

Оглавление

Занятие 1. Выборочные характеристики	1
Контрольные вопросы и задания	3
Аудиторные задачи	4
Домашнее задание	11
Занятие 2. Свойства оценок	25
Контрольные вопросы и задания	27
Аудиторные задачи	28
Домашнее задание	35
Занятие 3. Метод моментов построения оценок	47
Контрольные вопросы и задания	49
Аудиторные задачи	50
Домашнее задание	54
Занятие 4. Оценка максимального правдоподобия	62
Контрольные вопросы и задания	63
Аудиторные задачи	63
Домашнее задание	72
Занятие 5. Достаточные статистики	84
Контрольные вопросы и задания	85
Аудиторные задачи	85
Домашнее задание	90
Занятие 6. Теорема Колмогорова про улучшение оценок. Неравенство Рао-Крамера	95
Контрольные вопросы и задания	97
Аудиторные задачи	98
Домашнее задание	105
Занятие 7. Гауссовы системы	107
Контрольные вопросы и задания	109
Аудиторные задачи	110
Домашнее задание	112

Занятие 1. Выборочные характеристики

Контрольные вопросы и задания

Приведите определение выборки, вариационного ряда, статистики, порядковой статистики, эмпирической функции распределения.

x_1, \dots, x_n — наблюдаемые значения — независимые одинаково распределённые случайные величины с неизвестной функцией распределения $F(x)$.

Такой набор случайных величин называется выборкой из распределения F .

Вариационный ряд — последовательность $x_{(1)}, \dots, x_{(n)}$, полученная в результате расположения в порядке неубывания исходной последовательности независимых одинаково распределённых случайных величин x_1, \dots, x_n .

Статистикой называют функцию S от выборки $X = (x_1, x_2, \dots, x_n)$ такую, что $S(X) = S(x_1, x_2, \dots, x_n)$.

Вариационный ряд и его члены являются порядковыми статистиками.

Эмпирической (выборочной) функцией распределения, построенной по выборке x_1, \dots, x_n называется функция

$$F_n(x) = \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{x_k \leq x}, \quad x \in \mathbb{R}.$$

Какими свойствами обладает эмпирическая функция распределения?

Есть множество полной вероятности, на котором эмпирическая функция распределения аппроксимирует функцию распределения, то есть почти наверное $F_n \Rightarrow F, n \rightarrow \infty$.

Запишите выражения для выборочного среднего, выборочной дисперсии, выборочных моментов.

$$\frac{1}{n} \sum_{k=1}^n x_k$$

— выборочное среднее.

Выборочная дисперсия

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2.$$

Выборочные моменты в математической статистике — это оценка теоретических моментов распределения на основе выборки.

Выборочный момент порядка k — это случайная величина

$$a_n(k) = \frac{1}{n} \sum_{i=1}^n x_i^k.$$

Аудиторные задачи

1.4

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке $[0, \theta]$ с неизвестным параметром θ . Какие из приведённых ниже функций являются статистиками?

- a) \bar{X} ;
- b) $5X_{(n)}$;
- c) $\theta/2$;
- d) X_1/θ ;
- e) $X_{(1)} + X_1 + X_n$.

Решение.

- a) Да;
- b) да;
- c) нет, так как не функция от выборки;
- d) функция не только от выборки (зависит от неизвестного параметра). Отсюда следует, что это не статистика;
- e) да.

1.5

Задание. Пусть X_1, \dots, X_n — выборка из распределения Пуассона с параметром λ . Вычислите математическое ожидание и дисперсию статистики

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

Выясните, имеет ли статистика \bar{X} распределение Пуассона.

Решение. Все X_i одинаково распределены. Отсюда следует, что все математические ожидания одинаковы

$$M\bar{X} = \frac{1}{n} \sum_{i=1}^n MX_i = \frac{1}{n} \cdot nMX_1 = MX_1 = \lambda.$$

Для всякой выборки справедливо $M\bar{X} = MX_1$.

Из независимости X_i получаем

$$D\bar{X} = \frac{1}{n^2} \sum_{i=1}^n DX_i.$$

Так как X_i одинаково распределены, то все дисперсии одинаковы

$$D\bar{X} = \frac{DX_1}{n} = \frac{\lambda}{n}.$$

Математическое ожидание и дисперсия для распределения Пуассона совпадают. Отсюда следует, что эта случайная величина не имеет распределения Пуассона.

\bar{X} не обязательно будет принимать целые значения.

1.6

Задание. Вычислите математическое ожидание статистик:

a) $S^2 = \overline{X^2} - (\bar{X})^2$;

b) $S_0^2 = 1/(n-1) \cdot \sum_{i=1}^n (X_i - \bar{X})^2$.

Решение.

a) Распишем каждую из величин

$$S^2 = \overline{X^2} - (\bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2.$$

Распишем квадрат

$$\frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \frac{1}{n^2} \left(\sum_{i=1}^n X_i^2 + 2 \sum_{i,j=1, i < j}^n X_i X_j \right).$$

Берём слева и справа математическое ожидание. Из того, что случайные величины в выборке одинаково распределены

$$MS^2 = MX_1^2 - \frac{1}{n^2} [nMX_1 + 2C_n^2 (MX_1)^2].$$

Подставляем C_n^2 и группируем

$$MX_1^2 - \frac{1}{n^2} [nMX_1 + 2C_n^2 (MX_1)^2] = \frac{n-1}{n} \cdot MX_1^2 - \frac{n-1}{n} (MX_1)^2.$$

Вынесем общий множитель за скобки

$$\frac{n-1}{n} \cdot MX_1^2 - \frac{n-1}{n} (MX_1)^2 = \frac{n-1}{n} [MX_1^2 - (MX_1)^2] = \frac{n-1}{n} \cdot SX_1.$$

Эта оценка смещена асимптотически;

b) выразим S_0 через S . Раскроем квадрат

$$S_0^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i^2 - 2X_i\bar{X} + \bar{X}^2).$$

Имеем сумму n одинаковых слагаемых

$$S_0^2 = \frac{1}{n-1} (n\bar{X}^2 - 2(\bar{X})^2 n + n\bar{X}^2) = \frac{n}{n-1} [\bar{X}^2 - (\bar{X})^2] = \frac{n-1}{n} \cdot S^2.$$

Отсюда следует, что

$$MS_0^2 = \frac{n}{n-1} \cdot \frac{n-1}{n} \cdot DX_1 = DX_1.$$

1.7

Задание. Найдите в терминах функции распределения F выборки X_1, \dots, X_n :

- a) распределение k -ой порядковой статистики $X_{(k)}$;
- b) вероятность $P(X_{(k)} < y, X_{(k+1)} \geq y)$.

Решение.

- a) Сделали упорядочивание случайных величин

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(k)} \leq \dots \leq X_{(n)}.$$

По определению $F_{X_{(k)}}(y) = P(X_{(k)} \leq y) = P\{\text{хотя бы } k \text{ элементов выборки не превышает } y\} =$

$$= \sum_{i=k}^n P(A_i),$$

где $A_i = \{\text{ровно } i \text{ элементов выборки не превышают } y\}$.

Есть n испытаний, успех — $X_i \leq y$.

Вероятность успеха — это $F(y)$, вероятность неудачи — это $[1 - F(y)]$.

Это биномиальное распределение

$$F_{X_{(k)}} = \sum_{i=k}^n C_n^i F^i(y) [1 - F(y)]^{n-i};$$

б) согласно с предыдущим пунктом $P(X_{(k)} < y, X_{(k+1)} \geq y) = P\{\text{ровно } i \text{ элементов выборки не превышает } y\} = C_n^k F^k(y) [1 - F(y)]^{n-k}$.

1.8

Задание. Пусть $(-0.8; 2.9; 4.5; -5.7; 1.1; -3.2)$ — наблюдаемые значения выборки. Составьте вариационный ряд, постройте эмпирическую функцию распределения $F_6(x)$ и её график. Вычислите выборочное среднее и выборочную дисперсию.

Решение. Вариационный ряд: $(-5.7; -3.2; -0.8; 1.1; 2.9; 4.3)$.

Эмпирическая функция распределения (рис. 1).

$$F_6(y) = \frac{1}{6} \sum_{i=1}^6 \mathbb{1}\{x_i \leq y\} = \begin{cases} 0, & x < -5.7; \\ \frac{1}{6}, & -5.7 \leq x < -3.2, \\ \frac{2}{6}, & -3.2 \leq x < -0.8, \\ \frac{3}{6}, & -0.8 \leq x < 1.1, \\ \frac{4}{6}, & 1.1 \leq x < 2.9, \\ \frac{5}{6}, & 2.9 \leq x < 4.3, \\ 1, & x \geq 4.3. \end{cases}$$

Выборочное среднее

$$\bar{X} = \frac{1}{6} (-5.7 - 3.2 - 0.8 + 1.1 + 2.9 + 4.3) = \frac{1}{6} (-9.7 + 8.3) = -\frac{1}{6} \cdot 1.4 = -0.23.$$

Выборочная дисперсия

$$\hat{\sigma}^2 = \frac{1}{5} \sum_{i=1}^6 (X_i + 0.23)^2.$$

Она является несмещённой.



Рис. 1: Эмпирическая функция распределения

1.9

Задание. Вычислите вероятность $P(F_n(y) < F_n(z))$.

Решение.

- а) $y \geq z$. Событие невозможное, потому что $F_n(y) \geq F_n(z)$;
 б) рассмотрим случай, когда $y < z$.

Тогда искомая вероятность равна $P(F_n(y) < F_n(z)) = P\{\text{в } (y, z) \text{ попал хотя бы 1 элемент выборки}\} = 1 - P\{\text{в } (y, z) \text{ ни один элемент выборки не попал}\}$. Случайные величины одинаково распределены, поэтому $1 - P\{\text{в } (y, z) \text{ ни один элемент выборки не попал}\} =$
 $= [1 - P\{x_i \notin (y, z)\}]^n = 1 - [1 - P\{x_1 \in (y, z)\}]^n =$
 $= 1 - [1 - P(x_1 < z) + P(x_1 < y)]^n = 1 - [1 - F(z) + F(y)]^n.$

1.10

Задание. Пусть X_1, \dots, X_n — выборка из распределения F с плотностью f . Найдите совместную плотность распределения всех порядковых статистик, то есть плотность распределения случайного вектора $(X_{(1)}, \dots, X_{(n)})$.

Решение. $F_{(X_{(1)}, X_{(2)})}(y_1, y_2) = P(X_{(1)} \leq y_1, X_{(2)} \leq y_2)$. Воспользуемся формулой $P(A \cap B) = P(B) - P(\bar{A} \cap B)$. Получим

$$P(X_{(1)} \leq y_1, X_{(2)} \leq y_2) = P(X_{(1)} \leq y_2) - P(X_{(1)} > y_1, X_{(2)} \leq y_2).$$

Среди $X_{(1)}$ и $X_{(2)}$ случайная величина $X_{(2)}$ является максимальной.

$$\begin{aligned} & P(X_{(1)} \leq y_2) - P(X_{(1)} > y_1, X_{(2)} \leq y_2) = \\ & = P(X_1 \leq y_2, X_2 \leq y_2) - P(X_1 \in (y_1, y_2], X_2 \in (y_1, y_2]). \end{aligned}$$

Случайные величины X_1, X_2 — независимые и одинаково распределённые

$$\begin{aligned} & P(X_1 \leq y_2, X_2 \leq y_2) - P(X_1 \in (y_1, y_2], X_2 \in (y_1, y_2]) = \\ & = \begin{cases} [F(y_2)]^2, & y_1 \geq y_2, \\ [F(y_2)]^2 - [F(y_2) - F(y_1)]^2, & y_1 < y_2. \end{cases} \end{aligned}$$

Продифференцируем

$$f_{(X_{(1)}, X_{(2)})}(y_1, y_2) = \begin{cases} 0, & y_1 \geq y_2, \\ 2f(y_1)f(y_2), & y_1 < y_2. \end{cases}$$

Рассматриваем множество всех векторов, которые имеют упорядоченные координаты $\Delta = \{\vec{x} \in \mathbb{R}^n : z_1 < z_2 < \dots < z_n\}$, $\Gamma \subseteq \Delta$ — произвольное подмножество.

$$(X_{(1)}, \dots, X_{(n)}) \in \Delta.$$

Чтобы найти вероятность того, что данный вектор принадлежит Γ , должны проинтегрировать плотность этого вектора по этому множеству

$$P\{(X_{(1)}, \dots, X_{(n)}) \in \Gamma\} = \int_{\Gamma} f_{(X_{(1)}, \dots, X_{(n)})}(z_1, \dots, z_n) dz_1 \dots dz_n.$$

С другой стороны,

$$P\{(X_{(1)}, \dots, X_{(n)}) \in \Gamma\} = \sum_{\sigma \in S_n} P\{(X_{\sigma(1)}, \dots, X_{\sigma(n)}) \in \Gamma\}.$$

Учтём все перестановки

$$\sum_{\sigma \in S_n} P\{(X_{\sigma(1)}, \dots, X_{\sigma(n)}) \in \Gamma\} = n! P\{(X_1, \dots, X_n) \in \Gamma\}.$$

Подставим найденное выражение для вероятности

$$n! P\{(X_1, \dots, X_n) \in \Gamma\} = n! \cdot \int_{\Gamma} f(z_1) \cdot \dots \cdot f(z_n) dz_1 \dots dz_n.$$

Сравниваем полученные выражения

$$f_{(X_{(1)}, \dots, X_{(n)})}(z_1, \dots, z_n) = n! f(z_1) \cdot \dots \cdot f(z_n) \cdot \mathbb{1}\{z_1 < z_2 < \dots < z_n\}$$

— плотность вектора упорядоченных статистик.

1.11

Задание. Пусть задана выборка X_1, \dots, X_n из показательного распределения с параметром α .

- Докажите, что случайные величины $X_{(1)}, X_{(2)} - X_{(1)}, \dots, X_{(n)} - X_{(n-1)}$ являются независимыми;
- найдите распределение разности $X_{(k+1)} - X_{(k)}$ соседних порядковых статистик.

Решение. $\vec{\xi} = (\xi_1, \dots, \xi_n)$ — случайный вектор с плотностью распределения $f_{\vec{\xi}}(\vec{x})$.

Линейное преобразование этого вектора $\vec{\eta} = A\vec{\xi}$, где A — некоторая n -мерная матрица.

$$f_{A\vec{\xi}}(\vec{y}) = \frac{1}{|\det A|} \cdot f_{\vec{\xi}}(A^{-1}\vec{y}).$$

Составим вектор из величин $X_{(1)}, X_{(2)} - X_{(1)}, \dots, X_{(n)} - X_{(n-1)}$. Его плотность должна распадаться на произведение плотностей компонент.

Из задачи 1.10

$$f_{(X_{(1)}, \dots, X_{(n)})}(y_1, \dots, y_n) = n! f(y_1) \cdot \dots \cdot f(y_n) \cdot \mathbb{1}\{y_1 < y_2 < \dots < y_n\}.$$

Подставим плотность показательного распределения

$$\begin{aligned} & n! f(y_1) \cdot \dots \cdot f(y_n) \cdot \mathbb{1}\{y_1 < y_2 < \dots < y_n\} = \\ & = n! \alpha e^{-\alpha y_1} \cdot \mathbb{1}\{y_1 > 0\} \cdot \dots \cdot \alpha e^{-\alpha y_n} \cdot \mathbb{1}\{y_n > 0\} \cdot \mathbb{1}\{y_1 < y_2 < \dots < y_n\}. \end{aligned}$$

Перемножим

$$\begin{aligned} & n! \alpha e^{-\alpha y_1} \cdot \mathbb{1}\{y_1 > 0\} \cdot \dots \cdot \alpha e^{-\alpha y_n} \cdot \mathbb{1}\{y_n > 0\} \cdot \mathbb{1}\{y_1 < y_2 < \dots < y_n\} = \\ & = n! \alpha^n e^{-\alpha(y_1 + \dots + y_n)} \cdot \mathbb{1}\{0 < y_1 < y_2 < \dots < y_n\}. \end{aligned}$$

Нужно найти линейное преобразование

$$\begin{bmatrix} X_{(1)} \\ X_{(2)} - X_{(1)} \\ X_{(3)} - X_{(2)} \\ \dots \\ X_{(n)} - X_{(n-1)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \dots & & & & & & & \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} X_{(1)} \\ X_{(2)} \\ X_{(3)} \\ \dots \\ X_{(n)} \end{bmatrix},$$

где

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \dots & & & & & & & \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix} = A.$$

Определитель $\det A = 1$.

Ищем обратную матрицу

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & 0 \\ \dots & & & & & \\ 1 & 1 & 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} X_{(1)} \\ X_{(2)} - X_{(1)} \\ X_{(3)} - X_{(2)} \\ \dots \\ X_{(n)} - X_{(n-1)} \end{bmatrix} = \begin{bmatrix} X_{(1)} \\ X_{(2)} \\ X_{(3)} \\ \dots \\ X_{(n)} \end{bmatrix}.$$

Тогда имеем выражение

$$A^{-1}\vec{y} = \begin{bmatrix} y_1 \\ y_1 + y_2 \\ \dots \\ \sum_{i=1}^n y_i \end{bmatrix}.$$

Определим искомый вектор через

$$\vec{\eta} = (X_{(1)}, X_{(2)} - X_{(1)}, \dots, X_{(n)} - X_{(n-1)}).$$

Тогда

$$\begin{aligned} & f_{\vec{\eta}}(y_1, \dots, y_n) = \\ & = n! \alpha^n e^{-\alpha(ny_1 + (n-1)y_2 + \dots + y_n)} \cdot \mathbb{1}\{0 < y_1 < y_1 + y_2 < \dots < y_1 + y_2 + \dots + y_n\}. \end{aligned}$$

Разобьём на n множителей

$$\begin{aligned} & n! \alpha^n e^{-\alpha(ny_1 + (n-1)y_2 + \dots + y_n)} \cdot \mathbb{1}\{0 < y_1 < y_1 + y_2 < \dots < y_1 + y_2 + \dots + y_n\} = \\ & = [n\alpha e^{-\alpha ny_1} \cdot \mathbb{1}\{0 < y_1\}] \cdot [(n-1)\alpha e^{-\alpha(n-1)y_2} \cdot \mathbb{1}\{y_2 > 0\}] \cdot \dots \times \\ & \quad \times [\alpha e^{-\alpha y_n} \cdot \mathbb{1}\{y_n > 0\}]. \end{aligned}$$

Имеем произведение плотностей компонент, значит, элементы вектора независимы и показательно распределены с параметром $\alpha(n-k)$, то есть $X_{(k+1)} - X_{(k)} \sim \Pi(\alpha(n-k))$. Считаем, что $X_{(0)} = 0$.

Домашнее задание

1.15

Задание. Пусть X_1, \dots, X_n — выборка из равномерного на отрезке $[a, b]$ распределения. Вычислите математическое ожидание и дисперсию статистики

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

Выясните, имеет ли статистика \bar{X} равномерное распределение; нормальное распределение.

Решение. Все X_i одинаково распределены. Отсюда следует, что все математические ожидания одинаковы

$$M\bar{X} = \frac{1}{n} \sum_{i=1}^n MX_i = \frac{1}{n} \cdot nMX_1 = MX_1 = \frac{a+b}{2}.$$

Из независимости X_i получаем

$$D\bar{X} = \frac{1}{n^2} \sum_{i=1}^n DX_i.$$

Так как X_i одинаково распределены, то все дисперсии одинаковы

$$D\bar{X} = \frac{DX_1}{n} = \frac{(b-a)^2}{12n}.$$

Чтобы выяснить, распределена ли статистика \bar{X} по нормальному или равномерному распределению, найдём её характеристическую функцию $\varphi_{\bar{X}}(t)$. Учитывая независимость элементов выборки и то, что

$$\varphi_{X_1}(t) = \dots = \varphi_{X_n}(t) = \frac{e^{itb} - e^{ita}}{it(b-a)},$$

находим

$$\varphi_{\bar{X}}(t) = \varphi_{X_1}\left(\frac{t}{n}\right) \cdot \dots \cdot \varphi_{X_n}\left(\frac{t}{n}\right) = \left[\frac{(e^{itb} - e^{ita})n}{it(b-a)} \right]^n.$$

Отсюда следует, что \bar{X} не имеет указанных распределений.

1.16

Задание. Пусть X_1, \dots, X_n — выборка из некоторого распределения вероятностей, функция распределения которого F является непрерывной и строго возрастающей. Найдите распределение выборки Y_1, \dots, Y_n , где

$$Y_i = F(X_i).$$

Решение. По определению

$$F_{\eta_1, \dots, \eta_n}(X_1, \dots, X_n) = P(\eta_1 \leq X_1, \dots, \eta_n \leq X_n).$$

Воспользуемся независимостью

$$P(\eta_1 \leq X_1, \dots, \eta_n \leq X_n) = P(\eta_1 \leq X_1) \cdot \dots \cdot P(\eta_n \leq X_n).$$

Функция распределения i -й компоненты вектора равна

$$F_{\eta_i}(x) = P(F_{\xi_i}(X_i) \leq x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 1. \end{cases}$$

Рассмотрим $[0, 1]$.

Поскольку F — непрерывная и строго возрастающая, то существует $F^{-1}(x)$. Обозначим через z точку $F^{-1}(x)$ такую, что $F(z) = x$. Событие $\{\eta = F(\xi) < x\}$ происходит тогда и только тогда, когда происходит событие $\{\xi < z\}$.

Получаем на отрезке $[0, 1]$ равномерное распределение

$$F_{\eta}(x) = F_{\xi}(z) = F_{\xi}(F_{\xi}^{-1}(x)) = \begin{cases} 0, & x \leq 0, \\ x, & x \in (0, 1], \\ 1, & x > 1. \end{cases}$$

1.17

Задание. Пусть X_1, \dots, X_n — выборка из дискретного распределения с вероятностями $P(X_1 = m) = p_m$, где

$$\sum_{m=0}^N p_m = 1.$$

Найдите распределение k -й порядковой статистики $X_{(k)}$.

Решение. Сделали упорядочивание случайных величин

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(k)} \leq \dots \leq X_{(n)}.$$

По определению $F_{X_{(k)}}(y) = P(X_{(k)} \leq y) = P\{\text{хотя бы } k \text{ элементов выборки не превышает } y\} =$

$$= \sum_{i=k}^n P(A_i),$$

где $A_i = \{\text{ровно } i \text{ элементов выборки не превышают } y\}$.

Есть n испытаний, успех — $X_i \leq y$.

Вероятность успеха — это $F(y)$, вероятность неудачи — это $[1 - F(y)]$. Это биномиальное распределение

$$F_{X_{(k)}} = \sum_{i=k}^n C_n^i F^i(y) [1 - F(y)]^{n-i}.$$

Представим $F_{X_i}(y) = F(y)$ через m . Запишем по определению

$$F_{X_1}(y) = P(X_1 \leq y) = \sum_{m=1}^n P(X_1 = m) = \sum_{m=1}^n p_m.$$

Подставим полученное выражение в функцию распределения

$$F_{X_{(k)}} = \sum_{i=k}^n C_n^i \sum_{m=1}^n p_m \left(1 - \sum_{m=1}^n p_m\right)^{n-i}.$$

1.18

Задание. Пусть $(3, 0, 4, 3, 6, 0, 3, 1)$ — наблюдаемые значения выборки. Составьте вариационный ряд, постройте эмпирическую функцию распределения $F_8(x)$ и её график. Вычислите выборочное среднее и выборочную дисперсию.

Решение. Вариационный ряд: $(0, 0, 1, 3, 3, 3, 4, 6)$.

Эмпирическая функция распределение (рис. 2)

$$F_8(y) = \frac{1}{8} \sum_{i=1}^8 \mathbb{1}\{x_i \leq y\} = \begin{cases} 0, & x < 0, \\ \frac{2}{8} = \frac{1}{4}, & 0 \leq x < 1, \\ \frac{3}{8}, & 1 \leq x < 3, \\ \frac{6}{8} = \frac{3}{4}, & 3 \leq x < 4, \\ \frac{7}{8}, & 4 \leq x < 6, \\ 1, & x \geq 6. \end{cases}$$



Рис. 2: Эмпирическая функция распределения

Выборочное среднее

$$\bar{X} = \frac{1}{8} (0 + 0 + 1 + 3 + 3 + 4 + 6) = \frac{1}{8} \cdot 20 = \frac{10}{4} = 2.5.$$

Выборочная дисперсия

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{7} \sum_{i=1}^8 (X_i - 2.5)^2 = \\ &= \frac{1}{7} \left[2(0 - 2.5)^2 + (1 - 2.5)^2 + 3(3 - 2.5)^2 + (4 - 2.5)^2 + (6 - 2.5)^2 \right] = \\ &= \frac{1}{7} (12.5 + 2.25 + 0.75 + 2.25 + 12.25) = \frac{30}{7} \approx 4.29. \end{aligned}$$

1.19

Задание. По выборке объёма n из распределения Бернулли с параметром p постройте эмпирическую функцию распределения $F_n(y)$.

Решение. Случайная величина имеет распределение Бернулли, если она принимает всего два значения: 1 и 0 с вероятностями p и $(1 - p)$ соответственно. Таким образом: $P(x = 1) = p$, $P(x = 0) = 1 - p$.

Эмпирической функцией распределения, построенной по выборке

$$x_1, \dots, x_n,$$

называется функция

$$F_n(y) = \frac{1}{n} \sum_{k=1}^n \mathbb{1}(x_k \leq y).$$

Пусть есть набор из n чисел (нулей и единиц) — выборка из распределения Бернулли.

Для удобства выстроим числа в порядке их возрастания:

$$0, 0, \dots, 0, 1, 1, \dots, 1.$$

Видим, что слева от нуля эмпирическая функция распределения будет равна нулю.

В точке 0 произойдёт скачок на

$$\frac{n-k}{n},$$

где k — количество единиц, а $(n-k)$ — количество нулей в выборке.

В точке 1 будет скачок на

$$1 - \frac{n-k}{n} = \frac{n-n+k}{n} = \frac{k}{n},$$

а значение самой функции будет равно единице.

$$F_n(y) = \begin{cases} 0, & x < 0, \\ \frac{n-k}{n}, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases}$$

Эмпирическая функция распределения будет выглядеть так, как показано на рис. 3.



Рис. 3: Эмпирическая функция распределения

1.20

Задание. Пусть X_1, \dots, X_n — выборка из распределения F . Докажите, что для произвольных $y \in \mathbb{R}$ и $k \in \{0, 1, \dots, n\}$ справедливо равенство

$$P\left(F_n(y) = \frac{k}{n}\right) = C_n^k F^k(y) (1 - F(y))^{n-k}.$$

Решение. Эмпирической функцией распределения, построенной по выборке X_1, \dots, X_n , называется функция

$$F_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq y).$$

Посмотрим, что значит событие в указанной вероятности

$$F_n(y) = \frac{k}{n} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq y).$$

Сократим константы в знаменателях

$$k = \sum_{i=1}^n \mathbb{1}(X_i \leq y).$$

Это означает, что есть ровно k элементов выборки, не превышающих y . Следовательно, это биномиальное распределение с параметром $F(y)$, то есть

$$P\left(F_n(y) = \frac{k}{n}\right) = C_n^k F^k(y) (1 - F(y))^{n-k}.$$

1.21

Задание. Для выборки X_1, \dots, X_n из равномерного распределения на отрезке $[0, \theta]$ найдите плотность, математическое ожидание и дисперсию:

- а) максимального члена вариационного ряда $X_{(n)}$;
- б) минимального члена вариационного ряда $X_{(1)}$;
- в) совместную плотность распределения и ковариацию $X_{(n)}$ и $X_{(1)}$.

Решение.

$$\begin{aligned} \text{а) Найдем функцию распределения } n\text{-й порядковой статистики } F_{X_{(n)}}(y) &= \\ &= P(X_{(n)} \leq y) = P(n \text{ элементов выборки не превышают } y) = \\ &= C_n^n [F_{X_1}(y)]^n \cdot [1 - F_{X_1}(y)]^{n-n} = [F_{X_1}(y)]^n = [P(X_1 \leq y)]^n = \left(\frac{y}{\theta}\right)^n, \end{aligned}$$

при этом $y \in [0, \theta]$.

