

Решения задач по математической статистике

КОНСТАНТИН ЗЮБИН

СОДЕРЖАНИЕ

1. Дз 1

1

1. Дз 1

Приведём несколько стандартных утверждений.

Предложение 1. Пусть $X \sim \mathcal{N}(a, \sigma^2)$ — нормально распределённая случайная величина. Тогда

$$E X = a, E X^2 = \sigma^2 + a^2, D X = \sigma^2, E |X - a| = \sigma \cdot \sqrt{\frac{2}{\pi}}.$$

Доказательство. Для матожиданий имеем

$$E X = \int_{-\infty}^{+\infty} \frac{u}{\sqrt{2\pi}\sigma} e^{-\frac{(u-a)^2}{2\sigma^2}} du = \int_{-\infty}^{+\infty} \frac{t}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} dt + \int_{-\infty}^{+\infty} \frac{a}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} dt = 0 + \int_{-\infty}^{+\infty} \frac{a}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds = a \cdot 1 = a.$$

$$\begin{aligned} E X^2 &= \int_{-\infty}^{+\infty} \frac{u^2}{\sqrt{2\pi}\sigma} e^{-\frac{(u-a)^2}{2\sigma^2}} du = \int_{-\infty}^{+\infty} \frac{t^2}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} dt + \int_{-\infty}^{+\infty} \frac{2at}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} dt + \int_{-\infty}^{+\infty} \frac{a^2}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} dt = \\ &= \int_{-\infty}^{+\infty} \frac{t}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} d\frac{t^2}{2} + 0 + a^2 \cdot 1 = \int_{-\infty}^{+\infty} \frac{t}{\sqrt{2\pi}\sigma} (-\sigma^2) de^{-\frac{t^2}{2\sigma^2}} + a^2 = \\ &= -\frac{t\sigma}{\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} dt + a^2 = 0 + \int_{-\infty}^{+\infty} \frac{\sigma \cdot \sigma}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds + a^2 = \sigma^2 + a^2. \end{aligned}$$

$$\begin{aligned} E |X - a| &= \int_{-\infty}^{+\infty} \frac{|u-a|}{\sqrt{2\pi}\sigma} e^{-\frac{(u-a)^2}{2\sigma^2}} du = \int_{-\infty}^{+\infty} \frac{|t|}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} dt = 2 \int_0^{+\infty} \frac{t}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} dt = \\ &= 2 \int_0^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} d\frac{t^2}{2} = 2 \int_0^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} (-\sigma^2) de^{-\frac{t^2}{2\sigma^2}} = -\frac{2\sigma}{\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \Big|_0^{+\infty} = \frac{2\sigma}{\sqrt{2\pi}} = \sigma \cdot \sqrt{\frac{2}{\pi}}. \end{aligned}$$

Из формулы дисперсии получаем

$$D X = E X^2 - (E X)^2 = (\sigma^2 + a^2) - a^2 = \sigma^2.$$

□

Предложение 2. Пусть $X_1 \sim \mathcal{N}(a_1, \sigma_1^2)$, $X_2 \sim \mathcal{N}(a_2, \sigma_2^2)$ — две нормально распределённые случайные величины. Тогда их свёртка $X_1 + X_2 \sim \mathcal{N}(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$ имеет нормальное распределение.

Задача 1а

Пусть даны независимые случайные величины $X_1, \dots, X_n \sim \mathcal{N}(\theta_1, \theta_2^2)$, $\theta_2 > 0$. Проверим, что оценка $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ несмещена относительно θ_1 . Действительно, по линейности матожидания и так как $E_{\theta_1, \theta_2} X_i = \theta_1$ имеем

$$E_{\theta_1, \theta_2} \bar{X} = \frac{1}{n} \sum_{i=1}^n E_{\theta_1, \theta_2} X_i = \frac{1}{n} \cdot n \theta_1 = \theta_1.$$

Задача 1б

Проверим, что оценка

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

смещена относительно θ_2^2 . Имеем

$$\theta_2^2 = D_{\theta_1, \theta_2} X_i = E_{\theta_1, \theta_2} X_i^2 - (E_{\theta_1, \theta_2} X_i)^2 = E_{\theta_1, \theta_2} X_i^2 - \theta_1^2.$$

