ОТВЕТЫ НА ВОПРОСЫ

1.Что такое функция Лагранжа?

(в лекции 1 задача о линейной оценке с минимальной дисперсией)?

Решается задача на условный экстремум: найти минимум (или максимум)

$$\min_x f(x) = ? \qquad x = (x_1, x_2, ..., x_n)$$

при условии $\qquad g(x) = 0$

Решение (необходимое условие): составить функцию Лагранжа:

$$L(x,\lambda) = f(x) - \lambda g(x)$$

«двух» переменных (x,λ), и найти ее **безусловный** экстремум:

$$\min_{x,\lambda} L(x,\lambda) = ?$$

Необходимое условие:

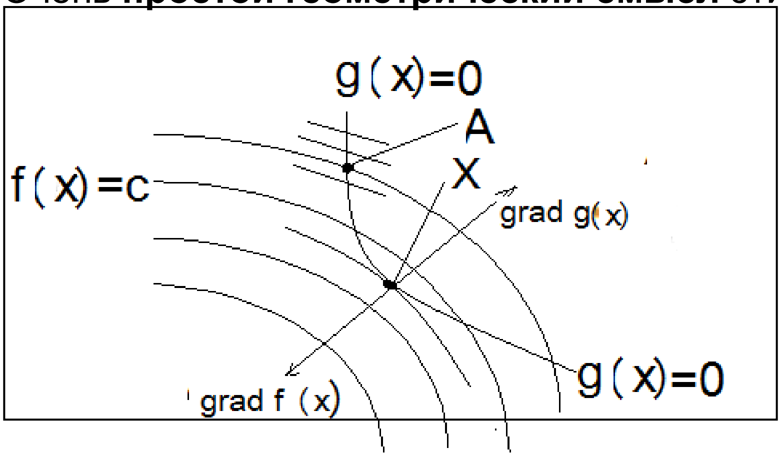
$$\frac{\partial}{\partial x} L(x,\lambda) = 0, \Rightarrow \qquad \frac{\partial}{\partial x} f(x) - \lambda \frac{\partial}{\partial x} g(x) = 0, \qquad (1)$$

$$\frac{\partial}{\partial \lambda} L(x,\lambda) = 0 \Rightarrow \qquad g(x) = 0, \qquad (2)$$

получаем, (1):

$$grad\, f(x) = \lambda grad\, g(x),$$
$$g(x) = 0$$

Очень **простой геометрический смысл** этих уравнений при n=2



Точка А – не искомая точка: двигаясь по $g(x) = 0$ пересекаем линии уровня $f(x)$. Точка X – искомая точка: двигаясь в ее окрестности, $f(x)$ не меяется, потому что касание кривой $g(x) = 0$ и $f(x) = c$

2. Что значит зависимость от параметра фиктивна?

(о неравенстве Рао-Крамера)

Утверждение. Эффективная оценка $\varphi(x)$ для $f(a)$, если она существует, может быть представлена формулой

$$\varphi(x) = f(a) + \frac{f'(a)}{I_\xi(a)} \cdot \frac{\partial \ln p(x;a)}{\partial a}, \tag{15}$$

причем зависимость правой части от параметра *a* фиктивна.

3. Чем отличается усиленный ЗБЧ от ЗБЧ в форме Чебышева ?

ЗБЧ в форме Чебышева:

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}M\xi_i \to 0 \quad \text{по вероятности при } n\to\infty, \qquad \text{(A)}$$

ЗБЧ усиленный

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}M\xi_i \to 0 \quad \text{с вероятностью 1 при } n\to\infty. \qquad \text{(B)}$$

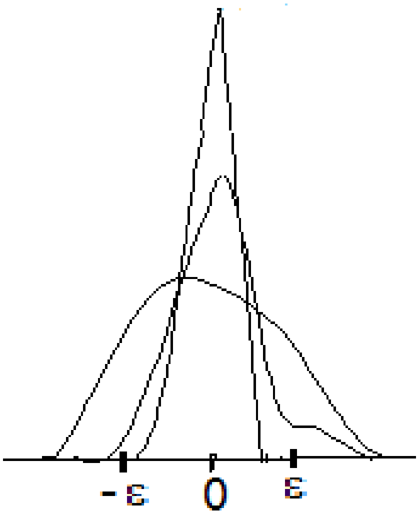
$\alpha_n \to 0$ по вероятности при $n\to\infty$,

т.е.

$$\forall\ \varepsilon > 0 \ \lim_{n\to\infty} P\{|\alpha_n| < \varepsilon\} = 1$$

Короткое обозначение

$$\alpha_n \xrightarrow{P} 0$$



$\alpha_n \to 0$ с вероятностью 1 при $n\to\infty$.