Продифференцируем

$$f_{X_{(n)}}(y) = \frac{\partial F_{X_{(n)}}(y)}{\partial y} = \frac{\partial}{\partial y} \left(\frac{y}{\theta}\right)^n = \frac{1}{\theta^n} \cdot \frac{dy^n}{dy} = \frac{ny^{n-1}}{\theta^n} \cdot \mathbb{1}\{y \in [0, \theta]\}.$$

По определению математического ожидания

$$MX_{(n)} = \int_0^\theta y dF^n(y) = \int_0^\theta y d\left(\frac{y}{\theta}\right)^n = \frac{1}{\theta^n} \int_0^\theta y dy^n = \frac{n}{\theta^n} \int_0^\theta y y^{n-1} dy.$$

Сложим степени сомножителей

$$\frac{n}{\theta^n} \int_0^\theta y y^{n-1} dy = \frac{n}{\theta^n} \int_0^\theta y^n dy = \frac{n}{\theta^n} \cdot \frac{y^{n+1}}{n+1} \Big|_0^\theta = \frac{\theta^{n+1} n}{\theta^n (n+1)} = \frac{\theta n}{n+1}.$$

Найдём второй момент

$$MX_{(n)}^2 = \int_0^\theta y^2 dF^n(y) = \frac{n}{\theta^n} \int_0^\theta y^2 y^{n-1} dy = \frac{n}{\theta^n} \int_0^\theta y^{n+1} dy.$$

Возьмём интеграл

$$\frac{n}{\theta^n} \int_0^\theta y^{n+1} dy = \frac{ny^{n+2}}{\theta^n (n+2)} \Big|_0^\theta = \frac{n\theta^{n+2}}{\theta^n (n+2)} = \frac{n\theta^2}{n+2}.$$

По свойствам дисперсии

$$DX_{(n)} = MX_{(n)}^2 - (MX_{(n)})^2 = \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2}.$$

Приведём к общему знаменателю

$$\begin{aligned} \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} &= \frac{n\theta^2 (n^2 + 2n + 1) - n^2\theta^2 (n+2)}{(n+2)(n+1)^2} = \\ &= \frac{n^3\theta^2 + 2n^2\theta^2 + n\theta^2 - n^3\theta^2 - 2n^2\theta^2}{(n+2)(n+1)^2} = \frac{n\theta^2}{(n+2)(n+1)^2}; \end{aligned}$$

b) найдём функцию распределения первой порядковой статистики

$$\begin{aligned} F_{X_{(1)}}(y) &= P(X_{(1)} \leq y) = P(\min(X_1, X_2, \dots, X_n) \leq y) = \\ &= P(\text{хотя бы 1 элемент выборки не превышает } y) = \end{aligned}$$

$$\begin{aligned} \sum_{k=1}^n P(\text{exactly } k \text{ elements of the sample does not exceed } y) &= \\ &= \sum_{k=1}^n C_n^k F^k(y) [1 - F(y)]^{n-k} = \\ &= \sum_{k=0}^n C_n^k F^k(y) [1 - F(y)]^{n-k} - C_n^0 F^0(y) [1 - F(y)]^{n-0}. \end{aligned}$$

Применяем формулу для бинома Ньютона

$$\begin{aligned} \sum_{k=0}^n C_n^k F^k(y) [1 - F(y)]^{n-k} - C_n^0 F^0(y) [1 - F(y)]^{n-0} = \\ = [F(y) + 1 - F(y)]^n - [1 - F(y)]^n = 1 - \left(1 - \frac{y}{\theta}\right)^n. \end{aligned}$$

Продифференцируем

$$f_{X_{(1)}}(y) = \frac{\partial F_{X_{(1)}}(y)}{\partial y} = \frac{\partial}{\partial y} \left(1 - \frac{(\theta - y)^n}{\theta^n}\right) = -\frac{1}{\theta^n} \cdot \frac{\partial}{\partial y} (\theta - y)^n.$$

Берём производную сложной функции

$$-\frac{1}{\theta^n} \cdot \frac{\partial}{\partial y} (\theta - y)^n - \frac{1}{\theta^n} \cdot n (\theta - y)^{n-1} (-1) = \frac{n (\theta - y)^{n-1}}{\theta^n}, \quad y \in [0, \theta].$$

Найдём математическое ожидание по определению

$$MX_{(1)} = \int_0^\theta y dF^n(y) = \int_0^\theta y d\left(1 - \frac{y}{\theta}\right)^n.$$

Сделаем замену

$$1 - \frac{y}{\theta} = z,$$

откуда $y = \theta(1 - z)$, при этом интегрирование происходит в пределах от одного до нуля

$$\int_0^\theta y d\left(1 - \frac{y}{\theta}\right)^n = \int_1^0 \theta(1 - z) dz^n = \theta n \int_1^0 (z - 1) z^{n-1} dz.$$

Разбиваем на 2 интеграла

$$\theta n \int_1^0 (z - 1) z^{n-1} dz = -\theta n \int_0^1 z^n dz + \theta n \int_0^1 z^{n-1} dz = -\theta n \cdot \frac{z^{n+1}}{n+1} \Big|_0^1 + \theta n \cdot \frac{z^n}{n} \Big|_0^1.$$

Подставляем пределы интегрирования

$$-\theta n \cdot \frac{z^{n+1}}{n+1} \Big|_0^1 + \theta n \cdot \frac{z^n}{n} \Big|_0^1 = -\theta n \cdot \frac{1}{n+1} + \theta n \cdot \frac{1}{n} = -\theta n \left(\frac{1}{n+1} - \frac{1}{n} \right).$$

Приводим к общему знаменателю

$$-\theta n \left(\frac{1}{n+1} - \frac{1}{n} \right) = -\theta n \cdot \frac{n - n - 1}{n(n+1)} = \frac{\theta n}{n(n+1)} = \frac{\theta}{(n+1)}.$$

Найдём второй момент

$$MX_{(1)}^2 = \int_0^\theta y^2 d\left(1 - \frac{y}{\theta}\right)^n.$$

Применяем такую же замену, как при поиске первого момента

$$\int_0^\theta y^2 d\left(1 - \frac{y}{\theta}\right)^n = \int_1^0 \theta^2 (1-z)^2 dz^n = \int_0^1 \theta^2 n (1-z)^2 z^{n-1} dz.$$

Выносим константу за знак интеграла и возводим скобку в квадрат

$$\begin{aligned} \int_0^1 \theta^2 n (1-z)^2 z^{n-1} dz &= \theta^2 n \int_0^1 (1-2z+z^2) z^{n-1} dz = \\ &= \theta^2 n \int_0^1 z^{n-1} dz - 2\theta^2 n \int_0^1 z^n dz + \theta^2 n \int_0^1 z^{n+1} dz = \\ &= \theta^2 n \cdot \frac{z^n}{n} \Big|_0^1 - 2\theta^2 n \cdot \frac{z^{n+1}}{n+1} \Big|_0^1 + \theta^2 n \cdot \frac{z^{n+2}}{n+2} \Big|_0^1 = \theta^2 - \frac{2\theta^2 n}{n+1} + \frac{\theta^2 n}{n+2} = \\ &= \theta^2 \left(1 - \frac{2n}{n+1} + \frac{n}{n+2}\right) = \\ &= \theta^2 \cdot \frac{(n+1)(n+2) - 2n(n+2) + n(n+1)}{(n+1)(n+2)} = \\ &= \frac{\theta^2 (n^2 + 3n + 2 - 2n^2 - 4n + n^2 + n)}{(n+1)(n+2)} = \frac{2\theta^2}{(n+1)(n+2)}. \end{aligned}$$

По свойствам дисперсии

$$DX_{(1)} = MX_{(1)}^2 - [MX_{(1)}]^2 = \frac{2\theta^2}{(n+1)(n+2)} - \frac{\theta^2}{(n+1)^2}.$$

Приводим к общему знаменателю

$$\begin{aligned} \frac{2\theta^2}{(n+1)(n+2)} - \frac{\theta^2}{(n+1)^2} &= \frac{2\theta^2(n+1) - \theta^2(n+2)}{(n+1)^2(n+2)} = \\ &= \frac{2\theta^2 n + 2\theta^2 - \theta^2 n - 2\theta^2}{(n+1)^2(n+2)} = \frac{\theta^2 n}{(n+1)^2(n+2)}; \end{aligned}$$

с) найдём функцию распределения вектора по определению

$$F_{(X_{(1)}, X_{(n)})}(y_1, y_n) = P(X_{(1)} \leq y_1, X_{(n)} \leq y_n).$$

Воспользовавшись формулой $P(A \cap B) = P(B) - P(\bar{A} \cap B)$, получим

$$P(X_{(1)} \leq y_1, X_{(n)} \leq y_n) = P(X_{(n)} \leq y_n) - P(X_{(1)} > y_1, X_{(n)} \leq y_n).$$

Если максимум меньше какого-то значения, то все элементы меньше него

$$\begin{aligned} P(X_{(1)} > y_1, X_{(n)} \leq y_n) &= \\ &= P(X_1 \leq y_n, \dots, X_n \leq y_n) - \\ &= P(X_1 \in (y_1, y_n], X_2 \in (y_1, y_n], \dots, X_n \in (y_1, y_n]) = \\ &= [F(y_n)]^n - [F(y_1)]^n \end{aligned}$$

при $y_1 < y_n$.

Продифференцируем

$$f_{(X_{(1)}, X_{(n)})}(y_1, y_n) = \begin{cases} 0, & y_1 \geq y_n, \\ n(n-1) \cdot [F(y_n) - F(y_1)]^{n-2} \times \\ \times p_{X_1}(y_1) p_{X_n}(y_n) = \\ = n(n-1) \cdot \left[\frac{y_n}{\theta} - \frac{y_1}{\theta} \right]^{n-2} \cdot \frac{1}{\theta^2} = \\ = n(n-1) \cdot \frac{(y_n - y_1)^n}{\theta^{n-2}} \cdot \frac{1}{\theta^2} = \frac{n(n-1)(y_n - y_1)^{n-1}}{\theta^n}. \end{cases}$$

Найдём математическое ожидание вектора

$$M(X_{(1)}, X_{(n)}) = \frac{n(n-1)}{\theta^n} \cdot \int_0^\theta \int_{y_n}^\theta y_1 y_n (y_n - y_1)^{n-2} dy_1 dy_n.$$

Заменяем разность величин величиной x . Якобиан преобразования равен

$$\frac{\partial(x, y_n)}{\partial(x, y_1)} = \left| \begin{bmatrix} \frac{\partial x}{\partial y_1} & \frac{\partial y_n}{\partial y_1} \\ \frac{\partial x}{\partial y_n} & \frac{\partial y_1}{\partial y_1} \end{bmatrix} \right| = \left| \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \right| = 1.$$

Получаем

$$\begin{aligned}
& \frac{n(n-1)}{\theta^n} \cdot \int_0^\theta \int_{y_n}^\theta y_1 y_n (y_n - y_1)^{n-2} dy_1 dy_n = \\
& = \frac{n(n-1)}{\theta^n} \cdot \int_0^\theta \int_0^{\theta-y_n} (x + y_n) y_n x^{n-2} dx dy_n = \\
& = \frac{n(n-1)}{\theta^n} \cdot \int_0^\theta y_n dy_n \cdot \int_0^{\theta-y_n} (x^{n-1} + y_n x^{n-2}) dx = \\
& = \frac{n(n-1)}{\theta^n} \cdot \int_0^\theta y_n dy_n \cdot \left(\frac{x^n}{n} + y_n \cdot \frac{x^{n-1}}{n-1} \right) \Big|_0^{\theta-y_n} = \\
& = \frac{1}{\theta^n} \cdot \int_0^\theta \left[(n-1)(\theta - y_n) + y_n n (\theta - y_n)^{n-1} \right] \cdot y_n dy_n = \\
& = \frac{1}{\theta^n} \cdot \int_0^\theta (\theta - y_n)^{n-1} \cdot [(n-1)(\theta - y_n) + y_n n] \cdot y_n dy_n = \\
& = \frac{1}{\theta^n} \cdot \int_0^\theta (\theta - y_n)^{n-1} [n\theta - (\theta - y_n)] y_n dy_n.
\end{aligned}$$

Заменяем первую скобку в интеграле на t . Получаем

$$\frac{1}{\theta^n} \cdot \int_0^\theta (\theta - y_n)^{n-1} [n\theta - (\theta - y_n)] y_n dy_n = \frac{1}{\theta^n} \cdot \int_0^\theta t^{n-1} (n\theta - t) (\theta - t) dt.$$

Перемножаем скобки

$$\begin{aligned}
& \frac{1}{\theta^n} \cdot \int_0^\theta t^{n-1} (n\theta - t) (\theta - t) dt = \frac{1}{\theta^n} \cdot \int_0^\theta t^{n-1} (n\theta^2 - (n+1)\theta t + t^2) dt = \\
& = \frac{1}{\theta^n} \cdot \int_0^\theta (n\theta^2 t^{n-1} - (n+1)\theta t^n + t^{n+1}) dt = \\
& = \frac{1}{\theta^n} \cdot \left[\theta^2 t^n - \theta t^{n+1} + \frac{t^{n+2}}{n+2} \right] \Big|_0^\theta = \frac{1}{\theta^n} \cdot \left[\theta^{n+2} - \theta^{n+2} + \frac{\theta^{n+2}}{n+2} \right] = \frac{\theta^2}{n+2}.
\end{aligned}$$

По определению ковариации

$$cov(X_{(1)}, X_{(n)}) = M(X_{(1)}X_{(n)}) - MX_{(1)}MX_{(n)} = \frac{\theta^2}{n+2} - \frac{\theta}{n+1} \cdot \frac{\theta n}{n+1}.$$

Выносим общий множитель за скобки

$$\frac{\theta^2}{n+2} - \frac{\theta}{n+1} \cdot \frac{\theta n}{n+1} = \theta^2 \left(\frac{1}{n+2} - \frac{n}{(n+1)^2} \right).$$

Приводим к общему знаменателю дроби в скобках

$$\theta^2 \left(\frac{1}{n+2} - \frac{n}{(n+1)^2} \right) = \theta^2 \cdot \frac{n^2 + 2n + 1 - n^2 - 2n}{(n+2)(n+1)^2} = \frac{\theta^2}{(n+2)(n+1)^2}.$$

1.22

Задание. Пусть задана выборка X_1, \dots, X_n из показательного распределения с параметром α . Докажите, что

$$MX_{(k)} = \frac{1}{\alpha} \left(\frac{1}{n-k+1} + \dots + \frac{1}{n} \right),$$

не находя распределения порядковой статистики $X_{(k)}$.

Решение. Преобразуем левую часть того, что нужно доказать

$$MX_{(k)} = MX_{(1)} + \sum_{i=1}^{k-1} M(X_{(i+1)} - X_{(i)}).$$

Элементы выборки имеют плотность распределения

$$p(y) = \alpha e^{-\alpha y} \cdot \mathbb{1}\{y \geq 0\}.$$

Интегрируя плотность, получим функцию распределения

$$F(y) = \int_{-\infty}^y p(x) dx = \int_{-\infty}^y \alpha e^{-\alpha x} \cdot \mathbb{1}\{x \geq 0\} dx = \int_0^y \alpha e^{-\alpha x} dx = -e^{-\alpha x} \Big|_0^y.$$

Подставляем пределы интегрирования

$$-e^{-\alpha x} \Big|_0^y = 1 - e^{-\alpha y}, \quad y \geq 0.$$

Находим функцию распределения минимальной порядковой статистики

$$F_{X_{(1)}}(y) = P(X_{(1)} \leq y) = 1 - P(X_{(1)} > y).$$

Если минимальный элемент выборки больше какого-то значения, то все элементы выборки больше него

$$1 - P(X_{(1)} > y) = 1 - P(X_1 > y, X_2 > y, \dots, X_n > y).$$

Из независимости случайных величин следует, что

$$1 - P(X_1 > y, X_2 > y, \dots, X_n > y) = 1 - P(X_1 > y) \cdot P(X_2 > y) \cdot \dots \cdot P(X_n > y).$$

Переходим к противоположным событиям и учитываем то, что случайные величины одинаково распределены

$$1 - P(X_1 > y) \cdot P(X_2 > y) \cdot \dots \cdot P(X_n > y) = 1 - [1 - F(y)]^n.$$

Находим плотность минимальной порядковой статистики

$$f_{X_{(1)}}(y) = \frac{\partial F_{X_{(1)}}(y)}{\partial y} = \frac{\partial \{1 - [1 - F(y)]^n\}}{\partial y} = -n[1 - F(y)]^{n-1} \cdot (-1) f(y).$$

Упрощаем и учитываем, что $f(y)$ — это плотность распределения элементов выборки

$$-n[1 - F(y)]^{n-1} \cdot (-1) f(y) = n[1 - F(y)]^{n-1} p(y).$$

Подставляем функцию и плотность распределения

$$n[1 - F(y)]^{n-1} p(y) = n[1 - (1 - e^{-\alpha y})]^n \alpha e^{-\alpha y} \cdot \mathbb{1}\{y \geq 0\}.$$

Упрощаем выражение в скобках

$$n[1 - (1 - e^{-\alpha y})]^n \alpha e^{-\alpha y} \cdot \mathbb{1}\{y \geq 0\} = n\alpha e^{-\alpha(n-1)y} e^{-\alpha y} \cdot \mathbb{1}\{y \geq 0\}.$$

Суммируем показатели экспонент

$$n\alpha e^{-\alpha(n-1)y} e^{-\alpha y} \cdot \mathbb{1}\{y \geq 0\} = n\alpha e^{-\alpha n y} \cdot \mathbb{1}\{y \geq 0\}.$$

Отсюда следует, что $X_{(1)} \sim \text{Exp}(n\alpha)$.

Из задачи 1.11 б) $X_{(i+1)} - X_{(i)} \sim \text{Exp}(\alpha(n-i))$.

Тогда такая разность имеет математическое ожидание

$$M(X_{(i+1)} - X_{(i)}) = \frac{1}{\alpha(n-i)}.$$

Подставляем в полученное в начале выражение

$$MX_{(1)} + \sum_{i=1}^{k-1} M(X_{(i+1)} - X_{(i)}) = \frac{1}{n\alpha} + \sum_{i=1}^{k-1} \frac{1}{\alpha(n-i)}.$$

Записываем сумму в явном виде

$$\frac{1}{n\alpha} + \sum_{i=1}^{k-1} \frac{1}{\alpha(n-i)} = \frac{1}{\alpha} \cdot \left(\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \dots + \frac{1}{n-k+1} \right).$$

1.23

Задание. Для выборки X_1, \dots, X_n из равномерного распределения на отрезке $[0, \theta]$ найдите:

- а) ковариацию $X_{(n)}$ и $X_{(1)}$;

б) совместную плотность распределения $X_{(k)}$ и $X_{(j)}$, $1 \leq k \leq j \leq n$.

Решение.

а) Данный пункт решён в задаче 1.21 с);

б) по определению функции распределения

$$F_{(X_{(k)}, X_{(j)})}(y, z) = P(X_{(k)} \leq y, X_{(j)} \leq z).$$

Воспользовавшись формулой $P(A \cap B) = P(B) - P(\bar{A} \cap B)$, получим $P(X_{(k)} \leq y, X_{(j)} \leq z) = P(X_{(j)} \leq z) - P(X_{(k)} > y, X_{(j)} \leq z)$. Аналогично задаче 1.10 разбиваем на 2 решения согласно коэффициентам

$$\begin{aligned} & P(X_{(j)} \leq z) - P(X_{(k)} > y, X_{(j)} \leq z) = \\ & = \begin{cases} P(X_{(j)} \leq z), & y \geq z, \\ P(X_{(j)} \leq z) - P(X_{(k)} > y, X_{(j)} \leq z), & y < z. \end{cases} \end{aligned}$$

Найдём первую вероятность $P(X_{(j)} \leq z) = P(\text{хотя бы } j \text{ элементов выборки не превышает } z) =$

$$= \sum_{i=j}^n C_n^i F^i(z) [1 - F(z)]^{n-i} = \sum_{i=j}^n C_n^i \left(\frac{z}{\theta}\right)^i \left(1 - \frac{z}{\theta}\right)^{n-i}.$$

Приводим к общему знаменателю

$$\sum_{i=j}^n C_n^i \left(\frac{z}{\theta}\right)^i \left(1 - \frac{z}{\theta}\right)^{n-i} = \sum_{i=j}^n C_n^i \frac{z^i}{\theta^i} \cdot \frac{(\theta - z)^{n-i}}{\theta^{n-i}} = \sum_{i=j}^n C_n^i \frac{z^i (\theta - z)^{n-i}}{\theta^n}.$$

Найдём вторую вероятность $P(X_{(j)} \leq z, X_{(k)} > y) = P(\text{хотя бы}$

$$(j - k + 1)$$

элементов выборки находится в интервале $(y, z]) =$

$$\begin{aligned} & = \sum_{i=j-k+1}^n C_n^i [F(z) - F(y)]^i \cdot \{1 - [F(z) - F(y)]\}^{n-i} = \\ & = \sum_{i=j-k+1}^n C_n^i \cdot \left(\frac{z}{\theta} - \frac{y}{\theta}\right)^i \cdot \left[1 - \left(\frac{z}{\theta} - \frac{y}{\theta}\right)\right]^{n-i} = \\ & = \sum_{i=j-k+1}^n C_n^i \cdot \frac{(z - y)^i}{\theta^i} \cdot \left(1 - \frac{z - y}{\theta}\right)^{n-i} = \\ & = \sum_{i=j-k+1}^n C_n^i \cdot \frac{(z - y)^i}{\theta^i} \cdot \frac{(\theta - z + y)^{n-i}}{\theta^{n-i}} = \\ & = \sum_{i=j-k+1}^n \frac{C_n^i}{\theta^n} (z - y)^i (\theta - z + y)^{n-i}. \end{aligned}$$

Продифференцируем. В первом случае будет 0. Найдём производную второй вероятности

$$\begin{aligned}
& -\frac{\partial^2 F}{\partial y \partial z} = \\
& = \frac{\partial}{\partial y} \left\{ \sum_{i=j-k+1}^n \frac{C_n^i}{\theta^n} \times \right. \\
& \times \left[i(z-y)^{i-1}(\theta-z+y)^{n-i} - (z-y)(n-i)(\theta-z+y)^{n-i-1} \right] \Big\} = \\
& = \sum_{i=j-k+1}^n \frac{C_n^i}{\theta^n} \times \\
& \times [-i(i-1)(z-y)^{i-2}(\theta-z+y)^{n-i} + \\
& + 2i(z-y)^{i-1}(n-i) \cdot (\theta-z+y)^{n-i-1} - \\
& - (z-y)^i(n-i)(n-i-1) \cdot (\theta-z+y)^{n-i-2}] = \\
& = \sum_{i=j-k+1}^n \frac{C_n^i(z-y)^{i-2} \cdot (\theta-z+y)^{n-i-2}}{\theta^n} \times \\
& \times [2i(z-y)(n-i)(\theta-z+y) - i(i-1)(\theta-z+y) - \\
& - (n-i)(n-i-1)(z-y)^2].
\end{aligned}$$

Тогда во втором случае плотность равна полученному выражению.

Занятие 2. Свойства оценок

Контрольные вопросы и задания

Что называют оценкой неизвестного параметра?

Статистику, значение которой заменяет неизвестный параметр, называют оценкой этого параметра.

Приведите определение оценки: несмещённой, асимптотически несмещённой, состоятельной, сильно состоятельной, оптимальной.

Оценка $\hat{\theta}$ несмещённая, если $\forall \theta \in \Theta : M_{\theta} \hat{\theta} = \theta$.

Асимптотически несмещённая оценка — такая оценка, математическое ожидание которой совпадает с оцениваемым параметром при $n \rightarrow \infty$.

Оценка $\hat{\theta}$ называется состоятельной, если стремится к истинному значению θ по вероятности $\hat{\theta} \xrightarrow{P} \theta, n \rightarrow \infty$.

Оценка $\hat{\theta}$ называется сильно состоятельной, если стремится к истинному значению θ почти наверное $\hat{\theta} \xrightarrow{a.s.} \theta, n \rightarrow \infty$.

Несмещённая оценка $\hat{\theta} \in K$ называется оптимальной в классе квадратично интегрируемых оценок K , если для всякой другой несмещённой оценки $\tilde{\theta} \in \Theta \forall \theta \in \Theta : D_{\theta} \hat{\theta} \leq D_{\theta} \tilde{\theta}$ или же $\forall \theta \in \Theta, M_{\theta} (\hat{\theta} - \theta)^2 \leq M_{\theta} (\tilde{\theta} - \theta)^2$.

Что называется среднеквадратическим отклонением оценки?

$M_{\theta} (\hat{\theta} - \theta)$ — среднеквадратическое отклонение.

Сформулируйте утверждение про поведение выборочных моментов.

Выборочный начальный момент M_k k -го порядка стремится к начальному моменту ν_k случайной величины X , то есть

$$\lim_{n \rightarrow \infty} P(|M_k - \nu_k| \geq \varepsilon) = 0,$$

для любого сколь угодно малого $\varepsilon > 0$, если моменты ν_{2k} и ν_k случайной величины X существуют и конечны.

Какая оценка является несмещённой и состоятельной для математического ожидания распределения выборки?

В качестве оценки для математического ожидания естественно предложить среднеарифметическое наблюдаемых значений

$$\tilde{m} = \frac{\sum_{i=1}^n X_i}{n}.$$

Какая статистика является несмещённой оценкой для дисперсии распределения выборки?

$$\frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2$$

— несмещённая оценка для $\sigma^2 = Dx_1$.

Аудиторные задачи

2.4

Задание. Для выборки равномерного распределения на отрезке $[\theta, 1]$ проверьте состоятельность и несмещённость оценки $X_{(1)}$ параметра θ .

Решение. θ — минимальное наблюдение. Проверяем, выполняется ли $X_{(1)} \xrightarrow{P} \theta, n \rightarrow \infty$.

По определению сходимости по вероятности

$$\forall \varepsilon > 0 \ P(|X_{(1)} - \theta| > \varepsilon) \rightarrow 0, \ n \rightarrow \infty.$$

Раскроем модуль

$$P\{X_{(1)} > \varepsilon + \theta\} = P(X_1 > \varepsilon + \theta, \dots, X_n > \varepsilon + \theta) = [P(X_1 > \varepsilon + \theta)]^n.$$

Подставим значение вероятности из геометрического эксперимента

$$[P(X_1 > \varepsilon + \theta)]^n = \left(\frac{1 - \theta - \varepsilon}{1 - \theta}\right)^n = \left(1 - \frac{\varepsilon}{1 - \theta}\right)^n \rightarrow 0, \ n \rightarrow \infty.$$

Число в скобках строго меньше единицы, так как $0 \leq \theta \leq 1$.

Отсюда следует, что оценка состоятельная.

Проверяем несмещённость оценки. Проверяем, выполняется ли

$$MX_{(1)} = \theta.$$

Нужно найти плотность

$$MX_{(1)} = \int_{\mathbb{R}} f_{X_{(1)}}(y) y dy.$$

Начинаем с функции распределения $F_{X_{(1)}}(y) = P(X_{(1)} \leq y)$. Переходим к противоположному событию

$$P(X_{(1)} \leq y) = 1 - P(X_{(1)} > y) = 1 - [P(X_1 > y)]^n.$$

Переходим к противоположному событию $1 - [P(X_1 > y)]^n = 1 - [1 - F(y)]^n$.

Продифференцируем

$$\frac{dF_{X_{(1)}}(y)}{dy} = n[1 - F(y)]^{n-1} f(y).$$

На отрезке $[\theta, 1]$ имеет равномерное распределение

$$n[1 - F(y)]^{n-1} f(y) = n \left[1 - \frac{y - \theta}{1 - \theta} \right]^{n-1} \cdot \mathbb{1}\{y \in [\theta, 1]\} \cdot \frac{1}{1 - \theta}.$$

Приведём к общему знаменателю

$$n \left[1 - \frac{y - \theta}{1 - \theta} \right]^{n-1} \cdot \mathbb{1}\{y \in [\theta, 1]\} \cdot \frac{1}{1 - \theta} = \frac{n}{(1 - \theta)^2} \cdot (1 - y)^{n-1} \cdot \mathbb{1}\{y \in [\theta, 1]\}.$$

Нашли плотность $X_{(1)}$ и теперь можем вычислить интеграл

$$MX_{(1)} = \int_{\theta}^1 y \cdot \frac{n}{(1 - \theta)^2} \cdot (1 - y)^{n-1} dy.$$

Замена:

$$1 - y = z, dy = -dz, y = 1 - z, y = 1 \Rightarrow z = 0, y = \theta \Rightarrow z = 1 - \theta.$$

Подставляя замену, получаем

$$\int_{\theta}^1 y \cdot \frac{n}{(1 - \theta)^2} \cdot (1 - y)^{n-1} dy = n \cdot \frac{1}{(1 - \theta)^n} \int_0^{1-\theta} (1 - z) z^{n-1} dz.$$

Вычислим интеграл

$$\begin{aligned} n \cdot \frac{1}{(1 - \theta)^n} \int_0^{1-\theta} (1 - z) z^{n-1} dz &= \frac{n}{(1 - \theta)^n} \left[\frac{(1 - \theta)^n}{n} - \frac{(1 - \theta)^{n+1}}{n + 1} \right] = \\ &= n \left(\frac{1}{n} - \frac{1 - \theta}{n + 1} \right) = 1 - \frac{n}{n + 1} (1 - \theta). \end{aligned}$$

Раскроем скобки

$$1 - \frac{n}{n + 1} (1 - \theta) = 1 - \frac{n}{n + 1} - \theta \cdot \frac{n}{n + 1} \neq \theta.$$

Отсюда следует, что оценка смещённая, но асимптотически несмещённая, потому что

$$1 - \frac{n}{n + 1} \rightarrow 0, n \rightarrow \infty$$

и

$$\frac{n}{n + 1} \rightarrow 1, n \rightarrow \infty.$$

2.5

Задание. Пусть X_1, \dots, X_n — выборка из распределения Пуассона с параметром $\lambda > 0$. Выясните, является ли статистика

$$\frac{1}{n} \sum_{i=1}^n X_i^2 :$$

- а) несмещённой оценкой для λ^2 ;
- б) состоятельной оценкой для λ^2 .

Решение.

- а) Нужно проверить, выполняется ли

$$M \frac{1}{n} \sum_{i=1}^n X_i^2 = \lambda^2.$$

Преобразуем левую часть

$$M \frac{1}{n} \sum_{i=1}^n X_i^2 = M X_1^2 = D X_1 + (M X_1)^2 = \lambda + \lambda^2 \neq \lambda^2.$$

Значит, оценка смещённая;

- б) проверяем, имеет ли место

$$\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} \lambda^2, n \rightarrow \infty.$$

По закону больших чисел

$$\frac{1}{n} \sum_{i=1}^n X_i^2 \rightarrow M X_1^2 = \lambda^2 + \lambda \neq \lambda^2,$$

значит, оценка не состоятельная.

