Занятие 3. Точечные оценки

3.1. Основные понятия

Рассмотрим вторую основную задачу математической статистики: X – случайная величина, для которой известен общий вид ее закона распределения, но который зависит от одного или нескольких неизвестных параметров; требуется оценить эти парметры. Для решения этой задачи используются два основных подхода:

1) построение точечных оценок;

2) построение доверительных интервалов.

Пусть X – случайная величина, закон распределения которой известен с точностью до параметра θ.

Определение 3.1. Точечной оценкой параметра θ называется статистика $\hat{\theta}(\overrightarrow{X})$, выборочное значение которой принимается в качестве значения этого параметра, т.е. принимается равенство:

$$\theta = \hat{\theta}(\overrightarrow{x}). \qquad \#$$

Пример 3.1. Пусть X – случайная величина, закон распределения которой неизвестен, $m = MX$ – ее математическое ожидание. В качестве оценки для m можно использовать статистики:

$$\begin{aligned}
&\hat{m}_1(\overrightarrow{X}) = \overline{X};\\
&\hat{m}_2(\overrightarrow{X}) = \frac{1}{2}\left[X_{(1)} + X_{(n)}\right];\\
&\hat{m}_3(\overrightarrow{X}) = \begin{cases} X_{(n/2)}, & \text{если } n \text{ – четное;}\\ \frac{1}{2}\left[X_{\left(\frac{n-1}{2}\right)} + X_{\left(\frac{n+1}{2}\right)}\right], & \text{если } n \text{ – нечетное,}\end{cases}\\
&\hat{m}_4(\overrightarrow{X}) = \exp\left(\sin \overline{X}^2 + 2X_{(1)}^3 - 3X_{(n)}^4\right).
\end{aligned}$$

При этом очевидно, что первые три статистики в целом являются более или менее адекватными оценками для m, в то время как последняя вряд ли даст удачный результат. #

Этот пример подчеркивает, что согласно определению (3.1) в качестве точечной оценки параметра может использоваться *любая* статистика, не обязательно принимающая значения, хоть сколько-нибудь близкие к теоретическому значению оцениваемого параметра. При этом для исследования качества точечной оценки используются следующие характеристики:

1) несмещенность,

2) состоятельность,

3) эффективность.

Определение 3.2. Оценка $\hat{\theta}$ неизвестного параметра θ называется несмещеннной, если ее математическое ожидание равно теоретическому значению оцениваемого параметра, то есть $\exists M\hat{\theta} = \theta$. #

Определение 3.3. Оценка $\hat{\theta}$ неизвестного параметра θ называется состоятельной, если при $n \to \infty$ (n – объем выборки) случайная величина $\hat{\theta}(\overrightarrow{X})$ сходится по вероятности к теоретическому значению оцениваемого параметра, то есть

$$\hat{\theta}\left(\overrightarrow{X}\right) \xrightarrow[n\to\infty]{\mathbf{P}} \theta. \qquad \#$$

Определение 3.4. Несмещенная оценка $\hat{\theta}$ неизвестного параметра θ называется эффективной, если она обладает наименьшей дисперсией среди всех несмещенных оценок этого параметра. #

Пример 3.2. Показать, что выборочная дисперсия

$$\hat{\sigma}^2(\overrightarrow{X}) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

является смещенной оценкой дисперсии.

Решение. Обозначим $m = MX$ – математическое ожидание случайной величины X. Тогда справедлива цепочка равенств:

$$\begin{aligned}
\hat{\sigma}^2(\overrightarrow{X}) &= \left|\begin{array}{l}\text{используем:}\\ \overline{X} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} X_j\end{array}\right| = \sum_{i=1}^{n}\left[X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right]^2 = \left|\begin{array}{l}\textit{добавим и вычтем}\\ \textit{m в квадратных}\\ \textit{скобках}\end{array}\right| =\\
&= \sum_{i=1}^{n}\left[(X_i - m) - \frac{1}{n}\sum_{j=1}^{n}(X_j - m)\right]^2 = \left|\begin{array}{l}\textit{раскроем квадрат}\\ \textit{на квадратных}\\ \textit{скобках}\end{array}\right| =\\
&= \sum_{i=1}^{n}\left[(X_i - m)^2 - \frac{2}{n}(X_i - m)\sum_{j=1}^{n}(X_j - m) + \frac{1}{n^2}\left[\sum_{j=1}^{n}(X_j - m)\right]^2\right] = \left|\begin{array}{l}\textit{раскроем квадрат на}\\ \textit{внутренних квадрат-}\\ \textit{ных скобках}\end{array}\right| =\\
&= \sum_{i=1}^{n}\left[A - \frac{2}{n}B + \frac{1}{n^2}\left[C + D\right]\right],
\end{aligned}$$

