1. Оценка параметров

1.1. Обобщенное распределение Пуассона

Случайная величина ζ имеет обобщенное Пуассоновское распределение $\mathcal{P}(\lambda|f(x))$, если она равна сумме τ независимых одинаково распределенных случайных величин ξ_i с плотностью распределения $f(x)$, где $\tau \sim \mathcal{P}(\lambda)$,

$$\zeta = \begin{cases} 0, & \text{при } \tau = 0 \text{ с} \quad \text{вероятностью} \quad p_0 = e^{-\lambda} \\ \xi_1 + \ldots + \xi_\tau, & \text{при } \tau > 0 \text{ с} \quad \text{вероятностью} \quad 1 - e^{-\lambda} \end{cases}.$$

Обозначим через $F_j(x) = P\{\xi_1 + \ldots + \xi_j \leq x\}$ функцию распределения и через $f_j(x)$ соответствующую плотность, через μ, σ^2 соответствующие МО и дисперсию. Применим формулу полной вероятности в

$$P\{\xi_1 + \ldots + \xi_\tau \leq x|\tau > 0\} = \frac{P\{\xi_1 + \ldots + \xi_\tau \leq x, \tau > 0\}}{P\{\tau > 0\}} =$$

$$= \frac{1}{1 - p_0} \sum_{j=1}^{\infty} p_j P\{\xi_1 + \ldots + \xi_j \leq x\} = \frac{1}{1 - p_0} \sum_{j=1}^{\infty} p_j F_j(x)\,,$$

таким образом, условная плотность имеет вид $\frac{1}{1-p_0} \sum_{j=1}^{\infty} p_j f_j(x)$. Обозначим через $g(t)$ ПФМ для СВ ξ_i. Вычислим ПФМ для СВ ζ.

$$\phi(t) = \mathbb{E}e^{t\zeta} = p_0 e^{t\cdot 0} + (1 - p_0) \int_{-\infty}^{\infty} \frac{1}{1 - p_0} \sum_{j=1}^{\infty} p_j f_j(x) e^{tx} dx =$$

$$= p_0 + \sum_{j=1}^{\infty} p_j \int_{-\infty}^{\infty} f_j(x) e^{tx} dx = p_0 + \sum_{j=1}^{\infty} p_j (g(t))^j =$$

$$= \sum_{j=0}^{\infty} p_j (g(t))^j = e^{-\lambda + \lambda g(t)}\,,$$

так как $p_j = \frac{\lambda^j}{j!} e^{-\lambda}$, и ПФ распределения Пуассона имеет вид

$$\sum_{j=0}^{\infty} p_j \nu^j = \sum_{j=0}^{\infty} \frac{(\lambda\nu)^j}{j!} e^{-\lambda} = e^{-\lambda + \lambda\nu}\,.$$

1.2. Метод моментов

Рассматривая количество моментов, равное количеству оцениваемых параметров, и приравнивая выборочные моменты к соответствующим моментам распределения, являющимся функциями от параметров, можно получить оценки параметров распределения.

В случае сложно-пуассоновского распределения $\mathcal{P}(\lambda|\mathcal{N}(\mu,\sigma))$ с ПФМ $g(t)$ внутреннего нормального распределения имеет вид $\phi(t) = e^{-\lambda+\lambda g(t)}$. Используя то, что $g'(0) = \mu$, $g''(0) = \mu^2 + \sigma^2$, можем вычислить $\mathbb{E}\zeta$, $\mathbb{D}\zeta$.

$$\phi'(t) = e^{-\lambda+\lambda g(t)}\lambda g'(t)\,, \quad \phi'(0) = \lambda\mu$$

$$\phi''(t) = e^{-\lambda+\lambda g(t)}(\lambda g'(t))^2 + e^{-\lambda+\lambda g(t)}\lambda g''(t)\,,\ \phi''(0) = \lambda^2\mu^2 + \lambda(\mu^2+\sigma^2),$$

$$\mathbb{E}\zeta = \lambda\mu, \quad \mathbb{D}\zeta = \lambda(\mu^2+\sigma^2)\,.$$

Если имеется выборка $x_1, \ldots, x_n$ и $\bar{x}$ выборочное среднее, S^2 несмещенная