2.6

Задание. Пусть X_1, \dots, X_n — выборка из показательного распределения с параметром $\alpha > 0$. Докажите, что статистика $1/\bar{X}$ является состоятельной оценкой для α .

Решение. Нужно показать, что

$$\frac{1}{\bar{X}} \xrightarrow{P} \alpha, n \rightarrow \infty.$$

Выборочное среднее

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

По закону больших чисел

$$\bar{X} \xrightarrow{P} MX_1 = \frac{1}{\alpha}.$$

Отсюда следует, что

$$\frac{1}{\bar{X}} \xrightarrow{P} \frac{1}{MX_1} = \alpha.$$

2.7

Задание. Пусть X_1, \dots, X_n — выборка из нормального распределения $N(a, \sigma^2)$. Докажите, что статистика

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

является несмещённой и состоятельной оценкой для σ^1 .

Решение. Нужно проверить условие $MS_n = \sigma^2$.

Разность двух соседних элементов выборки имеет распределение

$$X_{i+1} - X_i \sim N(0, 2\sigma^2).$$

Найдём математическое ожидание статистики

$$MS_n = M \frac{1}{2(n-1)} \cdot \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2 = \frac{1}{2(n-1)} \sum_{i=1}^n M(X_{i+1} - X_i)^2.$$

Случайные величины одинаково распределены

$$\frac{1}{2(n-1)} \sum_{i=1}^n M(X_{i+1} - X_i)^2 = \frac{n}{2(n-1)} \cdot M(X_2 - X_1)^2.$$

В данном случае второй момент равен дисперсии

$$\frac{n}{2(n-1)} \cdot M(X_2 - X_1)^2 = \frac{1}{2} \cdot 2\sigma^2 = \sigma^2.$$

Отсюда следует, что оценка несмещённая.

Проверим состоятельность, то есть $S_n \xrightarrow{P} \sigma^2, n \rightarrow \infty$.

Разобьём S_n на две суммы

$$S_n = \frac{1}{2(n-1)} \cdot \left[\sum_{\text{even } i} (X_{i+1} - X_i)^2 + \sum_{\text{odd } i} (X_{i+1} - X_i)^2 \right].$$

В каждой из сумм слагаемые независимы

$$\begin{aligned} & \frac{1}{2(n-1)} \cdot \left[\sum_{\text{even } i} (X_{i+1} - X_i)^2 + \sum_{\text{odd } i} (X_{i+1} - X_i)^2 \right] = \\ & = \frac{1}{2(n-1)} \cdot \left[\frac{m}{m} \sum_{\text{even } i} (X_{i+1} - X_i)^2 + \frac{n-1-m}{n-1-m} \sum_{\text{odd } i} (X_{i+1} - X_i)^2 \right]. \end{aligned}$$

По закону больших чисел

$$\begin{aligned} & \frac{1}{2(n-1)} \cdot \left[\frac{m}{m} \sum_{\text{even } i} (X_{i+1} - X_i)^2 + \frac{n-1-m}{n-1-m} \sum_{\text{odd } i} (X_{i+1} - X_i)^2 \right] \xrightarrow{P} \\ & \xrightarrow{P} \frac{1}{2} \cdot M(X_2 - X_1)^2 = \frac{1}{2} \cdot D(X_2 - X_1) = \sigma^2, n \rightarrow \infty. \end{aligned}$$

Отсюда следует, что оценка состоятельная.

2.8

Задание. Пусть X_1, \dots, X_n — выборка из показательного распределения с параметром $\alpha > 1$. Для какого параметра $\theta = \theta(\alpha)$ статистика

$$\hat{\theta}_n = e^{\bar{X}}$$

является состоятельной оценкой? Является ли $\hat{\theta}_n$ сильно состоятельной оценкой того же параметра? Является ли $\hat{\theta}_n$ несмещённой оценкой того же параметра? Асимптотически несмещённой?

Решение.

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

По закону больших чисел

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} MX_1, n \rightarrow \infty.$$

Случайные величины в выборке имеют показательное распределение

$$MX_1 = \frac{1}{\alpha},$$

значит,

$$\bar{X} \xrightarrow{P} \frac{1}{\alpha}, n \rightarrow \infty.$$

Применяем непрерывную функцию e^x . Получаем $e^{\bar{X}} \xrightarrow{P} e^{\frac{1}{\alpha}}, n \rightarrow \infty$.

Проверяем, является ли оценка $e^{\bar{X}}$ несмещённой к параметру $e^{\frac{1}{\alpha}}$, то есть выполняется ли $Me^{\bar{X}} = e^{\frac{1}{\alpha}}$.

Вычисляем $Me^{\bar{X}} = Me^{\frac{1}{n} \sum_{i=1}^n X_i}$. Случайные величины независимы и одинаково распределены, поэтому $Me^{\frac{1}{n} \sum_{i=1}^n X_i} = \left(Me^{\frac{X_1}{n}} \right)^n$. По определению характеристической функции $\varphi_{X_1} = Me^{itX_1}$ получаем

$$\left(Me^{\frac{X_1}{n}} \right)^n = \left[\varphi_{X_1} \left(\frac{1}{in} \right) \right]^n.$$

Характеристическая функция показательного распределения

$$\varphi_{X_1}(t) = Me^{itX_1} = \frac{\alpha}{\alpha - it}.$$

Подставляем

$$\left[\varphi_{X_1} \left(\frac{1}{in} \right) \right]^n = \left(\frac{\alpha}{\alpha - \frac{1}{n}} \right)^n.$$

Прибавим и отнимем в числителе $1/n$ и поделим числитель на знаменатель

$$\left(\frac{\alpha}{\alpha - \frac{1}{n}} \right)^n = \left(\frac{\alpha + \frac{1}{n} - \frac{1}{n}}{\alpha - \frac{1}{n}} \right)^n = \left[1 + \frac{1}{n(\alpha - \frac{1}{n})} \right]^n = e^{\frac{1}{\alpha}}.$$

Значит, оценка смещённая, но несмещённая асимптотически.

2.9

Задание. Пусть X_1, \dots, X_n — выборка из геометрического распределения с параметром p . Найдите состоятельные оценки для параметров

$$p, p^2, \ln p, p \sin(1-p), pe^{\frac{q^2}{2}}.$$

Решение. Всё это — непрерывные функции от p . Применение непрерывной функции не нарушает сходимости по вероятности.

Если параметр каким-то образом связан со средним, но нужно пробовать выборочные моменты.

Сформируем выборочное среднее и применим закон больших чисел

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} MX_1.$$

Для геометрического распределения

$$MX_1 = \frac{1-p}{p}.$$

Прибавим единицу слева и справа

$$1 + \bar{X} \xrightarrow{P} 1 + \frac{1-p}{p} = 1 + \frac{1}{p} - 1 = \frac{1}{p}.$$

Функция

$$f(x) = \frac{1}{x}$$

— это непрерывная функция, значит, можем применить эту функцию слева и справа, и сходимость сохранится

$$\frac{1}{1 + \overline{X}} \xrightarrow{P} p.$$

Состоятельной оценкой для параметра p будет

$$\hat{p} = \frac{1}{1 + \overline{X}}.$$

Состоятельной оценкой для p^2 будет $\hat{p}^2 = \hat{p}^2 \xrightarrow{P} p^2$, потому что $f(x) = x^2$ — это непрерывная функция.

Логарифм — это непрерывная функция

$$\ln \hat{p} = \ln \hat{p} = \ln \frac{1}{1 + \overline{X}} \xrightarrow{P} \ln p,$$

поскольку

$$\frac{1}{1 + \overline{X}} \xrightarrow{P} p.$$

2.12

Задание. Пусть X_1, \dots, X_n — выборка из распределения Пуассона с параметром $\lambda > 0$. Докажите, что не существует несмещённой оценки для параметра $1/\lambda$.

Решение. Будем действовать от противного.

Допустим, что такая оценка существует, то есть существует $\hat{\theta}$ такое, что

$$M\hat{\theta} = \frac{1}{\lambda}.$$

Это условие несмещённости.

$\hat{\theta}$ — функция от выборки, то есть $\hat{\theta} = f(X_1, \dots, X_n)$.

Распределение Пуассона дискретное

$$\begin{aligned} M\hat{\theta} &= Mf(X_1, \dots, X_n) = \\ &= \sum_{k_1=0}^{\infty} \dots \sum_{k_n=0}^{\infty} f(X_1, \dots, X_n) P(X_1 = k_1, X_2 = k_2, \dots, X_n = k_n). \end{aligned}$$

Воспользуемся независимостью

$$\begin{aligned} &\sum_{k_1=0}^{\infty} \dots \sum_{k_n=0}^{\infty} f(X_1, \dots, X_n) \cdot P(X_1 = k_1, X_2 = k_2, \dots, X_n = k_n) = \\ &= \sum_{k_1=0}^{\infty} \dots \sum_{k_n=0}^{\infty} f(X_1, \dots, X_n) \cdot P(X_1 = k_1) \cdot \dots \cdot P(X_n = k_n). \end{aligned}$$

Подставим в явном виде вероятности

$$\begin{aligned} \sum_{k_1=0}^{\infty} \dots \sum_{k_n=0}^{\infty} f(X_1, \dots, X_n) \cdot P(X_1 = k_1) \cdot \dots \cdot P(X_n = k_n) = \\ = \sum_{k_1=0}^{\infty} \dots \sum_{k_n=0}^{\infty} f(X_1, \dots, X_n) \cdot \frac{\lambda^{\sum_{i=1}^n k_i}}{\prod_{i=1}^n k_i!} \cdot e^{-\lambda n}. \end{aligned}$$

Обозначим

$$\sum k_i = k$$

и получим

$$e^{-\lambda n} \cdot \sum_{k=0}^{\infty} \lambda^k \cdot \sum_{k_i: k_1+k_2+\dots+k_n=k} \frac{f(k_1, \dots, k_n)}{k_1! \dots k_n!} = e^{-\lambda n} \sum_{k=0}^{\infty} \lambda^k c_k,$$

где c_k — число.

Допустим, что эта величина равна $1/\lambda$.

Посмотрим, возможно ли это

$$e^{-\lambda n} \sum_{k=0}^{\infty} \lambda^k c_k = \frac{1}{\lambda}.$$

Запишем так, чтобы с одной стороны было $e^{\lambda n}$, а всё остальное перенесём

$$e^{\lambda n} = \sum_{k=0}^{\infty} \lambda^{k+1} c_k.$$

Для экспоненты существует одно развитие в ряд. Из этого следует, что последнее равенство невозможно (развитие в ряд начинается со степени λ , равной единице).

Домашнее задание

2.17

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке $[0, \theta]$. Проверьте несмещённость, состоятельность и найдите среднеквадратическое отклонение следующих оценок параметра θ :

а) $X_{(1)} + X_{(n)}$;

б) $(n+1) X_{(1)}$.

Решение. θ — максимальное наблюдение.

а) Проверим, выполняется ли $X_{(1)} + X_{(n)} \xrightarrow{P} \theta, n \rightarrow \infty$.

По определению сходимости по вероятности $\forall \varepsilon > 0 P(X_{(1)} > \varepsilon) = P(X_1 > \varepsilon, X_2 > \varepsilon, \dots, X_n > \varepsilon) = [P(X_1 > \varepsilon)]^n$. Подставим значение вероятности из геометрического эксперимента

$$[P(X_1 > \varepsilon)]^n = \left(\frac{\theta - \varepsilon}{\theta}\right)^n \rightarrow 0, n \rightarrow \infty.$$

Значит, $X_{(1)} \xrightarrow{P} 0, n \rightarrow \infty$.

Остаётся проверить, выполняется ли $X_{(n)} \xrightarrow{P} \theta, n \rightarrow \infty$.

По определению сходимости по вероятности

$$P\{|X_{(n)} - \theta| > \varepsilon\} \rightarrow 0, n \rightarrow \infty.$$

Раскроем модуль

$$\begin{aligned} P\{\theta - X_{(n)} > \varepsilon\} &= P\{X_{(n)} - \theta < -\varepsilon\} = P\{X_{(n)} < \theta - \varepsilon\} = \\ &= P(X_1 < \theta - \varepsilon, X_2 < \theta - \varepsilon, \dots, X_n < \theta - \varepsilon) = [P(X_1 < \theta - \varepsilon)]^n. \end{aligned}$$

Подставим значение вероятности из геометрического эксперимента

$$[P(X_1 < \theta - \varepsilon)]^n = \left(\frac{\theta - \varepsilon}{\theta}\right)^n = \left(1 - \frac{\varepsilon}{\theta}\right)^n \rightarrow 0, n \rightarrow \infty.$$

Число в скобках строго меньше единицы, так как $0 \leq \theta \leq 1, \varepsilon > 0$.

Отсюда следует, что оценка состоятельная.

Проверим несмещённость оценки. Проверим, выполняется ли

$$M(X_{(1)} + X_{(n)}) = \theta.$$

Из задачи 1.21

$$MX_{(n)} = \frac{\theta n}{n+1}$$

и

$$MX_{(1)} = \frac{\theta}{n+1}.$$

Из свойства линейности математического ожидания

$$M(X_{(1)} + X_{(n)}) = MX_{(1)} + MX_{(n)} = \frac{\theta n}{n+1} + \frac{\theta}{n+1} = \frac{\theta(n+1)}{n+1} = \theta.$$

Отсюда следует, что оценка несмещённая.

Формула среднеквадратического отклонения имеет вид $\sigma = \sqrt{D\xi}$.

Найдём дисперсию оценки

$$\begin{aligned}
D(X_{(1)} + X_{(n)}) &= M(X_{(1)} + X_{(n)})^2 - [M(X_{(1)} + X_{(n)})]^2 = \\
&= M(X_{(1)}^2 + X_{(n)}^2 + 2X_{(1)}X_{(n)}) - (MX_{(1)} + MX_{(n)})^2 = \\
&= MX_{(1)}^2 + MX_{(n)}^2 + 2M(X_{(1)}X_{(n)}) - (MX_{(1)})^2 - 2MX_{(1)}MX_{(n)} - \\
&\quad - (MX_{(n)})^2 = DX_{(1)} + DX_{(n)} + 2cov(X_{(1)}, X_{(n)}).
\end{aligned}$$

Возьмём необходимые значения из задачи 1.21, а именно

$$DX_{(1)} = \frac{\theta^2 n}{(n+1)^2 (n+2)} = DX_{(n)}, \quad cov(X_{(1)}, X_{(n)}) = \frac{\theta^2}{(n+1)^2 (n+2)}.$$

Подставляя в найденное выражение, получаем

$$\begin{aligned}
DX_{(1)} + DX_{(n)} + 2cov(X_{(1)}, X_{(n)}) &= \\
&= \frac{\theta^2 n}{(n+1)^2 (n+2)} + \frac{\theta^2 n}{(n+1)^2 (n+2)} + \frac{2\theta^2}{(n+1)^2 (n+2)} = \\
&= \frac{2\theta^2 n + 2\theta^2}{(n+1)^2 (n+2)} = \frac{2\theta^2 (n+1)}{(n+1)^2 (n+2)} = \frac{2\theta^2}{(n+1)(n+2)}.
\end{aligned}$$

Извлекая корень, получаем

$$\sigma = \sqrt{\frac{2\theta^2}{(n+1)(n+2)}} = \theta \sqrt{\frac{2}{(n+1)(n+2)}};$$

b) проверим, выполняется ли $(n+1)X_{(1)} \xrightarrow{P} \theta, n \rightarrow \infty$.

По определению сходимости по вероятности

$$\forall \varepsilon > 0 \quad P(|(n+1)X_{(1)} - \theta| > \varepsilon) \rightarrow 0, \quad n \rightarrow \infty.$$

Перейдём к противоположному событию

$$P(|(n+1)X_{(1)} - \theta| > \varepsilon) = 1 - P\{|(n+1)X_{(1)} - \theta| \leq \varepsilon\}.$$

Раскроем модуль

$$\begin{aligned}
1 - P\{|(n+1)X_{(1)} - \theta| \leq \varepsilon\} &= \\
&= 1 - P\{-\varepsilon \leq (n+1)X_{(1)} - \theta \leq \varepsilon\} = \\
&= 1 - P\left\{\frac{-\varepsilon + \theta}{n+1} \leq X_{(1)} \leq \frac{\varepsilon + \theta}{n+1}\right\} = \\
&= 1 + P\left\{X_{(1)} \leq \frac{\theta - \varepsilon}{n+1}\right\} - P\left\{X_{(1)} < \frac{\varepsilon + \theta}{n+1}\right\} = \\
&= 1 + 1 - P\left(X_1 > \frac{\theta - \varepsilon}{n+1}, \dots, X_n > \frac{\theta - \varepsilon}{n+1}\right) - 1 + P\left(X_{(1)} > \frac{\varepsilon + \theta}{n+1}\right) = \\
&= 1 - \left[P\left(X_1 > \frac{\theta - \varepsilon}{n+1}\right)\right]^n + \left[P\left(X_1 > \frac{\theta + \varepsilon}{n+1}\right)\right]^n.
\end{aligned}$$

Подставим значения вероятностей из геометрического эксперимента

$$\begin{aligned} 1 - \left[P\left(X_1 > \frac{\theta - \varepsilon}{n+1}\right) \right]^n + \left[P\left(X_1 > \frac{\theta + \varepsilon}{n+1}\right) \right]^n &= \\ &= 1 - \left(\frac{\theta - \frac{\theta - \varepsilon}{n+1}}{\theta} \right)^2 + \left(\frac{\theta - \frac{\varepsilon + \theta}{n+1}}{\theta} \right)^n = \\ &= 1 - \left(1 - \frac{\theta - \varepsilon}{(n+1)\theta} \right)^n + \left(1 - \frac{\varepsilon + \theta}{(n+1)\theta} \right)^n \not\rightarrow 0, n \rightarrow \infty. \end{aligned}$$

Отсюда следует, что оценка несостоятельная.

Проверим несмещённость оценки. Проверим, выполняется ли

$$M(n+1)X_{(1)} = \theta.$$

Выносим константу из-под знака математического ожидания

$$(n+1)MX_{(1)} = (n+1) \cdot \frac{\theta}{n+1} = \theta.$$

Отсюда следует, что оценка несмещённая.

Найдём дисперсию оценки

$$D(n+1)X_{(1)} = (n+1)^2 DX_{(1)} = (n+1)^2 \cdot \frac{\theta^2 n}{(n+1)^2 (n+2)} = \frac{\theta^2 n}{n+2},$$

откуда среднееквадратическое отклонение равно

$$\sigma = \sqrt{D(n+1)X_{(1)}} = \sqrt{\frac{\theta^2 n}{n+2}} = \theta \sqrt{\frac{n}{n+2}}.$$

2.18

Задание. Пусть X_1, \dots, X_n — выборка из показательного распределения с параметром $1/\sqrt{\alpha}$. Выясните, является ли статистика $\hat{\alpha}_n = (\bar{X})^2$ несмещённой оценкой параметра α . Является ли эта оценка состоятельной?

Решение. Нужно проверить, выполняется ли $M(\bar{X})^2 = \alpha$.

Запишем, что означает выборочное среднее

$$M(\bar{X})^2 = M\left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2 = M\left(\frac{1}{n^2} \left(\sum_{i=1}^n X_i\right)^2\right) = \frac{1}{n^2} \cdot M\left(\sum_{i=1}^n X_i\right)^2.$$

Квадрат суммы запишем в виде двух сумм. Получим

$$\frac{1}{n^2} \cdot M\left(\sum_{i=1}^n X_i^2 + 2 \sum_{k < i} X_k X_i\right) = \frac{1}{n^2} \cdot M \sum_{i=1}^n X_i^2 + \frac{2}{n^2} M \sum_{k < i} X_k X_i.$$

Случайные величины независимы и одинаково распределены, поэтому

$$\frac{1}{n^2} \cdot M \sum_{i=1}^n X_i + \frac{2}{n^2} M \sum_{k < i} X_k X_i = \frac{1}{n^2} \sum_{i=1}^n M X_i^2 + \frac{2}{n^2} \sum_{k < i} M X_k M X_i.$$

Для показательного распределения

$$M X_i = \frac{1}{\lambda} = \frac{1}{\frac{1}{\sqrt{\alpha}}} = \sqrt{\alpha}, \quad M X_i^2 = \frac{2}{\lambda^2} = \frac{2}{\frac{1}{\alpha}} = 2\alpha.$$

Подставляем

$$\begin{aligned} \frac{1}{n^2} \sum_{i=1}^n M X_i^2 + \frac{2}{n^2} \sum_{k < i} M X_k M X_i &= \frac{1}{n^2} \cdot n \cdot 2\alpha + \frac{1}{n^2} \cdot (n-1)n (M X_1)^2 = \\ &= \frac{2\alpha}{n} + \frac{n-1}{n} \cdot \alpha = \frac{2\alpha}{n} + \alpha - \frac{\alpha}{n} = \frac{\alpha}{n} + \alpha \rightarrow \alpha, \quad n \rightarrow \infty, \end{aligned}$$

значит, оценка смещённая, но несмещённая асимптотически.

Проверим, имеет ли место $(\bar{X})^2 \xrightarrow{P} \alpha, n \rightarrow \infty$.

По закону больших чисел

$$(\bar{X})^2 = \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2 \rightarrow (M X_1)^2 = (\sqrt{\alpha})^2 = \alpha,$$

значит, оценка состоятельная.

2.21

Задание. Пусть X_1, \dots, X_n — выборка из биномиального распределения с параметрами 2 и p . Для какого параметра $\theta = \theta(p)$ статистика $\hat{\theta}_n = e^{\bar{X}}$ будет состоятельной? Является ли $\hat{\theta}_n$ сильно состоятельной оценкой того же параметра? Является ли $\hat{\theta}_n$ несмещённой оценкой того же параметра? Найдите среднеквадратическое отклонение этой оценки.

Решение.

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

По закону больших чисел

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} M X_1, \quad n \rightarrow \infty.$$

Случайные величины в выборке имеют биномиальное распределение с математическим ожиданием $M X_1 = np = 2p$, значит, $\bar{X} \xrightarrow{P} 2p, n \rightarrow \infty$.

Применим непрерывную функцию e^x . Получим $e^{\bar{X}} \xrightarrow{P} e^{2p}, n \rightarrow \infty$.

Проверим, является ли оценка несмещённой к параметру e^{2p} , то есть выполняется ли $M e^{\bar{X}} = e^{2p}$.

Вычисляем $Me^{\bar{X}} = Me^{\frac{1}{n} \sum_{i=1}^n X_i}$. Случайные величины независимы и одинаково распределены, поэтому $Me^{\frac{1}{n} \sum_{i=1}^n X_i} = \left(Me^{\frac{X_1}{n}}\right)^n$. По определению характеристической функции $\varphi_{X_1}(t) = Me^{itX_1}$, получим

$$\left(Me^{\frac{X_1}{n}}\right)^n = \left[\varphi_{X_1}\left(\frac{1}{in}\right)\right]^n.$$

Характеристическая функция биномиального распределения

$$\varphi_{X_1}(t) = Me^{itX_1} = [(e^{it} - 1)p + 1]^n.$$

Подставим и получим

$$\left[\varphi_{X_1}\left(\frac{1}{in}\right)\right]^n = \left[\left(e^{\frac{i}{in}} - 1\right)p + 1\right]^n = \left[\left(e^{\frac{1}{n}} - 1\right)p + 1\right]^n \neq e^{2p}.$$

Значит, оценка смещённая.

Проверим сильную состоятельность. По усиленному закону больших чисел

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{a.s.} MX_1 = 2p, n \rightarrow \infty.$$

Отсюда следует, что $e^{\bar{X}} \xrightarrow{a.s.} e^{2p}, n \rightarrow \infty$. Значит, оценка сильно состоятельная.

$$\text{Найдём дисперсию оценки } De^{\bar{X}} = Me^{2\bar{X}} - \left(Me^{\bar{X}}\right)^2.$$

$$\text{Нашли, что } Me^{\bar{X}} = \left[\left(e^{\frac{1}{n}} - 1\right)p + 1\right]^n.$$

Вычисляем $Me^{2\bar{X}} = Me^{\frac{2}{n} \sum_{i=1}^n X_i}$. Из независимости и одинаковой распределённости случайных величин следует, что $Me^{\frac{2}{n} \sum_{i=1}^n X_i} = \left(Me^{\frac{2X_1}{n}}\right)^n$. По определению характеристической функции $\varphi_{X_1}(t) = Me^{itX_1}$, получаем

$$\left(Me^{\frac{2X_1}{n}}\right)^n = \left[\varphi_{X_1}\left(\frac{2}{in}\right)\right]^n.$$

Подставляем характеристическую функцию биномиального распределения

$$\left[\varphi_{X_1}\left(\frac{2}{in}\right)\right]^n = \left[\left(e^{\frac{2i}{in}} - 1\right)p + 1\right]^n = \left[\left(e^{\frac{2}{n}} - 1\right)p + 1\right]^n.$$

Отсюда находим дисперсию оценки

$$De^{\bar{X}} = \left[\left(e^{\frac{2}{n}} - 1\right)p + 1\right]^n - \left[\left(e^{\frac{1}{n}} - 1\right)p + 1\right]^{2n}.$$

Извлекая корень, получим $\sigma = \sqrt{\left[\left(e^{\frac{2}{n}} - 1\right)p + 1\right]^n - \left[\left(e^{\frac{1}{n}} - 1\right)p + 1\right]^{2n}}$.

2.22

Задание. В партии из n изделий оказалось m бракованных. Незвестная вероятность p появления бракованного изделия оценивается величиной m/n . Проверьте состоятельность и несмещённость этой оценки.

Решение. Выборка имеет распределение Бернулли, то есть

$$X_i = \begin{cases} 1, & p, \\ 0, & 1 - p, \end{cases}$$

где событие $\{X_i = 1\}$ означает, что i -тое изделие браковано.

Тогда

$$\frac{m}{n} = \frac{X_1 + \dots + X_n}{n},$$

где $(X_1 + \dots + X_n)$ — количество бракованных изделий в партии.

Проверяем состоятельность, то есть проверяем, выполняется ли

$$\frac{m}{n} \xrightarrow{p} p, \quad n \rightarrow \infty.$$

По закону больших чисел

$$\frac{m}{n} = \frac{X_1 + \dots + X_n}{n} \xrightarrow{p} MX_1 = p, \quad n \rightarrow \infty,$$

так как это распределение Бернулли.

Это означает, что оценка состоятельная.

Проверим несмещённость оценки, то есть проверим, выполняется ли

$$M \frac{m}{n} = p.$$

Ищем математическое ожидание оценки

$$M \frac{m}{n} = M \frac{X_1 + \dots + X_n}{n} = \frac{1}{n} \cdot M(X_1 + \dots + X_n).$$

Пользуемся линейностью математического ожидания

$$\frac{1}{n} \cdot M(X_1 + \dots + X_n) = \frac{1}{n} \cdot nMX_1 = p,$$

то есть оценка несмещённая.

2.23

Задание. При каком значении k статистика

$$k \sum_{j=1}^n |X_j|,$$

которая построена по выборке из нормального распределения $N(0, \sigma^2)$, является несмещённой оценкой для параметра σ ?

Решение. Нужно найти такое k , для которого

$$M \left(k \sum_{j=1}^n |X_j| \right) = \sigma, \quad X_j \sim N(0, \sigma^2).$$

Выносим константу k за знак математического ожидания

$$M \left(k \sum_{j=1}^n |X_j| \right) = k M \sum_{j=1}^n |X_j|.$$

Пользуемся линейностью математического ожидания

$$k M \sum_{j=1}^n |X_j| = k \sum_{j=1}^n M |X_j|.$$

Из того, что случайные величины одинаково распределены, следует, что

$$k \sum_{j=1}^n M |X_j| = kn M |X_1|.$$

Найдём математическое ожидание модуля первого элемента выборки через плотность распределения

$$M |X_1| = \int_{\mathbb{R}} p(x) \cdot |x| dx = \int_{\mathbb{R}} |x| \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}} dx = 2 \cdot \int_0^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}} dx.$$

Внесём под знак дифференциала то, что стоит в степени экспоненты

$$2 \cdot \int_0^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}} dx = -2 \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot \frac{1}{2} \cdot \int_0^{+\infty} e^{-\frac{x^2}{2\sigma^2}} d \left(-\frac{x^2}{2\sigma^2} \right) \cdot 2\sigma^2 = \sqrt{\frac{2}{\pi}} \cdot \sigma.$$

Тогда математическое ожидание оценки

$$M \left(k \sum_{j=1}^n |X_j| \right) = kn\sigma \sqrt{\frac{2}{\pi}}$$

— условие несмещённости, откуда

$$k = \frac{1}{n} \sqrt{\frac{\pi}{2}}.$$

2.24

Задание. Пусть X_1, \dots, X_n — выборка из показательного распределения с параметром $\alpha > 0$. Найдите несмещённые состоятельные оценки для параметров

$$\alpha, \sin \frac{1}{\alpha}.$$

Решение. X_1, \dots, X_n — независимые одинаково распределённые. θ — неизвестный параметр.

$$\hat{\theta} = \frac{f(X_1) + \dots + f(X_n)}{n}.$$

Так как элементы выборки одинаково распределённые, то $M\hat{\theta} = Mf(X_1)$.

По закону больших чисел $\hat{\theta} \xrightarrow{P} Mf(X_1)$.

Хотим, чтобы

$$\begin{cases} M\hat{\theta} = \theta, \\ \hat{\theta} \xrightarrow{P} \theta. \end{cases}$$

Нужно подобрать f из условия, что $Mf(X_1) = \theta = \alpha$.