где использованы обозначения:

$$\begin{aligned}
A &= (X_i - m)^2,\\
B &= \sum_{j=1}^{n}(X_i - m)(X_j - m) = (X_i - m)^2 + \sum_{\substack{j=\overline{1,n}\\ j\neq i}}(X_i - m)(X_j - m),\\
C &= \sum_{j=1}^{n}(X_j - m)^2,\\
D &= 2\sum_{1\leqslant j<k\leqslant n}(X_j - m)(X_k - m).
\end{aligned}$$

Чтобы выяснить, является ли статистика $\hat{\sigma}^2$ смещенной оценкой дисперсии или нет, в соответствии с определением 3.2 найдем ее математическое ожидание:

$$M\left[\hat{\sigma}^2\left(\overrightarrow{X}\right)\right] = \sum_{i=1}^{n}\left[MA - \frac{2}{n}MB + \frac{1}{n^2}MC + \frac{1}{n^2}MD\right]. \qquad (3.1)$$

Согласно определению дисперсии и с учетом определения случайной выборки (все X_j имеют то же распределение, что и X) имеем:

$$\begin{aligned}
MA &= M\left[(X_i - m)^2\right] = \sigma^2,\\
MC &= \sum_{j=1}^{n} M\left[(X_j - m)^2\right] = \sum_{j=1}^{n}\sigma^2 = n\sigma^2.
\end{aligned}$$

Аналогично с использованием определения и свойств ковариации получаем:

$$MB = M\left[(X_i - m)^2\right] + \sum_{\substack{j=\overline{1,n}\\ j\neq i}} M\left[(X_i - m)(X_j - m)\right] = \sigma^2 + \sum_{\substack{j=\overline{1,n}\\ j\neq i}} \operatorname{cov}(X_i, X_j) =$$

$$= \left| \begin{array}{l} \textit{все элементы случайной выборки незави-} \\ \textit{симы} \Rightarrow \operatorname{cov}(X_i, X_j) = 0 \end{array} \right| = \sigma^2,$$

$$MD = 2\sum_{1\leqslant j<k\leqslant n} M\left[(X_j - m)(X_k - m)\right] = 2\sum_{1\leqslant j<k\leqslant n} \operatorname{cov}(X_j, X_k) = 0.$$

Используя полученные результаты с учетом (3.1) находим

$$M\left[\hat\sigma^2\left(\vec X\right)\right] = \sum_{i=1}^{n}\left[\sigma^2 - \frac{2}{n}\sigma^2 + \frac{n}{n^2}\sigma^2\right] = \left(1 - \frac{1}{n}\right)\sigma^2 = \frac{n-1}{n}\sigma^2 \neq \sigma^2,$$

откуда и следует, что статистика $\hat\sigma^2\left(\vec X\right)$ является смещенной оценкой дисперсии. #

Замечание 3.1. С учетом предыдущего примера можно утверждать, что оценка

$$S^2(\vec X) = \frac{n}{n-1}\hat\sigma^2\left(\vec X\right) = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \qquad (3.2)$$

является несмещенной оценкой дисперсии. Эта статистика называется ***исправленной выборочной дисперсией.*** #

Мы изучим два метода построения точечных оценок:

1) метод моментов;

2) метод максимального правдоподобия.