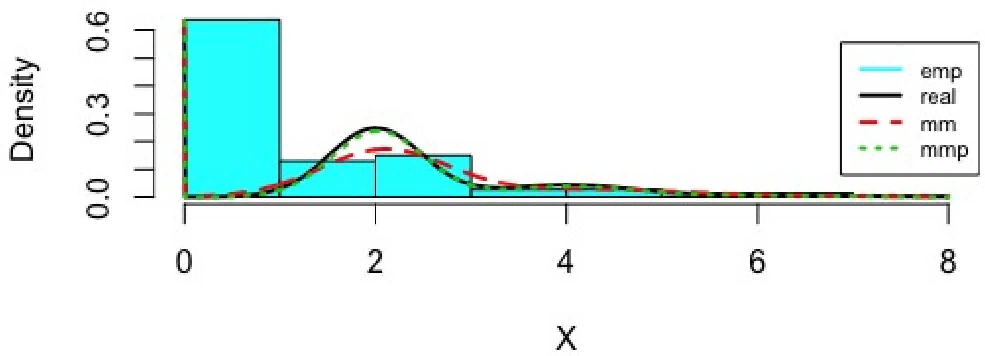

Рис. 1. Сложно-пуассоновское распределение, согласование гистограммы с плотностями при разных параметрах.

оценка дисперсии, то в $\hat{p}_0 = \frac{n_0}{n}$, где n_0 число нулей в выборке,

$$\hat{\lambda} = -ln(\hat{p}_0), \quad \hat{\mu} = \frac{\bar{x}}{\hat{\lambda}}, \quad \hat{\sigma^2} = \frac{S^2}{\hat{\lambda}} - \hat{\mu^2}\,. \tag{1}$$

1.3. Метод максимального правдоподобия

Пусть $f(x, \theta)$ — генеральная плотность распределения, зависящая от параметра θ, $\mathbf{x} = (x_1, \ldots, x_n)$ — случайная выборка наблюдений. Функцией правдоподобия называется вероятность появления данной выборки

$$\mathcal{L}(\mathbf{x}, \theta) = f(x_1, \theta) \cdot \ldots \cdot f(x_n, \theta)\,. \tag{2}$$

Метод максимального правдоподобия заключается в том, что в качестве оценки параметра принимается значение $\hat{\theta}$, при котором функция правдоподобия достигает своего максимального значения. Поскольку $\ln \mathcal{L}(\mathbf{x}, \theta)$ достигает максимума в той же точке, что и $\mathcal{L}(\mathbf{x}, \theta)$, следует решать относительно θ *уравнение правдоподобия.*

$$\frac{\partial \ln \mathcal{L}(\mathbf{x}, \theta)}{\partial \theta} = 0\,. \tag{3}$$

Оценкой максимального правдоподобия (ОМП) является каждое решение уравнения (3) .

Предложение 1. *При некоторых общих условиях уравнение правдоподобия имеет решение, сходящееся по вероятности к истинному значению параметра при $n \to \infty$. Оно является асимптотически нормальной и асимптотически эффективной оценкой для параметра θ.*

1. При каждом $\theta \in A$, где A невырожденный интервал, для почти всех x существуют производные $\frac{\partial \ln f}{\partial \theta}$, $\frac{\partial^2 \ln f}{\partial \theta^2}$, $\frac{\partial^3 \ln f}{\partial \theta^3}$.

2. При каждом $\theta \in A$ имеем $\left|\frac{\partial f}{\partial \theta}\right| < F_1(x)$, $\left|\frac{\partial^2 f}{\partial \theta^2}\right| < F_2(x)$, $\left|\frac{\partial^3 \ln f}{\partial \theta^3}\right| < H(x)$, $F_1(x)$ и $F_2(x)$ интегрируемы на $(-\infty; +\infty)$, $\int\limits_{-\infty}^{+\infty} H(x) f(x, \theta) dx < M$, причем M не зависит от θ.

3. При каждом $\theta \in A$ интеграл $\int\limits_{-\infty}^{+\infty} \left(\frac{\partial \ln f}{\partial \theta}\right)^2 f(x, \theta) dx$ конечен и положителен.