Кроме того,

$$\begin{aligned} E_{\theta_1, \theta_2} \bar{X}^2 &= \frac{1}{n^2} \left(\sum_{i=1}^n E_{\theta_1, \theta_2} X_i^2 + \sum_{i < j} E_{\theta_1, \theta_2} (2X_i X_j) \right) = \\ &= \frac{1}{n^2} (n(\theta_1^2 + \theta_2^2) + n(n-1) E_{\theta_1, \theta_2} X_i E_{\theta_1, \theta_2} X_j) = \frac{1}{n^2} (n(\theta_1^2 + \theta_2^2) + n(n-1)\theta_1^2) = \\ &= \frac{1}{n} \theta_2^2 + \theta_1^2. \end{aligned}$$

Отсюда

$$E_{\theta_1, \theta_2} S_n^2 = E_{\theta_1, \theta_2} \bar{X}^2 - E_{\theta_1, \theta_2} \bar{X}^2 = \frac{1}{n} \cdot n(\theta_1^2 + \theta_2^2) - (\frac{1}{n} \theta_2^2 + \theta_1^2) = \frac{n-1}{n} \theta_2^2 \neq \theta_2^2.$$

Задача 1в

В качестве чисел a_n (для $n > 1$) следует взять дроби $\frac{n}{n-1}$. Тогда, по доказанному выше

$$E_{\theta_1, \theta_2} \left(\frac{n}{n-1} \cdot S_n^2 \right) = \frac{n}{n-1} \cdot \frac{n-1}{n} \cdot \theta_2^2 = \theta_2^2.$$

Положим $T = \sum_{i=1}^{n-1} |X_i - X_{i+1}|$. Случайная величина $-X_{i+1}$ имеет распределение $\mathcal{N}(-\theta_1, \theta_2^2)$.

По формуле для свёртки нормально распределённых случайных величин случайная величина $X_i - X_{i+1}$ имеет распределение $\mathcal{N}(0, 2\theta_2^2)$. Первый момент такой случайной величины равен $\sqrt{2}\theta_2 \cdot \sqrt{\frac{2}{\pi}} = \frac{2\theta_2}{\sqrt{\pi}}$. Теперь мы можем вычислить матожидание T :

$$E_{\theta_1, \theta_2} T = (n-1) E_{\theta_1, \theta_2} |X_1 - X_2| = \frac{2(n-1)}{\sqrt{\pi}} \cdot \theta_2.$$

В качестве членов последовательности b_n следует взять числа $\frac{\sqrt{\pi}}{2(n-1)}$.

Задача 2а

Вычислим моменты случайной величины $X \sim E(\theta)$ с экспоненциальным распределением. Имеем

$$\begin{aligned} E X^k &= \int_0^{+\infty} \theta x^k e^{-\theta x} dx = - \int_0^{+\infty} x^k de^{-\theta x} = \\ &= -x^k e^{-\theta x} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-\theta x} dx^k = \int_0^{+\infty} k x^{k-1} e^{-\theta x} dx = \frac{k}{\theta} E X^{k-1}. \end{aligned}$$

Согласно этой формуле

$$E X^k = \frac{k!}{\theta^k} E X^0 = \frac{k!}{\theta^k}.$$

Пусть теперь $X_1, X_2, \dots, X_n \sim E(\frac{1}{\theta})$ — независимые одинаково распределённые случайные величины. Тогда имеем

$$E_{\theta} \overline{X^k} = \frac{1}{n} \cdot \sum_{i=1}^n E_{\theta} X_i^k = \frac{1}{n} \cdot n \cdot E_{\theta} X_1^k = k! \cdot \theta^k.$$

Подходящей последовательностью чисел c_k будет последовательность $\frac{1}{k!}$. Действительно, $E_{\theta} \overline{X^k} \cdot c_k = k! \cdot \theta^k \cdot \frac{1}{k!} = \theta^k$.

Задача 26

Пусть $P(\theta) = \sum_{k=0}^n a_k \theta^k$ — многочлен. Проверим, что оценка $\sum_{k=0}^n \frac{a_k}{k!} \overline{X^k}$ несмещена для $P(\theta)$.

