Тема 3. Методы построения оценок.

1. Метод моментов.

Пусть $\zeta_n = (\xi_1,..., \xi_n)$ – выборка из распределения $F_\xi(x\,|\,\theta)$, где $\theta = (\theta_1, \theta_2,..., \theta_k)$ – неизвестный векторный параметр, принимающий значения из множества допустимых значений параметра Θ, и требуется построить оценку некоторой величины $\tau(\theta) = \tau(\theta_1, \theta_2,..., \theta_k)$, зависящей от параметра θ.

Для решения поставленной задачи методом моментов, прежде всего, необходимо выразить оцениваемую величину $\tau(\theta)$ через моменты распределения $F_\xi(x\,|\,\theta)$, то есть построить такую функцию $G(\nu_1, \nu_2,..., \nu_m)$ что:

$$\tau(\theta_1, \theta_2,..., \theta_k) = G(\nu_1, \nu_2,..., \nu_m),$$

где ν_1, ν_2, …, ν_m – любые моменты распределения $F_\xi(x\,|\,\theta)$ (например, начальные и центральные и не обязательно по порядку). Если в функции $G(\nu_1, \nu_2,..., \nu_m)$ моменты ν_1, ν_2, …, ν_m заменить их оценками $\hat{\nu}_1(\zeta_n)$, $\hat{\nu}_2(\zeta_n)$, …, $\hat{\nu}_m(\zeta_n)$ (например, выборочными моментами), то в результате замены будет получена статистика $T(\zeta_n)$:

$$T(\zeta_n) = G(\hat{\nu}_1(\zeta_n), \hat{\nu}_2(\zeta_n)..., \hat{\nu}_m(\zeta_n)), \qquad (3.1)$$

которая может использоваться для оценки величины $\tau(\theta)$.

Определение 3.1.

Оценками, полученными по методу моментов (моментными оценками), называются статистики вида $T(\hat{\nu}_1,..., \hat{\nu}_k)$, где $\hat{\nu}_i(\zeta_n)$ $(i = \overline{1,k})$ – оценки моментов распределения $F_\xi(x\,|\,\theta)$.

Один из способов построения функции $G(\nu_1,..., \nu_m)$ заключается в том, чтобы выразить компоненты параметра θ_1, θ_2, …, θ_k через моменты ν_1, ν_2, …, ν_m. Поскольку функция распределения $F_\xi(x\,|\,\theta)$ зависит от параметра $\theta = (\theta_1, \theta_2,..., \theta_k)$, то выражения для моментов ν_1, ν_2, …, ν_m в общем случае содержат компоненты θ_1, θ_2, …, θ_k параметра, так что каждый момент ν_i представляет собой функцию $f_i(\theta_1, \theta_2,..., \theta_k)$:

$$\begin{cases} \nu_1 = f_1(\theta_1, \theta_2, ..., \theta_k) \\ \nu_2 = f_2(\theta_1, \theta_2, ..., \theta_k) \\ \quad ... \\ \nu_m = f_m(\theta_1, \theta_2, ..., \theta_k) \end{cases}.$$

Если представленная система разрешима относительно неизвестных компонент θ_1, θ_2, …, θ_k, тогда имеется возможность выразить компоненты θ_1, θ_2, …, θ_k через моменты ν_1, ν_2, …, ν_m:

$$\begin{cases} \theta_1 = g_1(\nu_1, \nu_2, ..., \nu_m) \\ \theta_2 = g_2(\nu_1, \nu_2, ..., \nu_m) \\ \quad ... \\ \theta_k = g_k(\nu_1, \nu_2, ..., \nu_m) \end{cases} \qquad (3.2)$$

В силу (3.2) функция $G(\nu_1,..., \nu_m)$ может быть получена непосредственно из величины $\tau(\theta_1,..., \theta_k)$:

$$G(\nu_1, \nu_2,..., \nu_m) = \tau(\theta_1, \theta_2,..., \theta_k) = \tau(g_1(\nu_1, \nu_2,..., \nu_m),..., g_k(\nu_1, \nu_2,..., \nu_m)).$$

Дополнительно из (3.2) непосредственно получаются моментные оценки $\hat{\theta}_i(\zeta_n)$ и самих компонент θ_i параметра θ:

$$\hat{\theta}_i(\zeta_n) = g_i(\hat{\nu}_1(\zeta_n),..., \hat{\nu}_m(\zeta_n)) . \qquad (3.3)$$

Моментные оценки $\hat{\theta}_1(\zeta_n)$, …, $\hat{\theta}_k(\zeta_n)$ (3.3) компонент θ_i параметра θ, как и моментная оценка $T(\zeta_n)$ (3.1) величины $\tau(\theta)$, в общем случае не обладают свойством несмещенности (тем не менее, в некоторых частных случаях моментные оценки оказываются несмещенными).