Пусть θ_0 — истинное значение параметра и θ_0 — внутренняя точка интервала A. Покажем, что уравнение правдоподобия имеет решение, сходящееся по вероятности к θ_0.

$$\frac{\partial \ln f}{\partial \theta} = \left(\frac{\partial \ln f}{\partial \theta}\right)_0 + (\theta - \theta_0)\left(\frac{\partial^2 \ln f}{\partial \theta^2}\right)_0 + \frac{1}{2}\Theta(\theta-\theta_0)^2 H(x),$$

где $|\Theta| < 1$, а индексом 0 указано, что для каждого $\theta \in A$ следует положить $\theta = \theta_0$. Умножим уравнение (3) на $\frac{1}{n}$.

$$\frac{1}{n}\frac{\partial \ln \mathcal{L}}{\partial \theta} = B_0 + B_1(\theta-\theta_0) + \frac{1}{2}\Theta B_2(\theta-\theta_0)^2 = 0, \tag{4}$$

где, обозначая через $f_i = f(x_i, \theta)$, имеем

$$B_0 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial \ln f_i}{\partial \theta}\right)_0, \quad B_1 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial^2 \ln f_i}{\partial \theta^2}\right)_0,$$
$$B_2 = \frac{1}{n}\sum_{i=1}^{n} H(x_i)\,.$$

Заметим, что по аналогии со свойствами информанта $s(\mathbf{x}, \theta)$

$$\mathbf{E}\left(\frac{\partial \ln f_i}{\partial \theta}\right)_0 = \int\limits_{-\infty}^{+\infty}\left(\frac{\partial \ln f_i}{\partial \theta}\right)_0 f(x_i, \theta_0)dx_i = 0\,,$$

$$\mathbf{E}\left(\frac{\partial^2 \ln f}{\partial \theta^2}\right)_0 = \int\limits_{-\infty}^{+\infty}\left(\frac{1}{f}\frac{\partial^2 f}{\partial \theta^2} - \left(\frac{1}{f}\frac{\partial f}{\partial \theta}\right)^2\right)_0 f(x_i, \theta_0)dx =$$
$$= -\mathbf{E}\left(\frac{\partial \ln f}{\partial \theta}\right)_0^2 = -k^2\,.$$

Итак, величина B_0 есть среднее арифметическое n независимых одинаково распределенных случайных величин с нулевым математическим ожиданием. Из теоремы Хинчина следует, что B_0 сходится по вероятности к нулю. Аналогично B_1 сходится по вероятности к $-k^2$, а B_2 сходится к неотрицательному значению $\mathbf{E}H(x) < M$.

Покажем, что для произвольно малых δ и ϵ уравнение правдоподобия имеет с вероятностью, превышающей $1-\epsilon$, корень, заключенный в пределах $\theta_0 \pm \delta$, если $n > n(\delta, \epsilon)$. Пусть S множество точек $(x_1, \ldots, x_n)$, для которых выполнены неравенства:

$$|B_0| < \delta^2, \quad B_1 < -\frac{k^2}{2}, \quad |B_2| < 2M\,.$$
$$\forall\; n > n(\delta, \epsilon) \quad P_0 = P\left\{|B_0| > \delta^2\right\} < \frac{\epsilon}{3},$$
$$P_1 = P\left\{B_1 > -\frac{k^2}{2}\right\} < \frac{\epsilon}{3}, \quad P_2 = P\left\{|B_2| > 2M\right\} < \frac{\epsilon}{3}\,.$$