Записывая математическое ожидание через интеграл от плотности получаем преобразование Лапласа

$$\int_0^{\infty} f(y) \alpha e^{-\alpha y} dy = \alpha.$$

Это способ построения одновременно несмещённой и состоятельной оценки.

Выносим α из-под знака интеграла

$$\alpha \int_0^{\infty} f(y) e^{-\alpha y} dy = \alpha.$$

Сокращая константы, получаем

$$\int_0^{\infty} f(y) e^{-\alpha y} dy = 1.$$

Из таблицы преобразований Лапласа $f(y) = \delta(y)$. Тогда оценка параметра α примет вид

$$\hat{\theta} = \frac{\delta(X_1) + \dots + \delta(X_n)}{n}.$$

Для второго параметра

$$\int_0^{\infty} f(y) \alpha e^{-\alpha y} dy = \sin \frac{1}{\alpha}.$$

Вынесем α из-под знака интеграла

$$\alpha \int_0^{\infty} f(y) e^{-\alpha y} dy = \sin \frac{1}{\alpha}.$$

Перенесём α вправо

$$\int_0^{\infty} f(y) e^{-\alpha y} dy = \frac{1}{\alpha} \cdot \sin \frac{1}{y} \sim \frac{1}{\alpha} \cdot \frac{1}{\alpha} = \frac{1}{\alpha^2},$$

откуда из таблицы преобразований Лапласа $f(y) = y$. Перешли к пределу, потому что иначе интеграл расходится. Оценка параметра в таком случае примет вид

$$\hat{\theta} = \frac{X_1 + \dots + X_n}{n} = \bar{X}.$$

2.25

Задание. Пусть X — количество успехов в серии из n испытаний Бернулли с вероятностью успеха p . Докажите, что не существует несмещённой оценки для параметра p^N при $N > n$.

Решение. Будем действовать от противного.

Допустим, такая оценка существует, то есть существует $\hat{\theta}$ такое, что $M\hat{\theta} = p^N$.

Это условие несмещённости.

$\hat{\theta}$ — функция от выборки X_1, \dots, X_n , где $X_1 + \dots + X_n = X$, то есть $\hat{\theta} = f(X_1, \dots, X_n)$.

Распределение Бернулли дискретное

$$\begin{aligned} M\hat{\theta} &= Mf(X_1, \dots, X_n) = \\ &= \sum_{k_1, \dots, k_n=1 \text{ or } 0} f(X_1, \dots, X_n) P(X_1 = k_1, \dots, X_n = k_n). \end{aligned}$$

Пользуясь независимостью, получаем

$$\begin{aligned} &\sum_{k_1, \dots, k_n=1 \text{ or } 0} f(X_1, \dots, X_n) P(X_1 = k_1, \dots, X_n = k_n) = \\ &= \sum_{k_1, \dots, k_n=1 \text{ or } 0} f(X_1, \dots, X_n) \cdot P(X_1 = k_1) \cdot \dots \cdot P(X_n = k_n). \end{aligned}$$

Подставляем вероятности в явном виде

$$\begin{aligned} &\sum_{k_1, \dots, k_n=1 \text{ or } 0} f(X_1, \dots, X_n) P(X_1 = k_1) \cdot P(X_n = k_n) = \\ &= \sum_{k_1, \dots, k_n=1 \text{ or } 0} f(X_1, \dots, X_n) p^{k_1} \cdot (1-p)^{1-k_1} \cdot \dots \cdot p^{k_n} \cdot (1-p)^{1-k_n} = \\ &= \sum_{k_1, \dots, k_n=1 \text{ or } 0} f(X_1, \dots, X_n) \cdot p^{k_1 + \dots + k_n} \cdot (1-p)^{n - (k_1 + \dots + k_n)}. \end{aligned}$$

Обозначим

$$\sum k_i = k.$$

Получим

$$\begin{aligned} & \sum_{k_1, \dots, k_n=1 \text{ or } 0} f(X_1, \dots, X_n) \cdot p^{k_1+\dots+k_n} \cdot (1-p)^{n-(k_1+\dots+k_n)} = \\ & = \sum_{k=0}^n \sum_{k_1, \dots, k_n=1 \text{ or } 0} f(k_1, \dots, k_n) p^k (1-p)^{n-k} = \sum_{k=0}^n c_k p^k (1-p)^{n-k}, \end{aligned}$$

где c_k — число.

Допустим, эта величина равна p^N .

Посмотрим, возможно ли это

$$\sum_{k=0}^n c_k p^k (1-p)^{n-k} = p^N.$$

Последнее равенство невозможно при $N > n$. Из этого следует, что не существует несмещённой оценки для параметра p^N при $N > n$.

2.26

Задание. Проводится n измерений неизвестного диаметра d круга. В первом приближении считается, что измерения $X_i = d + \varepsilon_i$ проводятся с независимыми случайными погрешностями ε_i , которые имеют нормальное распределение с нулевым математическим ожиданием и неизвестной дисперсией σ^2 . Проверьте несмещённость и состоятельность следующей оценки площади круга:

$$\hat{s}_n = \frac{\pi}{4} \left((\bar{X})^2 - \frac{S_0^2}{n} \right),$$

где

$$S_0^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Решение. Имеем случайные величины $\varepsilon_i \sim N(0, \sigma^2)$. Проверим несмещённость оценки, то есть проверим, выполняется ли

$$M\hat{s}_n = \frac{\pi d^2}{4}.$$

Преобразуем

$$S_0^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (d + \varepsilon_i - d - \bar{\varepsilon})^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2.$$

Математическое ожидание этой случайной величины равно

$$MS_0^2 = DS_0^2 = \sigma^2.$$

Найдём математическое ожидание квадрата выборочного среднего

$$M(\bar{X})^2 = M\left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2 = M\left[\frac{1}{n^2} \cdot \left(\sum_{i=1}^n X_i^2 + 2 \sum_{i,j=1, i \neq j}^n X_i X_j\right)\right].$$

Пользуемся тем, что случайные величины в выборке одинаково распределены

$$M\left[\frac{1}{n^2} \cdot \left(\sum_{i=1}^n X_i^2 + 2 \sum_{i,j=1, i \neq j}^n X_i X_j\right)\right] = \frac{1}{n^2} \cdot n M X_1^2 + \frac{1}{n^2} \cdot 2 C_n^2 (M X_1)^2.$$

Расписываем биномиальный коэффициент

$$\frac{1}{n^2} \cdot n M X_1^2 + \frac{1}{n^2} \cdot 2 C_n^2 (M X_1)^2 = \frac{1}{n} \cdot M X_1^2 + \frac{2}{n^2} \cdot \frac{n!}{2(n-2)!} \cdot (M X_1)^2.$$

Расписываем $n!$ через факториал, стоящий в знаменателе

$$\frac{1}{n} \cdot M X_1^2 + \frac{2}{n^2} \cdot \frac{n!}{2(n-2)!} \cdot (M X_1)^2 = \frac{1}{n} \cdot M X_1^2 + \frac{2}{n^2} \cdot \frac{(n-2)!(n-1)n}{2(n-2)!} \cdot (M X_1)^2.$$

Сокращаем

$$\frac{1}{n} \cdot M X_1^2 + \frac{2}{n^2} \cdot \frac{(n-2)!(n-1)n}{2(n-2)!} \cdot (M X_1)^2 = \frac{1}{n} \cdot M X_1^2 + \frac{n-1}{n} \cdot (M X_1)^2.$$

Найдём математическое ожидание измерения

$$M X_1 = M(d + \varepsilon_i) = M d + M \varepsilon_i = d.$$

Найдём математическое ожидание квадрата измерения

$$M X_1^2 = M(d + \varepsilon_i)^2 = M(d^2 + 2d\varepsilon_i + \varepsilon_i^2) = M d^2 + M(2d\varepsilon_i) + M \varepsilon_i^2.$$

Выносим константы за знак математического ожидания и учитываем то, что случайная величина ε_i имеет нулевой математическое ожидание

$$M d^2 + M(2d\varepsilon_i) + M \varepsilon_i^2 = d^2 + 2d M \varepsilon_i + M \varepsilon_i^2 = d^2 + \sigma^2.$$

Подставляя эти значения в полученное выражение для математического ожидания квадрата выборочного среднего, получаем

$$M(\bar{X})^2 = \frac{1}{n} \cdot (d^2 + \sigma^2) + \frac{n-1}{n} \cdot d^2 = \frac{d^2 + \sigma^2 + n d^2 - d^2}{n} = \frac{\sigma^2 + n d^2}{n} = d^2 + \frac{\sigma^2}{n}.$$

Получаем математическое ожидание оценки

$$M \hat{s}_n = M \left\{ \frac{\pi}{4} \cdot \left[(\bar{X})^2 - \frac{S_0^2}{n} \right] \right\} = \frac{\pi}{4} \cdot M \left[(\bar{X})^2 - \frac{S_0^2}{n} \right] = \frac{\pi}{4} \cdot M(\bar{X})^2 - \frac{\pi}{4} \cdot M \frac{S_0^2}{n}.$$

Подставляем найденные значения математических ожиданий и выносим за скобку общий множитель

$$\frac{\pi}{4} \cdot M(\bar{X})^2 - \frac{\pi}{4} \cdot M \frac{S_0^2}{n} = \frac{\pi}{4} \cdot \left(d^2 + \frac{\sigma^2}{n} \right) - \frac{\pi}{4n} \cdot MS_0^2 = \frac{\pi}{4} \cdot d^2 + \frac{\pi \sigma^2}{4n} - \frac{\pi \sigma^2}{4n} = \frac{\pi d^2}{4},$$

значит, оценка несмещённая.

Проверим состоятельность оценки, то есть выполняется ли

$$\hat{s}_n \xrightarrow{P} \frac{\pi d^2}{4}, \quad n \rightarrow \infty.$$

По закону больших чисел

$$(\bar{X})^2 \xrightarrow{P} (MX_1)^2 = d^2, \quad n \rightarrow \infty.$$

Остаётся проверить, стремится ли по вероятности второе слагаемое в оценке к нулю

$$P\left(\frac{S_0^2}{n} \geq \varepsilon\right) = P(S_0^2 \geq \varepsilon n).$$

Применим неравенство Чебышева

$$P(S_0^2 \geq \varepsilon n) \leq \frac{MS_0^2}{\varepsilon n} \rightarrow 0, \quad n \rightarrow \infty.$$

Остюда следует, что

$$\frac{S_0^2}{n} \xrightarrow{P} 0, \quad n \rightarrow \infty.$$

Значит,

$$\hat{s}_n \xrightarrow{P} \frac{\pi}{4} \cdot d^2,$$

то есть оценка состоятельная.

Занятие 3. Метод моментов построения оценок

Контрольные вопросы и задания

Приведите определение оценки: несмещённой, асимптотически несмещённой, состоятельной, сильно состоятельной, оптимальной.

Оценка $\hat{\theta}$ несмещённая, если $\forall \theta \in \Theta : M_{\theta} \hat{\theta} = \theta$.

Асимптотически несмещённая оценка — такая оценка, математическое ожидание которой совпадает с оцениваемым параметром при $n \rightarrow \infty$.

Оценка $\hat{\theta}$ называется состоятельной, если стремится к истинному значению θ по вероятности $\hat{\theta} \xrightarrow{P} \theta, n \rightarrow \infty$.

Оценка $\hat{\theta}$ называется сильно состоятельной, если стремится к истинному значению θ почти наверное $\hat{\theta} \xrightarrow{a.s.} \theta, n \rightarrow \infty$.

Несмещённая оценка $\hat{\theta} \in K$ называется оптимальной в классе квадратично интегрируемых оценок K , если для всякой другой несмещённой оценки $\tilde{\theta} \in \Theta \forall \theta \in \Theta : D_{\theta} \hat{\theta} \leq D_{\theta} \tilde{\theta}$ или же $\forall \theta \in \Theta, M_{\theta} (\hat{\theta} - \theta)^2 \leq M_{\theta} (\tilde{\theta} - \theta)^2$.

Что называется среднеквадратическим отклонением оценки?

$M_{\theta} (\hat{\theta} - \theta)$ — среднеквадратическое отклонение.

Сформулируйте утверждение про поведение выборочных моментов.

Выборочный начальный момент M_k k -го порядка стремится к начальному моменту ν_k случайной величины X , то есть

$$\lim_{n \rightarrow \infty} P(|M_k - \nu_k| \geq \varepsilon) = 0,$$

для любого сколь угодно малого $\varepsilon > 0$, если моменты ν_{2k} и ν_k случайной величины X существуют и конечны.

Сформулируйте основную идею метода моментов построения оценки неизвестного параметра.

x_1, \dots, x_n — выборка из распределения F_θ и $\theta : M_\theta f(x_1) = g(\theta)$.

Вычисляем математическое ожидание, считая, что x_1 имеет распределение с параметром θ , иными словами,

$$M_\theta f(x) = \int_{\mathbb{R}} x dF_\theta(x), \quad f, g \in C(\mathbb{R}), \quad g$$

— строго монотонная. Тогда в силу усиленного закона больших чисел

$$\theta \approx g^{-1} \left(\frac{1}{n} \sum_{k=1}^n f(x_k) \right) \stackrel{a.s.}{\rightarrow} \theta, \quad n \rightarrow \infty.$$

Если существуют непрерывные f и g , g — обратима и $g(\theta) = M_\theta f(x)$, то в качестве оценки можно выбрать

$$\hat{\theta} = g^{-1} \left(\int_{\mathbb{R}} f dF_n \right).$$

Аудиторные задачи

3.3

Задание. Пользуясь методом моментов, оцените параметр θ равномерного распределения на отрезке:

- a) $[0, \theta]$;
- b) $[\theta - 1, \theta + 1]$;
- c) $[0, 2\theta]$;
- d) $[-\theta, \theta]$.

Решение.

- a) $X_i \sim U([0, \theta])$. Записываем теоретический момент. Для равномерного распределения — это середина отрезка

$$MX_1 = \frac{\theta}{2}.$$

Должны приравнять

$$\frac{\theta^*}{2} = \bar{X},$$

откуда $\theta^* = 2\bar{X}$;

b) случайные величины имеют распределение $X_i \sim U([\theta - 1, \theta + 1])$.

Вычисляем теоретический момент $MX_1 = \theta$. Должны записать, что $\theta^* = \bar{X}$;

с) случайные величины имеют распределение $X_i \sim U([0, 2\theta])$.

Вычисляем теоретический момент $MX_1 = \theta$, откуда $\theta^* = \bar{X}$;

d) случайные величины имеют распределение $X_i \sim U([- \theta, \theta])$.

Вычисляем теоретический момент $MX_1 = 0$ — не подходит, потому что не является функцией от θ . Можем вычислить второй момент, который в данном случае совпадает с дисперсией $MX_1^2 = DX_1$. Для равномерного распределения

$$DX_1 = \frac{4\theta^2}{12} = \frac{\theta^2}{3}.$$

Должны записать уравнение

$$\frac{(\theta^2)^*}{3} = \overline{X^2},$$

откуда $(\theta^2)^* = 3\overline{X^2}$. Извлекая корень, получаем $\theta^* = \sqrt{3\overline{X^2}}$.

3.4

Задание. Пусть X_1, \dots, X_n — выборка из распределения Пуассона с параметром λ . Пользуясь методом моментов, постройте оценку параметра λ и убедитесь, что эта оценка является несмещённой и состоятельной.

Решение. Первый теоретический момент $MX_1 = \lambda$. Должны приравнять

$$\lambda^* = \frac{1}{n} \sum_{i=1}^n X_i.$$

По закону больших чисел $\lambda^* \xrightarrow{P} \lambda$. Отсюда следует состоятельность. $M\lambda^* = MX_1 = \lambda$. Отсюда следует несмещённость.

3.5

Задание. Пользуясь методом моментов с пробной функцией $g(y) = y$, оцените параметр сдвига $\beta \in \mathbb{R}$ показательного распределения с плотностью

$$f_\beta(y) = \begin{cases} e^{\beta-y}, & t \geq \beta, \\ 0, & y < \beta. \end{cases}$$

Решение. $Mg(X_1) = \overline{g(X)}$ — метод моментов.

Ищем теоретический момент, зная плотность распределения

$$MX_1 = \int_{\beta}^{\infty} ye^{\beta-y} dy.$$

Делаем замену $y - \beta = z$ и получаем

$$\int_{\beta}^{\infty} ye^{\beta-y} dy = \int_0^{\infty} (\beta + z) e^{-z} dz = \beta + 1.$$

Отсюда должны решить уравнение $\beta^* + 1 = \bar{X}$.

Отсюда $\beta^* = \bar{X} - 1$. Эта оценка несмещённая и состоятельная.

3.6

Задание. Пусть X_1, \dots, X_n — выборка из биномиального распределения с параметрами m и p . Пользуясь методом моментов, постройте оценку:

- а) параметра p , если параметр m известный;
- б) параметра m , если параметр p известный;
- с) векторного параметра (m, p) .

Решение.

- а) $MX_1 = mp$. Тогда уравнение имеет вид $mp^* = \bar{X}$. Отсюда

$$p^* = \frac{\bar{X}}{m};$$

- б) параметр m — это целое число.

$m^*p = \bar{X}$, откуда

$$m^* = \left[\frac{\bar{X}}{p} \right]$$

— целая часть;

- с) одного теоретического момента мало.

Нужно найти второй теоретический момент.

Выражаем его через дисперсию и первый момент

$$MX_1^2 = DX_1 + (MX_1)^2 = mp(1-p) + (mp)^2 = mp(1-p+mp).$$

Составляем уравнения с двумя неизвестными

$$\begin{cases} m^*p^* = \bar{X}, \\ m^*p^*(1-p^* + m^*p^*) = \overline{X^2}. \end{cases}$$

Во второе уравнение можем подставить \bar{X} вместо произведения $m^* p^*$ и решить $\bar{X}(1 - p^* + \bar{X}) = \bar{X}^2$, откуда

$$1 - p^* + \bar{X} = \frac{\bar{X}^2}{\bar{X}}.$$

Отсюда находим

$$p^* = 1 + \bar{X} - \frac{\bar{X}^2}{\bar{X}},$$

соответственно

$$m^* = \left[\frac{\bar{X}}{p^*} \right] = \left[\frac{\bar{X}}{1 + \bar{X} - \frac{\bar{X}^2}{\bar{X}}} \right] = \left[\frac{(\bar{X})^2}{\bar{X} + \bar{X}^2 - \bar{X}^2} \right].$$

3.7

Задание. Пусть X_1, \dots, X_n — выборка из распределения Парето с плотностью

$$f_{\beta, \theta}(y) = \begin{cases} \beta \theta y^{-(\beta+1)}, & y \geq \theta, \\ 0, & t < \theta, \end{cases} \quad \beta > 0, \theta > 0.$$

Постройте оценки по методу моментов

- а) параметра $\beta > 1$, если параметр $\theta > 0$ известный;
- б) параметра $\theta > 0$, если параметр $\beta > 1$ известный;
- в) векторного параметра (β, θ) , где $\beta > 2$, $\theta > 0$.

Решение.

- а) Вычисляем теоретический момент

$$MX_1 = \beta \int_{\theta}^{\infty} \theta^{\beta} y^{-\beta} dy = \beta \theta^{\beta} \cdot \frac{y^{-\beta+1}}{-\beta+1} = -\beta \theta^{\beta} \cdot \frac{\theta}{1-\beta} = \frac{\beta}{\beta-1} \cdot \theta.$$

Найдём второй момент

$$MX_1^2 = \beta \int_{\theta}^{\infty} \theta^{\beta} y^{-\beta+1} dy = \beta \cdot \frac{1}{\beta-2} \cdot \theta^2.$$

Приравниваем теоретический и выборочный момент, учитывая, что β — параметр

$$\frac{\beta^*}{\beta^* - 1} \cdot \theta = \bar{X},$$

откуда

$$\frac{\beta^*}{\beta^* - 1} = \frac{\overline{X}}{\theta}.$$

Решим пропорцию $\beta^* \theta = (\beta^* - 1) \overline{X} = \beta^* \overline{X} - \overline{X}$.

Получаем

$$\beta^* = \frac{\overline{X}}{\overline{X} - \theta};$$

b) считаем, что β — известно, ищем θ .

Приравниваем моменты

$$\frac{\beta}{\beta - 1} \cdot \theta = \overline{X},$$

откуда

$$\theta^* = \frac{\overline{X} (\beta - 1)}{\beta};$$

c) нужно 2 уравнения

$$\begin{cases} \frac{\beta^*}{\beta^* - 1} \cdot \theta^* = \overline{X}, \\ \frac{\beta^*}{\beta^* - 2} \cdot (\theta^*)^2 = \overline{X^2}. \end{cases}$$

Выражаем θ^* из первого уравнения и подставляем во второе

$$\frac{\beta^*}{\beta^* - 2} \cdot \left(\frac{\beta^* - 1}{\beta^*} \right)^2 \cdot (\overline{X})^2 = \overline{X^2}.$$

Умножим на знаменатель $(\overline{X})^2 (\beta^* - 1)^2 = \overline{X^2} \beta^* (\beta^* - 2)$.

Оставляем в левой части только квадрат выборочного момента

$$(\overline{X})^2 = \overline{X^2} \cdot \frac{(\beta^*)^2 - 2\beta^*}{(\beta^* - 1)^2}.$$

Делаем замену $(\beta^* - 1)^2 = t$.

Получаем

$$(\overline{X})^2 = \frac{t - 1}{t} \cdot \overline{X^2} = \left(1 - \frac{1}{t} \right) \cdot \overline{X^2} = \overline{X^2} - \frac{\overline{X^2}}{t}.$$

Отсюда

$$t = \frac{\overline{X^2}}{\overline{X^2} - (\overline{X})^2}.$$

Отсюда найдём $\beta^* = \sqrt{t} + 1$, соответственно находим θ^* .

Домашнее задание

3.10

Задание. Пользуясь методом моментов, оцените значение α по выборке из показательного распределения с параметром $1/\sqrt{\alpha}$.

Решение.

$$X_i \sim \text{Exp}\left(\frac{1}{\sqrt{\alpha}}\right).$$

Записываем теоретический момент. Для показательного распределения — это обратная величина к параметру распределения

$$MX_1 = \frac{1}{\frac{1}{\sqrt{\alpha}}} = \sqrt{\alpha}.$$

Должны приравнять $\sqrt{\alpha^*} = \bar{X}$, откуда $\alpha^* = (\bar{X})^2$.

3.11

Задание. Пусть X_1, \dots, X_n — выборка из показательного распределения с параметром $\alpha > 0$. Пользуясь методом моментов с пробной функцией $g(y) = y$, оцените параметр $\theta(\alpha) = P(X_1 > 1)$.

Решение. $Mg(X_1) = g(X)$ — метод моментов.

Ищем теоретический момент, зная плотность распределения

$$MX_1 = \frac{1}{\alpha}.$$

Из метода моментов следует, что

$$\frac{1}{\alpha^*} = \bar{X}.$$

Выражаем оценку

$$\alpha^* = \frac{1}{\bar{X}}.$$

Перейдём в параметре к противоположному событию

$$\theta(\alpha) = P(X_1 > 1) = 1 - P(X_1 \leq 1) = 1 - F_{X_1}(1).$$

Найдём функцию показательного распределения как интеграл от плотности

$$F_{X_1}(y) = \int_0^y \alpha e^{-\alpha x} dx = \alpha \cdot \left(-\frac{1}{\alpha}\right) \cdot \int_0^y e^{-\alpha x} d(-\alpha x) = -e^{-\alpha x} \Big|_0^y = 1 - e^{-\alpha y}.$$

Тогда параметр примет вид $\theta(\alpha) = 1 - (1 - e^{-\alpha}) = 1 - 1 + e^{-\alpha} = e^{-\alpha} = e^{-\frac{1}{\bar{X}}}$.

3.12

Задание. Пользуясь методом моментов с пробной функцией $g(y) = y$, оцените параметр p распределения Бернулли. Возможно ли с помощью метода моментов с некоторой пробной функцией $g(y)$ получить оценку параметра p , отличную от \bar{X} ?

Решение.

$$X_i = \begin{cases} 0, & 1-p, \\ 1, & p. \end{cases}$$

Теоретический момент $MX_1 = p$, $p^* = \bar{X}$.

Найдём теоретический момент с пробной функцией

$$Mg(X_1) = g(0)(1-p) + g(0)p = g(0) + p[g(1) - g(0)] = \overline{g(\bar{X})}.$$

Отсюда

$$p^* = \frac{g(\bar{X}) - g(0)}{g(1) - g(0)} = \frac{g(X_1) + \dots + g(X_n) - g(0)}{n[g(1) - g(0)]}.$$

Количество единиц равно $\bar{X}n$, соответственно $n - \bar{X}n$ — количество нулей.

$$\frac{g(X_1) + \dots + g(X_n) - g(0)}{n[g(1) - g(0)]} = \frac{g(1)\bar{X}n + g(0)(n - \bar{X}n) - g(0)n}{n[g(1) - g(0)]}.$$

Раскроем скобки и приведём подобные

$$\frac{g(1)\bar{X}n + g(0)(n - \bar{X}n) - g(0)n}{n[g(1) - g(0)]} = \frac{ng(1)\bar{X} - gn\bar{X}}{n[g(1) - g(0)]} = \bar{X},$$

откуда следует, что невозможно получить другую оценку параметра p .

3.13

Задание. Пользуясь методом моментов, оцените параметр $\lambda > 1$ по выборке из распределения Пуассона с параметром $\ln \lambda$.

Решение. Первый теоретический момент $MX_1 = \ln \lambda$. Должны приравнять $\ln \lambda = \bar{X}$. Отсюда следует, что $\lambda^* = e^{\bar{X}}$.

3.14

Задание. Пусть X_1, \dots, X_n — выборка из Γ -распределения с плотностью

$$f_{\alpha, \beta}(y) = \begin{cases} \frac{\alpha^\beta}{\Gamma(\beta)} \cdot y^{\beta-1} e^{-\alpha y}, & y \geq 0, \\ 0, & y < 0, \end{cases} \quad \alpha, \beta > 0.$$

Постройте оценки по методу моментов

- а) параметра α , если параметр β известный;
- б) параметра β , если параметр α известный;

с) векторного параметра (α, β) .

Решение.

а) Вычислим теоретический момент

$$MX_1 = \int_0^{+\infty} \frac{\alpha^\beta}{\Gamma(\beta)} \cdot y^{\beta-1} \cdot e^{-\alpha y} \cdot y dy = \frac{1}{\Gamma(\beta)} \cdot \int_0^{+\infty} \alpha^\beta y^\beta e^{-\alpha y} dy.$$

Сделаем замену $\alpha y = s$, откуда

$$y = \frac{s}{\alpha}, \quad dy = \frac{ds}{\alpha}.$$

Подставляем

$$\frac{1}{\Gamma(\beta)} \cdot \int_0^{+\infty} \alpha^\beta y^\beta e^{-\alpha y} dy = \frac{1}{\Gamma(\beta)} \cdot \int_0^{+\infty} s^\beta e^{-s} \cdot \frac{ds}{\alpha} = \frac{1}{\alpha \Gamma(\beta)} \cdot \int_0^{+\infty} s^\beta \cdot e^{-s} ds.$$

Интеграл равен Гамма-функции

$$\frac{1}{\alpha \Gamma(\beta)} \cdot \int_0^{+\infty} s^\beta \cdot e^{-s} ds = \frac{1}{\alpha \Gamma(\beta)} \cdot \Gamma(\beta + 1) = \frac{1}{\alpha \Gamma(\beta)} \cdot \beta \cdot \Gamma(\beta) = \frac{\beta}{\alpha}.$$

Найдём второй момент

$$MX_1^2 = \int_0^{+\infty} \frac{\alpha^\beta}{\Gamma(\beta)} \cdot y^{\beta-1} e^{-\alpha y} \cdot y^2 dy = \frac{1}{\Gamma(\beta)} \cdot \int_0^{+\infty} \alpha^\beta y^{\beta+1} e^{-\alpha y} dy.$$

Умножаем на α числитель и знаменатель

$$\frac{1}{\Gamma(\beta)} \cdot \int_0^{+\infty} \alpha^\beta y^{\beta+1} e^{-\alpha y} dy = \frac{1}{\Gamma(\beta) \alpha} \cdot \int_0^{+\infty} \alpha^{\beta+1} y^{\beta+1} e^{-\alpha y} dy.$$

Сделаем замену как и в предыдущем случае

$$\frac{1}{\Gamma(\beta) \alpha} \cdot \int_0^{+\infty} \alpha^{\beta+1} y^{\beta+1} e^{-\alpha y} dy = \frac{1}{\Gamma(\beta) \alpha^2} \cdot \int_0^{+\infty} s^{\beta+1} e^{-s} ds = \frac{\Gamma(\beta + 2)}{\alpha^2 \Gamma(\beta)}.$$

Пользуясь свойством Гамма-функции, записываем

$$\frac{\Gamma(\beta + 2)}{\alpha^2 \Gamma(\beta)} = \frac{(\beta + 1) \Gamma(\beta + 1)}{\alpha^2 \Gamma(\beta)} = \frac{\Gamma(\beta + 1) \beta + \Gamma(\beta + 1)}{\alpha^2 \Gamma(\beta)}.$$

Снова пользуемся свойством Гамма-функции и сокращаем её

$$\frac{\Gamma(\beta+1)\beta + \Gamma(\beta+1)}{\alpha^2\Gamma(\beta)} = \frac{\beta^2\Gamma(\beta) + \beta\Gamma(\beta)}{\alpha^2\Gamma(\beta)} = \frac{\beta^2 + \beta}{\alpha^2}.$$

Приравниваем теоретический момент и выборочный момент, учитывая, что α — параметр

$$\frac{\beta}{\alpha^*} = \bar{X},$$

откуда

$$\alpha^* = \frac{\beta}{\bar{X}};$$

б) считаем, что α — известно, ищем β^* .