Состоятельность каждой оценки $\hat{\theta}_i(\zeta_n)$ (3.3) зависит от свойств функции g_i и используемых оценок $\hat{\nu}_j(\zeta_n)$ ($j = \overline{1,m}$). Если функция g_i непрерывна в точке $(\nu_1, \nu_2,..., \nu_m)$ и все оценки $\hat{\nu}_j(\zeta_n)$ являются состоятельными оценками моментов ν_j, тогда оценка $\hat{\theta}_i(\zeta_n)$ является состоятельной оценкой θ_i. Действительно, в силу свойства сходимости по вероятности, непрерывная функция $\hat{\theta}_i(\zeta_n) = g_i(\hat{\nu}_1, \hat{\nu}_2,..., \hat{\nu}_k)$ от состоятельных оценок $\hat{\nu}_j(\zeta_n)$, имеющих пределом по вероятности ν_j, сходится по вероятности к величине $g_i(\nu_1, \nu_2,..., \nu_k) = \theta_i$, таким образом:

$$\hat{\theta}_i(\zeta_n) = g_i(\hat{\nu}_1(\zeta_n), \hat{\nu}_2(\zeta_n),..., \hat{\nu}_k(\zeta_n)) \xrightarrow{P} g_i(\nu_1, \nu_2,..., \nu_k) = \theta_i ,$$

$$\text{при } n \to \infty .$$

что означает состоятельность оценки $\hat{\theta}_i(\zeta_n)$.

Аналогично, если функция $G(\nu_1, \nu_2,..., \nu_m)$ является непрерывной в точке $(\nu_1, \nu_2,..., \nu_m)$ и все оценки $\hat{\nu}_j(\zeta_n)$ являются состоятельными оценками моментов ν_j, тогда оценка $T(\zeta_n)$ также является состоятельной оценкой величины $\tau(\theta)$:

$$T(\zeta_n) = G(\hat{\nu}_1(\zeta_n), \hat{\nu}_2(\zeta_n),..., \hat{\nu}_k(\zeta_n)) \xrightarrow{P} G(\nu_1, \nu_2,..., \nu_k) = \tau(\theta) ,$$

$$\text{при } n \to \infty .$$

Если в качестве моментов ν_j используются первые j начальных моментов m_j функции распределения $F_\xi(x \mid \theta)$, функции g_1, …, g_k непрерывно дифференцируемы и функция распределения $F_\xi(x \mid \theta)$ имеет $2k$ моментов, то моментные оценки $\hat{\theta}_i(\zeta_n)$ имеют асимптотически нормальное распределение:

$$\hat{\theta}_i(\zeta_n) \sim N\left(\theta_i, \frac{1}{n}\sum_{p=1}^{k}\sum_{q=1}^{k}(m_{p+q} - m_p m_q)\frac{\partial g_p}{\partial m_p}(m) \cdot \frac{\partial g_q}{\partial m_q}(m)\right),$$

где $m = (m_1,..., m_k)$ – вектор начальных моментов.

Моментные оценки в большом количестве случаев не являются оптимальными, тем не менее, применение метода моментов построения оценок во многих случаях не вызывает затруднений и сами выражения для моментных оценок (3.1), как правило, оказываются простыми для вычисления.

2. Метод максимального правдоподобия.

Пусть исходная серия наблюдений представлена совокупностью случайных величин $(\xi_1,..., \xi_n)$, которая имеет функцию плотности вероятности (либо функцию вероятности в дискретном случае) $p_\xi(x_1,..., x_n \mid \theta)$, зависящую от параметра $\theta = (\theta_1,..., \theta_k)$, принимающего значения из множества допустимых значений параметра Θ.