Дополнительное множество $\bar{S}$ состоит из точек, для которых справедливо хотя бы одно из этих неравенств, поэтому

$$P(\bar{S}) < P_0 + P_1 + P_2 < \epsilon,$$
$$P(S) > 1 - \epsilon\,.$$

При $\theta = \theta_0 \pm \delta$ правая часть уравнения (3) имеет вид

$$B_0 \pm B_1\delta + \frac{1}{2}\Theta B_2\delta^2\,.$$

В каждой точке $\mathbf{x} \in S$

$$|B_0 + \frac{1}{2}\Theta B_2 \delta^2| < \delta^2 + \frac{1}{2} \cdot 2M\delta^2 = (M+1)\delta^2,$$

$$B_1\delta < -\frac{k^2\delta}{2}\,.$$

При $\theta = \theta_0 - \delta$ и при $\delta^2(M+1) < \frac{k^2\delta}{2}$ имеем

$$\frac{1}{n}\frac{\partial \ln \mathcal{L}}{\partial \theta} = \left(B_0 + \frac{1}{2}\Theta B_2 \delta^2\right) + (-B_1\delta) > 0,$$

$$\text{так как} \quad (-B_1\delta) > \frac{k^2\delta}{2} > \delta^2(M+1), \quad \text{а при} \quad \theta = \theta_0 + \delta$$

$$\frac{1}{n}\frac{\partial \ln \mathcal{L}}{\partial \theta} = \left(B_0 + \frac{1}{2}\Theta B_2 \delta^2\right) + B_1\delta < 0,$$

$$\text{так как} \quad B_1\delta < -\frac{k^2\delta}{2} < -\delta^2(M+1) < -\left(B_0 + \frac{1}{2}\Theta B_2 \delta^2\right)\,.$$

Итак, существование решения θ^* уравнения правдоподобия установлено. Из уравнения (4)

$$B_0 + B_1(\theta - \theta_0) + \frac{1}{2}\Theta B_2(\theta - \theta_0)^2 = 0$$

выразим

$$B_0 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial \ln f_i}{\partial \theta}\right)_0 = (\theta^* - \theta)\left(-B_1 - \frac{1}{2}\Theta B_2(\theta^* - \theta)\right)\,,$$

$$k\sqrt{n}(\theta^* - \theta_0) = \frac{\frac{1}{k\sqrt{n}}\sum\limits_{i=1}^{n}\left(\frac{\partial \ln f_i}{\partial \theta}\right)_0}{-\frac{B_1}{k^2} - \frac{1}{2}\Theta B_2 \frac{\theta^* - \theta_0}{k^2}}\,.$$

Знаменатель дроби сходится по вероятности к 1. Так как

$$\left(\frac{\partial \ln f_i}{\partial \theta}\right)_0$$

имеет математическое ожидание 0 и стандартное отклонение k, сумма $\sum\limits_{i=1}^{n}\left(\frac{\partial \ln f_i}{\partial \theta}\right)_0$ асимптотически нормальна $\mathcal{N}(0, k\sqrt{n})$.

Следовательно, числитель дроби асимптотически нормален $\mathcal{N}(0,1)$, а оценка θ^* асимптотически нормальна $\mathcal{N}(0, \frac{1}{k\sqrt{n}})$, где

$$k^2 = \mathbf{E}\left(\frac{\partial \ln f}{\partial \theta}\right)_0^2\,.$$

Асимптотическая эффективность равна

$$e(\theta^*) = \left(\frac{1}{nk^2} \cdot I(\theta)\right)^{-1} = 1\,,$$

$$\text{так как} \quad I(\theta) = -\mathbf{E}\frac{\partial^2}{\partial \theta^2}\ln \mathcal{L}(\mathbf{x}, \theta) = nk^2\,.$$

1.4. Свойства оценок параметров

Помимо несмещенности оценок параметров по выборке, которая означает равенство математического ожидания оценки истинному значению параметра, используется свойство *состоятельности*, означающее сходимость по вероятности к истинному значению параметра и *эффективности* в смысле минимальности дисперсии оценки.

Определение 1. Оценка $\hat{\theta}(x_1, \ldots, x_n) = \hat{\theta}_n$ является *состоятельной*, если она по вероятности сходится к истинному значению параметра, т.е.

$$\lim_{n \longrightarrow \infty} P\{|\hat{\theta}_n - \theta| < \epsilon\} = 1 \quad \forall \epsilon > 0\,.$$

Теорема 1. *Если оценка несмещенная, и ее дисперсия стремится к нулю при увеличении объема выборки, то оценка состоятельна.*

