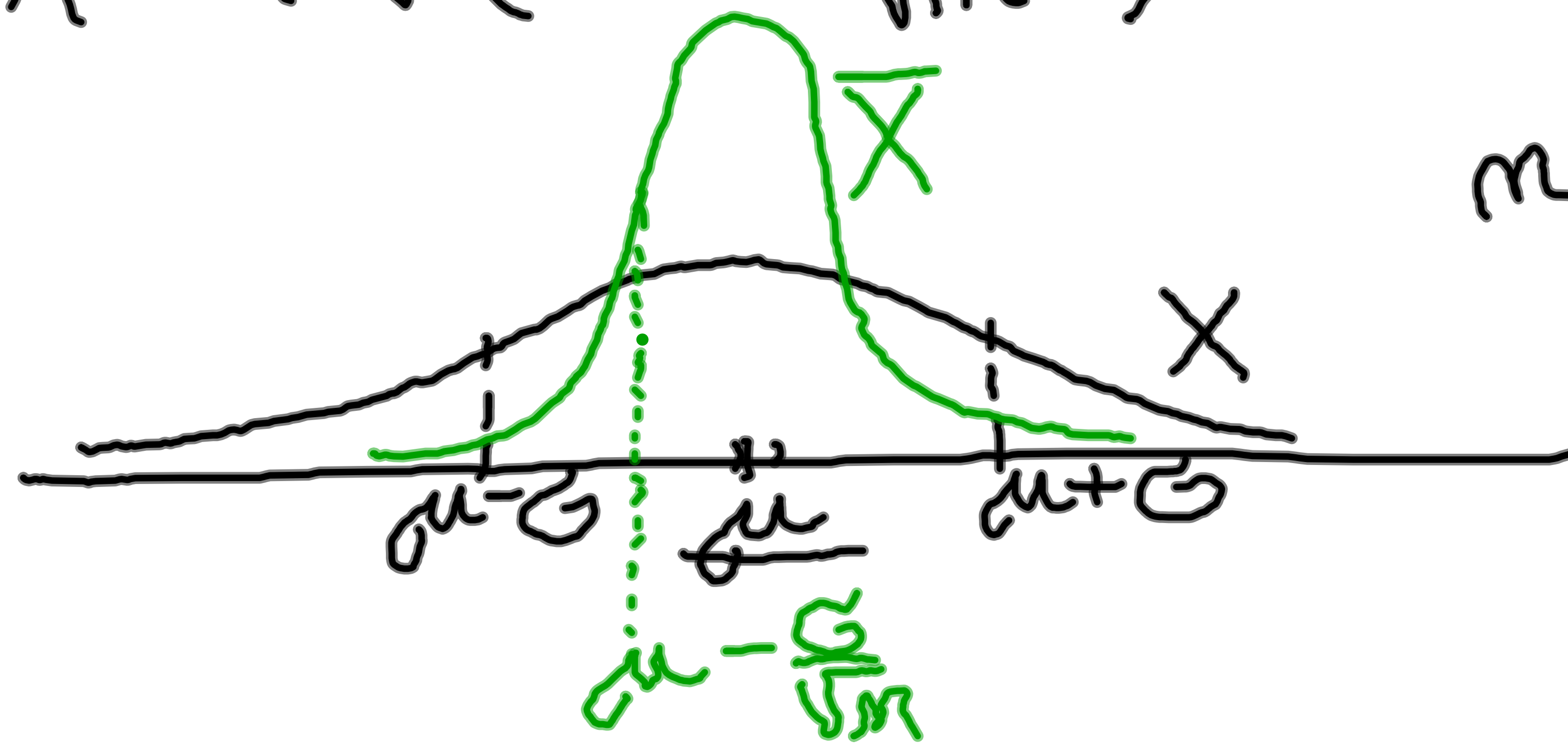
$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$X \sim N(\mu, \sigma)$$

$$\mathbb{X} = (X_1, \dots, X_n)$$

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$




$n \to \infty$

$$E(\overline{X}) = \mu$$

nestranný odhad

$$\lim_{n \to \infty} P\left(|\overline{X} - \mu| < \varepsilon\right) = 1 \quad \forall \varepsilon > 0$$

$\overline{X}$ konverguje k $\mu$ podle pravděpodobnosti

konzistentní odhad

$X$

$\overline{X}$

$E(\overline{X}) = E(X)$

$D(\overline{X}) = \frac{1}{n} D(X)$

$$n \to \infty$$

$$\overline{X} \sim N\left(\underline{E(X)}, \underline{\sqrt{\frac{D(X)}{n}}}\right)$$

plyne za centrální limitní věty

$\{X_n\}$, nezávislých, se stejným rozdělením

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n} X_i - nE(X)}{\sqrt{nD(X)}} < x\right) = \Phi(x)$$

$$\frac{1}{n}, \quad \frac{1}{n}$$

$$\frac{\frac{1}{n}\sum X_i - E(X)}{\sqrt{\frac{D(X)}{n}}} \quad N(0,1)$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(E(X), \sqrt{\frac{D(X)}{n}}\right)$$

**Tvrzení:** $(X_1, \dots, X_n)$ je náhodný výběr z rozdělení n. v. $X$ s distribuční funkcí $F(x)$ a střední hodnotou $E(X)$

$$E(\overline{X}) = E(X)$$

**D:** $E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \overbrace{E(X_i)}^{E(X)}$

$$= \frac{1}{n}\, n\, E(X) = E(X)$$

$$\boxed{D(\bar{X}) = \frac{D(X)}{n}}$$

$$D(\bar{X}) = D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) =$$

$$= \frac{1}{n^2} D\left(\sum_{i=1}^{n} X_i\right) =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{D(X_i)}_{D(X)} = \frac{1}{n^{\cancel{2}}} \cancel{n} D(X) =$$

$$= \frac{1}{n} D(X)$$

$\overline{X}$ konverguje podle pravděpodobnosti k $E(X)$

konsistentní

# Bodový odhad rozptylu

$$E(X-E(X))^2$$

$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

není nestranný

$$S^2 = \underbrace{\frac{1}{n-1}} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

nestranný a konzistentní odhad rozptylu,

$$E(S^2) = D(X)$$

<u>výběrový rozptyl</u>

$$E(S^2) = D(X)$$

$$(n-1)\underbrace{E(S^2)}_{\frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i-\overline{X})^2\right]} = E\left[\sum_{i=1}^{n}\underbrace{(X_i-\overline{X})^2}_{X_i^2-2X_i\overline{X}+\overline{X}^2}\right] =$$

$$= E\left[\sum X_i^2 \underbrace{-\,2\overline{X}\overbrace{\sum X_i}^{n\overline{X}} + n\overline{X}^2}_{-n\overline{X}^2}\right] =$$

$$= \sum_{i=1}^{n} \underline{E(X_i^2)} - n\,\underline{E(\overline{X}^2)} =$$

$$= \sum_{i=1}^{n} \left[D(X_i) + E^2(X_i)\right] - n\left[D(\overline{X}) + E^2(\overline{X})\right]$$

$$= nD(X) + nE^2(X) - n\frac{D(X)}{n} - nE^2(X)$$

$$= (n-1)D(X) \quad / :(n-1)$$

$$D(X) = E\left[(X-E(X))^2\right]$$

$$= E\left[X^2 - 2XE(X) + E^2(X)\right]$$

$$= E(X^2) \underbrace{- 2E^2(X) + E^2(X)}_{-E^2(X)}$$

$$E(X^2) = D(X) + E^2(X)$$

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

výběrová směrodatná odchylka