Лекция 2. Статистики первого типа. Точеченые оценки и их свойства

Грауэр Л.В., Архипова О.А.

CS center

Санкт-Петербург, 2015

Содержание

Статистики

Рассмотрим генеральную совокупность ξ и выборку $X_{[n]}$

Определение 1

Статистикой будем называть любую борелевскую функцию, заданную на выборочном пространстве, $S(X_{[n]})$.

Примеры

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

$$S(X_{[n]}) = \sup_{x} |F(x) - F_n^*(x)|$$

Статистики первого типа

Рассмотрим функционал $G(F)$, заданный на множестве функций распределения:

$$G(F) = h\left(\int\limits_{-\infty}^{\infty} g(x)dF(x)\right),$$

где $g : \mathbb{R} \to \mathbb{R}^m$ — заданная борелевская функция,
$h : \mathbb{R}^m \to \mathbb{R}^l$ некоторая борелевская функция, непрерывная в точке $a = \int_{-\infty}^{\infty} g(x)dF_\xi(x) \in \mathbb{R}^m$.

Назовем статистику $S(X_{[n]}) = G(F_n^*)$ *статистикой первого типа*.
Таким образом,

$$S(X_{[n]}) = G(F_n^*) = h\left(\int_{-\infty}^{\infty} g(x)dF_n^*(x)\right) = h\left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right).$$

Примеры

- $$a_k^* = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

$$h(t) \equiv t,\ \ g(x) = x^k, \quad G(F_n^*) = a_k^*$$

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

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$$

$$h(t_1, t_2) \equiv t_2 - t_1^2,\ \ g(x) = (g_1(x), g_2(x)) = (x, x^2)$$

Теорема 1

Пусть $S(X_{[n]}) = G(F_n^*)$ — статистика первого типа, тогда имеет место сходимость:

$$S(X_{[n]}) \xrightarrow[n\to\infty]{\text{п.н.}} h(a) = G(F_\xi).$$

Пусть у ξ существует $a_k = E\xi^k \in R$, тогда

$$a_k^* \xrightarrow[n\to\infty]{\text{п.н.}} a_k.$$

Пусть у ξ существует $a_k^0 = E(\xi - E\xi)^k \in R$, тогда

$$a_k^{0*} \xrightarrow[n\to\infty]{\text{п.н.}} a_k^0.$$

Рассмотрим случайный вектор $\xi = (\xi_1, \ldots, \xi_m)^T$ с m компонентами. Пусть имеется выборка:

$$(X_1, X_2, \ldots, X_n) = \begin{pmatrix} X_{11} & X_{12} & \ldots & X_{1n} \\ X_{21} & X_{22} & \ldots & X_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ X_{m1} & X_{m2} & \ldots & X_{nm} \end{pmatrix}.$$

Рассмотрим компоненты ξ_k и ξ_l, им соответствуют элементы выборки: $X_{k1}, \ldots, X_{kn}$ и $X_{l1}, \ldots, X_{ln}$.

$$\frac{1}{n}\sum_{i=1}^{n} X_{ki}X_{li} \xrightarrow[n\to\infty]{\text{п.н.}} E(\xi_k\xi_l), \quad \bar{X}_k\bar{X}_l \xrightarrow[n\to\infty]{\text{п.н.}} E\xi_k E\xi_l,$$

тогда имеет место сходимость почти наверное для выборочного коэффициента корреляции:

$$\hat{\varrho}(\xi_k\xi_l) = \frac{\widehat{cov}(\xi_k, \xi_l)}{\sqrt{s_k^2 s_l^2}} \xrightarrow[n\to\infty]{\text{п.н.}} \varrho(\xi_k\xi_l).$$

Теоремы непрерывности

Теорема 2

Пусть η *— случайная величина, заданная на вероятностном пространстве* $(\Omega, \mathcal{F}, P)$*, последовательность случайных величин* $\{\eta_n\}$ *также задана на* $(\Omega, \mathcal{F}, P)$*. Пусть борелевская функция* $H : \mathbb{R} \longrightarrow \mathbb{R}$ *непрерывна на борелевском множестве* $B \in \mathcal{B}(\mathbb{R})$*,* $P\{\eta \in B\} = 1$*, тогда справедливы утверждения:*