Действительно, по неравенству Чебышева,

$$P\{|\hat{\theta}_n - E\hat{\theta}_n| < \epsilon\} = P\{|\hat{\theta}_n - \theta| < \epsilon\} > 1 - \frac{D\hat{\theta}_n}{\epsilon^2} \longrightarrow 1.$$

Выборочное среднее $\hat{\mu} = \bar{x}$ является несмещенной и состоятельной оценкой математического ожидания μ.

$$\mathbf{E}\bar{x} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}x_i = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$$

$$\mathbf{D}\bar{x} = \mathbf{D}\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} \mathbf{D}x_k = \frac{1}{n^2}\sum_{k=1}^{n} \sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}.$$

Определение 2. Оценка с меньшей дисперсией называется *более эффективной*, то есть если $\mathbf{D}\hat{\theta}_1 < \mathbf{D}\hat{\theta}_2$, то $\hat{\theta}_1$ эффективнее $\hat{\theta}_2$.

Предложение 2. *(неравенство Рао-Крамера) Пусть $b(\theta) = \mathbf{E}\hat{\theta}_n - \theta$ смещение оценки $\hat{\theta}_n$ и информант второго рода $I(\theta) = -\mathbf{E}\frac{\partial^2}{\partial\theta^2}\ln\mathcal{L}(\mathbf{x}, \theta)$, тогда $I(\theta) = -\mathbf{E}\frac{\partial}{\partial\theta}s(\mathbf{x}, \theta)$ и*

$$\mathbf{D}\hat{\theta}_n \geq \frac{(1 + b'(\theta))^2}{I(\theta)}\,.$$

Выражение $I(\theta)/n$ называется информационным количеством Фишера. Производная по параметру θ логарифма функции правдоподобия (2) называется *информантом первого рода* $s(\mathbf{x},\theta)=\frac{\partial}{\partial\theta}\ln\mathcal{L}(\mathbf{x},\theta)$, для которого справедливы свойства:

$$1)\quad \mathbf{E}s(\mathbf{x},\theta)=\mathbf{E}\frac{\partial}{\partial\theta}\ln\mathcal{L}(\mathbf{x},\theta)=\int\limits_{\mathcal{X}^n}\frac{\partial}{\partial\theta}\ln\mathcal{L}(\mathbf{x},\theta)\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=$$

$$=\int\limits_{\mathcal{X}^n}\frac{\mathcal{L}'(\mathbf{x},\theta)}{\mathcal{L}(\mathbf{x},\theta)}\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=\int\limits_{\mathcal{X}^n}\frac{\partial}{\partial\theta}\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=\frac{\partial}{\partial\theta}\int\limits_{\mathcal{X}^n}\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=0$$

$$2)\quad \mathbf{cov}(\hat\theta_n,s(\mathbf{x},\theta))=\mathbf{E}(\hat\theta_n-\mathbf{E}\hat\theta_n)(s(\mathbf{x},\theta)-\mathbf{E}s(\mathbf{x},\theta))=\mathbf{E}(\hat\theta_n-\mathbf{E}\hat\theta_n)s(\mathbf{x},\theta)=$$

$$=\mathbf{E}\hat\theta_n s(\mathbf{x},\theta)-(\theta+b(\theta))\mathbf{E}s(\mathbf{x},\theta)=\mathbf{E}\hat\theta_n s(\mathbf{x},\theta)=\int\limits_{\mathcal{X}^n}\hat\theta_n\frac{\partial}{\partial\theta}\ln\mathcal{L}(\mathbf{x},\theta)\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=$$

$$=\int\limits_{\mathcal{X}^n}\hat\theta_n\mathcal{L}'(\mathbf{x},\theta)d(\mathbf{x})=\frac{\partial}{\partial\theta}\int\limits_{\mathcal{X}^n}\hat\theta_n\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=\frac{\partial}{\partial\theta}\mathbf{E}\hat\theta_n=1+b'(\theta)\,.$$

Для доказательства неравенства Рао-Крамера воспользуемся свойством коэффициента корреляции.