т.е.

$$P\left\{\lim_{n\to\infty} \alpha_n \to 0\right\} = 1$$

Сх-ть с вер 1 $\Rightarrow$ сх-ть по вер

т.е

$$P\left\{\lim_{n\to\infty} \alpha_n \to 0\right\} = 1 \quad \Rightarrow \quad \alpha_n \xrightarrow{P} 0,$$ обратное неверно

.

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
Лекция 4
§ 5. Методы построения оценок

Рассмотрим лишь три наиболее популярных метода.

5.1. Метод моментов.

Пусть $\xi_1, \xi_2 \ldots \xi_n$ - выборка,
т.е. n независимых наблюдений над случайной величиной, обозначим ее ξ_0,
$F(x; a_1, a_2 \ldots a_R)$ -функция распределения, зависящая от неизвестных параметров $a = (a_1, a_2 \ldots a_R)$, всего R штук.
Требуется оценить их.
Идея метода: *неизвестные параметры выразить через начальные моменты, а затем вместо моментов подставить несмещенные и состоятельные оценки моментов*.

Выразим R моментов через R параметров:

$$\begin{aligned} m_1 &= \int_{-\infty}^{\infty} x dF(x; a_1, a_2 \ldots a_R) = f_1(a_1, a_2 \ldots a_R), \\ m_2 &= \int_{-\infty}^{\infty} x^2 dF(x; a_1, a_2 \ldots a_R) = f_2(a_1, a_2 \ldots a_R), \quad , \quad j = 1, 2 \ldots R. \qquad (1) \\ &\ldots \\ m_R &= \int_{-\infty}^{\infty} x^R dF(x; a_1, a_2 \ldots a_R) = f_R(a_1, a_2 \ldots a_R). \end{aligned}$$

Пусть из этой системы равенств можно выразить параметры через моменты:

$$\begin{aligned} a_1 &= g_1(m_1, m_2 \ldots m_R), \\ a_2 &= g_2(m_1, m_2 \ldots m_R), \quad j = 1, 2 \ldots R. \qquad (2) \\ &\ldots \\ a_R &= g_R(m_1, m_2 \ldots m_R). \end{aligned}$$

Подставив вместо моментов $m_1, m_2 \ldots m_R$ оценки моментов $\hat{m}_1, \hat{m}_2 \ldots \hat{m}_R$, получаем:

$$\hat{a}_j = g_j(\hat{m}_1, \hat{m}_2 \ldots \hat{m}_R), \text{ где } \hat{m}_k = \frac{1}{n} \sum_{i=1}^{n} \xi_i^k, \; k = 1, 2 \ldots R.$$

Мы получили некоторые оценки $\hat{a}_1, \hat{a}_2 \ldots \hat{a}_R$; они называются оценками по ММ.

Справедливы следующие свойства (см., например, [1], [3]):

1) *если функции $g_j(\cdot)$, $j = 1, 2 \ldots R$, непрерывны, то оценки состоятельны;*

2) *если функции $g_j(\cdot)$, $j = 1, 2...R$, дифференцируемы, а распределение при любом а имеет 2R моментов, то оценки $\hat{a}_j$ асимптотически нормальн :* $\hat{a}_j \sim N\left(a_j, \frac{1}{n}\sum_{l,s=1}^{R}(m_{l+s} - m_l \cdot m_s)\frac{\partial g_j}{\partial m_l}\frac{\partial g_j}{\partial m_s}\right)$.

В справедливости этих свойств нетрудно убедиться.

Несмещённость оценок не гарантируется.

Что такое асимптотическая нормальность? Если закон распределения оценки с ростом n стремится к нормальному, то говорят, что оценка асимптотически нормальна.

$$\hat{a} = g(\hat{m}) = g(m + \delta_{\hat{m}}) \approx g(m) + \delta_{\hat{m}} g'(m) = a + c\delta_{\hat{m}} \sim N$$

$$\hat{m}_j = \frac{1}{n}\sum_{i=1}^{n}\xi^j{}_i = m_j + \delta_{\hat{m}},\ \delta_{\hat{m}} \to 0, \quad \delta_{\hat{m}} = \sum_{i=1}^{n} \quad \sim N$$

Замечания.

