Cvičení 8. Výběrové charakteristiky

Mějme náhodný výběr $\boldsymbol{X}$ z normálního rozdělení:

$$\boldsymbol{X} = (X_1, \ldots, X_n), \forall i = 1, \ldots, n : X_i \to N(\mu, \sigma).$$

Výběrová charakteristika	Rozdělení pravděpodobnosti	Poznámka
$\frac{\overline{X} - \mu}{\sigma}\sqrt{n}$	$N(0;1)$	viz CLV
$\frac{\overline{X} - \mu}{S}\sqrt{n}$	t_{n-1}	viz vlastnosti Studentova rozdělení
$\frac{S^2}{\sigma^2}(n-1)$	χ^2_{n-1}	viz vlastnosti χ^2 - rozdělení
$\frac{p - \pi}{\sqrt{\pi(1-\pi)}}\sqrt{n}$	$N(0;1)$	viz vlastnosti realtivní četnosti, předpoklad: $n > \frac{9}{p(1-p)}$

$\overline{X}$ .... výběrový průměr (odhad μ)

S ... výběrová směrodatná odchylka (odhad σ)

p .... relativní četnost (odhad π)

n ... velikost výběru

Rozdělení součtu a průměru

$$\frac{\bar{X}-\mu}{\sigma}\cdot\sqrt{n} \sim N(\mu=0;\sigma^2=1)$$

SOUČET: $\sum_{i=1}^{n} X_i = n\cdot\bar{X}$

$$\frac{\bar{X}-\mu}{\sigma}\cdot\sqrt{n} \sim N(0;1) \qquad \cdot\frac{\sigma}{\sqrt{n}}$$

$$\bar{X}-\mu \sim N\left(0;\left(\frac{\sigma}{\sqrt{n}}\right)^2\right) \qquad +\mu$$

$$\bar{X} \sim N\left(\mu;\frac{\sigma^2}{n}\right) \qquad \cdot n$$

$$n\cdot\bar{X} \sim N(n\cdot\mu;n\cdot\sigma^2)$$

$$\sum_{i=1}^{n} X_i \sim N(n\cdot\mu;n\cdot\sigma^2)$$

Průměr: $\bar{X} \sim N\left(\mu;\frac{\sigma^2}{n}\right)$

Příklad 1.

Zatížení letadla s 64 místy nemá překročit 6 000 kg. Jaká je pravděpodobnost, že při plném obsazení bude tato hodnota překročena, má-li hmotnost cestujícího střední hodnotu 90 kg a směrodatnou odchylku 10 kg?

X_i ... hmotnost cestujícího

$E(X_i) = 90$

$D(X_i) = 10^2$

$X = \sum_{i=1}^{64} X_i$

X ... váha 64 cestujících

$X \sim N(\mu = 64 \cdot 90 ; \sigma^2 = 64 \cdot 10^2)$

$P(X > 6000) = 1 - F_X(6000) \doteq 0{,}0013$

$\rightarrow 0{,}13\%$

Pravd., že letadlo bude přetíženo je 0,13%.

Příklad 2.

Zásilka obsahuje 300 výrobků určitého typu. Je známo, že pravděpodobnost zhotovení vadného výrobku tohoto typu je 0,04.

a)

Odhadněte pravděpodobnost, že absolutní odchylka podílu vadných výrobků v zásilce od pravděpodobnosti vyrobení vadného výrobku bude menší než 1 %.

b)

Jak se změní výsledek, jestliže zásilka bude obsahovat 3 000 výrobků?