3.2. Метод моментов

Пусть X – случайная величина, закон распределения которой известен с точностью до вектора $\vec\theta = (\theta_1, \ldots, \theta_r)$ неизвестных парметров. Предположим, что для X существуют r первых моментов, которые, очевидно, в общем случае зависят от неизвестных параметров:

$$MX^1 = m_1(\theta_1, \ldots, \theta_r),$$
$$\ldots \quad \ldots \quad \ldots$$
$$MX^r = m_r(\theta_1, \ldots, \theta_r).$$

Метод моментов основан на том, что выборочные моменты

$$\hat\mu_k\left(\vec X\right) = \frac{1}{n}\sum_{i=1}^{n} X_i^k, \qquad k = \overline{1,r},$$

являются состоятельными оценками своих теоретических аналогов. По этой причине в рассматриваемом методе принимают равенства

$$MX^k = \hat\mu_k\left(\vec X\right), \qquad k = \overline{1,r},$$

что приводит к системе (в общем случае нелинейной) относительно неизвестных параметров:

$$\begin{cases} m_1(\theta_1, \ldots, \theta_r) = \hat\mu_1\left(\vec X\right), \\ \quad \ldots \\ m_r(\theta_1, \ldots, \theta_r) = \hat\mu_r\left(\vec X\right). \end{cases} \qquad (3.3)$$

Решение этой системы

$$\theta_1 = \hat\theta_1\left(\vec X\right), \; \ldots, \; \theta_r = \hat\theta_r\left(\vec X\right)$$

дает искомые точечные оценки.

Замечание 3.2. Иногда некоторые уравнения системы (3.3) записываают относительно центральных, а не начальных моментов. В этом случае соответствующее уравнение будет иметь вид:

$$\overset{\circ}{m}_k(\theta_1, \ldots, \theta_r) = \hat\nu_k\left(\vec X\right),$$

где

$$\overset{\circ}{m}_k(\theta_1, \ldots, \theta_r) = M\left(X - MX\right)^k, \qquad \hat\nu_k\left(\vec X\right) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^k$$

— соответственно теоретический и выборочный k-е центральные моменты. #

Пример 3.3. Пусть $X \sim \mathrm{Exp}(\lambda, \alpha)$, то есть

С использованием метода момен

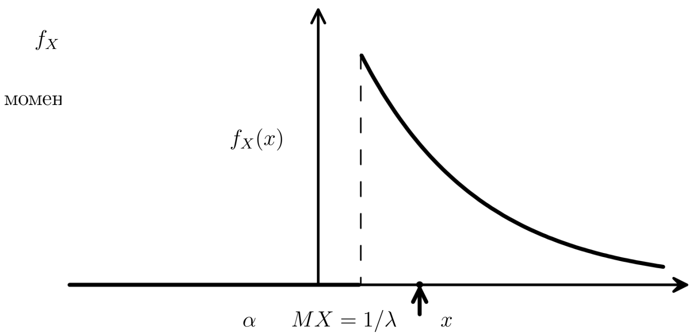

Рис. 3.1

Р е ш е н и е. В рассматриваемом примере закон распределения случайной величины X зависит от двух неизвестных параметров, поэтому система метода моментов будет содержать два уравнения, записанных относительно моментов 1-го и 2-го порядка. Запишем первое уравнение относительно начального момента, а второе относительно центрального. Первое уравнение в этом случае должно иметь вид

$$m_1(\lambda, \alpha) = \hat\mu_1\left(\vec X\right),$$

где $m_1(\lambda, \alpha) = MX$. Если вспомнить механическую интерпретацию математического ожидания как центра тяжести бесконечного стержня с плотностью $f_X(x)$, а также заметить, что случайная величина X из настоящего примера имеет "смещенный экспоенциальный" закон распределения (см. рис. 3.1), то становится очевидно, что координата центра масс величины X равна координате центра масс "обычной" экспоненциальной случайной величины плюс величина смещения, то есть

$$MX = \frac{1}{\lambda} + \alpha.$$

Поскольку $\hat\mu_1\left(\vec X\right) = \overline{X}$, первое уравнение принимает вид:

$$\frac{1}{\lambda} + \alpha = \overline{X}. \qquad (3.4)$$

Второе уравнение, записанное относительно второго центрального момента, должно иметь вид