$$\mathbf{cor}^2(\hat\theta_n,s(\mathbf{x},\theta))=\frac{\mathbf{cov}^2(\hat\theta_n,s(\mathbf{x},\theta))}{\mathbf{D}\hat\theta_n\mathbf{D}s(\mathbf{x},\theta)}\le 1\quad\Longleftrightarrow\mathbf{D}\hat\theta_n\ge\frac{(1+b'(\theta))^2}{\mathbf{D}s(\mathbf{x},\theta)},$$

$$I(\theta)=-\mathbf{E}\frac{\partial^2}{\partial\theta^2}\ln\mathcal{L}(\mathbf{x},\theta)=-\int\limits_{\mathcal{X}^n}\frac{\partial}{\partial\theta}\left(\frac{\partial}{\partial\theta}\ln\mathcal{L}(\mathbf{x},\theta)\right)\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=$$

$$=-\int\limits_{\mathcal{X}^n}\frac{\mathcal{L}''\mathcal{L}-\mathcal{L}'\mathcal{L}'}{\mathcal{L}^2}\mathcal{L}d(\mathbf{x})=\int\limits_{\mathcal{X}^n}\left(\frac{\mathcal{L}'}{\mathcal{L}}\right)^2\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=$$

$$=\int\limits_{\mathcal{X}^n}\left(\frac{\partial}{\partial\theta}\ln\mathcal{L}(\mathbf{x},\theta)\right)^2\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=\int\limits_{\mathcal{X}^n}(s(\mathbf{x},\theta))^2\,\mathcal{L}(\mathbf{x},\theta)d(\mathbf{x})=\mathbf{D}s(\mathbf{x},\theta)\,.$$

Определение 3. Несмещенная оценка, дисперсия которой достигает наименьшего значения, равного $I(\theta)^{-1}$, называется *эффективной*. *Эффективностью* несмещенной оценки, удовлетворяющей условиям регулярности, называется величина

$$e(\hat\theta_n)=\frac{1}{I(\theta)\mathbf{D}\hat\theta_n}\,. \tag{5}$$

Для эффективных оценок $e(\hat\theta_n)=1$. *Асимптотически эффективной* называется оценка $\hat\theta_n=\hat\theta_n$, у которой

$$\lim_{n\longrightarrow\infty}e(\hat\theta_n)=1\,.$$

1.5. Задачи для практики

Предложение 3.

$$\mathbf{D}m_2 = \frac{\mu_4 - \mu_2^2}{n} - \frac{2(\mu_4 - 2\mu_2^2)}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}\,. \tag{6}$$

Доказательство. Поскольку центральные моменты инвариантны относительно сдвига, поэтому, совмещая начало отсчета со средним значением, мы получим $\mu = 0$.

$$\mathbf{D}m_2 = \mathbf{E}m_2^2 - (\mathbf{E}m_2)^2\,.$$

Из (**??**) следует

$$\mathbf{E}m_2 = \mu_2 - \frac{\mu_2}{n}.$$

Далее $m_2^2 = a_2^2 - 2a_2\bar{x}^2 + \bar{x}^4$. Вычисляем $\mathbf{E}m_2^2$.

$$\mathbf{E}a_2^2 = \frac{1}{n^2}\mathbf{E}\left(\sum_{i=1}^n x_i^2\right)^2 = \frac{1}{n^2}\left(\sum_{i=1}^n \mathbf{E}x_i^4 + \sum_{i=1}^n \sum_{j=1, j\neq i}^n \mathbf{E}x_i^2 x_j^2\right) =$$

$$= \frac{n\mu_4 + n(n-1)\mu_2^2}{n^2} = \frac{\mu_4 + (n-1)\mu_2^2}{n} = \mu_2^2 + \frac{\mu_4 - \mu_2^2}{n}\,.$$

Перейдем к основной компоненте второго слагаемого.

$$\mathbf{E}a_2\bar{x}^2 = \frac{1}{n^3}\mathbf{E}\left(\sum_{i=1}^n x_i\right)^2 \sum_{i=1}^n x_i^2 = \frac{1}{n^3}\mathbf{E}\left(\sum_{i=1}^n x_i^2 \sum_{i=1}^n x_i^2\right) =$$

$$= \frac{1}{n^3}\sum_{i=1}^n \mathbf{E}x_i^4 + \sum_{i=1}^n \sum_{j=1, j\neq i}^n \mathbf{E}x_i^2 x_j^2 =$$