$n = 300$

$\pi = 0{,}04$

$$X \sim Bi(n = 300;\ \pi = 0{,}04)$$

$$\frac{p - \pi}{\sqrt{\pi \cdot (1-\pi)}} \cdot \sqrt{n} \sim N(0,1)$$

a) $P(|p - \pi| < 0{,}01) = ?$

$$P(-0{,}01 < p - \pi < 0{,}01) \qquad \cdot \frac{\sqrt{n}}{\sqrt{\pi \cdot (1-\pi)}}$$

$$\Rightarrow P\left(\frac{-0{,}01}{\sqrt{\pi \cdot (1-\pi)}} \cdot \sqrt{n} < \underbrace{\frac{p - \pi}{\sqrt{\pi \cdot (1-\pi)}} \cdot \sqrt{n}}_{Y} < \frac{0{,}01}{\sqrt{\pi \cdot (1-\pi)}} \cdot \sqrt{n}\right)$$

$-$

$Y \sim N(0,1)$

$P(a<Y<b) = F(b) - F(a)$

$P(\dots) = F_Y\left(\frac{0{,}01}{\sqrt{0{,}04 \cdot 0{,}96}} \cdot \sqrt{300}\right) - F_Y\left(\frac{-0{,}01}{\sqrt{0{,}04 \cdot 0{,}96}} \cdot \sqrt{300}\right)$

$\doteq 0{,}6232$

Pravd., že absol. odchylka $|p-\pi|$ bude menší než 0,01 je 62,32%

b) $n = 3000$

$P(\dots) = F_Y\left(\frac{0{,}01}{\sqrt{0{,}04 \cdot 0{,}96}} \cdot \sqrt{3000}\right) - F_Y\left(\frac{-0{,}01}{\sqrt{0{,}04 \cdot 0{,}96}} \cdot \sqrt{3000}\right)$

$\doteq 0{,}9948$

— " — 99,48%

$$\frac{p-\pi}{\sqrt{\pi\cdot(1-\pi)}}\cdot\sqrt{n} \sim N\overset{\mu,\,\sigma^2}{(0,1)} \quad \Big/ \cdot\frac{\sqrt{\pi\cdot(1-\pi)}}{\sqrt{n}}$$

$$p-\pi \sim N\left(0,\frac{\pi\cdot(1-\pi)}{n}\right) \quad \Big/ +\pi$$

$$p \sim N\left(\pi, \frac{\pi\cdot(1-\pi)}{n}\right)$$

Příklad 3.

Cestující pravidelně jezdí do zaměstnání a zpět MHD. Je známo, že doba čekání na příjezd MHD se pohybuje v mezích od 0 do 3 minut. Jaká je pravděpodobnost, že celková doba čekání zaměstnance na příjezd MHD během 23 pracovních dnů bude kratší než 80 minut?

$n = 23 \cdot 2 = 46$

X_i ... doba čekání na MHD

$X_i \sim Ro(a = 0;\ b = 3)$

$E(X_i) = \frac{0+3}{2} = 1{,}5$

$D(X_i) = \frac{(b-a)^2}{12} = \frac{9}{12} = \frac{3}{4} = 0{,}75$

$X = \sum_{i=1}^{46} X_i \qquad X \sim N(\mu = n \cdot E(X_i);\ \sigma^2 = n \cdot D(X_i))$

$P(X < 80) = F_X(80) \doteq 0{,}9694$

Pravd., že budeme čekat méně než 80 minut je 96,94%

Příklad 4.

Předpokládejme, že průměrná spotřeba elektrické energie domácností v určitém městě v lednu je 120 kWh a směrodatná odchylka spotřeby je 100 kWh. Určete pravděpodobnost, že průměrná spotřeba 100 náhodně vybraných domácností bude větší než 140 kWh.

$n = 100$

$E(X_i) = 120$

$D(X_i) = 100^2$

$\overline{X} = \dfrac{\sum_{i=1}^{100} X_i}{n}$

$\overline{X} \sim N\left(\mu = E(X_i);\ \sigma^2 = \dfrac{D(X_i)}{n}\right)$

$P(\overline{X} > 140) = 1 - P(\overline{X} \leq 140) = 1 - F_{\overline{X}}(140)$

$\doteq 0{,}0228$

Pravděp., že průměrná spotřeba 100 domácností bude větší než 140 kWh je 2,28 %.

Příklad 5.

Společnost Acme Battery Company vyvinula nový typ baterie mobilních telefonů. V průměru vydrží baterie 60 minut na jedno nabití. Směrodatná odchylka této doby je 4 minuty. Předpokládejme, že výrobní oddělení po 6 měsících spustí test kontroly kvality. Provedli dva náhodné výběry o rozsahu 10 baterií a v obou zjistili směrodatnou odchylku výdrže baterií větší než 6 minut. S jakou pravděpodobností takový výsledek mohli očekávat?