$$\overset{\circ}{m}_2(\lambda,\alpha)=\hat\nu\left(\vec X\right),$$

где $\overset{\circ}{m}_2(\lambda,\alpha)=DX$. Механическая интерпретация дисперсии как момента инерции вероятностной массы относительно ее центра тяжести позволяет утверждать, что дисперсия рассматриваемой случайной величины X совпадает с дисперсией "обычной" экспоненциальной случайной величины, то есть

$$DX=\frac{1}{\lambda^2}.$$

Поскольку $\hat\nu\left(\vec X\right)=\hat\sigma^2\left(\vec X\right)$, второе уравнение принмает вид:

$$\frac{1}{\lambda^2}=\hat\sigma^2\left(\vec X\right). \qquad (3.5)$$

Таким образом, в рассматриваемом примере метод моментов приводит к системе, содержащей уравнения (3.4) и (3.5), а ее решение относительно неизвестных параметров λ и α дает следующие точечные оценки:

$$\lambda=\hat\lambda\left(\vec X\right)=\frac{1}{\hat\sigma\left(\vec X\right)},\qquad \alpha=\hat\alpha\left(\vec X\right)=\overline{X}-\hat\sigma\left(\vec X\right).$$

Ответ: $\hat\lambda\left(\vec X\right)=\dfrac{\sqrt{n}}{\sqrt{\sum_{i=1}^n\left(X_i-\overline{X}\right)^2}}$, $\hat\alpha\left(\vec X\right)=\overline{X}-\sqrt{\frac{1}{n}\sum_{i=1}^n\left(X_i-\overline{X}\right)^2}$. #

Пример 3.4. Пусть $X\sim N(m,\theta^2)$. С использованием метода моментов построить оценки параметров m и θ^2.

Решение. Как и в предыдущем примере, закон распределения случайной величины X содержит два неизвестных параметра, поэтому система уравнений метода моментов также будет состоять из двух уравнений, записанных для моментов 1-го и 2-го порядков. Как и ранее, запишем первое уравнение относительно начального момента, а второе относительно центрального. Как известно, для случайной величины X из настоящего примера

$$MX=m_1=m,\qquad M\left(X-MX\right)^2=\overset{\circ}{m}_2=DX=\theta^2,$$

поэтому система будет иметь вид

$$\begin{cases} m=\overline{X},\\ \theta^2=\hat\sigma^2\left(\vec X\right).\end{cases}$$

Полученная система состоит из двух независимых уравнений, которые уже по сути разрешены относительно неизвестных параметров. Поэтому остается лишь записать

Ответ: $\hat m=\overline{X}$, $\hat\theta^2=\hat\sigma^2\left(\vec X\right)=\dfrac{1}{n}\sum_{i=1}^n\left(X_i-\overline{X}\right)^2$. #

Замечание 3.3. В обоих рассмотренных выше примерах 3.3 и 3.4 при составлении второго уравнения вместо статистики $\hat\sigma^2\left(\vec X\right)$ можно было использовать определенную соотношением (3.2) статистику $S^2\left(\vec X\right)$, которая является несмещенной оценкой дисперсии. В этом случае ответ примера (3.3) имел бы вид

$$\hat\lambda\left(\vec X\right)=\frac{1}{S\left(\vec X\right)},\qquad \hat\alpha\left(\vec X\right)=\overline{X}-S\left(\vec X\right),$$

а ответ примера (3.4) —

$$\hat m\left(\vec X\right)=\overline{X},\qquad \hat\theta^2\left(\vec X\right)=S^2\left(\vec X\right).$$

Впрочем, подобная щепетильность не имеет существенного значения, поскольку уже при сравнительно небольших n значения статистик S^2 и $\hat\sigma^2$ близки. #

Пример 3.5. Пусть $X\sim\Pi(\lambda)$. С использованием метода моментов построить оценку параметра λ.

Решение. Закон распределения случайной величины X зависит от одного неизвестного параметра, поэтому система уравнений метода моментов также будет состоять из одного уравнения, записанного для момента 1-го порядка:

$$m_1(\lambda)=\overline{X}.$$

Как известно, для случайной величины X из настоящего примера $MX=\lambda$, поэтому искомое уравнение примет вид

$$\lambda=\overline{X}.$$

Как и в предыдущем примере, это уравнение разрешено относительно неизвестного параметра, поэтому имеет место

Ответ: $\hat\lambda\left(\vec X\right)=\overline{X}$. #

3.3. Метод максимального правдоподобия

Пусть X – случайная величина, закон распределения которой зависит от вектора парметров $\vec\theta=(\theta_1,\ldots,\theta_r)$, $\vec X=(X_1,\ldots,X_n)$ – случайная выборка объема n из генеральной совокупности X.