Имеем

$$E_{\theta} \sum_{k=0}^n \frac{a_k}{k!} \overline{X^k} = \sum_{k=0}^n \frac{a_k}{k!} \cdot E_{\theta} \overline{X^k} = \sum_{k=0}^n \frac{a_k}{k!} \cdot k! \cdot \theta^k = P(\theta).$$

Задача 3а

Вычислим матожидание случайной величины $X \sim \text{Geom}(\theta)$ ($P(X = k) = \theta(1 - \theta)^{k-1}$) с геометрическим распределением:

$$\begin{aligned} E X &= \sum_{k=1}^{+\infty} k \theta (1 - \theta)^{k-1} = \theta \sum_{k=1}^{+\infty} k (1 - \theta)^{k-1} = \theta \left(- \sum_{k=1}^{+\infty} (1 - \theta)^k \right)' = \\ &= \theta \left(- \frac{1 - \theta}{1 - (1 - \theta)} - 1 \right)' = \theta \left(\frac{-1}{\theta} \right)' = \theta \cdot \frac{1}{\theta^2} = \frac{1}{\theta}. \end{aligned}$$

Пусть независимые случайные величины $X_1, \dots, X_n \sim \text{Geom}(\theta)$ имеют геометрическое распределение ($\theta \in (0, 1)$).

Пусть $g(X_1)$ — несмещённая оценка для функции $f(\theta)$. Тогда

$$f(\theta) = E_{\theta} g(X_1) = \sum_{k=1}^{+\infty} g(k) \theta (1 - \theta)^{k-1} = \sum_{k=0}^{+\infty} g(k+1) (-1)^k \theta (\theta - 1)^k.$$

или

$$\frac{f(\theta)}{\theta} = \sum_{k=0}^{+\infty} g(k+1) (-1)^k (\theta - 1)^k.$$

Далее будем рассматривать разложение в ряд Тейлора функции $\frac{f(\theta)}{\theta}$ около точки 1. Чтобы ряды сходились к одной и той же функции, их коэффициенты обязаны быть равными. Таким образом, сопоставляя коэффициенты рядов, мы вычислим функцию g .

Для $f(\theta) = \theta$ имеем $\frac{f(\theta)}{\theta} = 1 = 1 + \sum_{k=1}^{+\infty} 0 \cdot (\theta - 1)^k$. Тогда $g(1) = 1$ и $g(k) = 0$ для $k \neq 1$. Имеем $g(k) = I(k = 1)$ и $g(X_1) = I(X_1 = 1)$.

Для $f(\theta) = \frac{1-\theta}{\theta}$ имеем

$$\frac{f(\theta)}{\theta} = \frac{1-\theta}{\theta^2} = \frac{1}{\theta^2} - \frac{1}{\theta} = \sum_{k=0}^{+\infty} \frac{(-1)^k (k+1)!}{k!} (\theta - 1)^k - \sum_{k=0}^{+\infty} \frac{(-1)^k k!}{k!} (\theta - 1)^k = \sum_{k=0}^{+\infty} (-1)^k (k+1-1) (\theta - 1)^k.$$

Отсюда $g(k+1)(-1)^k = (-1)^k k$ и $g(k) = k - 1$.

Для $f(\theta) = \theta^2$ имеем

$$\frac{f(\theta)}{\theta} = \theta = 1 + (\theta - 1).$$

Отсюда $g(1)(-1)^0 = 1$, $g(2)(-1)^1 = 1$ и $g(k) = 0$ при $k \neq 1, 2$. Поэтому

$$g(k) = I(k = 1) - I(k = 2).$$

Задача 3б

Рассмотрим функцию $g(X_1, X_2) = I(X_1 = 1) \cdot I(X_2 = 1)$. Поскольку X_1 и X_2 независимы, то независимы $I(X_1 = 1)$ и $I(X_2 = 1)$. Поэтому

$$E_\theta g(X_1, X_2) = E_\theta I(X_1 = 1) \cdot E_\theta I(X_2 = 1) = \theta \cdot \theta = \theta^2.$$

Задача 3в

Предположим, что g — искомая оценка. Тогда, как было показано ранее, выполнено равенство

$$\frac{1}{1-\theta} = \sum_{k=1}^{+\infty} g(k) \theta (1-\theta)^k.$$

Отсюда

$$\frac{1}{\theta} = \sum_{k=1}^{+\infty} g(k) (-1)^k (\theta - 1)^k.$$

Разложим $\frac{1}{\theta}$ в ряд Тейлора около точки 1:

$$\frac{1}{\theta} = \frac{1}{1+(\theta-1)} = \sum_{k=0}^{+\infty} (-1)^k (\theta - 1)^k.$$

В разложении в ряд присутствует ненулевой свободный член, тогда как в сумме $\sum_{k=1}^{+\infty} g(k) (-1)^k (\theta - 1)^k$ он равен 0, поэтому эти ряды не могут сходиться к одной и той же функции.