1. В равенствах (1) вместо первых *R* моментов можно использовать любые *R* моментов; важно лишь, чтобы система была разрешима относительно параметров.

2. Моментные оценки не всегда обладают высокой точностью. Однако, обычно они достаточно просты в вычислительном отношении.

Пример 1. Оценим дисперсию σ^2 методом моментов. Дисперсия σ^2 выражается через первые два момента:

$$\sigma^2 = m_2 - m_1^2.$$

Подставив оценки моментов, получим оценку s^2 для дисперсии σ^2:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}\xi_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\xi_i^2 - (\bar{\xi})^2 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2. \qquad (3)$$

Последнее равенство нетрудно проверить:

$$\frac{1}{n}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2 = \frac{1}{n}\sum_{i=1}^{n}\xi_i^2 - 2\bar{\xi}\cdot\frac{1}{n}\sum_{i=1}^{n}\xi_i + \bar{\xi}^2 = \frac{1}{n}\sum_{i=1}^{n}\xi_i^2 - \bar{\xi}^2.$$

Оценка (3) совпадает с оценкой s^2, которая была проанализирована в разделе 2.3.

Пример 2. Оценка параметров равномерного распределения.

Пусть ξ_1, $\xi_2 \ldots \xi_n$ — выборка из совокупности, распределенной по равномерному закону R[*a*, *b*] на отрезке [*a*, *b*]. Оценим два неизвестных параметра *a* и *b*. Первые два момента выражаются через два параметра:

$$m_1 = (a + b) / 2,$$
$$m_2 - m_1^2 = \sigma^2 = (b - a)^2 / 12.$$

В этих уравнениях относительно *a* и *b* заменяем неизвестные моменты выборочными, при этом во втором уравнении слева, исходя из (3), имеем s^2. Получаем:

$$\bar{\xi} = (a + b) / 2,$$
$$s^2 = (b - a)^2 / 12.$$

$$\hat{a} + \hat{b} = 2\xi,$$
$$\hat{b} - \hat{a} = 2\sqrt{3}\,s.$$

Откуда:

$$\hat{a} = \xi - \sqrt{3}\,s, \;\; \hat{b} = \xi + \sqrt{3}\,s.$$

5.2. Метод максимального правдоподобия.

Пусть $\xi_1, \xi_2 \ldots \xi_n$ — выборка, $q(x_i;\, a)$ — плотность распределения одного i-го наблюдения (в дискретном случае $q(x_i;\, a)$ — вероятность принятия дискретного значения x_i), $a = (a_1, a_2 \ldots a_R)$ — неизвестный параметр,

$p_\xi(x;a) = \prod_{i=1}^{n} q(x_i; a)$ — распределение выборки $x = (x_1, x_2 \ldots x_n)$.

Функция $p_\xi(x;a)$, как функция параметра a, при фиксированном x, называется ***функцией правдоподобия***.

Оценкой максимального правдоподобия (МП оценкой) a^* параметра a называется такое значение, при котором функция правдоподобия $p_\xi(x;a)$ достигает максимума:

$$\boldsymbol{a^*}: \quad \boldsymbol{p_\xi(x;a^*)} = \max_a p_\xi(x;a). \tag{4}$$

Если максимум достигается во внутренней точке области определения функции, то a^* удовлетворяет системе уравнений:

$$\left.\frac{\partial \ln p_\xi(x;a)}{\partial a_i}\right|_{a=a^*} = 0,\; i = 1, 2 \ldots R. \tag{4a}$$

Использование логарифма не изменяет точки максимума, но упрощает выкладки при независимых наблюдениях. Оценка $a^* = a^*(x)$ является функцией наблюдений x. Чтобы подчеркнуть случайность аргумента, напишем $a^*(\xi)$.

Пример 1. МП оценка параметров нормального распределения.