Определение 3.5. Функцией правдоподобия случайной выборки $\vec X$ называется функция

$$\mathcal{L}\left(\vec X,\vec\theta\right)=p\left(X_1,\vec\theta\right)\cdot\ldots\cdot p\left(X_n,\vec\theta\right),$$

где

$$p\left(x,\vec\theta\right)=\begin{cases} P\{X=x\}, & \text{если } X \text{ – дискретная случайная величина},\\ f_X\left(x,\vec\theta\right), & \text{если } X \text{ – непрерывная случайная величина}\end{cases}$$

(здесь f_X – функция плотности случайной величины X). #

В методе максимального правдоподобия в качестве точечных оценок неизвестных параметров принимают такие их значения, которые максимизируют функцию правдоподобия, то есть полагают

$$\hat{\vec\theta}=\arg\max_{\vec\theta}\mathcal{L}\left(\vec X,\vec\theta\right).$$

При рассмотрении этой задачи оптимизации обычно используют необходимое условие экстремума функции нескольких переменных

$$\frac{\partial \mathcal{L}\left(\overrightarrow{X}, \theta_1, \ldots, \theta_r\right)}{\partial \theta_1} = 0,$$
$$\ldots$$
$$\frac{\partial \mathcal{L}\left(\overrightarrow{X}, \theta_1, \ldots, \theta_r\right)}{\partial \theta_r} = 0.$$

Указанные соотношения при этом называют ***уравнениями правдоподобия.***

Замечание 3.4. Функция правдоподобия представляет собой произведение n сомножителей, зависящих от вектора $\vec{\theta}$ неизвестных парметров, поэтому вычисление входящих в уравнения правдоподобия производных не очень удобно. По этой причине вместо задачи

$$\mathcal{L}\left(\overrightarrow{X}, \vec{\theta}\right) \longrightarrow \max_{\vec{\theta}}$$

обычно рассматривают эквивалентную задачу

$$\ln \mathcal{L}\left(\overrightarrow{X}, \vec{\theta}\right) \longrightarrow \max_{\vec{\theta}}$$

(задачи и в самом деле эквивалентны, поскольку логарифм является монотонно возрастающей функцией). В этом случае необходимое условие экстремума принимает вид

$$\frac{\partial \ln \mathcal{L}\left(\overrightarrow{X}, \theta_1, \ldots, \theta_r\right)}{\partial \theta_j} = 0, \qquad j = \overline{1, r},$$

а указанные соотношения также называют уравнениями правдоподобия. #

Пример 3.6. Пусть $X \sim \Pi(\lambda)$. С использованием метода максимального правдоподобия построить точечную оценку параметра λ.

Решение. Пуассоновкая случайная величина является дискретной и принимает целые неотрицательные значения с вероятностями

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k \in \mathbb{N}_0,$$

поэтому функция правдоподобия в рассматриваемом случае имеет вид

$$\mathcal{L}\left(\overrightarrow{X}, \lambda\right) = P\{X = X_1\} \cdot \ldots \cdot P\{X = X_n\} = \\ = \left(\frac{\lambda^{X_1}}{X_1!} e^{-\lambda}\right) \cdot \ldots \cdot \left(\frac{\lambda^{X_n}}{X_n!} e^{-\lambda}\right) = \frac{\lambda^{\sum_{i=1}^n X_i}}{X_1! \cdot \ldots \cdot X_n!} e^{-n\lambda},$$

а ее логарифм —

$$\ln \mathcal{L}\left(\overrightarrow{X}, \lambda\right) = \ln\left(\lambda^{\sum_{i=1}^n X_i}\right) - \ln\left(X_1! \cdot \ldots \cdot X_n!\right) + \ln e^{-n\lambda} = \ln \lambda \sum_{i=1}^n X_i - \ln\left(X_1! \cdot \ldots \cdot X_n!\right) - n\lambda.$$