Образуем так называемую функцию правдоподобия $L(\xi_1,..., \xi_n \mid \theta)$ путем подстановки величин ξ_i вместо величин x_i в функцию $p_\xi(x_1,..., x_n \mid \theta)$:

$$L(\xi_1,..., \xi_n \mid \theta) = p_\xi(\xi_1,..., \xi_n \mid \theta) .$$

Предположим, в результате проведения эксперимента получена реализация серии наблюдений $(\xi_1,..., \xi_n)$ – числовой вектор $(x_1^*,..., x_n^*)$. Если подставить вектор значений $(x_1^*,..., x_n^*)$ в функцию правдоподобия $L(\xi_1,..., \xi_n \mid \theta)$, то в результате получится функция $L(x_1^*,..., x_n^* \mid \theta)$, зависящая только от вектора параметров $\theta = (\theta_1,..., \theta_k)$. При одних значениях параметра θ значение функции правдоподобия $L(x_1^*,..., x_n^* \mid \theta)$ оказывается мало, при других значениях – велико. Поскольку значение функция правдоподобия $L(x_1^*,..., x_n^* \mid \theta)$ в некотором смысле отражает вероятность получения заданного вектора $L(x_1^*,..., x_n^* \mid \theta)$, то вполне естественно было бы подобрать такое значение параметра $\theta = (\theta_1,..., \theta_k)$, при котором вероятность получения наблюдаемого значения $(x_1^*,..., x_n^*)$ оказалась бы наибольшей.

Определение 3.2.

Оценкой, полученной по методу максимального правдоподобия (МП-оценкой), называется статистика $\theta^* = (\theta_1^*,..., \theta_k^*)$, доставляющая наибольшее значение функции правдоподобия $L(\xi_1,..., \xi_n \mid \theta)$ при каждой реализации величин $(\xi_1,..., \xi_n)$:

$$L(\xi_1,..., \xi_n \mid \theta^*) = \sup_{\theta=(\theta_1,...,\theta_k)} L(\xi_1,..., \xi_n \mid \theta).$$

Определение 3.2 следует понимать в следующем смысле: при каждом фиксированном элементарном событии $\omega \in \Omega$, случайные величины ξ_1, …, ξ_n принимают определенные числовые значения $x_i = \xi_i(\omega)$, а совокупность наблюдений $(\xi_1,..., \xi_n)$ при фиксированном ω принимает значение числового вектора $(x_1,..., x_n)$. Для заданного вектора $(x_1,..., x_n)$ согласно определению 3.2 вычисляется значение МП-оценки $\theta^* = \theta^*(x_1,..., x_n)$:

$$L(x_1,..., x_n \mid \theta^*) = \sup_{\theta=(\theta_1,...,\theta_k)} L(x_1,..., x_n \mid \theta).$$

Тем самым для каждого $\omega \in \Omega$ задан способ вычисления МП-оценки $\theta^*(x_1,..., x_n) = \theta^*(\xi_1(\omega),..., \xi_n(\omega))$, который и определяет функцию $\theta^* = \theta^*(\xi_1,..., \xi_n)$. Таким образом, МП-оценка как функция наблюдений является статистикой.

Если при каждой реализации совокупности $(\xi_1,..., \xi_n)$ наибольшее значение функции правдоподобия достигается во внутренней точке множества допустимых значений параметра Θ и функция правдоподобия $L(\xi_1,..., \xi_n \mid \theta)$ дифференцируема по параметру θ, то из необходимого условия экстремума следует равенство частных производных функции правдоподобия $L(\xi_1,..., \xi_n \mid \theta)$ нулю в точке МП-оценки θ^*:

$$\left.\frac{\partial}{\partial \theta_i} L(\xi_1,..., \xi_n \mid \theta_1,..., \theta_k)\right|_{\theta=\theta^*} = 0, \; i = \overline{1,k}.$$

Решение приведенной системы не всегда оказывается удобным, поэтому при выполнении определенных условий задачу нахождения наибольшего значения функции правдоподобия $L(x_1,..., x_n \mid \theta)$ заменяют задачей нахождения наибольшего значения логарифма функции правдоподобия $\ln L(x_1,..., x_n \mid \theta)$, поскольку логарифм функция монотонно возрастающая (и, следовательно, наибольшему значению логарифма функции правдоподобия будет соответствовать наибольшее значение аргумента, то есть функции правдоподобия):

$$\left.\frac{\partial}{\partial \theta_i} \ln L(\xi_1,..., \xi_n \mid \theta_1,..., \theta_k)\right|_{\theta=\theta^*} = 0, \; i = \overline{1,k}.$$

В случае одномерного параметра θ представленное уравнение имеет название *уравнения правдоподобия*.

Определение 3.3.

Статистика $T(\xi_1,..., \xi_n)$ называется *асимптотически нормальной с параметрами* m *и* σ^2:

$$T(\xi_1,..., \xi_n) \sim N(m, \sigma^2) \text{ при } n \to \infty,$$

если при каждом x:

$$\lim_{n \to \infty} F_{T_n}(x) = \Phi\left(\frac{x - m}{\sigma}\right),$$

где $F_{T_n}(x)$ – функция распределения $T(\xi_1,..., \xi_n)$, $\Phi(x)$ – функция Лапласа (функция распределения нормальной случайной величины с математическим ожиданием 0 и дисперсией 1).