Пусть $\xi_1, \xi_2 \ldots \xi_n$ — выборка из нормальной совокупности $N(m, \sigma^2)$, здесь $a \equiv (m, \sigma^2)$. Параметры m и σ^2 неизвестны. Плотность распределения выборки:

$$p_\xi(x;\, m, \sigma^2) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(x_i - m)^2}{2\sigma^2}\right\}.$$

Логарифм функции правдоподобия:

$$\ln p_\xi(x;\, m, \sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2.$$

Система уравнений для определения оценок:

$$\begin{cases} \dfrac{\partial \ln p_\xi(x;m,\sigma^2)}{\partial m} = \dfrac{2}{2}\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(x_i - m) = 0, \\ \dfrac{\partial \ln p_\xi(x;m,\sigma^2)}{\partial \sigma^2} = -\dfrac{n}{2}\dfrac{1}{\sigma^2} + \dfrac{1}{2(\sigma^2)^2}\sum\limits_{i=1}^{n}(x_i - m)^2 = 0. \end{cases}$$

Из первого уравнения находим

$$m^* = \frac{1}{n}\sum_{i=1}^{n} x_i \equiv \bar{\xi}. \tag{5}$$

Из второго уравнения находим

$$(\sigma^2)^* = \frac{1}{n}\sum_{i=1}^{n}(x_i - m^*)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{\xi})^2. \tag{6}$$

В данном случае оценки совпадают с выборочными средним и дисперсией.

Пример 2. МП оценка параметра равномерного распределения.

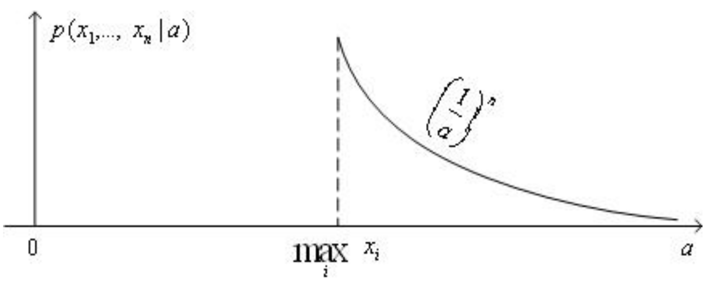
Пусть $\xi_1, \xi_2 \ldots \xi_n$ — выборка из совокупности, распределенной по равномерному закону R[0, a] с неизвестным правым концом $a > 0$. Плотнсть распределения для одного наблюдения с номером i:

$$p(x_i, a) = \begin{cases} \dfrac{1}{a}, 0 \le x_i \le a, \\ 0, \text{иначе}. \end{cases}$$

Плотность распределения выборки ξ

$$p_\xi(x_1, x_2 \ldots x_n; a) =$$

$$\begin{cases} (1/a)^n, 0 \le x_1 \le a, 0 \le x_2 \le a, \ldots, 0 \le x_n \le a, \\ 0, \text{ иначе}. \end{cases} = \begin{cases} (1/a)^n, 0 \le \min\limits_i x_i, \max\limits_i x_i \le a, \\ 0, \text{ иначе}. \end{cases}$$



При фиксированных $x_1, x_2 \ldots x_n$ функция правдоподобия убывает $(1/a)^n$ при $a \ge \max\limits_i x_i$ и равна 0 при $a < \max\limits_i x_i$ (рис. 2). Максимум достигается при $a^* = \max\limits_i x_i$.

Рис. 2. Функция правдоподобия

Проанализируем эту оценку. Ее функция распределения:

$$F_{a^*}(y) = P\{a^* = \max_i \xi_i < y\} = \prod_{i=1}^{n} P\{\xi_i < y\} = \left[F_\xi(y)\right]^n = \begin{cases} 0, \ y < 0, \\ (y/a)^n, \ y \in [0, a] \\ 1, \ y > a. \end{cases}$$

Плотность распределения:

$$p_{a*}(y) = [F_{a*}(y)]' = \frac{n}{a}(y/a)^{n-1},\ y \in [0,a],\ \text{иначе } 0.$$

Математическое ожидание:

$$Ma* = \int_0^a y p_{a*}(y)dy = \int_0^a y\frac{n}{a}(y/a)^{n-1}dy = an\int_0^a (y/a)^n d(y/a) = \frac{n}{n+1}a \neq a\ ,$$

т.е. оценка смещенная.