Для построения оценки максимального правдоподобия параметра λ нужно решить задачу оптимизации

$$\ln \mathcal{L}\left(\overrightarrow{X}, \lambda\right) \longrightarrow \max_{\lambda}.$$

Использование необходимого условия экстремума (мы рассматриваем $\ln \mathcal{L}\left(\overrightarrow{X}, \lambda\right)$ как функцию одного переменного λ) дает единственное уравнение правдоподобия

$$\frac{\partial \ln \mathcal{L}\left(\overrightarrow{X}, \lambda\right)}{\partial \lambda} = \frac{1}{\lambda} \sum_{i=1}^n X_i - n = 0,$$

из которого находим единственную критическую точку

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i = \overline{X}.$$

Осталось доказать, что найденное значение $\hat{\lambda}$ действительно доставляет функции $\ln \mathcal{L}\left(\overrightarrow{X}, \lambda\right)$ наибольшее значение. Для этого можно использовать одно из достаточных условий локального макисмума (вторая производная в стационарной точке должна быть отрицательной). В самом деле,

$$\left.\frac{\partial^2 \ln \mathcal{L}\left(\overrightarrow{X}, \lambda\right)}{\partial \lambda^2}\right|_{\lambda = \hat{\lambda}} = \left.\left[-\frac{1}{\lambda^2} \sum_{i=1}^n X_i\right]\right|_{\lambda = \hat{\lambda}} = \left.\left[-\frac{n}{\lambda^2} \overline{X}\right]\right|_{\lambda = \hat{\lambda}} = -\frac{n}{\overline{X}^2} \overline{X} = -\frac{n}{\overline{X}} < 0$$

(так как $X_i \sim \Pi(\lambda)$, то $X_i \geqslant 0$, поэтому $\overline{X} \geqslant 0$), что завершает доказательство.

Ответ: $\hat{\lambda}\left(\overrightarrow{X}\right) = \overline{X}$. #

Пример 3.7. Пусть $X \sim N(m, \theta)$. С использованием метода максимального правдоподобия построить точечные оценки параметров m и θ.

Решение. Нормальная случайная величина является непрерывной, а ее функция плотности имеет вид

$$f_X(x, m, \theta) = \frac{1}{\sqrt{2\pi}\sqrt{\theta}} e^{-\frac{(x-m)^2}{2\theta}}, \qquad x \in \mathbb{R},$$

поэтому в рассматриваемом случае функция правдоподобия принимает вид

$$\mathcal{L}\left(\overrightarrow{X}, m, \theta\right) = f_X(X_1, m, \theta) \cdot \ldots \cdot f_X(X_n, m, \theta) = \\ = \left(\frac{1}{\sqrt{2\pi}\sqrt{\theta}} e^{-\frac{(X_1-m)^2}{2\theta}}\right) \cdot \ldots \cdot \left(\frac{1}{\sqrt{2\pi}\sqrt{\theta}} e^{-\frac{(X_n-m)^2}{2\theta}}\right) = (2\pi)^{-n/2} \theta^{-n/2} \exp\left[-\frac{1}{2\theta} \sum_{i=1}^n (X_i - m)^2\right],$$

а ее логарифм —

$$\ln \mathcal{L}\left(\overrightarrow{X}, m, \theta\right) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \theta - \frac{1}{2\theta} \sum_{i=1}^n (X_i - m)^2.$$

Для построения оценки максимального правдоподобия нужно решить задачу

$$\ln \mathcal{L}\left(\overrightarrow{X}, m, \theta\right) \longrightarrow \max_{m, \theta}.$$

Как и ранее, воспользуемся необходимым условием экстремума функции теперь уже двух переменных m и θ:

$$\frac{\partial \ln \mathcal{L}\left(\overrightarrow{X}, m, \theta\right)}{\partial m} = \frac{2}{2\theta}\sum_{i=1}^{n}(X_i - m) = \frac{1}{\theta}\left[\sum_{i=1}^{n} X_i - nm\right] = 0,$$
$$\frac{\partial \ln \mathcal{L}\left(\overrightarrow{X}, m, \theta\right)}{\partial \theta} = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}(X_i - m)^2 = 0.$$