Замечание

Можно доказать, что при некоторых условиях МП-оценка $\theta^*(\xi_1,..., \xi_n)$ является:

1) состоятельной;
2) асимптотически нормальной.

Метод максимального правдоподобия может вызывать серьезные трудности в связи с необходимостью отыскания наибольшего значения функции правдоподобия, в частности решения уравнения правдоподобия, тем не менее, МП-оценки как правило являются оценками близкими к оптимальным оценкам.

3. Метод порядковых статистик.

Пусть $\zeta_n = (\xi_1,..., \xi_n)$ – выборка из распределения $F_\xi(x \mid \theta)$, где $\theta = (\theta_1, \theta_2,..., \theta_k)$ – неизвестный векторный параметр, принимающий значения из множества допустимых значений Θ, и требуется построить оценку величины $\tau(\theta) = \tau(\theta_1, \theta_2,..., \theta_k)$.

Метод порядковых статистик аналогичен методу моментов с той разницей что вместо моментов используются квантили x_{p_1}, x_{p_2}, …, x_{p_m} и их оценки $\tilde{x}_{p_1}(\zeta_n)$, $\tilde{x}_{p_2}(\zeta_n)$, …, $\tilde{x}_{p_m}(\zeta_n)$. Для построения оценки величины $\tau(\theta)$ по методу порядковых статистик достаточно выразить величину $\tau(\theta)$ через квантили x_{p_1}, x_{p_2},…, x_{p_m} функции распределения $F_\xi(x \mid \theta)$, то есть построить такую функцию $H(x_{p_1}, x_{p_2},..., x_{p_m})$ что:

$$\tau(\theta_1, \theta_2,..., \theta_k) = H(x_{p_1}, x_{p_2},..., x_{p_m}). \tag{3.4}$$

Если квантили x_{p_1}, x_{p_2},…, x_{p_m} заменить их оценками $\tilde{x}_{p_1}(\zeta_n)$, $\tilde{x}_{p_2}(\zeta_n)$, …, $\tilde{x}_{p_m}(\zeta_n)$, то будет получена статистика $T(\zeta_n)$:

$$T(\zeta_n) = H(\tilde{x}_{p_1}(\zeta_n), \tilde{x}_{p_2}(\zeta_n),..., \tilde{x}_{p_m}(\zeta_n)). \tag{3.5}$$

которая может быть использована в качестве оценки величины $\tau(\theta)$.

Один из способов построения функции $H(x_{p_1}, x_{p_2},..., x_{p_m})$, как и в методе моментов, заключается в том, чтобы при вычислении квантилей x_{p_1}, x_{p_2}, …, x_{p_m} выразить их через компоненты θ_1, θ_2, …, θ_k параметра:

$$\begin{cases} x_{p_1} = f_1(\theta_1, \theta_2,..., \theta_k) \\ x_{p_2} = f_2(\theta_1, \theta_2,..., \theta_k) \\ \qquad \ldots \\ x_{p_m} = f_m(\theta_1, \theta_2,..., \theta_k) \end{cases},$$

и, разрешив систему относительно неизвестных компонент θ_1, θ_2, …, θ_k, получить выражения для компонент θ_1, θ_2, …, θ_k через квантили x_{p_1}, x_{p_2}, …, x_{p_m}:

$$\begin{cases} \theta_1 = h_1(x_{p_1}, x_{p_2},..., x_{p_m}) \\ \theta_2 = h_2(x_{p_1}, x_{p_2},..., x_{p_m}) \\ \quad\quad ... \\ \theta_k = h_k(x_{p_1}, x_{p_2},..., x_{p_m}) \end{cases},$$

которые затем использовать при построении функции $H(x_{p_1}, x_{p_2},..., x_{p_m})$ непосредственно из величины $\tau(\theta_1,..., \theta_k)$:

$$H(x_{p_1}, x_{p_2},..., x_{p_m}) = \tau(\theta_1, \theta_2,..., \theta_k) = \tau(h_1(\tilde{x}_{p_1}, \tilde{x}_{p_2},..., \tilde{x}_{p_m}),..., h_k(\tilde{x}_{p_1}, \tilde{x}_{p_2},..., \tilde{x}_{p_m})) .$$

Теперь для построения оценки $T(\zeta_n)$ (3.4) остается лишь найти оценки $\tilde{x}_{p_i}(\zeta_n)$ ($i = \overline{1, m}$) квантилей x_{p_i}. Оценки $\tilde{x}_{p_i}(\zeta_n)$ могут быть получены различными способами, в частности с помощью порядковых статистик $\xi_{(k)}$ ($k = \overline{1, n}$) – величин вариационного ряда. Для построения оценок квантилей, прежде всего, рассмотрим определение квантили.

