Основные формулы

Математическое ожидание дискретной случайной величины:

$$\mathbb{E}\left(X\right) = \sum_{i=1}^{n} x_i \cdot p_i$$

Дисперсия дискретной случайной величины:

$$\mathbb{D}\left(X\right) = \mathbb{E}\left(X - \mathbb{E}\left(X\right)^2\right)$$

Среднее квадратическое отклонение:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}$$

Неравенство Маркова:

$$P_{\{\xi\{\omega\}\geq a\}} \leq \frac{\mathbb{E}f\left(\xi\right)}{b}$$

Неравенство Чебышева:

$$P_{\{\xi\{\omega\}\geq a\}} \leq \frac{\mathbb{E}f\left(\xi^2\right)}{a^2}$$

Задание 1: Среднее количество вызовов, поступающих на коммутатор завода в течение часа, равно 300. Оценить вероятность того, что в течение следующего часа число вызовов на коммутатор: а) превысит 400; б) будет не более 500.

Дано: $\mathbb{E}x = 300$

Найти: а) $P(x > 400)$ б) $P(x \leq 500)$

Решение:

Применим Неравенство Маркова

а) $P(x \geq 400) \leq \frac{300}{400} = 0.75 \Rightarrow$

$P(x > 400) \leq 0.75 - P(x = 400) < 0.75$

Ответ: $P(x > 400) < 0.75$

б) $P(x \leq 500) = 1 - P(x > 500)$

$P(x \geq 500) \leq \frac{300}{500} = 0.6$

$P(x > 500) \leq 0.6 - P(x = 500) < 0.6$

$P(x \leq 500) \geq 1 - 0.6 = 0.4$

Ответ: $P(x \leq 500) \geq 0.4$

Задание 2: В 1600 испытаниях Бернулли вероятность успеха в каждом испытании равна 0,3. С помощью неравенства Чебышёва оценить вероятность того, что разница между числом успехов в этих испытаниях и средним числом успехов будет меньше 50

Дано:

$n = 1600$

$\xi \sim Bernoulli(p)$

$p = 0.3$

$\varepsilon = 50$

Найти:

$P(|\xi - \mathbb{E}\xi| < \varepsilon)$

Решение:

$\mathbb{E}\xi = np = 1600 \cdot 0.3 = 480$

$\mathbb{D}\xi = np(1-p) = 336$

$P(|\xi - \mathbb{E}\xi| < \varepsilon) \geq 1 - \frac{\mathbb{D}\xi}{\varepsilon^2}$ - неравенство Чебышева

$P(|\xi - 480| < 50) \geq 1 - \frac{336}{50^2} = 0.866$

Ответ: $P(|\xi - \mathbb{E}\xi| < \varepsilon) = 0.866$

Задание 3: Дана выборка $9, 5, 7, 7, 4, 10$. Дисперсия $\mathbb{D} = 1$. Постройте 99% доверительный интервал

Дано:

$X \in \{9, 5, 7, 7, 4, 10\}$

$\mathbb{D}X = 1$

Найти:

99% доверительный интервал

Решение:

$\overline{X} = \frac{\sum_{i=1}^{6} x_i}{6} = \frac{9+5+7+7+4+10}{6} = \frac{42}{6} = 7$

$1 - \frac{\alpha}{2} = 1 - 0.005 = 0.995$

$z_{0.995} = 2.58$ -квантиль

$\Delta = \frac{\mathbb{D}X}{\sqrt{k}} z_{1-\frac{\alpha}{2}} = \frac{2.58}{\sqrt{6}} = 1.053$

$\overline{X} - \Delta = 5.947$

$\overline{X} + \Delta = 8.053$

99% доверительный интервал: $(\overline{X} - \Delta, \overline{X} + \Delta) = (5.947; 8.053)$

Ответ: $(5.947; 8.053)$

Задание 4: Пусть X_i подчиняется нормальному распределению $N(\mu, \sigma^2)$. Найти ОМП $\hat{\mu}$ и $\hat{\sigma}$.

Почему найденное решение будет точкой максимума функции прадоподобия, а не седловой точкой?

Дано:

$X_i \sim N(\mu, \sigma^2)$

Найти:

ОМП $\hat{\mu}$

ОМП $\hat{\sigma}$

Доказать:

$(\hat{\mu}, \hat{\sigma})$ – точка максимума функции правдоподобия

Решение:

$f_X(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$

$$f_X(x;\mu,\sigma^2) = \Pi_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x_i-\mu)^2}{2\sigma^2}\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left(\frac{-\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}\right)$$

$$\ln(L(X)) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \sum_{i=1}^n \left(\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

$$\frac{\partial \ln(L(X))}{\partial \mu} = -\frac{1}{2\sigma^2}\sum_{i=1}^n 2(x_i-\mu)(-1) = \sum_{i=1}^n \frac{x_i-\mu}{\sigma^2} = \frac{\sum_{i=1}^n x_i - n\mu}{\sigma^2} = \frac{n\overline{X} - n\mu}{\sigma^2} \Rightarrow \hat{\mu} = \overline{X}$$

$$\frac{\partial \ln(L(X))}{\partial \sigma} = -n\cdot\frac{1}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^n (x_i-\mu)^2 = \frac{\sum_{i=1}^n (x_i-\mu)^2 - n\sigma^2}{\sigma^3} = 0 \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i-\hat{\mu})^2}{n} \Rightarrow \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (x_i-\hat{\mu})^2}{n}}$$

Для того, чтобы найденная точка $(\hat{\mu}, \hat{\sigma})$ была точкой максимума, достаточно, чтобы все её вторые частные производные были отрицательны.

$$\frac{\partial^2 \ln(L(X))}{\partial \mu^2} = -\frac{n}{\sigma^2} < 0 \ \forall(\mu,\sigma)$$

$$\frac{\partial^2 \ln(L(X))}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{i=1}^n \left(x_i-\mu\right)^2 = \left.\frac{n\sigma^2 - 3\sum_{i=1}^n (x_i-\mu)^2}{\sigma^4}\right|_{\sigma=\hat{\sigma}} = \frac{\sum_{i=1}^n (x_i-\mu)^2 - 3\sum_{i=1}^n (x_i-\mu)^2}{\sigma^4} =$$
$$= -\frac{2\sum_{i=1}^n (x_i-\mu)^2}{\sigma^4} < 0$$

Следовательно, поскольку обе производные отрицательны для любой точки, то найденная точка будет точкой максимума, а не седловой точкой.

Ответ: $\hat{\mu} = \overline{X}, \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (x_i-\hat{\mu})^2}{n}}$