Решая полученные уравнения правдоподобия относительно m и θ, найдем координаты критической точки:

$$\hat{m} = \frac{1}{m}\sum_{i=1}^{n} X_i = \overline{X},$$
$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 = \hat{\sigma}^2\left(\overrightarrow{X}\right).$$

Можно показать, что для найденных значений $\hat{m}$ и $\hat{\theta}$ выполняются достаточные условия локального максимума функции двух переменных, поэтому записываем

Ответ: $\hat{m}\left(\overrightarrow{X}\right) = \overline{X}$, $\hat{\theta}\left(\overrightarrow{X}\right) = \hat{\sigma}^2\left(\overrightarrow{X}\right)$. #

Пример 3.8. Пусть $X \sim \mathrm{Exp}(\lambda, \alpha)$ (см. пример 3.3). С использованием метода максимального правдоподобия построить точечные оценки параметров λ и α.

Решение. Составим функцию правдоподобия:

$$\mathcal{L}(\vec{X}, \lambda, \alpha) = f_X(X_1, \lambda, \alpha) \cdot \ldots \cdot f_X(X_n, \lambda, \alpha) = \lambda^n e^{-n\lambda(\overline{X} - \alpha)}, \tag{3.6}$$

откуда

$$\ln \mathcal{L} = -n \ln \lambda - n\lambda(\overline{X} - \alpha).$$

Уравнения правдоподобия принимают вид

$$\begin{aligned} \frac{\partial \ln \mathcal{L}}{\partial \lambda} &= \frac{n}{\lambda} - n(\overline{X} - \alpha) = 0, \\ \frac{\partial \ln \mathcal{L}}{\partial \alpha} &= -n\lambda = 0, \end{aligned} \tag{3.7}$$

откуда следует противоречивый результат $\lambda = 0$.

Проблема в том, что в настоящем примере дифференцирование функции правдоподобия по параметру α некорректно, потому что границы области, в которой $f_X > 0$, зависят от значения α. Как следствие, правая часть соотношения (3.6) верна лишь в том случае, если все элементы выборки больше либо равны α. Если хотя бы один элемент X_i имеет значение, которое меньше α, то $f_X(X_i, \lambda, \alpha) = 0$ и, следовательно, $\mathcal{L}(\vec{X}, \lambda, \alpha) = 0$.

Чтобы функция правдоподобия имела положительное значение, необходимо потребовать выполнение условия $\alpha \leqslant X_i$, $i = \overline{1, n}$, или, что эквивалентно, условия

$$\alpha \leqslant X_{(1)}. \tag{3.8}$$

Из рис. 3.2 видно, что чем меньше значение α, тем меньшие значения принимает функция $\mathcal{L}$, поэтому значение α должно быть м
Таким образом, в качестве оцен
$\alpha = X_{(1)}$.

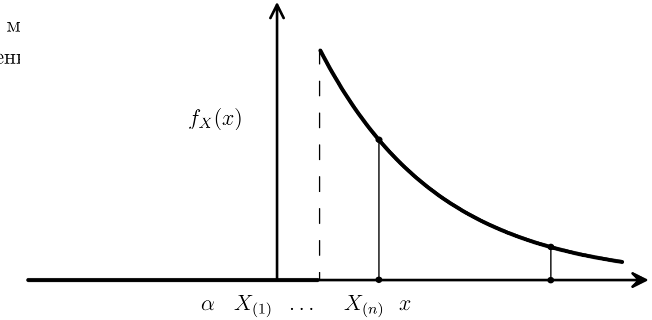

Рис. 3.2

Для нахождения оценки максимального правдоподобия параметра λ можно использовать стандартную технику. С учетом (3.8) справедливо соотношение (3.6) и, следовательно, первое уравнение из (3.7), из которого получаем

$$\lambda = \frac{1}{\overline{X} - X_{(1)}}.$$

Ответ: $\hat{\lambda}(\vec{X}) = \dfrac{1}{\overline{X} - X_{(1)}}$, $\hat{\alpha}(\vec{X}) = X_{(1)}$. #