Определение 3.4.

Пусть $F_\xi(x)$ – функция распределения, *квантиль уровня* p (p *-квантиль*) функции распределения $F_\xi(x)$ есть число x_p такое, что:

$$F_\xi(x_p) = p .$$

Для построения оценок квантилей необходимо некоторым образом заменить в определении квантили 3.4 неизвестную функцию распределения $F_\xi(x\,|\,\theta)$ известной эмпирической функцией распределения $F_n^*(x; \xi_1,..., \xi_n)$. Реализации эмпирической функции распределения являются кусочно-постоянными функциями, а приведенное выше определение квантили, как нетрудно заметить, не является подходящим для функций, имеющих разрывы. Действительно, если функция распределения $F_\xi(x)$ имеет разрывы, то вообще говоря, при некоторых p может не существовать такого числа x_p, при котором $F_\xi(x_p) = p$, поэтому определение 3.4 необходимо модифицировать.

Определением квантили, подходящим для более широкого класса функций, в том числе и кусочно-постоянных, может являться, например, следующее определение: квантилью уровня p (при $p > 0$) можно считать такое значение x_p, которое является точной нижней гранью множества всех значений x, при которых функция распределения $F_\xi(x)$ больше p,

$$x_p = \inf\{x : p < F_\xi(x)\},$$
$$p > 0 .$$

Введенное определения квантили для непрерывных функций распределений $F_\xi(x)$ совпадает с определением 3.4, но также подходит и для функций распределения $F_\xi(x)$, имеющих разрывы. Теперь, замена в этом определении функции распределения $F_\xi(x)$ известной эмпирической функцией распределения $F_n^*(x; \xi_1,..., \xi_n)$ приводит к оценке $\tilde{x}_p$ квантили x_p некоторого уровня $p > 0$:

$$\tilde{x}_p(\xi_1,..., \xi_n) = \inf\{x : p < F_n^*(x; \xi_1,..., \xi_n)\}.$$

Исходя из графиков реализаций эмпирической функции распределения $F_n^*(x; \xi_1,..., \xi_n)$ нетрудно заметить, что для значений p: $0 < p$, $\frac{k}{n} \le p < \frac{k+1}{n}$, ($k = \overline{0, n-1}$) оценка $\tilde{x}_p(\zeta_n)$

совпадает с $k+1$-ой порядковой статистикой $\xi_{(k+1)}$, причем если значение p таково, что $\frac{k}{n} \le p < \frac{k+1}{n}$, тогда $k = [np]$, где $[\cdot]$ означает целую часть числа.

Таким образом, оценкой $\tilde{x}_p(\zeta_n)$ квантили x_p некоторого уровня $p > 0$ является порядковая статистика $\xi_{([np]+1)}$ и статистика $T(\xi_1,..., \xi_n)$ (3.4), оценивающая величину $\tau(\theta)$, имеет вид:

$$T(\xi_1,..., \xi_n) = H(\xi_{([np_1]+1)},..., \xi_{([np_m]+1)})$$

Определение 3.5.

Оценками, полученными по методу порядковых статистик, называются оценки вида $T(\xi_{(i_1)},..., \xi_{(i_m)})$, где $\xi_{(i_k)}$ $(k = \overline{1,m})$ – порядковые статистики.

Свойства оценок, полученных по методу порядковых статистик, зависят от свойств самих порядковых статистик $\xi_{(1)}$, которые в свою очередь зависят от функции распределения $F_\xi(x \mid \theta)$ исходной выборки $(\xi_1,..., \xi_n)$.

Утверждение 3.6.