Приравниваем моменты

$$\frac{\beta^*}{\alpha} = \bar{X},$$

откуда $\beta^* = \alpha\bar{X}$;

с) нужно 2 уравнения

$$\begin{cases} \frac{\beta^*}{\alpha^*} = \bar{X}, \\ \frac{(\beta^*)^2 + \beta^*}{(\alpha^*)^2} = \bar{X}^2. \end{cases}$$

Выражаем α^* из первого уравнения и подставляем во второе

$$\frac{(\beta^*)^2 + \beta^*}{\beta^*} \cdot \bar{X} = \bar{X}^2.$$

Поделим числитель на знаменатель

$$\left(1 + \frac{1}{\beta^*}\right) \cdot (\bar{X})^2 = \bar{X}^2.$$

Оставляем в левой части только скобку

$$1 + \frac{1}{\beta^*} = \frac{\bar{X}^2}{(\bar{X})^2}.$$

Переносим единицу вправо

$$\frac{1}{\beta^*} = \frac{\bar{X}^2}{(\bar{X})^2} - 1.$$

Выражаем параметр

$$\beta^* = \frac{1}{\frac{\bar{X}^2}{(\bar{X})^2} - 1} = \frac{(\bar{X})^2}{\bar{X}^2 - (\bar{X})^2}.$$

Соответственно находим

$$\alpha^* = \frac{\beta^*}{\bar{X}} = \frac{\frac{(\bar{X})^2}{\bar{X}^2 - (\bar{X})^2}}{\bar{X}} = \frac{(\bar{X})^2}{(\bar{X}^2 - (\bar{X})^2) \bar{X}} = \frac{\bar{X}}{\bar{X}^2 - (\bar{X})^2}.$$

3.15

Задание. Пусть X_1, \dots, X_n — выборка из биномиального распределения с параметрами 2, p . Пользуясь методом моментов с пробной функцией

$$g(y) = y,$$

оцените параметр $\theta = e^{2p}$.

Решение. $Mg(X_1) = g(\bar{X})$ — метод моментов.

Ищем теоретический момент, зная распределение

$$MX_1 = \sum_{k=0}^2 k \cdot P(X_1 = k) = \sum_{k=0}^2 k \cdot C_2^k p^k (1-p)^{2-k}.$$

При $k = 0$ первое слагаемое зануляется, поэтому можно суммировать от единицы

$$\sum_{k=0}^2 k \cdot C_2^k p^k (1-p)^{2-k} = \sum_{k=1}^2 k \cdot C_2^k \cdot p^k (1-p)^{2-k} = 1 \cdot C_2^1 p (1-p) + 2 \cdot C_2^2 \cdot p^2.$$

Распишем биномиальные коэффициенты, раскроем скобки и приведём подобные

$$1 \cdot C_2^1 p (1-p) + 2 \cdot C_2^2 \cdot p^2 = 2p(1-p) + 2p^2 = 2p - 2p^2 + 2p^2 = 2p.$$

Должны решить уравнение $2p^* = \bar{X}$.

Отсюда

$$p^* = \frac{\bar{X}}{2}.$$

Оценим параметр $\theta^* = e^{2p^*} = e^{\frac{2\bar{X}}{2}} = e^{\bar{X}}$.

3.16

Задание. При нейтронном бомбардировании ядер урана начинается расщепление ядра, при котором ядро урана распадается на две части разного типа. В камере Вильсона это явление наблюдается в виде двух траекторий, которые выходят из одной точки. Эти траектории вскоре разделяются на несколько веток, которые получаются от столкновения частиц с молекулами газа в камере. Можно показать, что количество веток в одной траектории имеет так называемое «двойное» распределение Пуассона, то есть

$$P(X = k) = \frac{1}{2} \cdot \left(\frac{\lambda_1^k}{k!} \cdot e^{-\lambda_1} + \frac{\lambda_2^k}{k!} \cdot e^{-\lambda_2} \right), \quad k = 0, 1, \dots,$$

где $\lambda_1 < \lambda_2$ — некоторые положительные постоянные. Пользуясь методом моментов, оцените векторный параметр (λ_1, λ_2) .

Решение.

$$MX = \sum_{k=0}^{\infty} k \cdot P(X = k) = \sum_{k=0}^{\infty} k \cdot \frac{1}{2} \cdot \left(\frac{\lambda_1^k}{k!} \cdot e^{-\lambda_1} + \frac{\lambda_2^k}{k!} \cdot e^{-\lambda_2} \right).$$

Разобьём на две суммы

$$\sum_{k=0}^{\infty} k \cdot \frac{1}{2} \cdot \left(\frac{\lambda_1^k}{k!} \cdot e^{-\lambda_1} + \frac{\lambda_2^k}{k!} \cdot e^{-\lambda_2} \right) = \frac{1}{2} \cdot \sum_{k=0}^{\infty} k \cdot \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_1} + \frac{1}{2} \cdot \sum_{k=0}^{\infty} k \cdot \frac{\lambda_2^k}{k!} \cdot e^{-\lambda_2}.$$

Слагаемые при $k = 0$ равны нулю

$$\frac{1}{2} \cdot \sum_{k=0}^{\infty} k \cdot \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_1} + \frac{1}{2} \cdot \sum_{k=0}^{\infty} k \cdot \frac{\lambda_2^k}{k!} \cdot e^{-\lambda_2} = \frac{1}{2} \cdot \sum_{k=1}^{\infty} k \cdot \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_1} + \frac{1}{2} \cdot \sum_{k=1}^{\infty} k \cdot \frac{\lambda_2^k}{k!} \cdot e^{-\lambda_2}.$$

Вынесем λ_i за скобку и сократим k , где $i \in \{1, 2\}$. Получим

$$\begin{aligned} & \frac{1}{2} \cdot \sum_{k=1}^{\infty} k \cdot \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_1} + \frac{1}{2} \cdot \sum_{k=1}^{\infty} k \cdot \frac{\lambda_2^k}{k!} \cdot e^{-\lambda_2} = \\ &= \frac{\lambda_1}{2} \cdot \sum_{k=1}^{\infty} \frac{\lambda_1^{k-1}}{(k-1)!} \cdot e^{-\lambda_1} + \frac{\lambda_2}{2} \cdot \sum_{k=1}^{\infty} \frac{\lambda_2^{k-1}}{(k-1)!} \cdot e^{-\lambda_2} = \\ &= \frac{\lambda_1}{2} \cdot e^{\lambda_1} \cdot e^{-\lambda_1} + \frac{\lambda_2}{2} \cdot e^{\lambda_2} \cdot e^{-\lambda_2} = \frac{\lambda_1}{2} + \frac{\lambda_2}{2} = \frac{\lambda_1 + \lambda_2}{2}. \end{aligned}$$

Одного теоретического момента мало.

Нужно найти второй теоретический момент

$$MX^2 = \sum_{k=0}^{\infty} k^2 \cdot \frac{1}{2} \cdot \left(\frac{\lambda_1^k}{k!} \cdot e^{-\lambda_1} + \frac{\lambda_2^k}{k!} \cdot e^{-\lambda_2} \right).$$

Одно k сокращаем, записываем через две суммы

$$\begin{aligned}
& \sum_{k=0}^{\infty} k^2 \cdot \frac{1}{2} \cdot \left(\frac{\lambda_1^k}{k!} \cdot e^{-\lambda_1} + \frac{\lambda_2^k}{k!} \cdot e^{-\lambda_2} \right) = \\
& = \frac{1}{2} \cdot \sum_{k=1}^{\infty} k \cdot \frac{\lambda_1^k}{(k-1)!} \cdot e^{-\lambda_1} + \frac{1}{2} \cdot \sum_{k=1}^{\infty} k \cdot \frac{\lambda_2^k}{(k-1)!} \cdot e^{-\lambda_2} = \\
& = \frac{1}{2} \cdot \sum_{k=1}^{\infty} (k-1) \cdot \frac{\lambda_1^k}{(k-1)!} \cdot e^{-\lambda_1} + \frac{1}{2} \cdot \sum_{k=1}^{\infty} (k-1) \cdot \frac{\lambda_2^k}{(k-1)!} \cdot e^{-\lambda_2} + \\
& \quad + \frac{1}{2} \cdot \sum_{k=1}^{\infty} \frac{\lambda_1^k}{(k-1)!} \cdot e^{-\lambda_1} + \frac{1}{2} \cdot \sum_{k=1}^{\infty} \frac{\lambda_2^k}{(k-1)!} \cdot e^{-\lambda_2} = \\
& = \frac{1}{2} \cdot \sum_{k=2}^{\infty} (k-1) \cdot \frac{\lambda_1^k}{(k-1)!} \cdot e^{-\lambda_1} + \frac{1}{2} \cdot \sum_{k=2}^{\infty} (k-1) \cdot \frac{\lambda_2^k}{(k-1)!} \cdot e^{-\lambda_2} + \\
& \quad + \frac{1}{2} \cdot \lambda_1 \cdot e^{-\lambda_1} \cdot \sum_{k=1}^{\infty} \frac{\lambda_1^{k-1}}{(k-1)!} + \frac{1}{2} \cdot \lambda_2 \cdot e^{-\lambda_2} \cdot \sum_{k=1}^{\infty} \frac{\lambda_2^{k-1}}{(k-1)!} = \\
& = \frac{1}{2} \cdot \sum_{k=2}^{\infty} \frac{\lambda_1^k}{(k-2)!} \cdot e^{-\lambda_1} + \frac{1}{2} \cdot \sum_{k=2}^{\infty} \frac{\lambda_2^k}{(k-2)!} \cdot e^{-\lambda_2} + \frac{1}{2} \cdot \lambda_1 \cdot e^{-\lambda_1} \cdot e^{\lambda_1} + \\
& \quad + \frac{1}{2} \cdot \lambda_2 \cdot e^{-\lambda_2} \cdot e^{\lambda_2} = \\
& = \frac{1}{2} \cdot \lambda_1^2 \cdot e^{-\lambda_1} \cdot \sum_{k=2}^{\infty} \frac{\lambda_1^{k-2}}{(k-2)!} + \frac{1}{2} \cdot \lambda_1^2 \cdot e^{-\lambda_1} \cdot \sum_{k=2}^{\infty} \frac{\lambda_1^{k-2}}{(k-2)!} + \frac{1}{2} \cdot \lambda_1 + \frac{1}{2} \cdot \lambda_2 = \\
& = \frac{1}{2} \cdot \lambda_1^2 \cdot e^{-\lambda_1} \cdot e^{\lambda_1} + \frac{1}{2} \cdot \lambda_2^2 \cdot e^{-\lambda_2} \cdot e^{\lambda_2} + \frac{1}{2} (\lambda_1 + \lambda_2) = \frac{1}{2} (\lambda_1^2 + \lambda_2^2 + \lambda_1 + \lambda_2).
\end{aligned}$$

Составим уравнения с двумя неизвестными

$$\begin{cases} \frac{\lambda_1^* + \lambda_2^*}{2} = \overline{X}, \\ \frac{1}{2} \cdot ((\lambda_1^*)^2 + (\lambda_2^*)^2 + \lambda_1^* + \lambda_2^*) = \overline{X^2}. \end{cases}$$

Возведём первое уравнение в квадрат, предварительно умножив на 2. Получим $(\lambda_1^* + \lambda_2^*)^2 = (\lambda_1^*)^2 + 2\lambda_1^*\lambda_2^* + (\lambda_2^*)^2 = 4(\overline{X})^2$, откуда

$$(\lambda_1^*)^2 + (\lambda_2^*)^2 = 4(\overline{X})^2 - 2\lambda_1^*\lambda_2^*.$$

Подставим это выражение во второе уравнение

$$\frac{1}{2} \cdot [4(\overline{X})^2 - 2\lambda_1^*\lambda_2^* + \lambda_1^* + \lambda_2^*] = \overline{X^2}.$$

Раскроем скобки

$$2(\overline{X})^2 - \lambda_1^*\lambda_2^* + \frac{1}{2}(\lambda_1^* + \lambda_2^*) = \overline{X^2}.$$

Подставим сумму оценок из первого уравнения

$$2(\overline{X})^2 - \lambda_1^* \lambda_2^* + \frac{1}{2} \cdot 2\overline{X} = \overline{X^2}.$$

Выразим произведение оценок и сократим двойки $\lambda_1^* \lambda_2^* = 2(\overline{X})^2 + \overline{X} - \overline{X^2}$.

Из первого уравнения подставим $\lambda_2^* = 2\overline{X} - \lambda_1^*$ в полученное уравнение $\lambda_1^* (2\overline{X} - \lambda_1^*) = 2(\overline{X})^2 + \overline{X} - \overline{X^2}$.

Раскроем скобки и поменяем знаки в обеих частях

$$(\lambda_1^*)^2 - 2\overline{X}\lambda_1^* = -2(\overline{X})^2 - \overline{X} + \overline{X^2}.$$

Перенесём всё влево $(\lambda_1^*)^2 - 2\overline{X}\lambda_1^* + 2(\overline{X})^2 + \overline{X} - \overline{X^2} = 0$.

Получили квадратное уравнение относительно λ_1^* . Найдём его дискриминант

$$D = 4(\overline{X})^2 - 4(2(\overline{X})^2 + \overline{X} - \overline{X^2}) = 4(\overline{X})^2 - 8(\overline{X})^2 - 4\overline{X} + 4\overline{X^2}.$$

Приведём подобные

$$4(\overline{X})^2 - 8(\overline{X})^2 - 4\overline{X} + 4\overline{X^2} = 4\overline{X^2} - 4(\overline{X})^2 - 4\overline{X}.$$

Найдём корень

$$\lambda_1^* = \frac{2\overline{X} - 2\sqrt{\overline{X^2} - (\overline{X})^2 - \overline{X}}}{2} = \overline{X} - \sqrt{\overline{X^2} - (\overline{X})^2 - \overline{X}}.$$

Тогда оценка второго параметра равна $\lambda_2^* = \overline{X} + \sqrt{\overline{X^2} - (\overline{X})^2 - \overline{X}}$.

Занятие 4. Оценка максимального правдоподобия

Контрольные вопросы и задания

Как построить оценку максимального правдоподобия?

Плотность распределения выборки (X_1, \dots, X_n) имеет вид

$$L(\vec{X}, \theta) = \prod_{k=1}^n p(X_k, \theta),$$

как плотность вектора с независимыми координатами. $L(\vec{X}, \theta)$ — функция правдоподобия.

Оценка максимального правдоподобия $\hat{\theta}$ — такое значение параметра θ , при котором функция правдоподобия достигает своего максимального значения $\hat{\theta} = \underset{\theta}{\operatorname{argmax}} L(\vec{X}, \theta)$.

Сформулируйте свойства оценки максимального правдоподобия.

Оценка максимального правдоподобия, как правило, сильно состоятельная.

Аудиторные задачи

4.3

Задание. Постройте оценку максимального правдоподобия векторного параметра (a, σ^2) нормального распределения.

Решение. Выборка $x_1, \dots, x_n \sim N(a, \sigma^2)$. Сейчас θ будет векторным параметром, состоящим из двух элементов: a, σ^2 .

Плотность имеет вид

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Произведение этих плотностей даёт

$$L(\vec{x}, a, \sigma^2) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \cdot e^{-\frac{1}{2\sigma^2} \cdot \sum_{i=1}^n (x_i - a)^2}.$$

Обозначим $\sigma^2 = \sigma^*$, чтобы было удобней брать производную

$$\frac{1}{(2\pi)^{\frac{n}{2}} \sigma^2} \cdot e^{-\frac{1}{2\sigma^2} \cdot \sum_{i=1}^n (x_i - a)^2} = \frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^*)^{\frac{n}{2}}} \cdot e^{-\frac{1}{2\sigma^*} \cdot \sum_{i=1}^n (x_i - a)^2}.$$

Поскольку есть экспонента, стоит брать логарифм

$$l(\vec{x}, a, \sigma^*) = \ln L(\vec{x}, a, \sigma^*) = -\frac{n}{2} \cdot \ln(2\pi\sigma^*) - \frac{1}{2\sigma^*} \sum_{i=1}^n (x_i - a)^2.$$

Распишем логарифм произведения через сумму логарифмов

$$-\frac{n}{2} \cdot \ln(2\pi\sigma^*) - \frac{1}{2\sigma^*} \sum_{i=1}^n (x_i - a)^2 = -\frac{n}{2} \cdot \ln(2\pi) - \frac{n}{2} \cdot \ln\sigma^* - \frac{1}{2\sigma^*} \sum_{i=1}^n (x_i - a)^2.$$

Берём производную

$$\frac{\partial}{\partial \sigma^*} l = -\frac{n}{2} \cdot \frac{1}{\sigma^*} + \frac{1}{2(\sigma^*)^2} \sum_{i=1}^n (x_i - a)^2 = 0.$$

Теперь надо по a продифференцировать и тоже приравнять к нулю

$$\frac{\partial}{\partial a} l = \frac{1}{\sigma^*} \sum_{i=1}^n (x_i - a) = 0 = \frac{n\bar{X} - na}{\sigma^*}.$$

Распишем сумму в явном виде $(x_1 - a) + (x_2 - a) + \dots + (x_n - a) = 0$.

Отсюда

$$a = \frac{1}{n} \sum_{i=1}^n x_i.$$

Оценка равна $\hat{a} = \bar{x}$.

Тогда оценка второго параметра равна

$$\hat{\sigma}^* = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Проверили только необходимое условие для максимума. Теперь нужно проверить достаточность.

Имеем матрицу $n \times n$. Матрица положительно определена тогда и только тогда, когда каждый минор положительный. Матрица отрицательно определена тогда и только тогда, когда каждый минор отрицательный.

$$\begin{bmatrix} \frac{\partial^2}{\partial a^2} & \frac{\partial^2}{\partial a \partial \sigma^*} \\ \frac{\partial^2}{\partial a \partial \sigma^*} & \frac{\partial^2}{\partial (\sigma^*)^2} \end{bmatrix}.$$

Найдём последний элемент матрицы

$$\frac{\partial^2}{\partial (\sigma^*)^2} l = \frac{n}{2 (\sigma^*)^2} - \frac{1}{(\sigma^*)^3} \cdot \sum_{i=1}^n (x_i - a)^2.$$

Найдём смешанную производную

$$\frac{\partial^2}{\partial a \partial \sigma^*} l = - \sum_{i=1}^n (x_i - a) \cdot \frac{1}{(\sigma^*)^2} = \frac{n\bar{x} - na}{(\sigma^*)^2}.$$

Найдём первый элемент матрицы

$$\frac{\partial^2}{\partial a^2} l = - \frac{n}{\sigma^*}.$$

Теорема Шварца. Если

$$\exists \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}$$

и они непрерывны в точке (x_0, y_0) , то они совпадают в этой точке.

Первый определитель строго отрицательный.

Второй определитель

$$M_2 = - \frac{n^2}{2 (\sigma^*)^3} + \frac{n}{(\sigma^*)^4} \cdot \sum_{i=1}^n (x_i - a) - \frac{n}{(\sigma^*)^4} \cdot (\bar{x} - a)^2.$$

Последнее слагаемое равно нулю. Нужно проверять это значение в точке $\hat{a} = \bar{x}$. Имеем

$$\begin{aligned} & - \frac{n^2}{2 (\sigma^*)^3} + \frac{n}{(\sigma^*)^4} \cdot \sum_{i=1}^n (x_i - a) - \frac{n}{(\sigma^*)^4} \cdot (\bar{x} - a)^2 = \\ & = - \frac{n}{(\sigma^*)^3} \cdot \left(\frac{n}{2} - \frac{1}{\sigma^*} \cdot \sum_{i=1}^n (x_i - a) \right) > 0. \end{aligned}$$

Нужно проверить в точке $\sigma^* = \hat{\sigma}$. Имеем

$$\frac{n}{2} < \frac{\sum_{i=1}^n (x_i - a)^2}{\sigma^*}.$$

Отсюда выражаем значение оценки

$$\sigma^* = \hat{\sigma} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}).$$

Следовательно,

$$\frac{n}{2} < n,$$

то есть $M_2 < 0$. По критерию Сильвестра $\hat{A} < 0$ тогда и только тогда, когда $(\hat{a}, \hat{\sigma}^*)$ — точка максимума.

4.5

Задание. Постройте оценку максимального правдоподобия параметра θ показательного распределения.

Решение. $p(x, \theta) = \theta e^{-\theta x} \cdot \mathbb{1}\{x \geq 0\}$.

Нужно найти функцию правдоподобия как произведение плотностей $L(\vec{x}, \theta) = \theta^n \cdot e^{-n\theta \sum_{i=1}^n x_i} \cdot \mathbb{1}\{\vec{x} \in [0, +\infty)^n\}$. Предполагаем, что все $x_i \geq 0$, то есть индикатор можно отбросить. Если какой-то элемент выборки сильно отрицательный, то не имеет смысла говорить о равномерном распределении. Если какой-то элемент выборки близок к нулю и при этом отрицательный, то на него повлияла ошибка, и его не учитываем

$$\theta^n \cdot e^{-n\theta \sum_{i=1}^n x_i} \cdot \mathbb{1}\{\vec{x} \in [0, +\infty)^n\} = \theta^n e^{-n\theta \sum_{i=1}^n x_i}.$$

Удобнее будет рассмотреть

$$\ln L(\vec{x}, \theta) = l(\theta) = -\theta \sum_{i=1}^n x_i + n \ln \theta.$$

Нужно взять производную и приравнять к нулю

$$\frac{\partial}{\partial \theta} l(\theta) = -\sum_{i=1}^n x_i + \frac{n}{\theta} = 0.$$

Переносим сумму вправо

$$\frac{n}{\theta} = \sum_{i=1}^n x_i.$$

Получаем оценку

$$\hat{\theta} = \frac{n}{\sum_{i=1}^n x_i}.$$

Выразим через выборочное среднее

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

Остаётся проверить, что $\hat{\theta}$ будет действительно максимумом функции l . Нужно найти вторую производную

$$\frac{\partial^2}{\partial \theta^2} l(\theta) = -\frac{n}{\theta^3} < 0.$$

Отсюда следует, что функция выпуклая ввекх, соответственно это точка максимума.

4.6

Задание. Постройте оценку максимального правдоподобия параметра $\theta > 0$ равномерного распределения на отрезке $[0, \theta]$.

Решение. Плотность равномерного распределения на отрезке $[0, \theta]$ имеет вид

$$p_{\theta}(x) = \frac{\mathbb{1}\{x \in [0, \theta]\}}{\theta}.$$

По этой плотности строим функцию правдоподобия

$$L(\vec{x}, \theta) = \frac{\mathbb{1}\{\vec{x} \in [0, \theta]^n\}}{\theta^n} = \frac{\mathbb{1}\left\{\min_i x_i \geq 0, \max_i x_i \leq \theta\right\}}{\theta^n}.$$

Если какой-то элемент выборки сильно отрицательный, то не имеет смысла говорить о равномерном распределении, если близок к нулю, то на него повлияла ошибка, и его не учитываем

$$\frac{\mathbb{1}\left\{\min_i x_i \geq 0, \max_i x_i \leq \theta\right\}}{\theta^n} = \frac{\mathbb{1}\left\{\max_i x_i \leq \theta\right\}}{\theta}.$$

Нарисуем график получившейся функции как функции от θ (рис. 4).

Точка максимума такой функции — это $\max_i x_i$.

Значит, $\hat{\theta} = \max_{1 \leq i \leq n} x_i$.

4.7

Задание. Постройте оценку максимального правдоподобия параметра сдвига $\beta \in \mathbb{R}$ показательного распределения с плотностью

$$f_{\beta}(y) = \begin{cases} e^{\beta-y}, & y \geq \beta, \\ 0, & otherwise. \end{cases}$$



Рис. 4: График функции $L(\theta)$

Проверьте состоятельность этой оценки.

Решение. $f_\beta(y) = e^{\beta-y} \cdot \mathbb{1}\{y \geq \beta\}$.

Запишем функцию правдоподобия

$$L(\vec{x}, \beta) = e^{n\beta - \sum_{i=1}^n x_i} \cdot \mathbb{1}\{\vec{x} \in [\beta, +\infty)^n\} = e^{n\beta - \sum_{i=1}^n x_i} \cdot \mathbb{1}\left\{\min_{1 \leq i \leq n} x_i \geq \beta\right\}.$$

Рисуем схематический график (рис. 5).



Рис. 5: График функции $L(\beta)$

Из графика видно, что $\hat{\beta} = \min_{1 \leq i \leq n} x_i$. Проверяем состоятельность:

$$\hat{\beta} \xrightarrow{P} \beta, n \rightarrow \infty.$$

По определению $P\left(\left|\hat{\beta} - \beta\right| > \varepsilon\right) \rightarrow 0, n \rightarrow \infty, \forall \varepsilon > 0$. Подставляем оценку в функцию правдоподобия

$$e^{n\beta - \sum_{i=1}^n x_i} \cdot \mathbb{1}\left\{\min_{1 \leq i \leq n} x_i \geq \beta\right\} = P\left(\left|\min_{1 \leq i \leq n} x_i - \beta\right| > \varepsilon\right).$$

Минимальный член выборки не может быть меньше β из-за распределения.
 $P\left(\left|\min_{1 \leq i \leq n} x_i - \beta\right| > \varepsilon\right) = P\left(\min_{1 \leq i \leq n} x_i > \beta + \varepsilon\right)$. Запишем через вероятность пересечения событий $\{x_i > \beta + \varepsilon\}$. Получим

$$P\left(\min_{1 \leq i \leq n} x_i > \beta + \varepsilon\right) = P\left(\bigcap_{i=1}^n \{x_i > \beta + \varepsilon\}\right).$$

Случайные величины независимы и одинаково распределены

$$P\left(\bigcap_{i=1}^n \{x_i > \beta + \varepsilon\}\right) = [P(x_1 > \beta + \varepsilon)]^n \rightarrow 0, \quad n \rightarrow \infty,$$

так как такая вероятность меньше единицы.

Достаточно показать, что $P(x_1 > \beta + \varepsilon) \neq 1$. Тогда $l^n \rightarrow 0$.

Запишем интеграл от плотности со сдвигом

$$\int_{\beta+\varepsilon}^{+\infty} p_{\beta}(y) dy = 1 - \int_{\beta}^{\beta+\varepsilon} p_{\beta}(y) dy < 1,$$

так как второй интеграл больше нуля (забрали кусок $\beta + \varepsilon$).

4.8

Задание. Постройте оценку максимального правдоподобия параметра сдвига $\mu \in \mathbb{R}$ распределения Лапласа с плотностью $f_{\mu}(y) = e^{-\frac{|y-\mu|}{2}}$.

Решение. Записываем функцию правдоподобия

$$L(\vec{X}, \mu) = \prod_{i=1}^n f_{\mu}(X_i) = \prod_{i=1}^n \frac{1}{2} \cdot e^{-|X_i - \mu|} = \frac{1}{2^n} \cdot e^{-\sum_{i=1}^n |X_i - \mu|}.$$

Находим логарифм

$$l(\vec{X}, \mu) = c - \sum_{i=1}^n |X_i - \mu|,$$

где $c = \text{const}$. Дифференцировать по μ не можем. Хотим максимизировать эту функцию правдоподобия. Нужно минимизировать сумму X_i — точки на оси. $|X_i - \mu|$ — расстояние.

Вариационный ряд имеет понятие медианы. Это приблизительно середина вариационного ряда.

$$\mu^* = \begin{cases} X_{(m+1)}, & n = 2m + 1, \\ \frac{X_{(m)} + X_{(m+1)}}{2}, & n = 2m \end{cases}$$

— аналогия с центром масс треугольника, который задаётся как точка пересечения медиан.

4.9

Задание. Постройте оценку максимального правдоподобия параметра p распределения Бернулли.

Решение. Распределение Бернулли не имеет плотности. Есть выборка $x_1, \dots, x_n \sim B(1, p)$.

Сформулируем метод максимального правдоподобия в общем случае.

X_1, \dots, X_n — выборка с параметром θ .

X_1, \dots, X_n имеют плотность распределения $p_\theta(y)$. Тогда

$$L(\vec{X}, \theta) = \prod_{i=1}^n p_\theta(X_i).$$

Отсюда находим $\hat{\theta} = \operatorname{argmax} L(\vec{X}, \theta)$.

В дискретном случае плотность заменяем функцией $f(x) = P\{Y = x\}$.

Имеем распределение $P(X = k) = p^k \cdot (1 - p)^{1-k}$, $k = 0, 1$. Следовательно, но,

$$L(\vec{X}, p) = \prod_{i=1}^n p^{X_i} (1 - p)^{1-X_i} = p^{\sum_{i=1}^n X_i} \cdot (1 - p)^{\sum_{i=1}^n (1-X_i)} = p^{n\bar{X}} \cdot (1 - p)^{n(1-\bar{X})}.$$

Функцию непрерывна по p . Можем продифференцировать

$$l(\vec{X}, p) = n\bar{X} \ln p + n(1 - \bar{X}) \ln(1 - p).$$

Берём производную и приравниваем к нулю

$$\frac{\partial l}{\partial p} = \frac{n\bar{X}}{p} - \frac{n(1 - \bar{X})}{1 - p} = 0.$$

Сокращаем на n и переносим второе слагаемое вправо

$$\frac{\bar{X}}{p} = \frac{1 - \bar{X}}{1 - p}.$$

Перемножаем как пропорцию $\bar{X} - p\bar{X} = p(1 - \bar{X})$. Отсюда следует, что $p^* = \bar{X}$.