(так как слагаемые с нечетными степенями имеют нулевое математическое ожидание)

$$= \frac{n\mu_4 + n(n-1)\mu_2^2}{n^3} = \frac{\mu_4 + (n-1)\mu_2^2}{n^2} =$$

$$= \frac{\mu_2^2}{n} + \frac{\mu_4 - \mu_2^2}{n^2}.$$

Рассмотрим третье слагаемое.

$$\mathbf{E}\bar{x}^4 = \frac{1}{n^4}\mathbf{E}\left(\sum_{i=1}^n x_i\right)^4 =$$

$$= \frac{1}{n^4}\sum_{i1=1}^n \sum_{i2=1}^n \sum_{i3=1}^n \sum_{i4=1}^n \mathbf{E}x_{i1}x_{i2}x_{i3}x_{i4} =$$

так ненулевые слагаемые возникают только в случаях:

- $i_1 = i_2 = i_3 = i_4$;
- $i_1 = i_2, i_3 = i_4$;
- $i_1 = i_3, i_2 = i_4$;

- $i_1 = i_4, i_2 = i_3$.

$$= \frac{n\mu_4 + 3n(n-1)\mu_2^2}{n^4} = \frac{\mu_4 + 3(n-1)\mu_2^2}{n^3} = \frac{3\mu_2^2}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}.$$

Отсюда получаем

$$\mathbf{E}m_2^2 = \mu_2^2 + \frac{\mu_4 - \mu_2^2}{n} - \frac{2\mu_2^2}{n} - \frac{2\mu_4 - 2\mu_2^2}{n^2} + \frac{3\mu_2^2}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3} =$$

после приведения подобных членов

$$\mathbf{E}m_2^2 = \mu_2^2 + \frac{\mu_4 - 3\mu_2^2}{n} - \frac{2\mu_4 - 5\mu_2^2}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}\,.$$

Следовательно, дисперсия второго центрального момента m_2 равна

$$\begin{aligned} \mathbf{D}m_2^2 = \mathbf{E}m_2^2 - (\mathbf{E}m_2)^2 = \\ = \mu_2^2 + \frac{\mu_4 - 3\mu_2^2}{n} - \frac{2\mu_4 - 5\mu_2^2}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3} - \\ - \left(\mu_2^2 - \frac{2\mu_2^2}{n} + \frac{\mu_2^2}{n^2}\right) = \\ = \frac{\mu_4 - \mu_2^2}{n} - \frac{2(\mu_4 - 2\mu_2^2)}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}. \end{aligned}$$

Таким образом, дисперсия оценки m_2, а следовательно, и оценки S^2 имеет нулевой предел при увеличении объема выборки, и *оценка S^2 является состоятельной.*

ОМП параметра экспоненциального распределения

Экспоненциальное распределение с плотностью распределения $f(x,\alpha) = \alpha^{-1}e^{-x/\alpha}$. Математическое ожидание и дисперсия имеют вид $\mathbf{E}x_i = \alpha$, $\mathbf{D}x_i = \alpha^2$. В качестве оценки рассмотрим $\hat{\alpha} = \bar{x}$ с дисперсией $\mathbf{D}\hat{\alpha} = \frac{\alpha^2}{n}$.

$$\ln \mathcal{L}(\mathbf{x_1}, \ldots, \mathbf{x_n}) = \sum_{i=1}^{n} \left(-\ln\alpha - \frac{x_i}{\alpha}\right) = -n\ln\alpha - \frac{n\bar{x}}{\alpha}.$$

Производная логарифма функции правдоподобия имеет вид

$$\frac{\partial \ln \mathcal{L}(\mathbf{x})}{\partial(\alpha)} = -\frac{n}{\alpha} + \frac{n\bar{x}}{\alpha^2}. \tag{7}$$

Дифференцируя (7) по параметру, получаем

$$I(\alpha) = -\mathbf{E}\left(\frac{n}{\alpha^2} - \frac{2n\bar{x}}{\alpha^3}\right) = -\frac{n}{\alpha^2} + \frac{2n}{\alpha^2} = \frac{n}{\alpha^2}\,.$$

$$I(\alpha)\mathbf{D}\hat{\alpha} = \frac{n}{\alpha^2} \cdot \frac{\alpha^2}{n} = 1\,.$$

Оценка $\hat{\alpha} = \bar{x}$ параметра α экспоненциального распределения эффективна.