Оценку легко исправить, т.е. сделать несмещенной, умножив ее на $\frac{n+1}{n}$, в результате чего получим оценку

$$\hat{a} = \frac{n+1}{n}\max x_i\ .$$

Она уже несмещенная. Ее второй момент

$$M\hat{a}^2 = \left(\frac{n+1}{n}\right)^2(\max x_i)^2 = \left(\frac{n+1}{n}\right)^2\int_0^a y^2\frac{n}{a}(y/a)^{n-1}dy = a^2\left(\frac{n+1}{n}\right)^2 n\int_0^a (y/a)^{n+1} d(y/a)$$

$$= a^2\left(\frac{n+1}{n}\right)^2\frac{n}{n+2}$$

Дисперсия

$$D\hat{a} = M\hat{a}^2 - (M\hat{a})^2 = a^2\left(\frac{n+1}{n}\right)^2\frac{n}{n+2} - a^2 = \frac{a^2}{n(n+2)}$$

Из вышесказанного видно, что дисперсия убывает быстрее, чем $1/n$, что противоречит неравенству (14) раздела 3.2. Однако, в этом примере условия неравенства Рао-Крамера не выполняются, а именно, условие 1 о независимости носителя вероятности от параметра. Дисперсия может убывать быстрее, и это пример сверхэффективной оценки.

Свойства оценок максимального правдоподобия.

Пусть $\xi \equiv (\xi_1, \xi_2 \ldots \xi_n)$ — выборка объема n из совокупности, распределенной с плотностью $q(x;a)$, и выражение

$$p(x_1, x_2 \ldots x_n; a) = \prod_{i=1}^{n} q(x_i; a) \tag{7}$$

является плотностью распределения выборки.

При некоторых весьма широких условиях (см. ниже) оценки максимального правдоподобия:

— состоятельны;

— асимптотически эффективны;

— асимптотически нормальны.

Для одномерного случая:

$$Ma* \to a,\ \ Da* \to \left\{nM\left[\frac{\partial \ln q(\xi_i; a)}{\partial a}\right]^2\right\}^{-1} = \{nI(a)\}^{-1} \text{ при } n \to \infty\ . \tag{8}$$

Условия, при которых вышеприведенные свойства верны, совпадают с условиями неравенства Рао-Крамера:

а) независимость от параметра a множества $X = \{x: q(x/a) \neq 0\}$-носителя вер-ти;

б) существование производных $\frac{\partial q}{\partial a}$ и $\frac{\partial^2 q}{\partial a^2}$;

в) существование интеграла $\mathrm{M}\left[\frac{\partial \ln q(\xi; a)}{\partial a}\right]^2$.

Доказательство справедливости этих свойств можно найти, например, в [5]. Примем на веру состоятельность и покажем, как возникает асимптотическая эффективность и асимптотическая нормальность.

Рассмотрим случайную функцию от a

$$S_n(a,\xi) = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ln q_\xi(\xi_i; a)}{\partial a}. \qquad (9)$$

Учитывая (4а) $\left.\frac{\partial \ln p_\xi(x;a)}{\partial a_i}\right|_{a=a^*} = 0$ определение МП-оценки

и (7) $p(x_1, x_2 \ldots x_n; a) = \prod_{i=1}^{n} q(x_i; a)$ (7)

ясно, что оценка a^* является корнем этой случайной функции от параметра a

$$S_n(a^*,\xi) = 0.$$

Пусть a_0 — истинное значение параметра. Рассмотрим $S_n(a,\xi)$ - случайную величину в точке истинного значения параметра $a = a_0$. Учитывая состоятельность,

т.е. $a^* \to a_0$,

и гладкость функции $S_n(a,\xi)$, по теореме Лагранжа имеем:

$$S_n(a_0,\xi) = S_n(a^*,\xi) + (a_0 - a^*)\, S'_n(\tilde{a},\xi), \qquad (10)$$

где $\tilde{a}$ — промежуточная точка между a_0 и a^*, причем $\tilde{a} \to a_0$.

В силу предыдущего уравнения, справа первое слагаемое равно 0. Умножим это соотношение на $\sqrt{n}$:

$$\sqrt{n}\, S_n(a_0,\xi) = \sqrt{n}\,(a_0 - a^*)\, S'_n(\tilde{a},\xi), \qquad (11)$$

Слева имеем случайную величину

$$\zeta = \sqrt{n}\, S_n(a_0,\xi),$$

которая, учитывая суммирование случайных величин в (9), асимптотически нормальна $N(0, I(a_0))$ с параметрами: М.О.

$$\mathrm{M}\zeta = \sqrt{n}\,\mathrm{M}S_n(a_0,\xi) = \sqrt{n}\int_X \left.\frac{\partial \ln q_\xi(x;a)}{\partial a}\right|_{a=a_0} q_\xi(x;a_0)\,dx = 0.$$