Проверим это, взяв вторую производную и посмотрев на её знак

$$\frac{\partial^2 l}{\partial p^2} = -\frac{n\bar{X}}{p^2} - \frac{n(1 - \bar{X})}{(1 - p)^2} < 0,$$

следовательно, $p^* = \operatorname{argmax} l(\vec{X}, p)$.

4.10

Задание. Пусть задана выборка из нормального распределения с единичной дисперсией и математическим ожиданием a , которое может принимать только два значения: 1 или 2. Постройте оценку максимального правдоподобия параметра a .

Решение. Записываем общую функцию правдоподобия, как функцию параметра a .

Запишем для этого плотность распределения каждого элемента

$$p(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2}}.$$

Теперь

$$L(\vec{X}, a) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x_i-a)^2}{2}} = \frac{1}{(2\pi)^{\frac{n}{2}}} \cdot e^{-\frac{1}{2} \sum (x_i-a)^2}.$$

Записываем логарифм

$$l(\vec{X}, a) = -\frac{n}{2} \cdot \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n (X_i - a)^2.$$

Подставим возможные значения a и найдём разность логарифмов двух соответствующих функций правдоподобия

$$l(\vec{X}, 2) - l(\vec{X}, 1) = \frac{1}{2} \sum_{i=1}^n (X_i - 2)^2 - \frac{1}{2} \sum_{i=1}^n (X_i - 1)^2.$$

Запишем под одну сумму

$$\frac{1}{2} \sum_{i=1}^n (X_i - 2)^2 - \frac{1}{2} \sum_{i=1}^n (X_i - 1)^2 = \frac{1}{2} \cdot \sum_{i=1}^n [(X_i - 2)^2 - (X_i - 1)^2].$$

Разложим на 2 множителя

$$\frac{1}{2} \cdot \sum_{i=1}^n [(X_i - 2)^2 - (X_i - 1)^2] = \frac{1}{2} \cdot \sum_{i=1}^n (X_i - 2 - X_i + 1)(X_i - 2 + X_i - 1).$$

Приведём подобные

$$\frac{1}{2} \cdot \sum_{i=1}^n (X_i - 2 - X_i + 1)(X_i - 2 + X_i - 1) = \frac{1}{2} \sum_{i=1}^n (2X_i - 3)(-1).$$

Выносим n за скобки

$$\frac{1}{2} \sum_{i=1}^n (2X_i - 3)(-1) = n \left(\frac{3}{2} - \bar{X} \right).$$

Знак разности будет зависеть от значения \vec{X} .

Если

$$\vec{X} < \frac{3}{2},$$

то $a^* = 1$, если

$$\vec{X} \geq \frac{3}{2},$$

то $a^* = 2$.

Домашнее задание

4.14

Задание. Постройте оценку максимального правдоподобия дисперсии σ^2 нормального распределения, если математическое ожидание a известно. Выясните, не является ли оценка состоятельной оценкой параметра σ^2 .

Решение.

$$p_{\sigma^2}(y) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(y-a)^2}{2\sigma^2}}.$$

Запишем функцию правдоподобия

$$L(\vec{x}, \sigma^2) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \cdot e^{-\frac{\sum_{i=1}^n (x_i - a)^2}{2\sigma^2}} = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \cdot e^{-\frac{1}{2\sigma^2} \cdot \sum_{i=1}^n (x_i - a)^2}.$$

Обозначим $\sigma^2 = \sigma^*$, чтобы было удобнее брать производную

$$\frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \cdot e^{-\frac{1}{2\sigma^2} \cdot \sum_{i=1}^n (x_i - a)^2} = \frac{1}{(2\pi\sigma^*)^{\frac{n}{2}}} \cdot e^{-\frac{1}{2\sigma^*} \cdot \sum_{i=1}^n (x_i - a)^2}.$$

Поскольку есть экспонента, стоит брать логарифм

$$l(\vec{x}, \sigma^2) = \ln L(\vec{x}, \sigma^2) = -\frac{n}{2} \cdot \ln(2\pi\sigma^*) - \frac{1}{2\sigma^*} \sum_{i=1}^n (x_i - a)^2.$$

Распишем логарифм произведения через сумму логарифмов в первом слагаемом

$$-\frac{n}{2} \cdot \ln(2\pi\sigma^*) - \frac{1}{2\sigma^*} \sum_{i=1}^n (x_i - a)^2 = -\frac{n}{2} \cdot \ln(2\pi) - \frac{n}{2} \cdot \ln\sigma^* - \frac{1}{2\sigma^*} \sum_{i=1}^n (x_i - a)^2.$$

Берём производную и приравняем её к нулю

$$\frac{\partial l}{\partial \sigma^*} = -\frac{n}{2} \cdot \frac{1}{\sigma^*} + \frac{1}{2(\sigma^*)^2} \sum_{i=1}^n (x_i - a)^2 = 0.$$

Выносим в знаменателе σ^* за скобку, сокращаем и переносим первое слагаемое вправо

$$\frac{1}{2\sigma^*} \sum_{i=1}^n (x_i - a)^2 = \frac{n}{2}.$$

Выражаем оценку

$$\sigma^* = \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - a)^2}{n}.$$

Проверили необходимое условие для максимума. Теперь нужно проверить достаточное.

Найдём вторую производную

$$\frac{\partial^2 l}{\partial (\sigma^*)^2} = \frac{\partial}{\partial \sigma^*} \left[-\frac{n}{2} \cdot \frac{1}{\sigma^*} + \frac{1}{2(\sigma^*)^2} \sum_{i=1}^n (x_i - a)^2 \right] = \frac{n}{2(\sigma^*)^2} - \frac{1}{(\sigma^*)^3} \cdot \sum_{i=1}^n (x_i - a)^2.$$

Подставим найденное значение оценки

$$\frac{n}{2(\sigma^*)^2} - \frac{1}{(\sigma^*)^3} \cdot \sum_{i=1}^n (x_i - a)^2 = \frac{n \cdot n^2}{2 \left[\sum_{i=1}^n (x_i - a)^2 \right]^2} - \frac{n^3}{\left[\sum_{i=1}^n (x_i - a)^2 \right]^3} \cdot \sum_{i=1}^n (x_i - a)^2.$$

Упростим

$$\begin{aligned} & \frac{n \cdot n^2}{2 \left[\sum_{i=1}^n (x_i - a)^2 \right]^2} - \frac{n^3}{\left[\sum_{i=1}^n (x_i - a)^2 \right]^3} \cdot \sum_{i=1}^n (x_i - a)^2 = \\ & = \frac{n^3}{2 \left[\sum_{i=1}^n (x_i - a)^2 \right]^2} - \frac{n^3}{\left[\sum_{i=1}^n (x_i - a)^2 \right]^2} = \frac{n^3}{\left[\sum_{i=1}^n (x_i - a)^2 \right]^3} \cdot \left(\frac{1}{2} - 1 \right) = \\ & = - \frac{n^3}{2 \left[\sum_{i=1}^n (x_i - a)^2 \right]^3} < 0, \end{aligned}$$

следовательно, $\hat{\sigma}^2$ — точка максимума функции. Проверим состоятельность, то есть $\hat{\sigma}^2 \xrightarrow{P} \sigma^2, n \rightarrow \infty$.

По закону больших чисел

$$\frac{1}{n} \sum_{i=1}^n (x_i - a)^2 \xrightarrow{P} M(x_1 - a)^2 = Dx_1 = \sigma^2, n \rightarrow \infty,$$

то есть оценка состоятельная.

4.15

Задание. Постройте оценку максимального правдоподобия параметра $\lambda > 0$ распределения Пуассона.

Решение. Распределение Пуассона не имеет плотности. Есть выборка $x_1, \dots, x_n \sim Pois(\lambda)$, $\lambda > 0$.

Функция правдоподобия, когда есть плотность, выглядит так

$$L(\vec{x}, \lambda) = \prod_{i=1}^n p_{\lambda}(x_i).$$

Введём функцию $f_{\lambda}(x) = P\{x_1 = x\} = P\{\omega | x_1(\omega) = x\}$, где

$$x \in \{0, 1, \dots\}.$$

В этом случае функция правдоподобия примет вид

$$L(\vec{x}, \lambda) = \prod_{i=1}^n f_{\lambda}(x_i).$$

Ищем максимум функции правдоподобия. Нужно записать

$$f_{\lambda}(x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda}.$$

Подставляем в функцию правдоподобия

$$\prod_{i=1}^n f_{\lambda}(x_i) = \prod_{i=1}^n \left(\frac{\lambda^{x_i}}{x_i!} \cdot e^{-\lambda} \right) = \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \cdot e^{-\lambda n}.$$

Так как есть экспонента, возьмём логарифм

$$l(\vec{x}, \lambda) = \ln L(\vec{x}, \lambda) = \ln \left(\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \cdot e^{-\lambda n} \right) = -\lambda n + \sum_{i=1}^n x_i \cdot \ln \lambda - \ln \prod_{i=1}^n x_i!.$$

Берём производную по λ , приравниваем её к нулю

$$\frac{\partial l(\vec{x}, \lambda)}{\partial \lambda} = -n + \sum_{i=1}^n x_i \cdot \frac{1}{\lambda} = 0.$$

Ищем оценку

$$\sum_{i=1}^n x_i \cdot \frac{1}{\lambda} = n,$$

откуда

$$\lambda^* = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}.$$

Проверили необходимое условие максимума. Проверим достаточное, то есть найдём вторую производную логарифма функции правдоподобия и посмотрим на её знак

$$\frac{\partial^2 l(\vec{x}, \lambda)}{\partial \lambda^2} = \frac{\partial}{\partial \lambda} \left(-n + \sum_{i=1}^n x_i \cdot \frac{1}{\lambda} \right) = \sum_{i=1}^n x_i \cdot \frac{\partial \lambda^{-1}}{\partial \lambda} = - \sum_{i=1}^n x_i \cdot \frac{1}{\lambda^2} < 0,$$

то есть точка $\lambda^* = \bar{x}$ — точка максимума функции правдоподобия, то есть оценка максимального правдоподобия.

4.16

Задание. Постройте оценку максимального правдоподобия параметра

$$p \in (0, 1)$$

геометрического распределения.

Решение. Геометрическое распределение не имеет плотности.

Сформируем метод максимального правдоподобия в общем случае.

X_1, \dots, X_n — выборка с параметром θ .

Случайные величины X_1, \dots, X_n имеют плотность распределения $p_\theta(y)$.

Тогда

$$L(\vec{X}, \theta) = \prod_{i=1}^n p_\theta(X_i).$$

Отсюда находим $\hat{\theta} = \operatorname{argmax} L(\vec{X}, \theta)$.

В дискретном случае плотность заменяется функцией $f(x) = P\{Y = x\}$.

Имеем распределение $P(X = k) = p(1-p)^k$, где $k = 0, 1, 2, \dots$. Следовательно,

$$L(\vec{X}, p) = \prod_{i=1}^n p(1-p)^{X_i} = p^n (1-p)^{\sum_{i=1}^n X_i} = p^n (1-p)^{n\bar{X}}.$$

Функция непрерывна по p , можем дифференцировать.

$$l(\vec{X}, p) = n \ln p + n\bar{X} \ln(1-p).$$

Берём производную и приравняем к нулю

$$\frac{\partial l}{\partial p} = \frac{n}{p} - \frac{n\bar{X}}{1-p} = 0.$$

Сокращаем на n и переносим второе слагаемое вправо

$$\frac{1}{p} = \frac{\bar{X}}{1-p}.$$

Перемножаем как пропорцию $1 - p = p\bar{X}$. Отсюда слудует, что

$$1 = p + p\bar{X} = p(1 + \bar{X}),$$

и

$$p^* = \frac{1}{1 + \bar{X}}.$$

Проверим это, взяв вторую производную и посмотрев на её знак

$$\frac{\partial^2 l}{\partial p^2} = -\frac{n}{p^2} - \frac{n\bar{X}}{(1-p)^2} < 0,$$

следовательно, $p^* = \operatorname{argmax}_l(\vec{X}, p)$.

4.17

Задание. Пусть задана выборка из биномиального распределения с параметрами $p \in (0, 1)$ и m . Постройте оценку максимального правдоподобия параметра:

- а) p , если значение параметра m известно;
- б) m по выборке объёма $n = 1$, если значение p известно.

Решение. Биномиальное распределение не имеет плотности.

Сформируем метод максимального правдоподобия в общем случае.

X_1, \dots, X_n — выборка с параметром θ .

Случайные величины X_1, \dots, X_n имеют плотность распределения $p_\theta(y)$.

Тогда

$$L(\vec{X}, \theta) = \prod_{i=1}^n p_\theta(X_i).$$

Отсюда находим $\hat{\theta} = \operatorname{argmax}_L(\vec{X}, \theta)$.

В дискретном случае плотность заменяется функцией $f(x) = P\{Y = x\}$.

Имеем распределение $P(X = k) = C_m^k p^k (1-p)^{m-k}$, $k = \overline{0, m}$. Следовательно,

$$L(\vec{X}, p, m) = \prod_{i=1}^n C_m^{X_i} p^{X_i} (1-p)^{m-X_i} = \prod_{i=1}^n C_m^{X_i} \cdot p^{\sum_{i=1}^n X_i} \cdot (1-p)^{\sum_{i=1}^n (m-X_i)}.$$

Записываем через выборочное среднее

$$\prod_{i=1}^n C_m^{X_i} \cdot p^{\sum_{i=1}^n X_i} \cdot (1-p)^{\sum_{i=1}^n (m-X_i)} = \prod_{i=1}^n C_m^{X_i} \cdot p^{n\bar{X}} (1-p)^{n(m-\bar{X})}.$$

а) Функция непрерывна по p . Можем дифференцировать.

$$l(\vec{X}, p, m) = \sum_{i=1}^n C_m^{X_i} + n\bar{X} \ln p + n(m - \bar{X}) \ln(1 - p).$$

Берём производную и приравняем к нулю

$$\frac{\partial l}{\partial p} = \frac{n\bar{X}}{p} - \frac{n(m - \bar{X})}{1 - p} = 0.$$

Сокращаем на n и переносим вправо второе слагаемое

$$\frac{\bar{X}}{p} = \frac{m - \bar{X}}{1 - p}.$$

Перемножим как пропорцию $\bar{X} - p\bar{X} = p(m - \bar{X})$. Отсюда следует, что $p(m - \bar{X}) + p\bar{X} = \bar{X}$, и оценка равна

$$p^* = \frac{\bar{X}}{m - \bar{X} + \bar{X}} = \frac{\bar{X}}{m}.$$

Проверим это, взяв вторую производную и посмотрев на её знак

$$\frac{\partial^2 l}{\partial p^2} = -\frac{m\bar{X}}{p^2} - \frac{m(m - \bar{X})}{(1 - p)^2} < 0,$$

следовательно, $p^* = \operatorname{argmax} l(\vec{X}, p)$;

б) в данном случае $\vec{X} = X_1$. Имеем функцию правдоподобия

$$L_m = C_m^{X_1} p^{X_1} (1 - p)^{m - X_1} = L(X_1, m) = L(\vec{X}, m),$$

соответственно $L_{m-1} = C_{m-1}^{X_1} p^{X_1} (1 - p)^{m-1 - X_1}$.

Найдём отношение

$$\frac{L_m}{L_{m-1}} = \frac{C_m^{X_1} p^{X_1} (1 - p)^{m - X_1}}{C_{m-1}^{X_1} p^{X_1} (1 - p)^{m-1 - X_1}} = \frac{m! (1 - p) X_1! (m - X_1 - 1)!}{X_1! (m - X_1)! (m - 1)!}.$$

Сокращаем факториалы

$$\frac{m! (1 - p) X_1! (m - X_1 - 1)!}{X_1! (m - X_1)! (m - 1)!} = \frac{m(1 - p)}{m - X_1}.$$

Возможны 2 случая

$$\begin{cases} \frac{m(1-p)}{m-X_1} \geq 1, \\ \frac{m(1-p)}{m-X_1} \leq 1. \end{cases}$$

Переносим знаменатель вправо

$$\begin{cases} m(1-p) \geq m - X_1, \\ m(1-p) \leq m - X_1. \end{cases}$$

Переносим слагаемые с m в одну сторону

$$\begin{cases} m - m(1-p) \leq X_1, \\ m - m(1-p) \geq X_1. \end{cases}$$

Выносим m за скобки

$$\begin{cases} m(1-1+p) \leq X_1, \\ m(1-1+p) \geq X_1. \end{cases}$$

Получаем

$$\begin{cases} m \leq \frac{X_1}{p}, \\ m \geq \frac{X_1}{p}. \end{cases}$$

Отсюда оценка равна

$$m^* = \left\lceil \frac{X_1}{p} \right\rceil.$$

4.18

Задание. Постройте оценку максимального правдоподобия параметра θ равномерного распределения на отрезке:

- a) $[-\theta, 0]$, $\theta > 0$;
- b) $[-\theta, \theta]$, $\theta > 0$;
- c) $[\theta, \theta + 2]$, $\theta \in \mathbb{R}$;
- d) $[\theta, 2\theta]$, $\theta > 0$.

Решение.

- a) Плотность равномерного распределения на отрезке $[-\theta, 0]$, $\theta > 0$ имеет вид

$$p_\theta(x) = \frac{1}{\theta} \cdot \mathbb{1}\{x \in [-\theta, 0]\}.$$

По этой плотности строим функцию правдоподобия

$$L(\vec{x}, \theta) = \frac{\mathbb{1}\{\vec{x} \in [-\theta, 0]^n\}}{\theta^n} = \frac{\mathbb{1}\left\{\min_i x_i \geq -\theta, \max_i x_i \leq 0\right\}}{\theta^n}.$$

Если какой-то элемент выборки сильно положительный, то не имеет смысла говорить о равномерном распределении, если положительный, но близок к нулю, то на него повлияла ошибка, и его не учитываем

$$\frac{\mathbb{1} \left\{ \min_i x_i \geq -\theta, \max_i x_i \leq 0 \right\}}{\theta^n} = \frac{\mathbb{1} \left\{ \min_i x_i \geq -\theta \right\}}{\theta^n}.$$

Нарисуем график получившейся функции, как функции от θ (рис. 6).



Рис. 6: График функции $L(\theta)$

Точка максимума этой функции — это $\left(-\min_i x_i\right)$.

Значит, $\hat{\theta} = -\min_{1 \leq i \leq n} x_i$;

- b) плотность равномерного распределения на отрезке $[-\theta, \theta]$, $\theta > 0$ имеет вид

$$p_\theta(x) = \frac{\mathbb{1} \{x \in [-\theta, \theta]\}}{2\theta}.$$

По этой функции строим функцию правдоподобия

$$L(\vec{x}, \theta) = \frac{\mathbb{1} \{\vec{x} \in [-\theta, \theta]^n\}}{(2\theta)^n} = \frac{\mathbb{1} \left\{ \min_i x_i \geq -\theta, \max_i x_i \leq \theta \right\}}{(2\theta)^n}.$$

Меняем знак для минимального элемента выборки

$$\frac{\mathbb{1} \left\{ \min_i x_i \geq -\theta, \max_i x_i \leq \theta \right\}}{(2\theta)^n} = \frac{1}{(2\theta)^n} \cdot \mathbb{1} \left\{ \max_i x_i \leq \theta, -\min_i x_i \leq \theta \right\}.$$

Берём максимум

$$\begin{aligned} & \frac{1}{(2\theta)^n} \cdot \mathbb{1} \left\{ \max_i x_i \leq \theta, -\min_i x_i \leq \theta \right\} = \\ & = \frac{1}{(2\theta)^n} \cdot \mathbb{1} \left\{ \max \left(\max_i x_i, -\min_i x_i \right) \leq \theta \right\}. \end{aligned}$$



Рис. 7: График функции $L(\theta)$

Нарисуем график этой функции, как функции от θ (рис. 7).

Точка максимума такой функции — это $\max \left(\max_i x_i, -\min_i x_i \right)$.

Значит, $\hat{\theta} = \max \left(\max_{1 \leq i \leq n} x_i, -\min_{1 \leq i \leq n} x_i \right)$;

- с) плотность равномерного распределения на отрезке $[\theta, \theta + 2]$, $\theta \in \mathbb{R}$ имеет вид

$$p_\theta(x) = \frac{\mathbb{1}\{x \in [\theta, \theta + 2]\}}{2}.$$

По этой плотности строим функцию правдоподобия

$$L(\vec{x}, \theta) = \frac{\mathbb{1}\{\vec{x} \in [\theta, \theta + 2]^n\}}{2^n} = \frac{\mathbb{1}\left\{\min_i x_i \geq \theta, \max_i x_i \leq \theta + 2\right\}}{2^n}.$$

Переносим 2 влево

$$\frac{\mathbb{1}\left\{\min_i x_i \geq \theta, \max_i x_i \leq \theta + 2\right\}}{2^n} = \frac{\mathbb{1}\left\{\min_i x_i \geq \theta, \max_i x_i - 2 \leq \theta\right\}}{2^n}.$$

Значит, $\max_{1 \leq i \leq n} x_i - 2 \leq \hat{\theta} \leq \min_{1 \leq i \leq n} x_i$;

- д) плотность равномерного распределения на отрезке $[\theta, 2\theta]$, $\theta > 0$ имеет вид

$$p_\theta(x) = \frac{\mathbb{1}\{x \in [\theta, 2\theta]\}}{\theta}.$$

По этой функции строим функцию правдоподобия

$$L(\vec{x}, \theta) = \frac{\mathbb{1}\{\vec{x} \in [\theta, 2\theta]^n\}}{\theta^n} = \frac{\mathbb{1}\left\{\min_i x_i \geq \theta, \max_i x_i \leq 2\theta\right\}}{\theta^n}.$$

Делим неравенство с максимумом на 2

$$\frac{\mathbb{1} \left\{ \min_i x_i \geq \theta, \max_i x_i \leq 2\theta \right\}}{\theta^n} = \frac{\mathbb{1} \left\{ \min_i x_i \geq \theta, \frac{\max_i x_i}{2} \leq \theta \right\}}{\theta^n}.$$

Отсюда следует, что

$$\frac{\max_i x_i}{2} \leq \theta \leq \min_i x_i.$$

Нарисуем график полученной функции, как функции от θ (рис. 8).



Рис. 8: График функции $L(\theta)$

Точка максимума такой функции — это

$$\frac{\max_i x_i}{2}.$$

Значит,

$$\hat{\theta} = \frac{\max_{1 \leq i \leq n} x_i}{2}.$$

4.19

Задание. Пусть X_1, \dots, X_n — выборка из двухпараметрического показательного распределения с плотностью

$$f_{\alpha, \beta}(y) = \begin{cases} \alpha^{-1} e^{-\frac{y-\beta}{\alpha}}, & y \geq \beta, \\ 0, & \text{otherwise,} \end{cases}$$

где $\alpha > 0$, $\beta \in \mathbb{R}$. Постройте оценку максимального правдоподобия двумерного параметра (α, β) . Проверьте состоятельность этой оценки.

Решение. Сейчас θ будет векторным параметром, состоящим из двух элементов: α, β .

Плотность имеет вид

$$f_{\alpha, \beta}(y) = \frac{1}{\alpha} \cdot e^{-\frac{y}{\alpha} + \frac{\beta}{\alpha}} \cdot \mathbb{1}\{y \geq \beta\}.$$

Произведение этих плотностей даёт

$$L(\vec{X}, \alpha, \beta) = \frac{1}{\alpha^n} \cdot e^{-\frac{1}{\alpha} \sum_{i=1}^n (X_i - \beta)} \cdot \mathbb{1} \left\{ \vec{X} \in [\beta, +\infty)^n \right\}.$$

Все элементы выборки не меньше, чем β . Запишем это условие через минимальный элемент

$$\frac{1}{\alpha^n} \cdot e^{-\frac{1}{\alpha} \sum_{i=1}^n (X_i - \beta)} \cdot \mathbb{1} \left\{ \vec{X} \in [\beta, +\infty)^n \right\} = \frac{1}{\alpha^n} \cdot e^{-\frac{1}{\alpha} \sum_{i=1}^n (X_i - \beta)} \cdot \mathbb{1} \left\{ \min_{1 \leq i \leq n} X_i \geq \beta \right\}.$$

Найдём β^* . При фиксированном α нарисуем график функции правдоподобия как функции от β (рис. 9).



Рис. 9: График функции $L(\beta)$

Из графика видно, что точка максимума этой функции — это точка $\min_{1 \leq i \leq n} X_i = \beta^*$.

Так как есть экспонента, удобнее взять логарифм от функции правдоподобия

$$\ln(\vec{X}, \alpha, \beta) = l(\vec{X}, \alpha, \beta) = -n \ln \alpha + \frac{n\beta - \sum_{i=1}^n X_i}{\alpha}.$$

Возьмём производную по α и приравняем её к нулю

$$\frac{\partial l}{\partial \alpha} = -n \cdot \frac{1}{\alpha} - \frac{n\beta - \sum_{i=1}^n X_i}{\alpha^2} = 0.$$

Умножим на $(-\alpha^2)$ левую и правую части уравнения

$$n\alpha + n\beta - \sum_{i=1}^n X_i = 0.$$

Отсюда выражаем оценку

$$\alpha^* = \frac{1}{n} \sum_{i=1}^n X_i - \beta^* = \bar{X} - \min_{1 \leq i \leq n} X_i.$$

Проверим состоятельность. Нужно проверить, выполняется ли

$$(\alpha^*, \beta^*) \xrightarrow{P} (\alpha, \beta), n \rightarrow \infty.$$

По определению $P(|\beta^* - \beta| > \varepsilon) \rightarrow 0, n \rightarrow \infty, \forall \varepsilon > 0$.

Подставим в вероятность оценку

$$P(|\beta^* - \beta| > \varepsilon) = P\left(\left|\min_{1 \leq i \leq n} X_i - \beta\right| > \varepsilon\right).$$

Из распределения $\min_{1 \leq i \leq n} X_i \geq \beta$, следовательно, можем убрать модуль

$$P\left(\left|\min_{1 \leq i \leq n} X_i - \beta\right| > \varepsilon\right) = P\left(\min_{1 \leq i \leq n} X_i - \beta > \varepsilon\right) = P\left(\min_{1 \leq i \leq n} X_i > \beta + \varepsilon\right) =$$

$$= P\left(\bigcap_{i=1}^n \{X_i > \beta + \varepsilon\}\right). \text{ Случайные величины независимые и одинаково}$$

распределённые $P\left(\bigcap_{i=1}^n \{X_i > \beta + \varepsilon\}\right) = [P(X_1 > \beta + \varepsilon)]^n \rightarrow 0, n \rightarrow \infty$, так как $P(X_1 > \beta + \varepsilon) \neq 1$.

Значит, β^* — состоятельная оценка. Тогда по закону больших чисел

$$\alpha^* = \frac{1}{n} \sum_{i=1}^n X_i - \beta^* \xrightarrow{P} MX_1 - \beta = \int_{\beta}^{+\infty} \frac{1}{\alpha} \cdot e^{-\frac{y}{\alpha}} e^{\frac{\beta}{\alpha}} y dy - \beta.$$

Возьмём интеграл по частям

$$u = y, du = dy, dv = e^{-\frac{y}{\alpha}} dy, v = \int e^{-\frac{y}{\alpha}} dy = -\alpha e^{-\frac{y}{\alpha}}.$$

Получим

$$\int_{\beta}^{+\infty} \frac{1}{\alpha} \cdot e^{-\frac{y}{\alpha}} e^{\frac{\beta}{\alpha}} y dy - \beta = \frac{1}{\alpha} \cdot e^{\frac{\beta}{\alpha}} (-\alpha) y e^{-\frac{y}{\alpha}} \Big|_{\beta}^{+\infty} + \frac{1}{\alpha} \cdot e^{\frac{\beta}{\alpha}} \cdot \alpha \cdot \int_{\beta}^{+\infty} e^{-\frac{y}{\alpha}} dy - \beta.$$

Подставим пределы интегрирования и возьмём интеграл

$$\frac{1}{\alpha} \cdot e^{\frac{\beta}{\alpha}} (-\alpha) y e^{-\frac{y}{\alpha}} \Big|_{\beta}^{+\infty} + \frac{1}{\alpha} \cdot e^{\frac{\beta}{\alpha}} \cdot \alpha \cdot \int_{\beta}^{+\infty} e^{-\frac{y}{\alpha}} dy - \beta = e^{\frac{\beta}{\alpha}} \beta \cdot e^{-\frac{\beta}{\alpha}} - e^{\frac{\beta}{\alpha}} \cdot \alpha e^{-\frac{y}{\alpha}} \Big|_{\beta}^{+\infty} - \beta.$$

Сократим экспоненты и подставим пределы интегрирования во втором слагаемом

$$e^{\frac{\beta}{\alpha}} \beta \cdot e^{-\frac{\beta}{\alpha}} - e^{\frac{\beta}{\alpha}} \cdot \alpha e^{-\frac{y}{\alpha}} \Big|_{\beta}^{+\infty} - \beta = \beta + \alpha e^{\frac{\beta}{\alpha}} \cdot e^{-\frac{\beta}{\alpha}} - \beta = \alpha.$$

Получили, что оценка α^* — тоже состоятельная.