1. *Если* $\{\eta_n\} \xrightarrow[n\to\infty]{\text{п.н.}} \eta$*, тогда* $H(\eta_n) \xrightarrow[n\to\infty]{\text{п.н.}} H(\eta)$*.*
2. *Если* $\{\eta_n\} \xrightarrow[n\longrightarrow\infty]{p} \eta$*, тогда* $H(\eta_n) \xrightarrow[n\longrightarrow\infty]{p} H(\eta)$*.*

Теорема 3

Пусть борелевская функция $H : \mathbb{R}^m \to \mathbb{R}^k$ *непрерывна на* $B \in B(\mathbb{R}^m)$ *и* $P\{\eta \in B\} = 1$. *Пусть* $\eta_n^T = (\eta_{1n}, \ldots, \eta_{mn}) \xrightarrow[n\to\infty]{d} \eta$, *тогда* $H(\eta_n) \xrightarrow[n\to\infty]{d} H(\eta)$.

Теорема 4

Пусть последовательность случайных величин η_n *сходится по распределению к случайной величине* η, $\eta_n \xrightarrow[n\longrightarrow\infty]{d} \eta$. *Пусть функция* $H : \mathbb{R} \to \mathbb{R}$ — *борелевская функция. Числовая последовательность* $b_n \xrightarrow[n\to\infty]{} 0$, *причем* $b_n \neq 0$ *для любого* n. *Тогда справедливы утверждения:*

1. *Если функция* H *дифференцируема в точке* $a \in \mathbb{R}$, *то*

$$\frac{H(a + b_n\eta_n) - H(a)}{b_n} \xrightarrow[n\longrightarrow\infty]{d} H'(a)\eta.$$

2. *Если функция* H *дифференцируема в некоторой окрестности точки* a, $H'(a) = 0$, *и существует* $H''(a)$, *то*

$$\frac{H(a + b_n\eta_n) - H(a)}{b_n^2} \xrightarrow[n\longrightarrow\infty]{d} \frac{1}{2}H''(a)\eta^2.$$

Аналогичная теорема выполняется и для многомерного случая

Предельное распределение статистик первого типа

Пусть задана статистика первого типа:

$$S(X_{[n]}) = G(F_n^*) = h\left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right),$$

где $G(F) = h\left(\int_{-\infty}^{\infty} g(x)dF(x)\right)$, $\int_{-\infty}^{\infty} g(x)dF_\xi(x) = a \in \mathbb{R}^m$, т. е. $G(F_\xi) = h(a)$.

Теорема 5

Пусть имеется выборка $X_{[n]}$ из генеральной совокупности ξ с функцией распределения F_ξ и $S(X_{[n]}) = h\left(\frac{1}{n}\sum\limits_{i=1}^{n} g(X_i)\right)$ — статистика I типа, борелевские функции $h : \mathbb{R} \longrightarrow \mathbb{R}$, $g : \mathbb{R} \longrightarrow \mathbb{R}$, тогда справедливы утверждения:

1. *Если существует $h'(a)$, то*

$$\sqrt{n}\left(S(X_{[n]}) - h(a)\right) \xrightarrow[n\longrightarrow\infty]{d} h'(a)\zeta,$$

где ζ — случайная величина, распределенная нормально с параметрами $(0, Dg(\xi))$, $\zeta \sim N(0, \sqrt{Dg(\xi)})$.

2. *Если $h'(a) = 0$ и существует $h''(a)$, то*

$$n\left(S(X_{[n]}) - h(a)\right) \xrightarrow[n\longrightarrow\infty]{d} \frac{1}{2}h''(a)\zeta^2.$$

Теорема 6

Пусть задана статистика I типа $S(X_{[n]}) = h\left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right)$ *и борелевские функции* $h : \mathbb{R}^m \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}^m$, *тогда справедливы утверждения:*

1. *Если существует* $h'(a) = \left(\frac{\partial h}{\partial t_1}, \ldots, \frac{\partial h}{\partial t_m}\right)\Big|_{t=a}$, *где* $a = Eg(\xi) = (Eg_1(\xi), \ldots, Eg_m(\xi))$, *то*

$$\sqrt{n}\left(S(X_{[n]}) - h(a)\right) \xrightarrow[n\to\infty]{d} h'(a)\zeta^T,$$

   где случайный вектор $\zeta = (\zeta_1, \ldots, \zeta_m)$ *подчиняется многомерному нормальному распределению с параметрами* $(0, Dg(\xi))$, $\zeta \sim N(0, \sqrt{Dg(\xi)})$.