При вычислении интеграла учтено, что

$$\int_X \frac{1}{q_\xi(x;a)} \frac{\partial q_\xi(x;a)}{\partial a}\bigg|_{a=a_0} q_\xi(x;a_0)dx = \frac{\partial}{\partial a}\int_X q_\xi(x;a)dx\bigg|_{a=a_0} = 0.$$

Что касается дисперсии, то она равна информации Фишера в одном наблюдении в точке a_0: действительно

$$D\zeta = \frac{n\cdot n}{n^2} D\frac{\partial \ln q_\xi(\xi_i;a)}{\partial a}\bigg|_{a=a_0} = M\left[\frac{\partial \ln q_\xi(\xi_i;a)}{\partial a}\bigg|_{a=a_0}\right]^2 = I(a_0).$$

Теперь определим параметры случайной величины $S'_n(\tilde{a},\xi)$ в правой части (11) при $n\to\infty$ с учетом того, что $\tilde{a}\to a_0$:

$$MS'_n(\tilde{a},\xi) = \int_X \frac{\partial^2 \ln q_\xi(x;a)}{\partial a^2}\bigg|_{a=\tilde{a}} q_\xi(x;a_0)dx \to -I(a_0),$$

$$DS'_n(\tilde{a},\xi) = \frac{1}{n} D\frac{\partial^2 \ln q_\xi(\xi;a)}{\partial a^2}\bigg|_{a=\tilde{a}} \to 0.$$

Это означает, что $S'_n(\tilde{a},\xi)$ сходится к константе $I(a_0)$. Из (11) в пределе получаем

$$\zeta = -\sqrt{n}\,(a_0 - a^*)\,I(a_0),$$

что означает, выразив a^* через ζ:

$$a^* = a_0 + \frac{\zeta}{\sqrt{nI(a_0)}}.$$

Из этого следует, что оценка a^* асимптотически нормальна, а дисперсия $\{nI(a_0)\}^{-1}$. Это значение совпадает с границей Рао-Крамера.

Замечания.

1. *Эффективная оценка, если она существует, является оценкой максимального правдоподобия*.

Действительно, если $\varphi(x)$ — эффективная оценка для параметра a, то по лемме из раздела 3.3 имеем

$$\frac{\partial \ln p(x;a)}{\partial a} = I(a)[\varphi(x) - a],$$

откуда, приравнивая производную к нулю, получаем $a^* = \varphi(x)$.

2. *Оценка максимального правдоподобия является функцией достаточной статистики, если последняя существует*.

Действительно, если $T(x)$ — достаточная статистика, то в силу критерия факторизации в разделе 4.2 справедливо представление

$$p(x;a) = g(T(x), a)h(x),$$

и потому

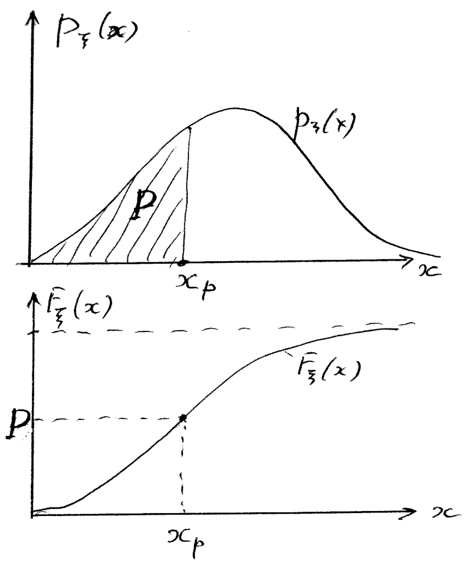
$$\max_a p(x;a) = h(x)\cdot \max_a g(T(x), a),$$

откуда экстремальная точка $a^* = a^*[(T(x)]$.

5.3. Метод порядковых статистик.

В статистике широко используется система числовых характеристик, называемых *квантилями*.

*Значение x_p случайной величины ξ называется **p-квантилью**, если*



$$P\{\xi < x_p\} = p,$$

где x_p — это корень уравнения

$F_\xi(x_p) = p$

(рис. 3).