4.20

Задание. Пусть задана выборка из двухпараметричного распределения Лапласа с плотностью

$$f_{\mu,\sigma}(y) = \frac{e^{-\frac{|y-\mu|}{\sigma}}}{2\sigma}, \quad \mu \in \mathbb{R}, \sigma > 0.$$

Постройте оценку максимального правдоподобия двумерного параметра (μ, σ) .

Решение. Записываем функцию правдоподобия

$$L(\vec{X}, \mu, \sigma) = \prod_{i=1}^n f_{\mu,\sigma}(X_i) = \prod_{i=1}^n \frac{1}{2\sigma} \cdot e^{-\frac{|X_i-\mu|}{\sigma}} = \frac{1}{(2\sigma)^n} \cdot e^{-\frac{1}{\sigma} \sum_{i=1}^n |X_i-\mu|}.$$

Находим логарифм

$$l(\vec{X}, \mu, \sigma) = -n \ln(2\sigma) - \frac{1}{\sigma} \sum_{i=1}^n |X_i - \mu| = -n \ln 2 - n \ln \sigma - \frac{1}{\sigma} \sum_{i=1}^n |X_i - \mu|.$$

Дифференцировать по μ не можем. Хотим максимизировать функцию правдоподобия. Нужно минимизировать сумму $|X_i - \mu|$ — точки на оси. $|X_i - \mu|$ — расстояние.

Вариационный ряд имеет понятие медианы. Это приблизительно середина вариационного ряда

$$\mu^* = \begin{cases} X_{(m+1)}, & n = 2m + 1, \\ \frac{X_{(m)} + X_{(m+1)}}{2}, & n = 2m \end{cases}$$

— аналог с центром масс треугольника, который задаётся как точка пересечения медиан.

Найдём оценку второго параметра. Нужно взять производную и приравнять к нулю

$$\frac{\partial l(\vec{X}, \mu, \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^2} \sum_{i=1}^n |X_i - \mu| = \frac{1}{\sigma} \cdot \left(\frac{1}{\sigma} \sum_{i=1}^n |X_i - \mu| - n \right) = 0.$$

Приравниваем выражение в скобках к нулю

$$\frac{1}{\sigma} \sum_{i=1}^n |X_i - \mu| = n.$$

Отсюда оценка равна

$$\sigma^* = \frac{\sum_{i=1}^n |X_i - \mu^*|}{n} = |\bar{X} - \mu^*|.$$

Занятие 5. Достаточные статистики

Контрольные вопросы и задания

Приведите определение условного распределения, определение достаточной статистики.

$$P(\xi \in \Delta \mid \mathcal{F}') = M[\mathbb{1}_\Delta(\xi) \mid \mathcal{F}'].$$

Статистика T называется достаточной для параметра θ , если условное распределение при известном T не зависит от параметра θ .

Сформулируйте теорему про характеризацию достаточной статистики.

Пусть x_1, \dots, x_n — выборка из распределения с плотностью

$$p(x, \theta), \theta \in \Theta.$$

Статистика T является достаточной тогда и только тогда, когда функция правдоподобия $L(\vec{x}, \theta)$ допускает факторизацию, то есть может быть представлена произведением двух функций следующего вида

$$L(\vec{x}, \theta) = h(T, \theta) \cdot g(\vec{x}).$$

Аудиторные задачи

5.3

Задание. Пусть X_1, \dots, X_n — выборка из распределения Бернулли с параметром p . Выясните, является ли \bar{X} достаточной статистикой для параметра p .

Решение. Записываем условное распределение.

Пусть $k_i = 0$ или 1 .

По определению условной вероятности

$$\begin{aligned} & P\{X_1 = k_1, X_2 = k_2, \dots, X_n = k_n | \bar{X} = y\} = \\ & = \frac{P\{X_1 = k_1, X_2 = k_2, \dots, X_n = k_n, \bar{X} = y\}}{P\{\bar{X} = y\}}. \end{aligned}$$

Выборочное среднее \bar{X} выражаем через сумму

$$\begin{aligned} & \frac{P\{X_1 = k_1, X_2 = k_2, \dots, X_n = k_n, \bar{X} = y\}}{P\{\bar{X} = y\}} = \\ & = \frac{P\left\{X_1 = k_1, X_2 = k_2, \dots, X_n = k_n, \sum_{i=1}^n X_i = ny\right\}}{P\left\{\sum_{i=1}^n X_i = ny\right\}}. \end{aligned}$$

Все X_i — независимы, следовательно, будет произведение вероятностей событий

$$\begin{aligned} & \frac{P\left\{X_1 = k_1, X_2 = k_2, \dots, X_n = k_n, \sum_{i=1}^n X_i = ny\right\}}{P\left\{\sum_{i=1}^n X_i = ny\right\}} = \\ & = \frac{\mathbb{1}\left\{\sum_{i=1}^n k_i = ny\right\} P\{X_1 = k_1\} \cdot \dots \cdot P\{X_n = k_n\}}{P\left\{\sum_{i=1}^n X_i = ny\right\}}. \end{aligned}$$

Сумма распределена по биномиальному распределению с параметрами n, p . Подставляем значения вероятностей

$$\begin{aligned} & \frac{\mathbb{1}\left\{\sum_{i=1}^n k_i = ny\right\} P\{X_1 = k_1\} \cdot \dots \cdot P\{X_n = k_n\}}{P\left\{\sum_{i=1}^n X_i = ny\right\}} = \\ & = \frac{\mathbb{1}\left\{\sum_{i=1}^n k_i = ny\right\} p^{k_1} (1-p)^{1-k_1} \cdot \dots \cdot p^{k_n} (1-p)^{1-k_n}}{C_n^{ny} p^{ny} (1-p)^{n+ny}}. \end{aligned}$$

Воспользуемся индикатором

$$\begin{aligned} & \frac{\mathbb{1}\left\{\sum_{i=1}^n k_i = ny\right\} \cdot p^{k_1} (1-p)^{1-k_1} \cdot \dots \cdot p^{k_n} (1-p)^{1-k_n}}{C_n^{ny} p^{ny} (1-p)^{n+ny}} = \\ & = \frac{\mathbb{1}\left\{\sum_{i=1}^n k_i = ny\right\} \cdot p^{ny} (1-p)^{n-ny}}{C_n^{ny} p^{ny} (1-p)^{n-ny}}. \end{aligned}$$

Всё, что связано с p , пропадает

$$\frac{\mathbb{1} \left\{ \sum_{i=1}^n k_i = ny \right\} \cdot p^{ny} (1-p)^{n-ny}}{C_n^{ny} p^{ny} (1-p)^{n-ny}} = \frac{\mathbb{1} \left\{ \sum_{i=1}^n k_i = ny \right\}}{C_n^{ny}}.$$

Зависимости от p нет. Отсюда следует, что \bar{X} — достаточная статистика.

5.4

Задание. Пусть X_1, \dots, X_n — выборка из распределения Пуассона с параметром λ . Найдите условное распределение выборки при условии

$$X_1 + \dots + X_n = k.$$

Выясните, является ли \bar{X} достаточной статистикой для параметра λ .

Решение. Это дискретное распределение

$$\begin{aligned} P \left\{ X_1 = k_1, \dots, X_n = k_n \mid \sum_{i=1}^n X_i = k \right\} = \\ \frac{P \left\{ X_1 = k_1, \dots, X_n = k_n, \sum_{i=1}^n X_i = k \right\}}{P \left\{ \sum_{i=1}^n X_i = k \right\}}. \end{aligned}$$

Случайные величины X_i — независимы

$$\begin{aligned} \frac{P \left\{ X_1 = k_1, \dots, X_n = k_n, \sum_{i=1}^n X_i = k \right\}}{P \left\{ \sum_{i=1}^n X_i = k \right\}} = \\ \frac{\mathbb{1} \left\{ \sum_{i=1}^n k_i = k \right\} P \{X_1 = k_1\} \cdot \dots \cdot P \{X_n = k_n\}}{P \left\{ \sum_{i=1}^n X_i = k \right\}}. \end{aligned}$$

Имеем распределение Пуассона. Сумма независимых случайных величин имеет распределение Пуассона с параметром λn . Подставляем значения вероятностей

$$\begin{aligned} \frac{\mathbb{1} \left\{ \sum_{i=1}^n k_i = k \right\} P \{X_1 = k_1\} \cdot \dots \cdot P \{X_n = k_n\}}{P \left\{ \sum_{i=1}^n X_i = k \right\}} = \\ = \mathbb{1} \left\{ \sum_{i=1}^n k_i = k \right\} \cdot \frac{\frac{\lambda^{k_1}}{k_1!} \cdot e^{-\lambda} \cdot \dots \cdot \frac{\lambda^{k_n}}{k_n!} \cdot e^{-\lambda}}{\frac{(\lambda n)^k}{k!} \cdot e^{-\lambda}} = \frac{\mathbb{1} \left\{ \sum_{i=1}^n k_i = k \right\} k!}{k_1! \dots k_n! n^k}. \end{aligned}$$

Видим, что нет зависимости от λ , следовательно, выборочное среднее \bar{X} — достаточная статистика для неизвестного λ .

5.5

Задание. Пусть X_1, \dots, X_n — выборка из нормального распределения со средним a и единичной дисперсией. Найдите условное совместное распределение выборки при условии $X_1 + \dots + X_n = y$. Выясните, является ли \bar{X} достаточной статистикой для параметра a .

Решение. Нужно записать функцию правдоподобия

$$L(X_1, \dots, X_n, a) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(X_i - a)^2}{2}} = \frac{1}{(\sqrt{2\pi})^n} \cdot e^{-\sum_{i=1}^n \frac{(X_i - a)^2}{2}}.$$

Откроем квадрат и запишем в статистических обозначениях

$$\frac{1}{(\sqrt{2\pi})^n} \cdot e^{-\sum_{i=1}^n \frac{(X_i - a)^2}{2}} = \frac{1}{(\sqrt{2\pi})^n} \cdot e^{-\frac{n\bar{X}^2}{2} - an\bar{X} - \frac{a^2 n}{2}} = \frac{1}{(\sqrt{2\pi})^n} \cdot e^{-\frac{n\bar{X}^2}{2}} \cdot e^{an\bar{X} - \frac{a^2 n}{2}},$$

где

$$\frac{1}{(\sqrt{2\pi})^n} \cdot e^{-\frac{n\bar{X}^2}{2}} = h(X_1, \dots, X_n), \quad e^{an\bar{X} - \frac{a^2 n}{2}} = g(a, \bar{X}).$$

Предъявляем функции, о которых идёт речь в теореме.

По теореме про характеризацию достаточной статистики \bar{X} — достаточная статистика.

5.6

Задание. Пусть X_1, \dots, X_n — выборка из нормального распределения с нулевым средним и дисперсией σ^2 . Найдите достаточную статистику со значениями в \mathbb{R} для параметра σ^2 .

Решение.

$$L(\vec{X}, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{x_i^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \cdot e^{-\frac{\sum_{i=1}^n x_i^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi}\sigma)^n} \cdot e^{-\frac{n\bar{X}^2}{2\sigma^2}}.$$

Достаточной статистикой будет второй выборочный момент.

$h(X_1, \dots, X_n) = 1$, тогда

$$g(\sigma^2, \bar{X}^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \cdot e^{-\frac{n\bar{X}^2}{2\sigma^2}}.$$

Для σ^2 достаточная статистика — $\hat{\sigma}^2 = \bar{X}^2$.

5.7

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке $[0, \theta]$. Найдите достаточную статистику со значениями в \mathbb{R} для параметра θ .

Решение.

$$L(\vec{X}, \theta) = \prod_{i=1}^n \frac{1}{\theta} \cdot \mathbb{1}\{X_i \in [0, \theta]\}.$$

Все X_i должны быть не меньше нуля и не больше θ . Получаем

$$\prod_{i=1}^n \frac{1}{\theta} \cdot \mathbb{1}\{X_i \in [0, \theta]\} = \frac{1}{\theta^n} \cdot \mathbb{1}\{X_{(1)} \geq 0, X_{(n)} \leq \theta\}.$$

Разделим индекатор на 2

$$\frac{1}{\theta^n} \cdot \mathbb{1}\{X_{(1)} \geq 0, X_{(n)} \leq \theta\} = \mathbb{1}\{X_{(1)} \geq 0\} \cdot \frac{1}{\theta^n} \cdot \mathbb{1}\{X_{(n)} \leq \theta\},$$

где первый индикатор — это функция $h(\vec{X})$, всё остальное — $g(\theta, X_{(n)})$.

Вывод: $X_{(n)}$ — достаточная статистика для θ .

5.8

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке $[a, b]$. Выясните, какие из статистик \bar{X} , $X_{(n)}$, $(X_{(1)}, X_{(n)})$ являются достаточными для двумерного параметра (a, b) .

Решение. Для оценки (a, b) будет достаточно $(X_{(1)}, X_{(n)})$.

Информацию про середину отрезка даёт \bar{X} . Этой информации недостаточно для оценки концов.

Зная $X_{(n)}$, будем знать, где правая граница, следовательно, этой информации недостаточно.

Ищем функцию правдоподобия

$$L(\vec{X}, a, b) = \prod_{i=1}^n \frac{1}{b-a} \cdot \mathbb{1}\{X_i \in [a, b]\} = \frac{1}{(b-a)^n} \cdot \mathbb{1}\{X_{(1)} \geq a, X_{(n)} \leq b\}.$$

Записываем через функции, которые используются в теореме о характеристизации достаточной статистики

$$\frac{1}{(b-a)^n} \cdot \mathbb{1}\{X_{(1)} \geq a, X_{(n)} \leq b\} = g(a, b, X_{(1)}, X_{(n)}), \quad h(\vec{X}) = 1.$$

По теореме о характеристизации, $(X_{(1)}, X_{(n)})$ — достаточная статистика для (a, b) .

Допустим, \bar{X} — достаточная статистика, значит должно выполняться $L(\vec{X}, a, b) = h(\vec{X}) \cdot g(a, b, \bar{X})$.

Пусть имеем две выборки: X_1^1, \dots, X_n^1 и X_1^2, \dots, X_n^2 , причём

$$\overline{X^1} = \overline{X^2} = c.$$

Допустим, что $X_{(1)}^1 = a$, $X_{(n)}^1 = b$, следовательно, $X_i^1 \sim U([a, b])$.

Вторую выборку подбираем так, чтобы наблюдаемое значение выходило за пределы $[a, b]$, например, $X_{(1)}^2 < a$, но $X_{(i)}^2 \in (a, b)$ при $i \neq 1$.

Ищем функции правдоподобия для этих двух выборок

$$L(\vec{X}^1) = \frac{1}{(b-a)^n} = h(\vec{X}^1) \cdot g(a, b, c), \quad L(\vec{X}^2) = 0 = h(\vec{X}^2) g(a, b, c).$$

Ноль — реакция на индикатор.

Если бы взяли выборку внутри (a, b) , то ничего бы не поменялось.

Но как только пересекаем границу, то есть реакция, то есть h уже зависит от концов отрезка, h зависит от a, b .

Для статистики $X_{(n)}$ — аналогично.

5.9

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке $[a(\theta), b(\theta)]$. Докажите, что если $a(\theta) \uparrow$, $b(\theta) \downarrow$ с возрастанием θ , то статистика $T = \min(a^{-1}(X_{(1)}), b^{-1}(X_{(n)}))$ достаточная для параметра θ .

Решение. Записываем функцию правдоподобия

$$L(\vec{X}, a(\theta), b(\theta)) = \prod_{i=1}^n \frac{1}{b(\theta) - a(\theta)} \cdot \mathbb{1}\{X_i \in [a(\theta), b(\theta)]\}.$$

Вычисляем произведение

$$\begin{aligned} & \prod_{i=1}^n \frac{1}{b(\theta) - a(\theta)} \cdot \mathbb{1}\{X_i \in [a(\theta), b(\theta)]\} = \\ & = \frac{1}{[b(\theta) - a(\theta)]^n} \cdot \mathbb{1}\{X_{(1)} \geq a(\theta), X_{(n)} \leq b(\theta)\}. \end{aligned}$$

Это строго монотонные функции. Можем решить неравенства относительно θ . Получим

$$\begin{aligned} & \frac{1}{[b(\theta) - a(\theta)]^n} \cdot \mathbb{1}\{X_{(1)} \geq a(\theta), X_{(n)} \leq b(\theta)\} = \\ & = \frac{1}{[b(\theta) - a(\theta)]^n} \cdot \mathbb{1}\{\theta \leq a^{-1}(X_{(1)}), \theta \leq b^{-1}(X_{(n)})\} = \\ & = \frac{1}{[b(\theta) - a(\theta)]^n} \cdot \mathbb{1}\{\theta \leq T\} = g(\theta, T), \end{aligned}$$

а $h \equiv 1$.

Отсюда T — достаточная статистика для θ .

Домашнее задание

5.11

Задание. Пусть X_1, \dots, X_n — выборка из биномиального распределения с параметрами m и p . Найдите условное распределение выборки при условии $X_1 + \dots, X_n = k$. Выясните, является ли \bar{X} достаточной статистикой для параметра p при известном m .

Решение. Это дискретное распределение

$$\begin{aligned} P\left(X_1 = k_1, \dots, X_n = k_n \mid \sum_{i=1}^n X_i = k\right) &= \\ &= \frac{P\left(X_1 = k_1, \dots, X_n = k_n, \sum_{i=1}^n X_i = k\right)}{P\left(\sum_{i=1}^n X_i = k\right)}. \end{aligned}$$

Имеем биномиальное распределение. Сумма независимых случайных величин имеет биномиальное распределение с параметрами mn и p . Подставляем значения вероятностей

$$\begin{aligned} &\frac{P\left(X_1 = k_1, \dots, X_n = k_n, \sum_{i=1}^n X_i = k\right)}{P\left(\sum_{i=1}^n X_i = k\right)} = \\ &= \frac{\mathbb{1}\left\{\sum_{i=1}^n k_i = k\right\} C_m^{k_1} p^{k_1} q^{m-k_1} \dots C_m^{k_n} p^{k_n} q^{m-k_n}}{C_{mn}^k p^k q^{mn-k}} = \\ &= \frac{\mathbb{1}\left\{\sum_{i=1}^n k_i = k\right\} C_m^{k_1} \dots C_m^{k_n}}{C_{mn}^k}. \end{aligned}$$

Видим, что нет зависимости от p , следовательно, выборочное среднее \bar{X} — достаточная статистика для неизвестного p .

5.12

Задание. Пусть X_1, \dots, X_n — выборка из нормального распределения со средним a и дисперсией σ^2 . Найдите достаточную статистику со значениями в \mathbb{R}^2 для двумерного параметра (a, σ^2) .

Решение. Ищем функцию правдоподобия

$$L(\vec{X}, a, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(X_i - a)^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi}\sigma)^n} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - a)^2}.$$

Раскрываем квадрат

$$\frac{1}{(\sqrt{2\pi}\sigma)^n} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - a)^2} = \frac{1}{(\sqrt{2\pi}\sigma)^n} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i^2 - 2aX_i + a^2)}.$$

Записываем через статистические величины

$$\frac{1}{(\sqrt{2\pi}\sigma)^n} \cdot e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i^2 - 2aX_i + a^2)} = \frac{1}{(\sqrt{2\pi}\sigma)^n} \cdot e^{-\frac{n\overline{X^2}}{2\sigma^2}} \cdot e^{\frac{n\overline{X}a}{2\sigma^2}} \cdot e^{-\frac{a^2n}{2\sigma^2}}.$$

Эта величина равна $g(a, \sigma^2, \overline{X}, \overline{X^2})$, $h(\vec{X}) = 1$.

По теореме о характеристизации, $(\overline{X}, \overline{X^2})$ — достаточная статистика для (a, σ^2) .

5.13

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке

- a) $[\theta, \theta + 1]$;
- b) $[\theta, 2\theta]$.

Найдите достаточную для параметра θ статистику со значениями в \mathbb{R}^2 .

Решение.

- a) Ищем функцию правдоподобия

$$L(\vec{X}, \theta) = \prod_{i=1}^n \mathbb{1}\{X_i \in [\theta, \theta + 1]\}.$$

Все X_i должны быть не меньше θ и не больше $(\theta + 1)$. Получаем

$$\prod_{i=1}^n \mathbb{1}\{X_i \in [\theta, \theta + 1]\} = \mathbb{1}\{X_{(1)} \geq \theta, X_{(n)} \leq \theta + 1\}.$$

Это равно $g(\theta, X_{(1)}, X_{(n)} - 1)$, $h(\vec{X}) = 1$.

Вывод: $(X_{(1)}, X_{(n)} - 1)$ — достаточная статистика для θ ;

- b) ищем функцию правдоподобия

$$L(\vec{X}, \theta) = \prod_{i=1}^n \frac{1}{\theta} \cdot \mathbb{1}\{X_i \in [\theta, 2\theta]\}.$$

Все X_i должны быть не меньше чем θ , не больше 2θ . Получим

$$\prod_{i=1}^n \frac{1}{\theta} \cdot \mathbb{1}\{X_i \in [\theta, 2\theta]\} = \frac{1}{\theta^n} \cdot \mathbb{1}\{X_{(1)} \geq \theta, X_{(n)} \leq 2\theta\}.$$

Это равно

$$g\left(\theta, X_{(1)}, \frac{X_{(n)}}{2}\right), h(\vec{X}) = 1.$$

Вывод:

$$\left(X_{(1)}, \frac{X_{(n)}}{2}\right)$$

— достаточная статистика для θ .

5.14

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке $[-\theta, \theta]$. Найдите достаточную статистику для параметра θ со значениями в \mathbb{R} .

Решение.

$$L(\vec{X}, \theta) = \prod_{i=1}^n \frac{1}{2\theta} \cdot \mathbb{1}\{X_i \in [-\theta, \theta]\}.$$

Все X_i должны быть не меньше $(-\theta)$ и не больше θ . Получим

$$\prod_{i=1}^n \frac{1}{2\theta} \cdot \mathbb{1}\{X_i \in [-\theta, \theta]\} = \frac{1}{(2\theta)^n} \cdot \mathbb{1}\{X_{(1)} \geq -\theta, X_{(n)} \leq \theta\}.$$

Умножим первое неравенство на (-1) . Получим

$$\begin{aligned} & \frac{1}{(2\theta)^n} \cdot \mathbb{1}\{X_{(1)} \geq -\theta, X_{(n)} \leq \theta\} = \\ & = \frac{1}{(2\theta)^n} \cdot \mathbb{1}\{-X_{(1)} \leq \theta, X_{(n)} \leq \theta\}. \end{aligned}$$

По теореме о характеристизации это равно $h(\vec{X}) \cdot g(\theta, \max(-X_{(1)}, X_{(n)}))$.

Вывод: $\max(-X_{(1)}, X_{(n)})$ — достаточная статистика для θ .

5.15

Задание. Пусть X_1, \dots, X_n — выборка из Γ -распределения с плотностью

$$f_{\alpha, \beta}(y) = \begin{cases} \frac{\alpha^\beta}{\Gamma(\beta)} \cdot y^{\beta-1} e^{-\alpha y}, & y \geq 0, \\ 0, & y < 0, \end{cases} \quad \alpha, \beta > 0.$$

Найдите достаточную статистику для параметра (α, β) .

Решение. Запишем функцию правдоподобия данной выборки

$$L(X_1, \dots, X_n, \alpha, \beta) = \prod_{i=1}^n f_{\alpha, \beta}(X_i) = \frac{\alpha^{n\beta}}{\Gamma^n(\beta)} \cdot \prod_{i=1}^n \left(X_i^{\beta-1} e^{-\alpha X_i}\right) \cdot \mathbb{1}\{X_{(1)} \geq 0\}.$$

Раскроем скобки в произведении

$$\frac{\alpha^{n\beta}}{\Gamma^n(\beta)} \cdot \prod_{i=1}^n \left(X_i^{\beta-1} e^{-\alpha X_i} \right) \cdot \mathbb{1} \{X_{(1)} \geq 0\} = \frac{\alpha^{n\beta}}{\Gamma^n(\beta)} \cdot \prod_{i=1}^n X_i^{\beta-1} \cdot e^{-\alpha \bar{X}n} \cdot \mathbb{1} \{X_{(1)} \geq 0\}.$$

Видно, что функция правдоподобия подаётся в виде произведения функции $h(X_1, \dots, X_n) = \mathbb{1} \{X_{(1)} \geq 0\}$, которая не зависит от неизвестных параметров, и функции

$$g(X_1, \dots, X_n, \alpha, \beta) = \frac{\alpha^{n\beta}}{\Gamma^n(\beta)} \cdot \prod_{i=1}^n X_i^{\beta-1} \cdot e^{-\alpha \bar{X}n},$$

которая зависит от неизвестных параметров и от функции выборки

$$T_1(X_1, \dots, X_n) = \bar{X}$$

и

$$T_2(X_1, \dots, X_n) = \prod_{i=1}^n X_i.$$

Тогда, по теореме о характеристизации достаточной статистики, двумерная функция

$$T(X_1, \dots, X_n) = (T_1(X_1, \dots, X_n), T_2(X_1, \dots, X_n)) = \left(\bar{X}, \prod_{i=1}^n X_i \right)$$

является достаточной статистикой для параметра (α, β) .

5.16

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке $[a(\theta), b(\theta)]$. Докажите, что если $a(\theta) \downarrow$, $b(\theta) \uparrow$ с возрастанием θ , то статистика $T = \max(a^{-1}(X_{(1)}), b^{-1}(X_{(n)}))$ достаточная для параметра θ . По выборке с равномерным распределением на $[-\theta, \theta]$ найдите достаточную статистику для параметра θ со значениями в \mathbb{R} .

Решение. Запишем функцию правдоподобия

$$L(\vec{X}, a(\theta), b(\theta)) = \prod_{i=1}^n \frac{1}{b(\theta) - a(\theta)} \cdot \mathbb{1} \{X_i \in [a(\theta), b(\theta)]\}.$$

Вычисляем произведение

$$\begin{aligned} & \prod_{i=1}^n \frac{1}{b(\theta) - a(\theta)} \cdot \mathbb{1} \{X_i \in [a(\theta), b(\theta)]\} = \\ & = \frac{1}{[b(\theta) - a(\theta)]^n} \cdot \mathbb{1} \{X_{(1)} \geq a(\theta), X_{(n)} \leq b(\theta)\}. \end{aligned}$$

Это строго монотонные функции. Можем решить неравенства относительно θ . Получим

$$\begin{aligned} & \frac{1}{[b(\theta) - a(\theta)]^n} \cdot \mathbb{1} \left\{ X_{(1)} \geq a(\theta), X_{(n)} \leq b(\theta) \right\} = \\ &= \frac{1}{[b(\theta) - a(\theta)]^n} \cdot \mathbb{1} \left\{ \theta \geq a^{-1}(X_{(1)}), \theta \geq b^{-1}(X_{(n)}) \right\} = \\ &= \frac{1}{[b(\theta) - a(\theta)]^n} \cdot \mathbb{1} \{ \theta \geq T \} = g(\theta, T), \end{aligned}$$

а $h \equiv 1$.

Значит, T — достаточная статистика для θ .

Для отрезка $[-\theta, \theta]$ получаем $b(\theta) = \theta$, $a(\theta) = -\theta$. Имеем неравенства $X_{(1)} \geq -\theta$, $X_{(n)} \leq \theta$, откуда $\theta \geq -X_{(1)}$, $\theta \geq X_{(n)}$, то есть

$$T = \max(-X_{(1)}, X_{(n)}).$$

Занятие 6. Теорема Колмогорова про улучшение оценок. Неравенство Рао-Крамера

Контрольные вопросы и задания

Приведите определение достаточной статистики.

Статистика T называется достаточной для параметра θ , если условное распределение при известном T не зависит от параметра θ .

Сформулируйте теорему про характеризацию достаточной статистики.

Пусть x_1, \dots, x_n — выборка из распределения с плотностью

$$p(x, \theta), \theta \in \Theta.$$

Статистика T является достаточной тогда и только тогда, когда функция правдоподобия $L(\vec{x}, \theta)$ допускает факторизацию, то есть может быть представлена произведением двух функций следующего вида

$$L(\vec{x}, \theta) = h(T, \theta) \cdot g(\vec{x}).$$

Сформулируйте теорему Колмогорова про улучшение оценки с помощью достаточной статистики.

Оптимальная оценка единственная (в том случае, когда она существует).

Что называется количеством информации Фишера?

$$D_\theta U(\vec{x}, \theta) = n D_\theta \left(\frac{\partial}{\partial \theta} \ln p(x, \theta) \right) = n I_0,$$

где nI_0 — количество информации Фишера.

Запишите неравенство Рао-Крамера.

Пусть $\hat{\theta}$ — несмещённая оценка для параметра θ . Тогда

$$\forall \theta \in \Theta : D_{\theta} \hat{\theta} \geq \frac{1}{nI_0}.$$

Как угодно лучшую оценку взять нельзя.

Какая оценка называется эффективной?

Если $\hat{\theta}$ такова, что $D_{\theta} \hat{\theta} = \frac{1}{nI_0}$ — несмещённая и

$$D_{\theta} \hat{\theta} = \frac{1}{nI_0}, \theta \in \Theta,$$

то $\hat{\theta}$ называется эффективной оценкой.

Приведите определение экспоненциальной семьи распределений.

Аудиторные задачи

6.3

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке $[0, \theta]$. Найдите несмещённую оценку неизвестного параметра $\tau(\theta, y) = P(X_1 > y)$ и улучшите её усреднением по достаточной для параметра θ статистике.

Решение. $X_{(n)}$ — достаточная статистика для θ .

