# Лекция 2 (от 9.09)

## Глава 2. Точечные оценки параметров

### 2.1. Статистики и оценки

Пусть $(\mathscr{X}, \mathcal{B}_{\mathscr{X}}, \mathcal{P})$ — вероятностно-статистическая модель, $\mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$ — параметрическое семейство распределений.

*Задача*: оценить $\theta$.

Пусть $X = (X_1, \ldots, X_n)$ — выборка из неизвестного распределения $P \in \mathcal{P}$.

> **Определение**: Пусть $(E, \mathcal{E})$ — измеримое пространство. Тогда измеримая функция $S : \mathscr{X}^n \to E$ называется *статистикой*.
>
> Если $E = \Theta$, то $S(X)$ называется *оценкой* $\theta$.

**Примеры статистик:**

Пусть $X = (X_1, \ldots, X_n)$ — действительная выборка, т. е. $\mathscr{X} = \mathbb{R}$.

1. Выборочные характеристики:
   - $\overline{g(X)} = \frac{1}{n}\sum\limits_{i=1}^{n} g(X_i)$ — *выборочная характеристика* функции $g$ ($g$ борелевская).
   - $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ — *выборочное среднее*.
   - $\overline{X^k} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i^k$ — *выборочный $k$-ый момент*.
2. Функции от выборочных характеристик (т.е $h(\overline{g_1(X)}, \ldots, \overline{g_k(X)});\ h, g_i$ — борелевские):
   - $g_1(x) = x^2, g_2(x) = x, h(x, y) = x - y^2$
     $h(\overline{g_1(X)}, \overline{g_2(X)}) = \overline{X}^2 - \overline{X}^2 = S^2$ — *выборочная дисперсия*.

   **Утверждение:** $S^2 = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2$.
3. Порядковые статистики:

   Упорядочим выборку по возрастанию: $(X_{(1)}, \ldots, X_{(n)})$ — *вариационный ряд*.

   $X_{(k)}$ — *$k$-я порядковая статистика*.

**Пример:**

$(X_1, X_2, X_3) = (2, 5, 1)$.

$\overline{X} = 8/3$

$\overline{X^2} = 10$
$S^2 = 10 - 64/9 = 26/9.$

Вариационный ряд: $(X_{(1)}, X_{(2)}, X_{(3)}) = (1, 2, 5)$.

## 2.2. Свойства оценок

**Замечание:** для распределения $P_\theta$ будем обозначать: $E_\theta$ — матожидание, $D_\theta$ — дисперсия, $P_\theta$-п.н., $d_\theta$.

Пусть $X = (X_1, \ldots, X_n)$ — выборка из неизвестного распределения $P \in \{P_\theta \mid \theta \in \Theta\}, \Theta \in \mathbb{R}^d$.

> **Определение:** оценка $\hat{\theta}$ называется *несмещенной оценкой* $\tau(\theta)$, если $E_\theta \hat{\theta}(X) = \tau(\theta) \ \ \forall \theta \in \Theta$.

**Примеры:**

- $\hat{\theta}_1 = X_1, \ \hat{\theta}_2 = \overline{X}$ — несмещенные оценки для $\tau(\theta) = E_\theta X_1$.
- $\mathcal{P} = \{Bern(\theta) \mid \theta \in (0,1)\} : \overline{X}, X_1$ — несмещенные оценки $\theta$.
- $\mathcal{P} = \{Exp(\theta) \mid \theta > 0\} : \overline{X}, X_1$ — несмещенные оценки $\frac{1}{\theta}$.

**Асимптотические свойства**

Пусть $X = (X_1, \ldots)$ — выборка неограниченного размера из $P \in \{P_\theta \mid \theta \in \Theta\}, \Theta \in \mathbb{R}^d$.

> **Определение:**
>
> 1. Оценка $\hat{\theta_n}(X_1, \ldots, X_n)$ называется *состоятельной оценкой* $\theta$, если $\hat{\theta_n}(X_1, \ldots, X_n) \xrightarrow{P_\theta} \theta \quad \forall \theta \in \Theta$.
> 2. Оценка $\hat{\theta_n}(X_1, \ldots, X_n)$ называется *сильно состоятельной оценкой* $\theta$, если $\hat{\theta_n}(X_1, \ldots, X_n) \xrightarrow{P_\theta - \text{п.н.}} \theta \quad \forall \theta \in \Theta$.
> 3. Оценка $\hat{\theta_n}(X_1, \ldots, X_n)$ называется *асимптотически нормальной оценкой* $\theta$, если $\sqrt{n}(\hat{\theta_n}(X_1, \ldots, X_n) - X) \xrightarrow{d_\theta} \mathcal{N}(0, \Sigma(\theta)) \quad \forall \theta \in \Theta$., где $\Sigma(\theta)$ — *асимптотическая матрица ковариаций*. Если $d = 1$, то $\Sigma(\theta) = \sigma^2(\theta)$ — *асимптотическая дисперсия*.