Пусть $(\xi_1,..., \xi_n)$ – выборка из распределения $F_\xi(x \mid \theta)$, тогда функция распределения k-ой порядковой статистики $F_{(k)}(x \mid \theta)$:

$$F_{(k)}(x \mid \theta) = \sum_{m=k}^{n} C_n^m F_\xi^m(x \mid \theta)(1 - F_\xi(x \mid \theta))^{n-m} = \begin{cases} 1 - (1 - F_\xi(x \mid \theta))^n & ,k = 1 \\ \sum_{m=k}^{n} C_n^m F_\xi^m(x \mid \theta)(1 - F_\xi(x \mid \theta))^{n-m} & ,1 < k < n \\ F_\xi^n(x \mid \theta) & ,k = n \end{cases}$$

Доказательство:

Выберем произвольным образом и зафиксируем значение x, определим на основе выборки $(\xi_1,..., \xi_n)$ вектор бинарных случайных величин $(\varepsilon_1,..., \varepsilon_n)$:

$$\varepsilon_i(\omega) = \begin{cases} 1 & ,\xi_i(\omega) < x \\ 0 & ,\xi_i(\omega) \ge x \end{cases}.$$

Случайные величины ε_i независимы в совокупности (поскольку случайные величины ξ_i независимы в совокупности) и имеют одинаковое распределение P_ε (поскольку случайные величины ξ_i имеют одинаковую функцию распределения):

$$P_\varepsilon(1) = P\{\omega : \varepsilon_i(\omega) = 1\} = P\{\omega : \xi_i(\omega) < x\} = F_\xi(x \mid \theta),$$

$$P_\varepsilon(0) = P\{\omega : \varepsilon_i(\omega) = 0\} = P\{\omega : \xi_i(\omega) \ge x\} = 1 - F_\xi(x \mid \theta).$$

Пусть $\xi_{(k)}$ – k-ая порядковая статистика, по определению функция распределения $F_{(k)}(x)$:

$$F_{(k)}(x) = P\{\omega : \xi_{(k)}(\omega) < x\}.$$

Порядковая статистика $\xi_{(k)}$ меньше величины x тогда и только тогда, когда среди величин выборки $(\xi_1,..., \xi_n)$ m $(k \le m \le n)$ величин меньше x и $n-m$ величин не меньше x, то есть тогда и только тогда, когда в векторе бинарных случайных величин $(\varepsilon_1,..., \varepsilon_n)$ m величин равны 1 и $n-m$ величин равны 0, что эквивалентно тому, что случайная величина $\sum_{i=1}^{n} \varepsilon_i$ больше или равна k:

$$F_{(k)}(x) = P\{\omega : \xi_{(k)}(\omega) < x\} = P\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) \ge k\right\} \quad (3.6)$$

Событие $\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) \geq k\right\}$ можно представить как объединение непересекающихся событий $\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) = m\right\}$ при $m = \overline{k,n}$:

$$\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) \geq k\right\} = \bigcup_{m=k}^{n}\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) = m\right\},$$

тогда

$$P\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) \geq k\right\} = P\left\{\bigcup_{m=k}^{n}\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) = m\right\}\right\} = \sum_{m=k}^{n} P\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) = m\right\}, \qquad (3.7)$$

где последнее равенство получено с учетом того, что события $\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) = m\right\}$ не пересекаются при различных m. Поскольку все случайные величины ε_i независимы и имеют одинаковое бинарное распределение, то случайная величина $\sum_{i=1}^{n} \varepsilon_i$ имеет распределение Бернулли с параметрами n и $F_\xi(x \mid \theta)$, то есть:

$$P\left\{\omega : \sum_{i=1}^{n} \varepsilon_i(\omega) = m\right\} = C_n^m F_\xi^m(x \mid \theta)(1 - F_\xi(x \mid \theta))^{n-m} \qquad (3.8)$$

Таким образом, из (3.6)-(3.8) окончательно получим:

$$F_{(k)}(x) = \sum_{m=k}^{n} C_n^m F_\xi^m(x \mid \theta)(1 - F_\xi(x \mid \theta))^{n-m}. \qquad (3.9)$$

Заметим, что полученное равенство справедливо для любого x, поскольку величина x была выбрана произвольным образом.

При $k = 1$ из (3.9) получим:

$$F_{(1)}(x) = \sum_{m=1}^{n} C_n^m F_\xi^m(x \mid \theta)(1 - F_\xi(x \mid \theta))^{n-m} = \sum_{m=1}^{n} C_n^m F_\xi^m(x \mid \theta)(1 - F_\xi(x \mid \theta))^{n-m} +$$

$$+ C_n^0(1 - F_\xi(x \mid \theta))^n - C_n^0(1 - F_\xi(x \mid \theta))^n = \sum_{m=0}^{n} C_n^m F_\xi^m(x \mid \theta)(1 - F_\xi(x \mid \theta))^{n-m} - C_n^0(1 - F_\xi(x \mid \theta))^n =$$

$$= (F_\xi(x \mid \theta) + (1 - F_\xi(x \mid \theta)))^n - (1 - F_\xi(x \mid \theta))^n = 1 - (1 - F_\xi(x \mid \theta))^n.$$

При $k = n$ из (3.9) получим:

$$F_{(n)}(x) = \sum_{m=n}^{n} C_n^m F_\xi^m(x \mid \theta)(1 - F_\xi(x \mid \theta))^{n-m} = F_\xi^n(x \mid \theta).$$

Утверждение доказано.