ОМП параметров нормального распределения

Пусть имеется выборка $x_1, \ldots, x_n$ с генеральной плотностью $\mathcal{N}(\mu, \sigma)$. Функция правдоподобия имеет вид

$$\mathcal{L}(x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}},$$

$$\ln \mathcal{L}(x_1, \ldots, x_n) = -\frac{n}{2}\left(\ln(2\pi) + \ln \sigma^2\right) - \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}\,.$$

В результате дифференцирования по параметрам получаем систему нормальных уравнений.

$$\begin{cases} \frac{\partial \ln \mathcal{L}(\mathbf{x})}{\partial \mu} = -\sum\limits_{i=1}^{n} \frac{2(x_i-\mu)\cdot(-1)}{2\sigma^2} = \sum\limits_{i=1}^{n} \frac{(x_i-\mu)}{\sigma^2} = 0 \\ \frac{\partial \ln \mathcal{L}(\mathbf{x})}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \sum\limits_{i=1}^{n} \frac{(x_i-\mu)^2}{2(\sigma^2)^2} = 0 \end{cases}$$

Решением этой системы уравнений являются оценки вида

$$\begin{cases} \hat{\mu} = \bar{x}, \\ \hat{\sigma}^2 = \frac{1}{n}\sum\limits_{i=1}^{n}(x_i - \hat{\mu})^2 = m_2. \end{cases}$$

Нормальное распределение с логарифмом плотности

$$\ln f(x) = -\ln\sqrt{2\pi} - \frac{1}{2}\ln\sigma^2 - \frac{(x-\mu)^2}{2\sigma^2},$$

$$\ln \mathcal{L}(x_1, \ldots, x_n) = -n\ln\sqrt{2\pi} - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.$$

$$\frac{\partial \ln \mathcal{L}(\mathbf{x})}{\partial(\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{\sum\limits_{i=1}^{n}(x_i-\mu)^2}{2(\sigma^2)^2},$$

$$\frac{\partial^2 \ln \mathcal{L}(\mathbf{x})}{\partial(\sigma^2)^2} = \frac{n}{2\sigma^4} - \frac{\sum\limits_{i=1}^{n}(x_i-\mu)^2}{\sigma^6},$$

$$I(\sigma^2) = -\mathbf{E}\frac{\partial^2 \ln \mathcal{L}(\mathbf{x})}{\partial(\sigma^2)^2} = -\frac{n}{2\sigma^4} + \frac{n\sigma^2}{\sigma^6} = \frac{n}{2\sigma^4}\,.$$

В случае нормально распределенной генеральной совокупности, согласно (6), получаем

$$\begin{aligned}\mathbf{D}m_2 = \frac{3\sigma^4 - \sigma^4}{n} - \frac{2(3\sigma^4 - 2\sigma^4)}{n^2} + \frac{3\sigma^4 - 3\sigma^4}{n^3} = \\ = \frac{2\sigma^4}{n} - \frac{2\sigma^4}{n^2} = \frac{2(n-1)\sigma^4}{n^2}, \\ \mathbf{D}S^2 = \frac{n^2}{(n-1)^2}\mathbf{D}m_2 = \frac{2\sigma^4}{n-1}\,.\end{aligned}$$

Вычисляем эффективность оценки S^2.

$$e(S^2) = \left(I(\sigma^2)\mathbf{D}S^2\right)^{-1} = \left(\frac{n}{2\sigma^4} \cdot \frac{2\sigma^4}{n-1}\right)^{-1} = \frac{n-1}{n},$$

следовательно, $\lim\limits_{n\longrightarrow\infty} e(S^2) = 1$ и S^2 является асимптотически эффективной оценкой дисперсии σ^2.

Сравнивая одинаково асимптотически эффективные оценки S^2 и m_2, заметим, что S^2 обладает несмещенностью, а m_2 имеет меньшую дисперсию, то есть более эффективна.