**Смысл:**

1. *Состоятельность:* при больших $n$ вероятность большого отклонения оценки $\hat{\theta_n}$ от $\theta$ мала, но нет численной характеристики степени отклонения.
2. *Асимпт. нормальность:* дает численную характеристику степени отклонения

Пусть $\hat{\theta_n}$ — а. н. о $\theta$ с а. д $\sigma^2(\theta)$. Тогда при больших $n$ $\quad \hat{\theta_n} \sim_{\text{прибл.}} \mathcal{N}\left(\theta, \frac{\sigma^2(\theta)}{n}\right)$.

3. *Сильная состоятельность* важна тогда, когда данные поступают последовательно.

**Пример:** Пусть $X_1, \ldots, X_n$ – выборка из распределения Лапласа со сдвигом $\theta$.

$$p_\theta(x) = \frac{1}{2} e^{-|x-\theta|}. \quad E_\theta X_1 = \theta, D_\theta X_1 = 2.$$

*УЗБЧ*: $\overline{X} \xrightarrow{P_\theta-\text{п.н.}} \theta \implies \overline{X}$ — (сильно) состоятельная оценка $\theta$.

*ЦПТ:* $\sqrt{n}(\overline{X}-\theta) \xrightarrow{d_\theta} \mathcal{N}(0,2) \implies \overline{X} \sim_{\text{прибл.}} \mathcal{N}(0,\frac{2}{n})$. По свойствам нормального распределения, с вероятностью $> 0.99$:

$$\theta - 3\sqrt{\frac{2}{n}} < \overline{X} < \theta + 3\sqrt{\frac{2}{n}}$$
$$\overline{X} - 3\sqrt{\frac{2}{n}} < \theta < \overline{X} + 3\sqrt{\frac{2}{n}}$$

(*доверительный интервал*).

Пусть $n = 200, \overline{X} = 1$. Тогда неравенство имеет вид

$$0.7 < \theta < 1.3$$

(*реализация доверительного интервала*).

> **Утверждение**:
>
> $$\begin{array}{ccc} \text{Сильная состоятельность} & & \\ & \searrow & \\ & & \text{Состоятельность} \\ & \nearrow & \\ \text{Асимпт. нормальность} & & \end{array}$$
>
> Других следствий нет.

> **Утверждение**: Пусть $X_1, \ldots, X_n$ — выборка, т. ч. $E_\theta |X_1|^{2k} < +\infty$. Тогда $\overline{X^k}$ — несмещенная сильно состоятельная асимптотически нормальная оценка $E_\theta X^k$.

## 2.3 Наследование свойств

**Цель:** получить оценку для $\tau(\theta)$, обладающие некоторым свойством, если имеется оценка для $\psi(\theta)$ с тем же свойством.

> **Теорема** (о наследовании сходимостей): Пусть $\{\xi_n, n \in \mathbb{N}\}, \xi$ — случайные векторы размерности $d$. Тогда:
>
> 1. Если $\xi_n \xrightarrow{P} \xi$ и $h : \mathbb{R}^d \to \mathbb{R}^k$, т. ч. $h$ непрерывна на $B : P(\xi \in B) = 1$. Тогда $h(\xi_n) \xrightarrow{P} h(\xi)$.
> 2. Аналогично для сходимости п. н.
> 3. Если $\xi_n \xrightarrow{d} \xi$ и $h : \mathbb{R}^d \to \mathbb{R}^k$ непрерывна, то $h(\xi_n) \xrightarrow{d} h(\xi)$.

**Пример:** Пусть $\{\xi_n, n \in \mathbb{N}\}$ — н.о.р.с.в., т.ч. $E\xi_1 = a \neq 0$, $D\xi_n$ ограничена.

Из ЗБЧ: $\dfrac{S_n}{n} \xrightarrow{P} a, \quad S_n = \sum \xi_i$. Рассмотрим $h(x) = 1/x$ и применим теорему:

$$h\left(\frac{S_n}{n}\right) = \frac{n}{S_n} \xrightarrow{P} h(a) = \frac{1}{a}.$$

> **Утверждение:** Пусть $\hat{\theta}$ — (сильно) состоятельная оценка $\theta$. Пусть $\tau$ непрерывна на $\Theta$. Тогда $\tau(\hat{\theta})$ — (сильно) состоятельная оценка $\tau(\theta)$.