Примеры р-квантили:

$x_{0,5}$ — медиана — характеристика среднего значения случайной величины;

$x_{0,98}$ — максимальное, в некотором смысле, значение случайной величины, т.к. $P\{\xi < x_{0,98}\} = 0,98$;

$x_{0,02}$ — минимальное, в некотором смысле, значение случайной величины, т.к. $P\{\xi \geq x_p\} = 1 - P\{\xi < x_p\} = 1 - p = 0,98$;

$x_{3/4}$ и $x_{1/4}$ — верхняя и нижняя квартили; их разность ($x_{0,75} - x_{0,25}$) — межквартильная широта — служит характеристикой разброса.

Рис. 3. Графическая иллюстрация квантили x_p

Оценка p-квантилей. Неизвестные *p*-квантили легко оцениваются по выборке. Действительно, пусть x_1, $x_2 ... x_n$ — результаты n независимых наблюдений над случайной величиной ξ с функцией распределения $F(x)$. Упорядочив их по возрастанию, получаем вариационный ряд

$$x_{(1)} \leq x_{(2)} \leq \ \ldots \leq x_{(n)}.$$

Чтобы подчеркнуть случайность ряда, запишем его греческими символами

$$\xi_{(1)} \leq \xi_{(2)} \leq \ \ldots \leq \xi_{(n)}.$$

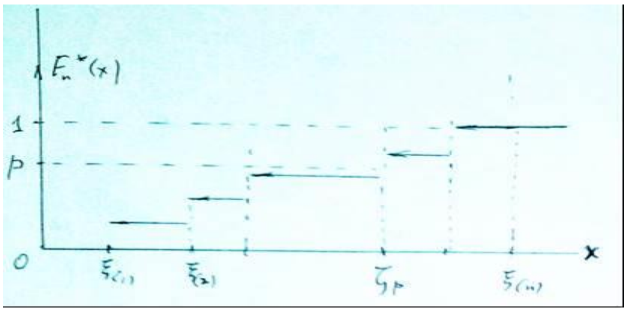
Член вариационного ряда $\xi_{(i)}$ с номером i (заметим, что это случайная величина) называется *i-й порядковой статистикой*. По вариационному ряду построим функцию $F_n^*(x) \equiv F_n^*(x;\, \xi_{(1)} \leq \xi_{(2)} \leq \ldots \leq \xi_{(n)})$ эмпирического распределения, и, согласно общему принципу о том, что выборочные характеристики являются состоятельными оценками характеристик распределения генеральной совокупности, рассмотрим в качестве оценки для *p*-квантили x_p выборочную квантиль ζ_p, т.е. корень уравнения

$$F_n^*(\zeta_p) = p. \qquad (8)$$

Поскольку $F_n^*(x)$ — функция кусочно-постоянная, то корнем является одна из порядковых статистик

$$\zeta_p = \xi_{([np]+1)}, \qquad (9)$$

с номером $[np]+1$, т.е. целая часть числа np плюс 1(рис. 4).



Нетрудно показать, что ζ_p *является состоятельной оценкой для* x_p. Кроме того, известна **теорема Крамера**, которая гласит, что *для непрерывных распределений с плотностью $q(x)$ оценка ζ_p асимптотически нормальна с параметрами*:

Рис. 4. Графическая иллюстрация выборочной квантили

$$\mathrm{M}\zeta_p = x_p,\ \ \mathrm{D}\zeta_p = \frac{1}{n}\frac{p(1-p)}{q^2(x_p)}. \qquad (10)$$

Метод оценки параметров основан на оценках ζ_p при разных p. Пусть $\xi_1, \xi_2 \ldots \xi_n$ — выборка с функцией распределения $F(x;a)$, зависящей от параметра a, значение которого требуется оценить. Выберем p так, чтобы квантиль x_p зависела от параметра:

$$x_p = f(a).$$

Выразим параметр a через квантиль x_p:

$$a = g(x_p),$$

и вместо x_p подставим выборочную квантиль $\zeta_p = \xi_{([np]+1)}$, в результате чего получим *состоятельную оценку*

$$\hat{a} = g(\xi_{([np]+1)}).$$

Таким же образом можно построить оценки и для неодномерного параметра.

Основное и очень важное преимущество оценок, основанных на порядковых статистиках, — *их устойчивость к засорению наблюдений и к изменениям закона распределения*.

Примеры оценок параметров нормального распределения. Пусть $\xi_1, \xi_2 \ldots \xi_n$ — выборка из нормальной совокупности $N(m, \sigma^2)$.