Задача 4а

Пусть $X \sim R[0, \theta]$ — случайная величины. Вычислим её моменты:

$$E_\theta \xi^k = \int_0^\theta \frac{x^k}{\theta - x} dx = \frac{1}{\theta} \cdot \frac{\theta^{k+1} - 0^{k+1}}{k+1} = \frac{1}{k+1} \cdot \frac{\theta^{k+1}}{\theta} = \frac{\theta^k}{k+1}.$$

Из вычисления выше следует, что оценка $(k+1)X^k$ несмещена для θ^k .

Пусть теперь $X_1, \dots, X_n \sim R[0, \theta]$ — независимые случайные величины. Тогда

$$E_\theta (k+1) \overline{X^k} = \frac{k+1}{n} \cdot n \cdot E_\theta X_1^k = (k+1) \cdot \frac{\theta^k}{k+1} = \theta^k.$$

Вычислим среднеквадратичный риск:

$$\begin{aligned} D_{\theta}(k+1)\overline{X^k} &= (k+1)^2 \cdot \frac{1}{n^2} \cdot \sum_{m=1}^n D X_m^k = (k+1)^2 \cdot \frac{n}{n^2} D X_1^k = \\ &= \frac{(k+1)^2}{n} \cdot (E X_1^{2k} - (E X_1^k)^2) = \frac{(k+1)^2}{n} \cdot \left(\frac{\theta^{2k}}{2k+1} - \frac{\theta^{2k}}{(k+1)^2} \right) = \frac{\theta^{2k}}{n} \cdot \frac{k^2}{2k+1}. \end{aligned}$$

Задача 4б

Из определения функции распределения, независимости и одинаковой распределённости имеем:

$$\begin{aligned} F_{X_{(n)},\theta}(x) &= P_{\theta}(\max(X_1, \dots, X_n) \leq x) = P_{\theta}(X_1 \leq x \wedge \dots \wedge X_n \leq x) = \\ &= P_{\theta}(X_1 \leq x) \cdot \dots \cdot P_{\theta}(X_n \leq x) = F_{X_1,\theta}(x) \cdot \dots \cdot F_{X_n,\theta}(x) = \left(\frac{x}{\theta}\right)^n \cdot I(x \in [0, \theta]) + I(x \in (\theta, +\infty)). \end{aligned}$$

Поскольку в точках непрерывности функции плотности, она совпадает с производной от функции распределения, то $f_{X_{(n)},\theta}(x) = \frac{nx^{n-1}}{\theta^n} \cdot I(x \in [0, \theta])$.

Задача 4в

Найдём несмещённую оценку для $\frac{1}{\theta^3}$ от $X_{(n)}$ (здесь $n > 3$). Предположим, что g — искомая оценка. Тогда мы имеем равенство

$$\frac{1}{\theta^3} = E_{\theta} g(X_{(n)}) = \int_{-\infty}^{+\infty} \frac{g(t) \cdot nt^{n-1}}{\theta^n} \cdot I(t \in [0, \theta]) dt = \int_0^{\theta} \frac{g(t) \cdot nt^{n-1}}{\theta^n} dt.$$

Отсюда

$$\theta^{n-3} = \int_0^{\theta} g(t) \cdot nt^{n-1} dt.$$

Дифференцируя по θ , получаем

$$(n-3)\theta^{n-4} = ng(\theta)\theta^{n-1}$$

и, наконец,

$$g(\theta) = \frac{n-3}{n\theta^3}, g(X_{(n)}) = \frac{n-3}{nX_{(n)}^3}.$$

задача 4г

Предположим, что существует оценка g такая, что $E_{\theta} g(X_1) = \frac{1}{\theta^3}$. Тогда выполнено равенство

$$\frac{1}{\theta^3} = \int_0^{\theta} \frac{g(t)}{\theta} dt.$$

Следовательно,

$$\frac{1}{\theta^2} = \int_0^{\theta} g(t) dt$$

и после дифференцирования по θ получаем

$$\frac{-2}{\theta^3} = g(\theta).$$

Однако, интеграл $\int_0^{\theta} \frac{-2}{t^3} dt$ расходится, что противоречит предположению о существовании такой оценки g .