2. *Если* $h'(a) = 0$ *и существует* $h''(a)$, *то*

$$n\left(S(X_{[n]}) - h(a)\right) \xrightarrow[n\longrightarrow\infty]{d} \frac{1}{2}\zeta h''(a)\zeta^T.$$

Пример

Рассмотрим ξ, для которой $E\xi = \alpha > 0$, $D\xi = \sigma^2$.
Рассмотрим статистику $1/\bar{X}$.
Покажем, что $1/\bar{X}$ — статистика первого типа.

$$h(t) = 1/t,\ g(x) = x$$

$$1/\bar{X} = h\left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right).$$

$$a = \int\limits_{-\infty}^{+\infty} g(x)dF_\xi(x) = Eg(\xi) = E\xi = \alpha,$$

Тогда $1/\bar{X} = h(\frac{1}{n}\sum_{i=1}^{n} g(x_i))$ является статистикой первого типа.
Из теоремы 5 следует, что

$$\sqrt{n}\left(\frac{1}{\bar{X}} - \frac{1}{\alpha}\right) \xrightarrow[n\to\infty]{\mathrm{d}} \zeta\left(\frac{-1}{\alpha^2}\right) = -\zeta\frac{1}{\alpha^2},\quad \text{где}\quad \zeta \sim N(0,\sigma).$$

Точечные оценки

Рассмотрим генеральную совокупность ξ и ее функцию распределения $F_\xi(x;\theta)$, где $\theta = (\theta_1, \ldots, \theta_m)$ — неизвестные параметры в распределении случайной величины ξ.

По имеющейся выборке $X_{[n]} = (X_1, \ldots, X_n)$ необходимо построить оценку для этих параметров.

Определение 2

Пусть $\theta \in \Theta \subset \mathbb{R}$. Под оценкой понимается статистика $\hat{\theta}(X_{[n]})$ такая, что получившееся значение можно рассматривать как точечную оценку параметра θ ($\hat{\theta}(X_{[n]}) \sim \theta$).

Пример

$$\xi \sim N(a, \sigma), \quad a \text{ неизвестно}$$

Возможные оценки параметра a

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

$$\bar{X}_{0.1} = \frac{1}{[0.9n]} \sum_{i=[0.05n]+1}^{[0.95n]} X^{(i)}$$

$$x^*_{med} = \begin{cases} X_{(k+1)}, & n = 2k+1 \\ \frac{X_{(k)}+X_{(k+1)}}{2}, & n = 2k \end{cases}$$

Какую оценку считать «хорошей»?

Свойства точечных оценок

1. **Несмещенность.**

Определение 3

Пусть параметр $\theta \in \Theta \subset \mathbb{R}$. *Говорят, что оценка* $\hat{\theta}(X_{[n]})$ *является несмещенной оценкой параметра* θ, *если*

$$E\hat{\theta}(X_{[n]}) = \theta \tag{1}$$

для любого $\theta \in \Theta$.

Определение 4

Говорят, что оценка $\hat{\theta}(X_{[n]})$ *является асимптотически несмещенной оценкой параметра* θ, *если*

$$E\hat{\theta}(X_{[n]}) \xrightarrow[n\to\infty]{} \theta \tag{2}$$

для любого $\theta \in \Theta$.