Утверждение 3.7.

Пусть выполнены условия утверждения 3.6 и функция распределения $F_\xi(x \mid \theta)$ дифференцируема по x при всех x и θ, тогда плотность вероятности k-ой порядковой статистики $f_{(k)}(x)$:

$$f_{(k)}(x \mid \theta) = kC_n^k F_\xi^{k-1}(x \mid \theta)(1 - F_\xi(x \mid \theta))^{n-k} f_\xi(x \mid \theta),$$

где $f_\xi(x \mid \theta) = \dfrac{d}{dx} F_\xi(x \mid \theta)$ – плотность вероятности.

Доказательство:

Действительно, для того, чтобы получить выражение для плотности вероятности $f_{(k)}(x \mid \theta)$ достаточно продифференцировать функцию распределения $F_{(k)}(x \mid \theta)$ по x:

При $k = 1$ получим:

$$f_{(k)}(x \mid \theta) = \frac{d}{dx}\left(1 - (1 - F_{\xi}(x \mid \theta))^{n}\right) = n(1 - F_{\xi}(x \mid \theta))^{n-1} f_{\xi}(x \mid \theta).$$

При $k = n$ получим:

$$f_{(k)}(x \mid \theta) = \frac{d}{dx}\left(F_{\xi}^{n}(x \mid \theta)\right) = nF_{\xi}^{n-1}(x \mid \theta) f_{\xi}(x \mid \theta).$$

При $1 < k < n$ получим (для краткости опускаем аргумент функций):

$$f_{(k)}(x \mid \theta) = \frac{d}{dx}\left(\sum_{m=k}^{n} C_n^m F_{\xi}^m (1 - F_{\xi})^{n-m}\right) = \frac{d}{dx}\left(\sum_{m=k}^{n-1} C_n^m F_{\xi}^m (1 - F_{\xi})^{n-m} + C_n^n F_{\xi}^n\right) =$$

$$= \sum_{m=k}^{n-1}\left(C_n^m mF_{\xi}^{m-1} f_{\xi}(1 - F_{\xi})^{n-m} + C_n^m F_{\xi}^m (n-m)(1 - F_{\xi})^{n-m-1}(-f_{\xi})\right) + C_n^n nF_{\xi}^{n-1} f_{\xi} =$$

$$= \sum_{m=k}^{n-1} C_n^m mF_{\xi}^{m-1}(1 - F_{\xi})^{n-m} f_{\xi} - \sum_{m=k}^{n-1} C_n^m (n-m) F_{\xi}^m (1 - F_{\xi})^{n-m-1} f_{\xi} + C_n^n nF_{\xi}^{n-1} f_{\xi} =$$

$$= C_n^k kF_{\xi}^{k-1}(1 - F_{\xi})^{n-k} f_{\xi} + \sum_{m=k+1}^{n-1} C_n^m mF_{\xi}^{m-1}(1 - F_{\xi})^{n-m} f_{\xi} -$$

$$- \sum_{m=k}^{n-2} C_n^m (n-m) F_{\xi}^m (1 - F_{\xi})^{n-m-1} f_{\xi} - C_n^{n-1}(n - (n-1)) F_{\xi}^{n-1}(1 - F_{\xi})^{n-(n-1)-1} f_{\xi} + C_n^n nF_{\xi}^{n-1} f_{\xi} =$$

$$= C_n^k kF_{\xi}^{k-1}(1 - F_{\xi})^{n-k} f_{\xi} + \sum_{m=k}^{n-2} C_n^{m+1}(m+1) F_{\xi}^m (1 - F_{\xi})^{n-m-1} f_{\xi} -$$

$$- \sum_{m=k}^{n-2} C_n^m (n-m) F_{\xi}^m (1 - F_{\xi})^{n-m-1} f_{\xi} - C_n^{n-1} F_{\xi}^{n-1} f_{\xi} + C_n^{n-1} F_{\xi}^{n-1} f_{\xi} =$$

$$= C_n^k kF_{\xi}^{k-1}(1 - F_{\xi})^{n-k} f_{\xi} + \sum_{m=k}^{n-2}\left(C_n^{m+1}(m+1) F_{\xi}^m (1 - F_{\xi})^{n-m-1} f_{\xi} - C_n^m (n-m) F_{\xi}^m (1 - F_{\xi})^{n-m-1} f_{\xi}\right) =$$