X_i … výdrž baterie

$E(X_i) = 60$

$D(X_i) = 4^2$

$n = 10$ (dva výběry)

$$P(S_1 > 6 \wedge S_2 > 6) = P(S_1 > 6) \cdot P(S_2 > 6) = P(S > 6)^2 = ?$$

$$\underbrace{\frac{S^2}{\sigma^2} \cdot (n-1)}_{Y} \sim \chi^2_{n-1}$$

$$P(S^2 > 36) \qquad / \cdot \frac{n-1}{\sigma^2}$$

$$P\Big(\underbrace{\frac{S^2}{\sigma^2}\cdot(n-1)}_{Y} > \frac{36}{D(X_i)}\cdot(n-1)\Big)$$

$$\sigma^2 = D(X_i)$$

$$Y \sim \chi^2_9$$

$$P\left(Y > \frac{36}{16}9\right) = 1 - P\left(Y \leq \frac{36 \cdot 9}{16}\right)$$

$$= 1 - F_Y\left(\frac{36 \cdot 9}{16}\right)$$

$$\doteq 0{,}0104$$

$$P(S_1 > 6 \wedge S_2 > 6) \doteq 0{,}00027$$

Pravd., že oba výběry budou mít $S > 6$ je $0{,}027\%$

Příklad 6.

Z úmrtnostních tabulek vyplývá pravděpodobnost 0,99, že se 35 - letý muž dožije dalšího roku. Roční pojistné této věkové skupiny činí 2 000 Kč, v případě úmrtí pojišťovna vyplatí 100 000 Kč. Jaká je pravděpodobnost, že zisk z 500 pojištěných mužů ve věku 35 let bude alespoň 500 000 Kč? (Řešte dvěma způsoby - pomocí binomického rozdělení a pomoci aproximace binomického rozdělení rozdělením normálním.)

X ... počet vyplacených pojistek

$X \sim Bi(n = 500; \pi = 0{,}01)$

Z ... zisk pojišťovny

$Z = \underbrace{500 \cdot 2000}_{10^6} - X \cdot 100\,000$

$P(Z \geq 500\,000) =$

$P(500 \cdot 2000 - X \cdot 100\,000 \geq 500\,000)$

$= P(X \cdot 100\,000 \leq 500\,000) =$

$= \underline{P(X \leq 5)} \doteq 0{,}6160$

$X \sim Bi(n, \pi) \approx N(\mu = n \cdot \pi,$
$\sigma^2 = n \cdot \pi \cdot (1 - \pi))$

$P(Z \geq 500\,000) = \underline{P(X \leq 5)} \approx P(X < 5{,}5)$

OPRAVA NA SPOJITOST

$$\left[\begin{array}{l} P(X = 5) = 0 \\ \quad \Rightarrow \approx P(4{,}5 < X < 5{,}5) \end{array}\right]$$

$\doteq \underline{\underline{0{,}5884}}$

Pravd., že pojišťovna získá více než 500.000 Kč je

61,60% / 58,84%.

Příklad 7.

Předpokládejme, že v populaci má přibližně 60 % mladých mužů vyšší než doporučenou hladinu cholesterolu v séru. S jakou pravděpodobností bude mít v náhodném výběru 200 mladých mužů více než 120 z nich vyšší než doporučenou hladinu cholesterolu v séru?

X ... počet mužů s vyšším chol.

$X \sim Bi(n = 200; \pi = 0{,}6)$

$P(X > 120) = \underline{1 - P(X \leq 120)}$

$\doteq 0{,}4732$

$X \approx N(\mu = n \cdot \pi; \sigma^2 = n \cdot \pi \cdot (1-\pi))$

$P(X > 120) = \underline{1 - P(X \leq 120)}$

$\approx 1 - P(X < 120{,}5) \doteq 0{,}4712$

Pravd., že ve výběru bude více než 120 mužů s vyšším chol. je 47,32% / 47,12%.