Нужно найти $\hat{\tau}$ из условия, что она была бы несмещённой, что

$$M\hat{\tau} = \tau = P(X_1 > y).$$

Берём в качестве $\hat{\tau} = \mathbb{1}\{X_1 > y\}$ — это несмещённая оценка для τ . Теперь улучшаем эту оценку. $\tau^* = M(\hat{\tau} \mid X_{(n)}) = M\{\mathbb{1}\{X_1 > y\} \mid X_{(n)}\}$. Индикатор может принимать значения 0 и 1, значит

$$\begin{aligned} & M\{\mathbb{1}\{X_1 > y\} \mid X_{(n)}\} = \\ &= 0 \cdot P(\mathbb{1}\{X_1 > y\} = 0 \mid X_{(n)}) + 1 \cdot P(\mathbb{1}\{X_1 > y\} = 1 \mid X_{(n)}). \end{aligned}$$

Первое слагаемое пропадает

$$\begin{aligned} & 0 \cdot P(\mathbb{1}\{X_1 > y\} = 0 \mid X_{(n)}) + 1 \cdot P(\mathbb{1}\{X_1 > y\} = 1 \mid X_{(n)}) = \\ &= P(X_1 > y \mid X_{(n)}) = f(X_{(n)}). \end{aligned}$$

Ищем $f(z) = P(X_1 > y \mid X_{(n)} = z)$. Выборка имеет равномерное распределение, то есть распределение, которое имеет плотность. Вероятность попадания в точку равна нулю

$$P(X_1 > y \mid X_{(n)} = z) = \lim_{h \rightarrow 0} P\{X_1 > y \mid X_{(n)} \in [z, z+h]\}.$$

Воспользуемся определением условной вероятности

$$\lim_{h \rightarrow 0} P\{X_1 > y \mid X_{(n)} \in [z, z+h]\} = \lim_{h \rightarrow 0} \frac{P\{X_1 > y, X_{(n)} \in [z, z+h]\}}{P\{X_{(n)} \in [z, z+h]\}}.$$

Это нужно рассматривать при $y < z$. Разбиваем на разность двух вероятностей в числителе и знаменателе

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{P\{X_1 > y, X_{(n)} \in [z, z+h]\}}{P\{X_{(n)} \in [z, z+h]\}} = \\ &= \lim_{h \rightarrow 0} \left[\frac{P\{X_1 > y, X_{(n)} \leq z+h\}}{P\{X_{(n)} \leq z+h\} - P\{X_{(n)} \leq z\}} - \frac{P\{X_1 > y, X_{(n)} \leq z\}}{P\{X_{(n)} \leq z+h\} - P\{X_{(n)} \leq z\}} \right]. \end{aligned}$$

Переходим к функции распределения выборки

$$\begin{aligned} & \lim_{h \rightarrow 0} \left[\frac{P\{X_1 > y, X_{(n)} \leq z+h\}}{P\{X_{(n)} \leq z+h\} - P\{X_{(n)} \leq z\}} - \frac{P\{X_1 > y, X_{(n)} \leq z\}}{P\{X_{(n)} \leq z+h\} - P\{X_{(n)} \leq z\}} \right] = \\ &= \lim_{h \rightarrow 0} \left\{ \frac{[F(z+h) - F(y)][F(z+h)]^{n-1}}{[F(z+h)]^n - [F(z)]^n} - \frac{[F(z) - F(y)][F(z)]^{n-1}}{[F(z+h)]^n - [F(z)]^n} \right\}. \end{aligned}$$

Знаменатель и числитель стремятся к нулю. Дифференцируем по h числитель и знаменатель. Это правило Лопиталя

$$\begin{aligned} & \lim_{h \rightarrow 0} \left\{ \frac{[F(z+h) - F(y)][F(z+h)]^{n-1}}{[F(z+h)]^n - [F(z)]^n} - \frac{[F(z) - F(y)][F(z)]^{n-1}}{[F(z+h)]^n - [F(z)]^n} \right\} = \\ &= \lim_{h \rightarrow 0} \left\{ \frac{p(z+h)[F(z+h)]^{n-1}}{n[F(z+h)]^{n-1}p(z+h)} + \frac{(n+1)[F(z+h)]^{n-2}p(z+h)[F(z+h) - F(y)]}{n[F(z+h)]^{n-1}p(z+h)} \right\} = \\ &= \lim_{h \rightarrow 0} \frac{F(z+h) + (n-1)[F(z+h) - F(y)]}{nF(z+h)}. \end{aligned}$$

Переходим к пределу

$$\lim_{h \rightarrow 0} \frac{F(z+h) + (n-1)[F(z+h) - F(y)]}{nF(z+h)} = \frac{F(z) + (n+1)[F(z) - F(y)]}{nF(z)}.$$

Приводим подобные

$$\frac{F(z) + (n+1)[F(z) - F(y)]}{nF(z)} = \frac{nF(z) - (n-1)F(y)}{nF(z)}.$$

Почленно поделим числитель на знаменатель

$$\frac{nF(z) - (n-1)F(y)}{nF(z)} = 1 - \frac{n-1}{n} \cdot \frac{F(y)}{F(z)}.$$

Подставляем значения функции равномерного распределения на отрезке $[0, \theta]$. Получаем

$$1 - \frac{n-1}{n} \cdot \frac{F(y)}{F(z)} = 1 - \frac{n-1}{n} \cdot \frac{y}{z}.$$

Окончательно

$$f(X_{(n)}) = 1 - \frac{n-1}{n} \cdot \frac{y}{X_{(n)}}.$$

6.4

Задание. Пусть X_1, \dots, X_n — выборка из распределения Пуассона с параметром λ . В качестве оценки параметра $\theta = e^{-\lambda}$ рассматривают статистику $\hat{\theta} = \mathbb{1}\{X_1 = 0\}$. Вычислите смещение $b_n(\theta) = M\hat{\theta} - \theta$ этой оценки и улучшите её усреднением по достаточной статистике для параметра θ статистикой.

Решение. Неизвестный параметр — это $e^{-\lambda} = \theta$.

Ищем смещение $M\hat{\theta} = M\mathbb{1}\{X_1 = 0\} = P\{X_1 = 0\} = e^{-\lambda} = \theta$, значит, смещение $b_n(\theta) = 0$. Это несмещённая оценка.

Займёмся поиском достаточной статистики

$$L(\vec{X}, \lambda) = \prod_{i=1}^n \frac{\lambda^{X_i}}{X_i!} \cdot e^{-\lambda} = \frac{\lambda^{n\bar{X}}}{X_1! \dots X_n!} \cdot e^{-\lambda n} = \frac{1}{X_1! \dots X_n!} \cdot \lambda^{n\bar{X}} e^{-\lambda n}.$$

Вывод: \bar{X} — достаточная статистика для λ , а значит и для θ , как функции от λ .

Нужно найти условное математическое ожидание $M(\hat{\theta} | \bar{X}) = f(\bar{X})$. Начнём с $f(y) = M(\hat{\theta} | \bar{X} = y) = M(\mathbb{1}\{X_1 = 0\} | \bar{X} = y)$. Индикатор — это дискретная случайная величина, которая принимает 2 значения

$$M(\mathbb{1}\{X_1 = 0\} | \bar{X} = y) = 1 \cdot P(X_1 = 0 | \bar{X} = y) = \frac{P\left(X_1 = 0, \sum_{i=1}^n X_i = ny\right)}{P\left(\sum_{i=1}^n X_i = ny\right)}.$$

В числителе стоит случайное событие $X_1 = 0$. Оно означает, что суммировать можно от двух

$$\frac{P\left(X_1 = 0, \sum_{i=1}^n X_i = ny\right)}{P\left(\sum_{i=1}^n X_i = ny\right)} = \frac{P(X_1 = 0) P\left(\sum_{i=2}^n X_i = ny\right)}{P\left(\sum_{i=1}^n X_i = ny\right)}.$$

Сумма независимых пуассоновских случайных величин имеет распределение Пуассона с параметром, который равен сумме параметров

$$\frac{P(X_1 = 0) P\left(\sum_{i=2}^n X_i = ny\right)}{P\left(\sum_{i=1}^n X_i = ny\right)} = \frac{e^{-\lambda} \cdot \frac{((n-1)\lambda)^{ny}}{(ny)!} \cdot e^{-\lambda(n-1)}}{\frac{(n\lambda)^{ny}}{(ny)!} \cdot e^{-\lambda n}} = \left(\frac{n-1}{n}\right)^{ny}.$$

Чтобы записать ответ, нужно вместо y записать \bar{X} . Получим

$$\left(\frac{n-1}{n}\right)^{n\bar{X}}.$$

Когда $n \rightarrow \infty$, эта оценка стремится к $e^{-\bar{X}}$.

6.5

Задание. Пользуясь неравенством Рао-Крамера, выясните, является ли эффективной оценка \bar{X} параметра p распределения Бернулли.

Решение. X_1, \dots, X_n — выборка.

Несмещённой оценкой для параметра является $\hat{p} = \bar{X}$. Формально в этом убедимся. Итак,

$$M\hat{p} = M\bar{X} = \frac{1}{n} \cdot M \sum_{i=1}^n X_i.$$

Все случайные величины одинаково распределены, следовательно, все математические ожидания совпадают

$$\frac{1}{n} \cdot M \sum_{i=1}^n X_i = \frac{1}{n} \cdot n M X_1 = M X_1 = p.$$

Ищем дисперсию

$$D\hat{p} = D\bar{X} = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot D \sum_{i=1}^n X_i.$$

Дисперсии равны. Случайные величины независимы, следовательно, дисперсия суммы — это сумма дисперсий

$$\frac{1}{n^2} \cdot D \sum_{i=1}^n X_i = \frac{1}{n^2} \cdot n D X_1 = \frac{1}{n} \cdot p(1-p).$$

Сначала нужна функция правдоподобия для одного наблюдения

$$L(X_1, p) = p^{X_1} (1-p)^{1-X_1}.$$

Логарифмируем $\ln L(X_1, p) = X_1 \ln p + (1 - p) \ln(1 - p)$. Ищем первую производную

$$\frac{\partial}{\partial p} \ln L(X_1, p) = \frac{X_1}{p} - \frac{1 - X_1}{1 - p}.$$

Ищем вторую производную

$$\frac{\partial^2}{\partial p^2} \ln L(X_1, p) = -\frac{X_1}{p^2} - \frac{1 - X_1}{(1 - p)^2}.$$

Пишем количество информации. Берём математическое ожидание со знаком «минус»

$$I(\theta) = M\left(\frac{X_1}{p} + \frac{1 - X_1}{(1 - p)^2}\right) = \frac{1}{p} + \frac{1 - p}{(1 - p)^2} = \frac{1}{p} + \frac{1}{1 - p}.$$

Приводим к общему знаменателю

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

Ищем обратную величину, делённую на n . Получаем

$$\frac{1}{nI(\theta)} = \frac{1}{n[p(1 - p)]^{-1}} = \frac{p(1 - p)}{n} = D\hat{p},$$

следовательно, \hat{p} — эффективная оценка параметра p .

6.6

Задание. Пусть X_1, \dots, X_n — выборка из равномерного распределения с параметром p . Выясните, является ли эффективной оценка $\hat{\theta} = 1 + \bar{X}$ параметра

$$\theta = \frac{1}{p}.$$

Решение. Берём дисперсию

$$D\hat{\theta} = D(1 + \bar{X}) = D\bar{X} = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot D \sum_{i=1}^n X_i = \frac{1}{n^2} \cdot nDX_1.$$

Сократим n и подставим значение дисперсии равномерного распределения

$$\frac{1}{n^2} \cdot nDX_1 = \frac{1}{n} \cdot \frac{1 - p}{p^2}.$$

Можем записать это как функцию от θ . Получим

$$\frac{1}{n} \cdot \frac{1 - p}{p^2} = \frac{1}{n} \cdot (\theta^2 - \theta).$$

Можно θ вынести за скобки

$$\frac{1}{n} \cdot (\theta^2 - \theta) = \frac{\theta}{n} \cdot (\theta - 1).$$

Записываем функцию правдоподобия для одного наблюдения

$$L(X_1, p) = p(1-p)^{X_1}.$$

Подставим оценку параметра

$$L(X_1, \theta) = \frac{1}{\theta} \cdot \left(1 - \frac{1}{\theta}\right)^{X_1} = \frac{1}{\theta^{X_1+1}} \cdot (\theta - 1)^{X_1}.$$

Логарифмируем $\ln L(X_1, \theta) = -(X_1 + 1) \ln \theta + X_1 \ln(\theta - 1)$. Берём первую производную

$$\frac{\partial}{\partial \theta} \ln L(X_1, \theta) = -\frac{X_1 + 1}{\theta} + \frac{X_1}{\theta - 1}.$$

Берём вторую производную

$$\frac{\partial^2}{\partial \theta^2} \ln L(X_1, \theta) = \frac{X_1 + 1}{\theta^2} - \frac{X_1}{(\theta - 1)^2}.$$

Берём математическое ожидание

$$MX_1 = \frac{1-p}{p}.$$

Почленно поделим числитель на знаменатель

$$\frac{1-p}{p} = \theta - 1.$$

Берём математическое ожидание второй производной функции правдоподобия со знаком «минус»

$$-M \left[\frac{\partial^2}{\partial \theta^2} \ln L(X_1, \theta) \right] = -\frac{\theta}{\theta^2} + \frac{\theta - 1}{(\theta - 1)^2} = -\frac{1}{\theta} + \frac{1}{\theta - 1} = \frac{-\theta + 1 + \theta}{\theta(\theta - 1)}.$$

Приводим подобные

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

Формируем обратную величину и делим на n . Значит,

$$\frac{1}{nI(\theta)} = \frac{\theta(\theta - 1)}{n} = D\hat{\theta}.$$

Получили, что оценка эффективная.

6.7

Задание. Пусть X_1, \dots, X_n — выборка из распределения $N(\theta, \sigma^2)$ с неизвестным параметром θ и известной дисперсией. Выясните, является ли статистика $\theta^* = \bar{X}$ эффективной оценкой параметра θ .

Решение. Проверим несмещённость

$$M\theta^* = M\bar{X} = M\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \cdot M \sum_{i=1}^n X_i = \frac{1}{n} \cdot nMX_1 = \theta.$$

Несмещённость есть. Ищем дисперсию

$$D\theta^* = D\bar{X} = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot D \sum_{i=1}^n X_i.$$

Воспользуемся независимостью случайных величин в выборке

$$\frac{1}{n^2} \cdot D \sum_{i=1}^n X_i = \frac{1}{n^2} \cdot nDX_1.$$

Случайные величины в выборке одинаково распределены, значит

$$\frac{1}{n^2} \cdot nDX_1 = \frac{\sigma^2}{n}.$$

Переходим к функции правдоподобия и оценке количества информации. Смотрим на одно наблюдение

$$L(X_1, \theta) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(X_1 - \theta)^2}{2\sigma^2}}.$$

Логарифмируем

$$\ln L(X_1, \theta) = \ln\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - (X_1 - \theta)^2 \cdot \frac{1}{2\sigma^2}.$$

Ищем первую производную

$$\frac{\partial}{\partial \theta} \ln L(X_1, \theta) = 2(X_1 - \theta) \cdot \frac{1}{2\sigma^2} = \frac{X_1 - \theta}{\sigma^2}.$$

Ищем вторую производную

$$\frac{\partial}{\partial \theta^2} \ln L(X_1, \theta) = -\frac{1}{\sigma^2}.$$

Таким образом,

$$I(\theta) = -M\left(-\frac{1}{\sigma^2}\right) = \frac{1}{\sigma^2}.$$

Формируем величину

$$\frac{1}{nI(\theta)} = \frac{\sigma^2}{n} = D\theta^*,$$

значит, оценка эффективная.

6.8

Задание. Пусть X_1, \dots, X_n — выборка из экспоненциального распределения с плотностью

$$f_{\theta}(x) = \begin{cases} \frac{1}{\theta} \cdot e^{-\frac{x}{\theta}}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Убедитесь, что оценка $\theta^* = \bar{X}$ — эффективная по Рао-Крамеру оценка неизвестного параметра θ .

Решение.

$$M\theta^* = M\bar{X} = M\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \cdot M \sum_{i=1}^n X_i = \frac{1}{n} \cdot nMX_1 = MX_1 = \theta,$$

значит, имеем несмещённость.

Теперь ищем дисперсию

$$D\bar{X} = \frac{1}{n} \cdot DX_1 = \frac{1}{n} \cdot \frac{1}{\left(\frac{1}{\theta}\right)^2} = \frac{\theta^2}{n}.$$

Ищем количество информации. Начнём с функции правдоподобия

$$L(X_1, \theta) = \frac{1}{\theta} \cdot e^{-\frac{X_1}{\theta}} \cdot \mathbb{1}\{X_1 \geq 0\}.$$

Пусть $X_1 \geq 0$. Тогда берём логарифм

$$\ln L(X_1, \theta) = -\ln \theta - \frac{X_1}{\theta}.$$

Берём первую производную от логарифма функции правдоподобия

$$\frac{\partial L}{\partial \theta} = -\frac{1}{\theta} + \frac{X_1}{\theta^2}.$$

Берём вторую производную от логарифма функции правдоподобия

$$\frac{\partial^2 L}{\partial \theta^2} = \frac{1}{\theta^2} - \frac{2X_1}{\theta^3}.$$

Ищем количество информации

$$I(\theta) = -M\left(-\frac{1}{\theta^2} - \frac{2X_1}{\theta^3}\right) = -\frac{1}{\theta^2} + \frac{2}{\theta^3} \cdot MX_1 = -\frac{1}{\theta^2} + \frac{2}{\theta^2} = \frac{1}{\theta^2}.$$

Ищем обратную величину к количеству информации и делим на n . Получаем

$$\frac{1}{nI(\theta)} = \frac{\theta^2}{n} = D\theta^*.$$

Значит, выборочное среднее — эффективная оценка.

Домашнее задание

6.10

Задание. Пусть X_1, \dots, X_n — выборка из нормального распределения с параметрами $(a, 1)$. Улучшите оценку $\hat{a} = X_1$ усреднением по фиксированному значению достаточной статистики \bar{X} . Найдите распределение, математическое ожидание и дисперсию улучшенной оценки.

Решение. Улучшаем оценку

$$a^* = M(\hat{a} \mid \bar{X}) = M(X_1 \mid \bar{X}) = \frac{1}{n} \sum_{i=1}^n M(X_i \mid \bar{X}) = M\left(\frac{1}{n} \sum_{i=1}^n X_i \mid \bar{X}\right).$$

Запишем через выборочное среднее

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

Математическое ожидание равно

$$Ma^* = M\bar{X} = M\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n MX_i = \frac{1}{n} \cdot na = a,$$

дисперсия —

$$Da^* = D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot nDX_1 = \frac{1}{n} \cdot 1 = \frac{1}{n}.$$

Она имеет нормальное распределение

$$a^* \sim N\left(a, \frac{1}{n}\right).$$

Так как $Ma^* = a$, то оценка несмещённая.

6.11

Задание. Пусть X_1, \dots, X_n — выборка из распределения Пуассона с параметром λ . В качестве оценки параметра $\theta = e^{-\lambda}$ рассматривают статистику

$$\hat{\theta} = \overline{\mathbb{1}\{X = 0\}} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{X_i = 0\}.$$

Вычислите смещение $b_n(\theta) = M\hat{\theta} - \theta$ этой оценки и выясните, является ли она эффективной.

Решение.

$$M\hat{\theta} = M\left(\frac{1}{n} \sum_{i=1}^n \mathbb{1}\{X_i = 0\}\right) = \frac{1}{n} M\left(\sum_{i=1}^n \mathbb{1}\{X_i = 0\}\right) = \frac{1}{n} \cdot \sum_{i=1}^n P(X_i = 0).$$

Подставляем значение вероятности для распределения Пуассона

$$\frac{1}{n} \cdot \sum_{i=1}^n P(X_i = 0) = \frac{1}{n} \cdot \sum_{i=1}^n \frac{\lambda^0}{0!} \cdot e^{-\lambda} = \frac{1}{n} \cdot e^{-\lambda} \cdot n = e^{-\lambda}.$$

Подставляем полученное значение в формулу для смещения

$$b_n(\theta) = M\hat{\theta} - \theta = e^{-\lambda} - e^{-\lambda} = 0.$$

Занятие 7. Гауссовы системы

Контрольные вопросы и задания

Приведите определение гауссового случайного вектора, ковариационной матрицы гауссового случайного вектора.

$\vec{\xi} = (\xi_1, \dots, \xi_n)$ — гауссовый, если

$$\forall \alpha_1, \dots, \alpha_n : \sum_{k=1}^n \alpha_k \xi_k$$

— гауссовская случайная величина.

$$A = \text{cov}_{\vec{\xi}\vec{\xi}}, A_{ij} = \text{cov}(\xi_i, \xi_j) = M(\xi_i \xi_j) - M\xi_i M\xi_j.$$

Какими свойствами владеет ковариационная матрица?

$A_{ii} = D\xi_i \geq 0$ — на диагонали — неотрицательные числа.

Матрица симметрична: $A_{ij} = A_{ji}$.

Матрица неотрицательно определена.

Как изменяются характеристики гауссового случайного вектора при действии на него линейного оператора?

Пусть $\vec{\xi}$ случайный n -элементный вектор, имеющий гауссовское распределение с параметрами \vec{a} и A , $\vec{\xi} \sim N(\vec{a}, A)$, $\vec{b} \in \mathbb{R}^m$, $T \in \mathbb{R}^{m \times n}$. Тогда $T\vec{\xi} + \vec{b} \sim N(T\vec{a} + \vec{b}, TAT^*)$.

При каких условиях гауссовский случайный вектор имеет плотность распределения?

Плотность можно записать только в случае, когда $\det A \neq 0$.

Запишите плотность распределения гауссового случайного вектора.

$$p_{\vec{\xi}}(\vec{x}) = \left(\frac{1}{\sqrt{2\pi}} \right)^n \cdot \frac{1}{\sqrt{\det A}} \cdot e^{-\frac{1}{2}[\vec{x}-\vec{a}, A^{-1}(\vec{x}-\vec{a})]}.$$

Сформулируйте теорему про нормальную корреляцию.

Есть гауссовский вектор $\vec{\xi} \circ \vec{\eta}$ с ненулевой ковариацией $cov_{\vec{\xi} \circ \vec{\eta}, \vec{\xi} \circ \vec{\eta}} \neq 0$.
Определитель ковариационной матрицы вектора η положителен

$$\det cov_{\vec{\eta}, \vec{\eta}} \geq 0.$$

Тогда вектор $\vec{\xi}$ при условии $\vec{\eta}$ — гауссовский случайный вектор

$$\vec{\xi} \Big| \vec{\eta} \sim N(\vec{m}, D).$$

Параметры \vec{m} и D имеют следующий вид

$$\vec{m} = M\vec{\xi} + cov_{\vec{\xi}, \vec{\eta}} cov_{\vec{\eta}, \vec{\eta}}^{-1} (\eta - M\vec{\eta}), \quad D = cov_{\vec{\xi}, \vec{\xi}} - cov_{\vec{\xi}, \vec{\eta}} cov_{\vec{\eta}, \vec{\eta}}^{-1} cov_{\vec{\eta}, \vec{\xi}}.$$

Аудиторные задачи

7.3

Задание. Может ли матрица A быть ковариационной матрицей гауссового случайного вектора, если:

a) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};$

b) $A = \begin{bmatrix} 1 & 2 \\ 3 & 9 \end{bmatrix};$

c) $A = \begin{bmatrix} 4 & 2 \\ 2 & -1 \end{bmatrix};$

d) $A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix};$

e) $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & -1 \\ 1 & -1 & 12 \end{bmatrix};$

f) $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 0 \\ 1 & 0 & 10 \end{bmatrix}.$

Если так, то предъявите такой вектор.

Решение.

- a) Нет (матрица не квадратная);
- b) нет (матрица не симметрична);
- c) нет (на диагонали отрицательные числа);
- d) матрица квадратная, симметричная, неотрицательно определённая, на диагонали — неотрицательные числа;
- e) матрица квадратная, симметричная, на диагонали — неотрицательные числа. Первый минор $M_1 = 1 \geq 0$, второй минор $M_2 = 6 - 4 = 2 \geq 0$, третий минор

$$\det \begin{pmatrix} 1 & 2 & 1 \\ 2 & 6 & -1 \\ 1 & -1 & 12 \end{pmatrix} = 1 \cdot (72 - 1) - 2 \cdot (24 + 1) + 1 \cdot (-2 - 6) = 13 \geq 0,$$

значит, матрица неотрицательно определена, то есть может быть ко-вариационной матрицей гауссового случайного вектора;

- f) матрица квадратная, симметричная, на диагонали — неотрицательные числа. Первый минор $M_1 = 1 \geq 0$, второй минор $M_2 = 5 - 2 = 3 \geq 0$, третий минор

$$M_3 = 1 \cdot (50 - 0) - 2 \cdot (20 - 0) + 1 \cdot (0 - 5) = 50 - 40 - 5 = 5 \geq 0,$$

значит, матрица неотрицательно определена, то есть может быть ко-вариационной матрицей гауссового случайного вектора.

7.5

Задание. Пусть $\xi = (\xi_1, \xi_2, \xi_3)$ — гауссовый вектор с математическим ожиданием $(-1, 0, 2)$ и матрицей ковариаций из задачи 7.3 е).

- a) Выпишите плотность распределения и характеристическую функцию для $\xi_1, (\xi_1, \xi_2), \xi$.
- b) Найдите матрицу ковариаций и математическое ожидание для вектора (η_1, η_2, η_3) , где $\eta_1 = \xi_1 - \xi_2, \eta_2 = \xi_1 + 2\xi_2 + 3\xi_3, \eta_3 = \xi_3$.
- c) Вычислите $M\eta_2^2, M\eta_2^3, M\eta_2^4$.
- d) Выясните, являются ли случайные величины η_1 и η_2 независимыми.
- e) Найдите $M(\xi | (\xi_1, \xi_3)), M(\xi_1 | \xi_2)$.

Решение.

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & -1 \\ 1 & -1 & 12 \end{bmatrix}$$

- а) Начинаем со случайной величины ξ_1 . Запишем её как линейную комбинацию гауссового вектора $\xi_1 = 1 \cdot \xi_1 + 0 \cdot \xi_2 + 0 \cdot \xi_3$. Следовательно, $\xi_1 \sim N(-1, 1)$, где -1 — это первый элемент \vec{a} , а 1 — элемент a_{11} матрицы ковариации.

Записываем плотность

$$p_{\xi_1}(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x+1)^2}{2}}$$

и характеристическую функцию $\varphi_{\xi_1}(t) = e^{-it - \frac{t^2}{2}}$.

Переходим к вектору (ξ_1, ξ_2) . Это гауссовый вектор, потому что произвольная линейная комбинация $\lambda_1 \xi_1 + \lambda_2 \xi_2 = \lambda_1 \xi_1 + \lambda_2 \xi_2 + 0 \cdot \xi_3$ — это случайная величина, которая должна иметь нормальное распределение. Отсюда следует, что $(\xi_1, \xi_2) = \vec{\xi}$ — гауссовый вектор.

$M\vec{\xi} = (-1, 0)$, а ковариационная матрица данного вектора — это сужение исходной ковариационной матрицы на соответствующие вектора

$$R = \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}.$$

Записываем характеристическую функцию и плотность. Матрица ковариации невырождена, так что плотность может записать

$$p_{(\xi_1, \xi_2)}(\vec{x}) = \frac{1}{(2\pi)^{\frac{3}{2}} \sqrt{2}} \cdot e^{-\frac{1}{2}(x_1+1 \ x_2) \cdot \begin{bmatrix} 3 & -1 \\ -1 & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} x_1+1 \\ x_2 \end{bmatrix}}.$$

Определитель этой матрицы $\det R = 6 - 4 = 2$. Обратная матрица

$$R^{-1} = \frac{1}{2} \cdot \begin{bmatrix} 6 & -2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -1 & \frac{1}{2} \end{bmatrix}.$$

Подставим полученное в выражение для плотности

$$\begin{aligned} & \frac{1}{(2\pi)^{\frac{3}{2}} \sqrt{2}} \cdot e^{-\frac{1}{2}(x_1+1 \ x_2) \cdot \begin{bmatrix} 3 & -1 \\ -1 & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} x_1+1 \\ x_2 \end{bmatrix}} = \\ & = \frac{1}{(2\pi)^{\frac{3}{2}} \sqrt{2}} \cdot \exp \left\{ -\frac{1}{4} (x_1+6-2 \cdot x_2-2x_1-2+2x_2) \cdot \begin{bmatrix} x_1+1 \\ x_2 \end{bmatrix} \right\} = \\ & = \frac{1}{(2\pi)^{\frac{3}{2}} \sqrt{2}} \cdot \exp \left\{ -\frac{1}{4} \cdot [6x_1(x_1+1) - 2x_1(x_1-1) - 2x_1x_2 + 2x_2^2] \right\}. \end{aligned}$$

Для вектора ξ математическое ожидание равно $(-1, 0, 2)$, а матрица ковариаций совпадает с исходной.

- b) Найдите матрицу ковариаций и математическое ожидание для вектора (η_1, η_2, η_3) , где $\eta_1 = \xi_1 - \xi_2$, $\eta_2 = \xi_1 + 2\xi_2 + 3\xi_3$, $\eta_3 = \xi_3$.
- c) Вычислите $M\eta_2^2$, $M\eta_2^3$, $M\eta_2^4$.
- d) Выясните, являются ли случайные величины η_1 и η_2 независимыми.
- e) Найдите $M(\xi \mid (\xi_1, \xi_3))$, $M(\xi_1 \mid \xi_2)$.

Домашнее задание