Семинар №4
Методы нахождения оценок.

Метод моментов.

Пусть параметр $\theta = (\theta_1, \ldots, \theta_k) \in \mathbb{R}^k$, а $g_1(x), \ldots, g_k(x)$ – борелевские функции. Пусть $X_1, \ldots, X_n$ – выборка из неизвестного распределения P_θ, $\theta \in \Theta$.
Рассмотрим функции $m_i(\theta) = E_\theta g_i(X_1)$, $i = 1 \ldots k$.
Составим систему уравнений относительно θ : $\begin{cases} m_1(\theta^*) = \overline{g_1(X)} \\ \quad\ldots \\ m_k(\theta^*) = \overline{g_k(X)} \end{cases}$
Определение. Решение такой системы (относительно θ^*) $\theta^*(X)$ называется оценкой по методу моментов с пробными функциями $g_1, \ldots, g_k$.
Замечание. Естественно, что нужно подбирать такие пробные функции, чтобы весь вектор θ однозначно выражался из введённой нами системы. Функции $g_i(x) = x^i$, $i = 1 \ldots k$, называются стандартными пробными функциями.
Задача. Найти оценки по методу моментов со стандартными пробными функциями для $N(a, \sigma^2)$.
Решение.
Здесь $\theta = (a, \sigma^2)$. Выберем стандартные пробные функции $g_1(x) = x$ и $g_2(x) = x^2$.
Тогда $E_\theta g_1(X_1) = E_\theta X_1 = a$, $E_\theta g_2(X_1) = E_\theta X_1^2 = D_\theta X_1 + (E_\theta X_1)^2 = \sigma^2 + a^2$. Тем самым, получаем систему: $\begin{cases} a^* = \overline{X} \\ (a^*)^2 + (\sigma^2)^* = \overline{X^2} \end{cases}$,
решая которую, получаем ответ: оценкой для вектора $\theta = (a, \sigma^2)$ служит $(\overline{X}, s^2)$. ■
Оценка методом моментов не обязательно является несмещённой. Однако верно следующее утверждение.

Лемма 1. *Если $\overrightarrow{m}(\theta) = (m_1(\theta), \ldots, m_k(\theta))$ биективна на Θ, $(\overrightarrow{m})^{-1}$ является непрерывным отображением и, кроме того, $E_\theta|g_i(X_1)| < \infty$, то оценка методом моментов сильно состоятельна.*

При выполнении некоторых более узких условий на вектор-функцию $\overrightarrow{m}$ оценка методом моментов является асимптотически нормальной.

Метод максимального правдоподобия.

Пусть $\{P_\theta, \theta \in \Theta\}$ – параметрическое семейство распределений. Если все P_θ – абсолютно непрерывные распределения, то положим $p_\theta(x)$ равной плотности P_θ. Если же все $\{P_\theta, \theta \in \Theta\}$ – дискретные распределения, то положим $p_\theta(x) = P_\theta(X = x)$.
Определение. Функцией правдоподобия выборки $X_1, \ldots, X_n$ из распределения P_θ, $\theta \in \Theta$, называется случайная величина $f_\theta(X_1, \ldots, X_n) = p_\theta(X_1)p_\theta(X_2)\ldots p_\theta(X_n)$. Величина $L_\theta(X_1, \ldots, X_n) = \ln f_\theta(X_1, \ldots, X_n)$ называется логарифмической функцией правдоподобия.
Определение. Оценкой максимального правдоподобия (ОМП) параметра θ называется

$$\widehat{\theta}(X_1, \ldots, X_n) = \arg\max_{\theta \in \Theta} f_\theta(X_1, \ldots, X_n),$$

т.е. то значение $\theta \in \Theta$, при котором достигается максимум функции правдоподобия при фиксированных $X_1, \ldots, X_n$.
Задача. Найти оценку максимального правдоподобия для $X_i \sim Bern(\theta)$.
Решение.
Поскольку бернуллиевское распределение дискретно, то $p_\theta(x) = P_\theta(X = x)$, где $x \in \{0, 1\}$. В свою очередь,

$P_\theta(X=x) = \begin{cases} \theta, \text{ если } x=1, \\ 1-\theta, \text{ если } x=0. \end{cases} = \theta^{I\{x=1\}}(1-\theta)^{I\{x=0\}} = \theta^x(1-\theta)^{1-x} I\{x \in \{0,1\}\}.$

Получаем, что функция правдоподобия равна $f_\theta(X_1,\ldots,X_n) = \theta^{\sum X_i}(1-\theta)^{n-\sum X_i}$, а логарифмическая функция правдоподобия (которую, как правило, легче дифференцировать, чтобы найти точку максимума) – $L_\theta(X_1,\ldots,X_n) = \sum X_i \ln\theta + (n-\sum X_i)\ln(1-\theta)$. Продифференцируем $L_\theta(X_1,\ldots,X_n)$ по θ :
$\frac{\sum X_i}{\theta} - \frac{n-\sum X_i}{1-\theta} = 0$, откуда видим, что $\theta^* = \overline{X}$. $\square$

Замечание. Если у уравнения правдоподобия $L'_\theta(X_1,\ldots,X_n) = 0$ только одно решение, то оно максимизирует функцию правдоподобия и, тем самым, является ОМП и состоятельной оценкой искомого параметра. Когда решений несколько, то даже то решение, которое максимизирует функцию правдоподобия, не обязательно является состоятельной оценкой искомого параметра.

Метод выборочных квантилей.

Определение. Пусть $F(x)$ – функция распределения на $\mathbb{R}$. Пусть $p \in (0,1)$, тогда p-квантилью функции распределения $F(x)$ называют величину $z_p = \inf\{x : F(x) \geq p\}$.

Определение. Пусть $X_1,\ldots,X_n$ – выборка из неизвестного распределения P. Статистику

$$z_{n,p} = \begin{cases} X_{([np]+1)}, & \text{если } np \notin \mathbb{Z}; \\ X_{(np)}, & \text{если } np \in \mathbb{Z} \end{cases}$$

называют выборочной $p-$ квантилью.

Теорема. *(Об асимптотической нормальности выборочной квантили)*
Пусть $X_1,\ldots X_n$ – выборка из распределения P с плотностью $f(x)$. Пусть z_p – p-квантиль распределения P, причём $f(x)$ непрерывно дифференцируема в окрестности z_p и $f(z_p) > 0$. Тогда

$$\sqrt{n}(z_{n,p} - z_p) \xrightarrow{d} N\left(0, \frac{1}{4f^2(z_p)}\right).$$

Определение. Медианой распределения P называют его $\frac{1}{2}$-квантиль.

Определение. Выборочной медианой выборки $X_1,\ldots X_n$ называют статистику $\widehat{\mu} = \begin{cases} X_{(k+1)}, & \text{если } n = 2k+1; \\ \frac{X_{(k)}+X_{(k+1)}}{2}, & \text{если } n = 2k. \end{cases}$

Асимптотическое поведение $\widehat{\mu}$ схоже с поведением $z_{n,0,5}$: в условиях теоремы о выборочной квантили $\sqrt{n}(\widehat{\mu} - z_{0.5}) \xrightarrow{d} N\left(0, \frac{1}{4f^2(z_{0.5})}\right)$ (этот факт носит название теоремы о выборочной медиане).