**Замечание:** Условие непрерывности на $\Theta$ нельзя ослабить.

> **Теорема:** (лемма Слуцкого)
> Пусть $\{\xi_n, n \in \mathbb{N}\},\ \{\eta_n, n \in \mathbb{N}\},\ \xi$ — случайные величины, $C \in \mathbb{R}$. Пусть $\xi_n \xrightarrow{d} \xi,\ \eta_n \xrightarrow{d} C$. Тогда $\xi_n + \eta_n \xrightarrow{d} \xi + C$,
> $\xi_n \cdot \eta_n \xrightarrow{d} \xi C$.

> **Теорема:** (о производной)
> Пусть $\{\xi_n, n \in \mathbb{N}\},\ \xi$ — случайные векторы размерности $d$, т.ч. $\xi_n \xrightarrow{d} \xi, h :$ $\mathbb{R}^d \to \mathbb{R}^k$ непрерывно дифференцируема в точке $a \in \mathbb{R}^d$,
> $\{b_n\} : b_n > 0, b_n \to 0$ — числовая последовательность. Тогда
> $$\frac{h(a + \xi_n b_n) - h(a)}{b_n} \xrightarrow{d} \frac{\partial h}{\partial x}\Big|_a \cdot \xi,$$
> где $\frac{\partial h}{\partial x}\Big|_a$ — матрица Якоби функции $h$ в точке $a$.

**Док-во** $(d = 1)$ **:**

Определим функцию $\quad H(x) = \begin{cases} \dfrac{h(x+a) - h(a)}{x}, & \text{если } x \neq 0 \\ h'(a), & \text{если } x = 0 \end{cases}.$

Функция $H$ непрерывна в нуле. Тогда по лемме Слуцкого $\xi_n b_n \xrightarrow{d} \xi \cdot 0 = 0 \implies$ $\implies \xi_n b_n \xrightarrow{p} 0$. Применим теорему о наследовании сходимостей:

$$H(\xi_n b_n) = \frac{h(\xi_n b_n + a) - h(a)}{\xi_n b_n} \xrightarrow{p} H(0) = h'(a) \implies$$
$$\implies \frac{h(\xi_n b_n + a) - h(a)}{\xi_n b_n} \xrightarrow{d} h'(a).$$

Применим еще раз лемму Слуцкого:

$$\xi_n H(\xi_n b_n) \xrightarrow{d} h'(a)\xi.$$

Следовательно, $\dfrac{h(\xi_n b_n + a) - h(a)}{b_n} \xrightarrow{d} h'(a) \qquad \square.$

**Пример:** Пусть $\{\xi_n, n \in \mathbb{N}\}$ — н.о.р.с.в, т.ч. $E\xi_1 = a \neq 0,\ , D\xi_1 = \sigma^2$.

$\sqrt{n}\left(\dfrac{n}{S_n} - \dfrac{1}{a}\right) \xrightarrow{d} ?$

$\triangle$ ЦПТ: $\sqrt{n}(\frac{S_n}{n} - a) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$.

Воспользуемся теоремой о производной с $\xi_n = \sqrt{n}(\frac{S_n}{n} - a)$, $\xi \sim \mathcal{N}(0, \sigma^2),\ h(x) = \frac{1}{x},\ b_n = \frac{1}{\sqrt{n}}$ :

$$\frac{h(\xi_n b_n + a) - h(a)}{b_n} = \sqrt{n}\left[h\left(a + \left(\frac{S_n}{n} - a\right)\right) - h(a)\right] = \sqrt{n}\left(\frac{n}{S_n} - \frac{1}{a}\right) \xrightarrow{d}$$
$$\xrightarrow{d} \xi \cdot \left(\frac{1}{x}\right)\Bigg|_a = -\xi \cdot \frac{1}{a^2} \sim \mathcal{N}\left(0, \frac{\sigma^2}{a^4}\right) \qquad \square.$$

**Замечание**: Если мы рассмотрим $\xi_n$ как выборку $(X_1, X_2, \dots)$, то $1/\overline{X}$ — а. н. о. для $1/a$ с асимптотической дисперсией $\sigma^2/a^4$.