$$= C_n^k kF_{\xi}^{k-1}(1 - F_{\xi})^{n-k} f_{\xi} + \sum_{m=k}^{n-2}\left(C_n^{m+1}(m+1) - C_n^m (n-m)\right) F_{\xi}^m (1 - F_{\xi})^{n-m-1} f_{\xi} =$$

$$= C_n^k kF_{\xi}^{k-1}(1 - F_{\xi})^{n-k} f_{\xi} + \sum_{m=k}^{n-2}\left(\frac{n!}{(m+1)!(n-(m+1))!}(m+1) - \frac{n!}{m!(n-m)!}(n-m)\right) F_{\xi}^m (1 - F_{\xi})^{n-m-1} f_{\xi} =$$

$$= C_n^k kF_{\xi}^{k-1}(1 - F_{\xi})^{n-k} f_{\xi} + \sum_{m=k}^{n-2}\left(\frac{n!}{m!(n-m-1)!} - \frac{n!}{m!(n-m-1)!}\right) F_{\xi}^m (1 - F_{\xi})^{n-m-1} f_{\xi} =$$

$$= kC_n^k F_{\xi}^{k-1}(x \mid \theta)(1 - F_{\xi}(x \mid \theta))^{n-k} f_{\xi}(x \mid \theta).$$

Утверждение доказано.

Теорема 3.8. (Крамер)

Пусть $(\xi_1, ..., \xi_n)$ – выборка из распределения $F_{\xi}(x)$, x_p – p-квантиль распределения $F_{\xi}(x)$ и в некоторой окрестности точки x_p плотность вероятности $f_{\xi}(x) = F_{\xi}'(x)$ непрерывно дифференцируема и положительна, $f_{\xi}(x) > 0$, тогда порядковая статистика $\xi_{([np]+1)}$ имеет асимптотически нормальное распределение:

$$\xi_{([np]+1)} \sim N\left(x_p, \frac{p(1-p)}{nf_{\xi}^2(x_p)}\right), \text{ при } n \to \infty.$$

Следствие

При выполнении условий теоремы 3.8 статистика $\xi_{([np]+1)}$ является состоятельной оценкой p-квантиля x_p, поскольку математическое ожидание $M[\xi_{([np]+1)}] \to x_p$ и дисперсия $D[\xi_{([np]+1)}] \to 0$ при $n \to \infty$.

Из теоремы Крамера следует, что при некоторых условиях статистики $\tilde{x}_{p_i}(\zeta_n) = \xi_{([np_i]+1)}$ являются состоятельными оценками квантилей x_{p_i}, то есть имеет место сходимость по вероятности:

$$\tilde{x}_{p_i}(\zeta_n) \xrightarrow{P} x_{p_i},$$
$$\text{при } n \to \infty.$$

Если функция $H(x_{p_1}, x_{p_2},..., x_{p_m})$ из (3.4), связывающая оцениваемую величину $\tau(\theta)$ с квантилями x_{p_i}, непрерывна в точке $(x_{p_1}, x_{p_2},..., x_{p_m})$, тогда по свойству сходимости по вероятности статистика $H(\tilde{x}_{p_1}(\zeta_n), x_{p_2}(\zeta_n),..., x_{p_m}(\zeta_n))$ (3.5) сходится по вероятности к величине $H(x_{p_1}, x_{p_2},..., x_{p_m})$:

$$T(\zeta_n) = H(\tilde{x}_{p_1}(\zeta_n), x_{p_2}(\zeta_n),..., x_{p_m}(\zeta_n)) \xrightarrow{P} H(x_{p_1}, x_{p_2},..., x_{p_m}) = \tau(\theta),$$
$$\text{при } n \to \infty,$$

и статистика $T(\zeta_n)$ по определению является состоятельной оценкой $\tau(\theta)$.

Оценки, полученные методом порядковых статистик, как правило, имеют дисперсию больше, чем дисперсии оценок, полученные другими методами. Тем не менее, оценки, полученные методом порядковых статистик, могут обладать дополнительными положительными свойствами, например, устойчивостью к «засорению» выборки («засорение» выборки означает наличие в выборке ошибочных значений, полученных, например, в результате неверного измерения). К тому же метод порядковых статистик является достаточно общим, то есть применим в тех случаях, когда получение оценок другими методами оказывается затруднительным или вовсе невозможным.