Свойство несмещенности позволяет агрегировать информацию, накопленную в различных научных центрах.
Пусть $\hat{\theta}_1$ — несмещенная оценка параметра θ, полученная в некотором научном центре, $\hat{\theta}_2$ — несмещенная оценка того же параметра, полученная в другом научном центре. Предполагая, что техническая оснащенность научных центров одинаковая, будем считать, что дисперсии оценок одинаковы:

$$D(\hat{\theta}_i) = E(\hat{\theta}_i - \theta)^2 = \sigma^2(\theta),$$

$$E(\hat{\theta}_i) = \theta, \quad i = 1, 2.$$

Рассмотрим новую оценку:

$$\hat{\theta} = \frac{\hat{\theta}_1 + \hat{\theta}_2}{2}, \quad E\hat{\theta} = \frac{E\hat{\theta}_1 + E\hat{\theta}_2}{2} = \theta,$$

тогда имеют место равенства:

$$D\hat{\theta} = E(\hat{\theta} - \theta)^2 = \frac{1}{4}E\{(\hat{\theta}_1 - \theta) + (\hat{\theta}_2 - \theta)\}^2 = \frac{\sigma^2(\theta)}{2}.$$

Примеры

Рассмотрим **выборочную дисперсию**, проверим, выполнено ли свойство несмещенности:

$$Es^2 = E\left\{\frac{1}{n}\sum_{k=1}^{n}\left(X_k - a_1 - \frac{1}{n}\sum_{i=1}^{n}(X_i - a_1)\right)\right\}^2 =$$

$$= E\left\{\frac{1}{n}\sum_{k=1}^{n}(X_k - a_1)^2 - \left(\frac{1}{n}\sum_{k=1}^{n}(X_k - a_1)\right)^2\right\} =$$

$$= \frac{1}{n}\sum_{k=1}^{n}E(X_k - a_1)^2 - \frac{1}{n^2}\sum_{k=1}^{n}\sigma^2 = \sigma^2 - \frac{1}{n}\sigma^2 = \frac{n-1}{n}\sigma^2$$

Следовательно, s^2 — смещенная оценка, однако она является асимптотически несмещенной оценкой: $Es^2 \xrightarrow[n\to\infty]{} \sigma^2$.

Рассмотрим **исправленную оценку дисперсии**:

$$\tilde{s}^2 = \frac{n}{n-1}s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2,$$

что доказывает, что $\tilde{s}^2$ — несмещенная оценка дисперсии.

Выборочное среднее является несмещенной оценкой для математического ожидания:

$$E\bar{X} = E\left\{\frac{1}{n}\sum_{k=1}^{n}X_k\right\} = \frac{1}{n}\sum_{k=1}^{n}EX_k = E\xi = a_1.$$

2. **Состоятельность.**

Определение 5

Пусть параметр $\theta \in \Theta \subset \mathbb{R}$. *Говорят, что оценка* $\hat{\theta}(X_{[n]})$ *состоятельна, если*

$$\hat{\theta}(X_{[n]}) \xrightarrow[n\to\infty]{p} \theta \tag{3}$$

для любого $\theta \in \Theta$.

Определение 6

Оценка $\hat{\theta}(X_{[n]})$ *называется сильно состоятельной оценкой параметра* θ, *если*

$$\hat{\theta}(X_{[n]}) \xrightarrow[n\to\infty]{\text{п.н.}} \theta \tag{4}$$

для любого $\theta \in \Theta$.

Пусть существует $E\xi^k$, тогда a_k^* — статистика первого рода, где a_k^* — эмпирический момент порядка k. Тогда $a_k^* \to a_k = E\xi^k$, то есть, a_k^* является сильно состоятельной оценкой.

Пусть существует $E(\xi - E\xi)^k$, тогда $a_k^{0*} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k$ — сильно состоятельная оценка для теоретического момента, то есть:

$$a_k^{0*} \xrightarrow[n\to\infty]{\text{п.н.}} a_k^0 = E(\xi - E\xi)^k.$$

$$\bar{X} \xrightarrow[n\to\infty]{\text{п.н}} a_1 = E\xi,$$

$$s^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{X})^2 \xrightarrow[n\to\infty]{\text{п.н}} a_2^0 = D\xi.$$

3. **Эффективность.**
Рассмотрим некоторый класс оценок $K = \{\hat{\theta}(X_{[n]})\}$ параметра θ.

Определение 7

Говорят, что оценка $\theta^*(X_{[n]}) \in K$ является эффективной оценкой параметра θ в классе K, если для любой другой оценки $\hat{\theta} \in K$ имеет место неравенство:

$$E(\theta^* - \theta)^2 \leqslant E(\hat{\theta} - \theta)^2 \tag{5}$$

для любого $\theta \in \Theta$.

