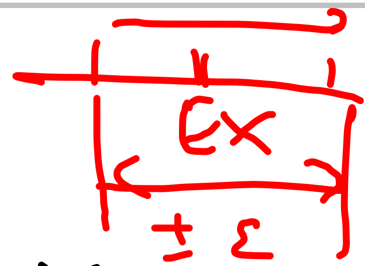
1) Čebyševova nerovnost



X, $\operatorname{var} X < \infty$, $\varepsilon > 0$:

$$P(|X - EX| \geq \varepsilon) \leq \frac{\operatorname{var} X}{\varepsilon^2}$$

Důkaz:

$$\operatorname{var} X = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$$

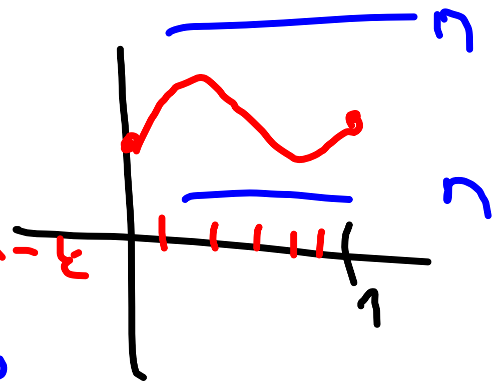
$$= \int_{|x-\mu| \geq \varepsilon} (x-\mu)^2 f(x)\, dx + \underbrace{\int_{|x-\mu| < \varepsilon} (x-\mu)^2 f(x)\, dx}_{\geq 0}$$

$$\geq \int_{|x-\mu| \geq \varepsilon} \varepsilon^2 f(x)\, dx$$

$$= \varepsilon^2 P(|X-\mu| \geq \varepsilon) \qquad \square$$

Pozn. (Weierstrassova věta)

$f(p)$ spojitá na $\langle 0, 1\rangle$



$$B_n(p) = \sum_{k=0}^{n} f(k/n) \underbrace{\binom{n}{k} p^k (1-p)^{n-k}}_{Bi(n,p)}$$

stejnoměrně aproximuje f na $\langle 0, 1\rangle$

Dk. $|f(p) - B_n(p)| \leq \ldots$

$$\leq \varepsilon + 2M\, P(|X - EX| \geq \delta)$$

$$\leq \varepsilon + \frac{2M}{n\delta^2}$$

Centrální limitní věta

$Bi(n,p)$ ... $\mu = np$, $\operatorname{var} X = np(1-p)$

Moivre–Laplace: $p \in (0,1)$, $-\infty < a < b < \infty$

X_n ... $Bi(n,p)$. Pak

$$\lim_{n\to\infty} P\Big(a < \underbrace{\frac{X_n - np}{\sqrt{npq}}}_{\text{přenormování na } \mu=0,\ \sigma^2=1} < b\Big) = \Phi(b) - \Phi(a)$$

(Φ je distribuční funkce pro $N(0,1)$)

rychlost konvergence je $O(n^{-1/2})$

$X_1, X_2, \dots$ náhodné veličiny, $\underset{0}{\mu}$, $\underset{1}{\sigma^2}$

$$S_n = \frac{1}{\sqrt{n}} \sum X_i$$

$$\boxed{\lim_{n\to\infty} S_n \text{ má rozdělení } N(0,1)}$$

$Y_1, Y_2, \dots \quad \mu, \sigma^2$

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{(Y_i - \mu)}{\sigma} \xrightarrow[n\to\infty]{} N(0,1)$$

$$X = bi(n, p)$$

neznámé

$$P\left(\left(\frac{1}{n}X - p\right) < 0{,}05\right) \cong 0{,}9$$

$$\Phi\left(\frac{0{,}05\, n}{\sqrt{np(1-p)}}\right) \cong (1+0{,}9)/2 = 0{,}95 \Rightarrow \underline{\underline{n > 270}}$$

$=$

$N(0,1)$

Popisná statistika

soubor hodnot $x_1, x_2, \ldots, x_n$

– tzv. statistického souboru

Měřítka: 1) nominální

2) ordinální (pořadí + uspořádání)

3) intervalová ←

4) poměrová měřítka

Četnost hodnot

$x_{(1)} \le x_{(2)} \le \ldots$

uspořádaný soubor hodnot

Míry polohy / variability

průměr („lineární")

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum_{j=1}^{m} n_j a_j$$

$x = (x_1, \dots, x_n)$

$$\overline{(a+bx)} = a + b\bar{x}$$

$$\bar{x}_G = \sqrt[n]{x_1 \cdot \dots \cdot x_n} \qquad (\ln \bar{x}_G = \overline{\ln x})$$

(výše inflace 20%, 50%, 30%, 20%, 5%, $\sqrt[5]{1{,}2 * 1{,}5 * 1{,}3 * 1{,}2 * 1{,}05}$)

$\bar{x}_G$ je lepší průměrná inflace

medián, kvantily, percentily

hodnota „v polovině" — po setřídění

[illegible] po čtvrtinách

Rozptyl (výběru) $\quad (x = (x_1, \dots, x_n))$

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{j=1}^{m} n_j (a_j - \bar{x})^2$$

Směrodatná odchylka $\quad s_x$

3. kvartil $\quad$ 1. kvartil

Kvartilové rozpětí $\quad R_Q = Q_3 - Q_1$

Statistika nad úplným souborem

× nad výběrem

A) výběr bez vracení z konečné populace

$x = (x_1, x_2, \dots, x_N)$ statistický soubor

$\mu = \frac{1}{N} \sum x_i \qquad \sigma^2 = \frac{1}{N} \sum (x_i - \mu)^2$

výběr n statistických jednotek – výběrový soubor

$= \binom{N}{n}$ možností

$(i_1, i_2, \dots, i_n)$ indexy vybraných hodnot

↑ pravděpodobnost $\frac{1}{N} \cdot \frac{1}{N-1} \cdots = \frac{(N-n)!}{N!}$

elementární jev ω odpovídající náhodné veličině

$(X_1(\omega), X_2(\omega), \dots, X_n(\omega)) = (x_{i_1}, \dots, x_{i_n})$

Nechť $(x_1, \dots, x_N)$ uvažujme $(w_1, \dots, w_N)$

$$w_i = w_i(s) = \begin{cases} 1 & \text{pro } i \in s \\ 0 & \text{pro } i \notin s \end{cases}$$

kde $s = (i_1, \dots, i_n)$.

$$\bar{X} = \frac{1}{n} \sum_{i \in s} x_i = \frac{1}{n} \sum_{i=1}^{N} x_i w_i$$

$$\Rightarrow \quad E\bar{X} = \mu$$

μ je nestranný odhad pro $E\bar{X}$

$$\operatorname{var} \bar{X} = \frac{N-n}{N-1} \frac{\sigma^2}{n}$$

$$\Downarrow$$

$$ES^2 = \frac{N}{N-1} \sigma^2$$