1) Оценка среднего m. Известно или нет значение σ — безразлично. В силу симметрии нормального распределения параметр m является медианой, т.е. квантилью уровня ½, и потому может быть оценен выборочной медианой:

$$\hat{m} = \zeta_{1/2} = \xi_{([n/2]+1)}.$$

Можно сравнить по точности эту оценку с эффективной оценкой

$$\hat{m} = \sum_{i=1}^{n} \xi_i / n$$

с дисперсией $\mathrm{D}\hat{m}^* = \sigma^2/n$.

Согласно n (10), теореме Крамера, $\mathrm{D}\hat{m} \approx \frac{1}{n}\cdot\frac{0{,}5(1-0{,}5)}{(2\pi\sigma^2)^{-1}} = \frac{\sigma^2}{n}\cdot\frac{\pi}{2}$, т.е. очень простая и устойчивая к засорению оценка $\hat{m}$ уступает по точности оценке $\hat{m}^*$ в $\sqrt{\pi/2} \approx 1{,}25$ раза, т.е. 25 %.

2) Оценка стандартного уклонения σ.

Легко проверить, что верхняя и нижняя квартили равны соответственно $x_{3/4} = m + 0{,}675\sigma$ и $x_{1/4} = m - 0{,}675\sigma$, т.к.

$$P\{\xi < m + 0.675\sigma\} = 3/4, P\{\xi < m - 0.675\sigma\} = 1/4$$

И потому

$$\sigma = (x_{3/4} - x_{1/4}) / 1{,}35,$$

и потому оценивать σ можно следующим образом:

$$\hat{\sigma} = \frac{\zeta_{3/4} - \zeta_{1/4}}{1{,}35} = \frac{\xi_{([3n/4]+1)} - \xi_{([n/4]+1)}}{1{,}35}.$$

3) Оценка стандартного уклонения σ *по размаху*.

Пусть $\xi_{(1)}$ и $\xi_{(n)}$ — минимальный и максимальный член выборки, разность которых называется размахом w:

$$w = \xi_{(n)} - \xi_{(1)}.$$

Ясно, что $Mw = c(n)\sigma$, и потому оценкой для σ может служить

$$\hat{\sigma} = w/c(n) = k(n)w,$$

где $k(n)$ берем из статистических таблиц [4]. Ниже приведены значения коэффициента $k(n)$ и коэффициента эффективности

$$eff = \frac{\sigma_0^2}{D\hat{\sigma}},$$ где σ_0^2 — нижняя граница Рао-Крамера,

а также потеря точности

$$(1 - \sqrt{eff}) \cdot 100,$$

измеряемая в процентах, по сравнению с нижней границей Рао-Крамера.

Табл. 1. Значение коэффициентов *k* и *n*

n	2	5	10
$k(n)$	0,866	0,430	0,325
eff	1,000	0,955	0,855
потеря точности, $(1 - \sqrt{eff})100$, %	0	2,5	7

Для устойчивости оценки к засорению используют подразмахи w_m порядка m, где $m = 1, 2, 3\ldots$:

$$w_m = \xi_{(n-m+1)} - \xi_{(m)},$$

так что оценка имеет вид:

$$\hat{\sigma} = k_m(n)\, w_m.$$

Значение коэффициента $k_m(n)$ берется из таблиц.

4) Распределение порядковых статистик. При анализе оценок, получаемых рассматриваемым методом, необходимо знать распределения порядковых статистик. Если распределение одного наблюдения ξ непрерывно с плотностью $p(x) = F'(x)$, то плотность распределения для k-й порядковой статистики $\xi_{(k)}$ выражается следующей формулой:

$$p_{\xi_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} F^{k-1}(x)[1-F(x)]^{n-k}\, p(x)\,,$$

которая получается вычислением вероятности события

$$p_{\xi_{(k)}}(x)\Delta x = P\left\{\xi_{(k)} \in (x-\Delta x, x)\right\},$$

означающего, что при n-кратном испытании случайной величины ξ событие $\left\{\xi < x-\Delta x\right\}$,вероятность которого $F(x - \Delta x)$, появится (k-1) раз, событие $\left\{\xi \geq x\right\}$, вероятность которого $(1 - F(x))$, появится (n-k) раз, и событие $\left\{\xi \in (x-\Delta x, x)\right\}$, вероятность которого $p(x)\Delta x$, появится 1 раз.