Класс несмещенных оценок обозначим через

$$K_0 = \left\{\hat{\theta}(X_{[n]}) : E\hat{\theta} = \theta, \forall\theta \in \Theta\right\}.$$

Рассмотрим случай, когда $m > 1$, то есть, $\theta = (\theta_1, \ldots, \theta_m)$.
Для любого $y \in \mathbb{R}^m$ определим $\alpha_y = (\theta, y) = \theta_1 y_1 + \ldots + \theta_m y_m$.
Тогда $\alpha_y^* = (\theta^*, y)$ — оценка параметра α_y.

Определение 8

Будем говорить, что оценка $\theta^* \in K$ является эффективной оценкой параметра $\theta = (\theta_1, \ldots, \theta_m)$ в классе K, если для любой другой оценки $\hat{\theta} \in K$ и любого $y \in \mathbb{R}^m$ при любом допустимом значении $\theta \in \Theta$ имеет место неравенство:

$$E(\alpha_y^* - \alpha_y)^2 \leqslant E(\hat{\alpha}_y - \alpha_y)^2, \tag{6}$$

где $\hat{\alpha}_y = (\hat{\theta}, y)$.

Теорема 7

Пусть несмещенные оценки $\hat{\theta}_1$ и $\hat{\theta}_2$ параметра $\theta \in \Theta \subset \mathbb{R}$ являются эффективными, тогда оценки $\hat{\theta}_1$ и $\hat{\theta}_2$ почти наверное совпадают.

Определение 9

Оценка $\hat{\theta}$ эффективна в классе K_0, или просто эффективна, если $D\tilde{\theta} - D\hat{\theta} \succeq 0$ (неотрицательно определенная матрица), где $\tilde{\theta} \in K_0$ для любого $\theta \in \Theta \subset \mathbb{R}^k$.

4. **Асимптотическая нормальность.**

Определение 10

Пусть оценивается параметр $\theta \in \Theta \subset \mathbb{R}$. Оценка $\hat{\theta}$ называется асимптотически нормальной оценкой параметра θ с коэффициентом рассеивания $\sigma^2(\theta)$, если

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow[n\to\infty]{d} \zeta \sim N(0, \sigma(\theta)). \tag{7}$$

Из этого определения следует, что для любого $x \in \mathbb{R}$ имеет место сходимость:

$$P\left\{\sqrt{n}(\hat{\theta} - \theta) \leqslant x\right\} \xrightarrow[n\longrightarrow\infty]{} \frac{1}{\sqrt{2\pi}\sigma(\theta)} \int\limits_{-\infty}^{x} e^{-\frac{y^2}{2\sigma^2(\theta)}} dy.$$

Определение 11

Пусть оценивается параметр $\theta \in \Theta \subset \mathbb{R}^m$. *Оценка* $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_m)$ *называется асимптотически нормальной с матрицей рассеивания* $\Sigma(\theta)$, *если имеет место сходимость по распределению:*

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow[n \to \infty]{d} \eta \sim N(0, \Sigma(\theta)).$$

5. **Асимптотическая эффективность.**

Определение 12

Оценка $\hat{\theta}$ *называется асимптотически эффективной в класс* K *оценок параметра* $\theta \in \Theta \subset \mathbb{R}$, *если*

$$\overline{\lim_{n\to\infty}} \frac{E(\hat{\theta}-\theta)^2}{E(\tilde{\theta}-\theta)^2} \leqslant 1$$

для любого параметра $\theta \in \Theta \subset \mathbb{R}$ *и любой оценки* $\tilde{\theta} \in K$.

Статистическая оценка считается «хорошей», если она обладает хотя бы некоторыми из свойств 1-5.

Методы построения точечных оценок

Метод моментов

Пусть требуется оценить параметр $\theta \in \Theta \subset \mathbb{R}$ по имеющейся выборке $X_{[n]} = (X_1, \ldots, X_n)$. Рассмотрим борелевскую функцию $g(x) : \mathbb{R} \to \mathbb{R}$ и определим функцию $m(\theta) = \int\limits_{-\infty}^{\infty} g(x) dF_{\xi}(x; \theta)$.

