BODOVÉ ODHADY STŘEDNÍ HODNOTY A ROZPTYLU

... Pravděp., že výrobek je vadný, je 0,05 ...

? Jak to víme ?

$$X = \begin{cases} 0, & \text{je-li výrobek v pořádku} \\ 1, & \text{je-li výrobek vadný} \end{cases}$$

$X_i \quad (i = 1, 2, \dots, 100)$

$(X_1, \dots, X_{100})$ – náhodný výběr

$(x_1, \dots, x_{100})$ – realizace náhodného výběru

$$\mu \approx \underbrace{\frac{1}{100}\sum_{i=1}^{100} x_i \sim \frac{1}{100}\sum_{i=1}^{100} X_i}$$

Nechť $X_1, X_2, \dots, X_n$ jsou nezávislé náhodné proměnné na $(\Omega, \mathcal{S}, P)$. Nechť jsou stejně rozdělené. Pak je $(X_1, \dots, X_n)$ <u>náhodný výběr</u> s rozsahem n.

<u>Realizace náhodného výběru</u> je $(x_1, x_2, \dots, x_n)$, kde x_i je realizace X_i.

Bodové odhady střední hodnoty.

1.) Aritmetický průměr

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

je nazýván výběrovým průměrem.

Věta. Výběrový průměr je nevychýlený odhad střední hodnoty.

Proč? $E(X_i) = \mu, \; i = 1, \ldots, n$

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

$$= \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} \cdot n \cdot \mu$$

Poznámka. $\sqrt[n]{x_1 \cdot x_2 \cdots x_n}$ - geom. průměr

$$\frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$ – harmonický průměr

medián – „prostřední" hodnota x_i

modus – „nejčetnější" —"—

hrubé odhady střední hodnoty

Bodové odhady rozptylu

$$E(X_i) = \mu$$

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

Platí: $S^2(\mu)$ je nestranný odhad rozptylu σ^2.

Proč? $X_1, \dots, X_n$ je náhodný výběr s rozdělením náh. veličiny X

$$E(X_i) = \mu, \quad D(X_i) = \frac{1}{n} E(X_i - \mu)^2 = \sigma^2$$

Pak je

$$\underline{E(S^2(\mu))} = E\left(\frac{1}{n} \sum (X_i - \mu)^2\right) =$$

$$= \frac{1}{n} \sum E(X_i - \mu)^2 = \frac{1}{n} \cdot n \cdot \underline{\sigma^2}$$

Problém: Hodnota μ je neznámá.

$$\mu \sim \frac{1}{n}\sum X_i = \overline{X}$$

Zkusme

$$M_2 = \frac{1}{n}\sum \left(X_i - \overline{X}\right)^2$$

výběrový rozptyl

Jak je to teď se střední hodnotou

$$E(M_2) = ?$$

Platí: $E(M_2) = \left(1 - \frac{1}{n}\right)\sigma^2$

Důkaz: Po úpravě

$$M_2 = \sum X_i^2 - n\overline{X}^2$$

$$E(M_2) = \frac{1}{n}\Big(\underbrace{E\left(\sum X_i^2\right)}_{n(\sigma^2+\mu^2)} - \underbrace{E\left(n\overline{X}^2\right)}_{n\left(\frac{\sigma^2}{n}+\mu^2\right)}\Big)$$

Je dobře

$$E\left(\sum X_i^2\right) = D(X_i) + E(X_i^2)$$

$$E(n\bar{X}^2) = n\left(D(\bar{X}) + \left(E(\bar{X})\right)^2\right)$$

Místo $S^2(\mu)$ je třeba vzít

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

Platí: $E(S^2) = \sigma^2$, protože

$$S^2 = \frac{n}{n-1}\cdot M_2$$

$$E\left(\frac{n}{n-1}\cdot M_2\right) = \frac{n}{n-1}E(M_2) = \frac{n}{n-1}\cdot\frac{n-1}{n}\cdot\sigma^2$$

Realizace S^2 je s^2.

Pro praktický výpočet je vhodné použít

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$$

Pozn. variační rozpětí

$$x_{max} - x_{min}$$

Statistické zpracování naměřených dat

Hodnoty x_i	Četnosti	Relativní četnost	Kumulativní relativní četnost
1	n_1	p_1	p_1
2	n_2	p_2	p_1+p_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	n_k	p_k	$\sum_{i=1}^{k} p_i = 1$
$\sum$	n	1,000	

Grafické znázornění
histogramy, frekvenční křivky...
Zajímavost. Existují způsoby
čitelného zápisu statistik. Dat,
které jsou současně grafické!

Příklad · Archeologie – 35 nálezů

s průměry

48 57 . – – .

. – – –

27

stem and leaf

leaves

stem	leaves
2	5 7 8
3	0 4 5 8 8 8 9
4	0 0 0 2 3 3 3 4 5 7 7 7 ...
5	0 0 3 7 7 8
6	6