Далее положим, что

$$\int\limits_{-\infty}^{\infty} g(x) dF_n^*(x) = \frac{1}{n} \sum_{i=1}^{n} g(X_i) = \overline{g}. \tag{8}$$

Получим уравнение

$$m(\theta) = \overline{g} = \frac{1}{n} \sum_{i=1}^{n} g(X_i). \tag{9}$$

Предположим, что уравнение (9) имеет единственное решение $\hat{\theta}(X_{[n]})$, тогда будем это решение называть оценкой $\hat{\theta}$ неизвестного параметра θ, полученной по методу моментов:

$$\hat{\theta}(X_{[n]}) = m^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right).$$

Метод моментов легко обобщить на многомерный случай, при этом, $g(x) = (g_1(x), \dots, g_k(x))$, где k — число неизвестных параметров, то есть, $\theta = (\theta_1, \dots, \theta_k)^T \in \Theta \subset \mathbb{R}^k$.

Свойства оценок, построенных по методу моментов:

1. Если функция $m^{-1}(y)$ непрерывна на всей области определения, то оценка по методу моментов сильно состоятельна.
2. Если $m'(\theta) \neq 0$ для всех $\theta \in \Theta$, тогда оценка по методу моментов асимптотически нормальна с коэффициентом рассеяния $\frac{Dg(\xi)}{(m'(\theta))^2}$, где θ — истинное значение параметра.

Пример

Пусть $\xi \sim N(a, \sigma)$, тогда $\theta = (a, \sigma^2)^T \in \Theta = \mathbb{R} \times \mathbb{R}_+$. Выберем $g(x) = (x, x^2)$, тогда

$$Eg(\xi) = \begin{pmatrix} E\xi \\ E\xi^2 \end{pmatrix} = \begin{pmatrix} a \\ \sigma^2 + a^2 \end{pmatrix},$$

Нетрудно показать, что

$$\bar{g} = \begin{pmatrix} \frac{1}{n} \sum\limits_{k=1}^{n} X_k \\ \frac{1}{n} \sum\limits_{k=1}^{n} X_k^2 \end{pmatrix} = \begin{pmatrix} \bar{X} \\ s^2 + \bar{X}^2 \end{pmatrix},$$

Следовательно, оценки по методу моментов имеют следующий вид:

$$\begin{cases} \hat{a} = \bar{X}, \\ \hat{\sigma}^2 = s^2. \end{cases}$$

Пример

Рассмотрим равномерно распределенную случайную величину ξ с плотностью распределения:

$$f_\xi(x) = \begin{cases} \frac{1}{\theta}, & x \in [0, \theta]; \\ 0, & x \notin [0, \theta]. \end{cases}$$

Так как неизвестный параметр один, то $g(x) = x$. Вычислим математическое ожидание:

$$Eg(\xi) = E\xi = \int\limits_0^\theta x \frac{1}{\theta} dx = \frac{1}{2\theta} x^2 = \frac{\theta}{2}.$$

Уравнение имеет вид:

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

откуда получаем оценку:

$$\hat{\theta} = 2\bar{X} = 2\frac{1}{n}\sum_{k=1}^{n} X_i.$$

Метод максимального правдоподобия построения точечных оценок.

- Пусть ξ — непрерывная случайная величина с плотностью распределения $f_\xi(x, \theta)$.

Совместная плотность распределения выборки имеет вид:

$$f_{X_{[n]}}(X_{[n]}|\theta) = \prod_{i=1}^{n} f_\xi(X_i; \theta), \text{ где } \theta = (\theta_1, \ldots, \theta_m) \in \Theta.$$

- Пусть ξ — дискретная случайная величина с распределением вероятностей $p_\xi(z, \theta)$.

Совместное вероятностное распределение выборки имеет вид:

$$p_{X_{[n]}}(X_{[n]}|\theta) = \prod_{i=1}^{n} p_\xi(X_i; \theta) \text{ где } \theta = (\theta_1, \ldots, \theta_m) \in \Theta.$$

Определение 13

Если генеральная совокупность имеет плотность распределения f_ξ, то функцией правдоподобия выборки $X_{[n]}$ будем называть функцию

$$L(X_{[n]}, \theta) = \prod_{i=1}^{n} f_\xi(X_i; \theta).$$

Определение 14

Если генеральная совокупность ξ — дискретная случайная величина с возможными значениями $\{z_i\}$ и соответствующими вероятностями $p_\xi(z_i; \theta)$, то функцией правдоподобия выборки $X_{[n]}$ будем называть функцию

$$L(X_{[n]}, \theta) = \prod_{i=1}^{n} p_\xi(X_i; \theta).$$

Для нахождения оценки параметра θ решаем задачу:

$$\max_{\theta\in\Theta} L(X_{[n]},\theta).$$

Определение 15

Оценкой максимального правдоподобия параметра θ называется оценка

$$\hat{\theta}(X_{[n]}) = \arg\max_{\theta\in\Theta} L(X_{[n]},\theta), \tag{10}$$

если решение задачи максимизации существует и единственно.

Часто вместо функции $L(X_{[n]},\theta)$ рассматривают функцию $\ln L(X_{[n]},\theta)$, поскольку функция $\ln(t)$ является строго возрастающей функцией своего аргумента t, и данный переход правомерен.

Свойства оценок максимального правдоподобия:

Если функция правдоподобия непрерывно дифференцируема, и выполнены некоторые условия гладкости, то можно доказать, что оценки метода максимального правдоподобия —

- сильно состоятельны,
- асимптотически эффективны и
- асимптотически нормальны.

Пример

Рассмотрим случайную величину $\xi \sim N(a, \sigma)$ с плотностью распределения

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

Функция правдоподобия имеет вид:

$$L(X_{[n]}, a, \sigma^2) = \prod_{i=1}^{n} f_\xi(X_i; a, \sigma^2) = \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} e^{\frac{-\sum\limits_{i=1}^{n}(X_i-a)^2}{2\sigma^2}}.$$

Тогда

$$\ln L = \ln \frac{1}{((2\pi)^{\frac{1}{2}}\sigma)^n} - \frac{\sum\limits_{i=1}^{n}(X_i - a)^2}{2\sigma^2},$$

продифференцируем по a: $\partial \ln L/\partial a = 0$, или $\sum_{i=1}^{n} X_i - an = 0$, откуда $\hat{a} = \bar{X}$.

Продифференцируем по σ:

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum\limits_{i=1}^{n}(X_i - a)^2}{\sigma^3} = 0,$$

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

откуда находим решение:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - a)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = s^2.$$

Нетрудно проверить, что $\bar{X}$ и s^2 доставляют максимум функции правдоподобия.

Пример

Пусть случайная величина ξ подчиняется равномерному распределению с плотностью:

$$f(x,\theta) = \begin{cases} \frac{1}{\theta}, & x \in [0,\theta]; \\ 0, & x \notin [0,\theta]. \end{cases}$$

Запишем функцию правдоподобия:

$$L(X_{[n]},\theta) = \prod_{i=1}^{n} f(X_i;\theta) = \begin{cases} \frac{1}{\theta^n}, & \text{если для } \forall i : X_i \in [0,\theta]; \\ 0, & \text{если } \exists i : X_i \notin [0,\theta]. \end{cases}$$

Построим вариационный ряд $X_{(1)} \leq \ldots \leq X_{(n)}$. Таким образом, получаем:

$$L(X_{[n]},\theta) = \begin{cases} \frac{1}{\theta^n}, & X_{(n)} \in [0,\theta]; \\ 0, & \exists k : X_{(k)} \notin [0,\theta]. \end{cases}$$

Очевидно, что оценка максимального правдоподобия $\hat{\theta}(X_{[n]}) = X_{(n)}$.

Неравенство Рао-Крамера

Пусть имеется выборка $X_{[n]}$ из генеральной совокупности ξ с функцией распределения $F_\xi(x;\theta)$ и плотностью распределения $f_\xi(x;\theta)$, где $\theta \in \Theta \subset \mathbb{R}$ — неизвестный параметр. Функция правдоподобия имеет вид:

$$L(X_{[n]},\theta) = \prod_{i=1}^{n} f_\xi(X_i;\theta),$$

совместная плотность выборки имеет вид:

$$f_{X_{[n]}}(x,\theta) = L(x,\theta) = \prod_{i=1}^{n} f_\xi(x_i;\theta).$$

Имеет место следующее равенство:

$$\int\limits_{\mathbb{R}^n} L(x,\theta)dx = 1. \tag{11}$$

Пусть имеется оценка $\hat{\theta}(X_{[n]})$ неизвестного параметра θ, и справедливо следующее равенство:

$$E\hat{\theta} = \int\limits_{\mathbb{R}^n} \hat{\theta}(x_1,\ldots,x_n)L(x,\theta)dx = h(\theta). \tag{12}$$

Обозначим через $I_n(\theta)$ математическое ожидание:

$$I_n(\theta) = E\left(\frac{\partial \ln L(X_{(n)},\theta)}{\partial\theta}\right)^2 = \int\limits_{\mathbb{R}^n} \left(\frac{\partial \ln L(x,\theta)}{\partial\theta}\right)^2 L(x,\theta)dx.$$

Определение 16

Величина $I_n(\theta)$, если математическое ожидание существует и конечно, называется информационным количеством Фишера (соответствующим выборке объема n).

Будем предполагать, что выполнены *условия регулярности*:

- Для информационного количества Фишера выполнено неравенство $0 < I_n(\theta) < \infty$ для любого $\theta \in \Theta$.
- Равенства (11) и (12) можно продифференцировать и получить следующие уравнения:

$$\int\limits_{\mathbb{R}^n} \frac{\partial L(x,\theta)}{\partial \theta} dx = 0, \tag{13}$$

$$\int\limits_{\mathbb{R}^n} \hat{\theta}(x_1,\ldots,x_n) \frac{\partial L(x,\theta)}{\partial \theta} dx = h'(\theta). \tag{14}$$

- Множество $N = \{x \in R^n : L(x,\theta) = 0\}$ не зависит от θ.

Теорема 8 (Неравенство Рао-Крамера)

Пусть имеется генеральная совокупность ξ с функцией распределения $F_\xi(y, \theta)$, где $\theta \in \Theta \subset \mathbb{R}$. Задана выборка $X_{[n]}$ из генеральной совокупности ξ, и выполнены условия регулярности, тогда имеет место неравенство:

$$D\hat{\theta} \geqslant \frac{(h'(\theta))^2}{I_n(\theta)}. \tag{15}$$

Если $E\hat{\theta} = \theta$ для любого $\theta \in \Theta$ (то есть, оценка — несмещенная), то справедливо неравенство:

$$D\hat{\theta} \geq \frac{1}{I_n(\theta)}.$$

По определению $I_n(\theta) = E\left(\frac{\partial \ln L}{\partial \theta}\right)^2$.

$$I_n(\theta) = nI_1(\theta),$$

где $I_1(\theta)$ — информационное количество Фишера, соответствующее одному наблюдению. Как видим, наблюдается линейный рост информации.

- *Из неравенства Рао-Крамера можно сделать вывод, что в регулярном случае дисперсия не может убывать быстрее чем* $1/n$.
- *Для «хороших» оценок дисперсия должна убывать, разброс должен становиться меньше с ростом* n.
- *Для несмещенных оценок при выполнении условий регулярности оценка эффективна, если неравенство Рао-Крамера выполнено как равенство.*

Пример

Пусть $\xi \sim N(a, \sigma^2)$, методы максимального правдоподобия и моментов дали оценку $\hat{a} = \bar{X}$, которая является сильно состоятельной, несмещенной, асимптотически нормальной оценкой. Выясним, обращается ли неравенство Рао-Крамера в равенство. Действительно, можно заметить, что

$$\frac{\partial \ln L}{\partial a} = \frac{\sum\limits_{i=1}^{n}(X_i - a)}{\sigma^2} = \frac{n}{\sigma^2}(\bar{X} - a),$$

где

$$L(X_{[n]}, a) = \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} e^{-\frac{\sum\limits_{i=1}^{n}(X_i - a)^2}{2\sigma^2}}.$$

Случайные величины $\bar{X}$ и $\frac{n}{\sigma^2}(\bar{X} - a)$ пропорциональны, следовательно, $\bar{X}$ — эффективная оценка.

- Биллингсли П.Сходимость вероятностных мер. — М.: Изд. Наука, 1977.
- Боровков А.А.Теория вероятностей. — М.: Изд. Наука, 1986.
- Боровков А.А.Математическая статистика. — М.: Изд. Наука, 1984.