

Seminar 2: Статистики и оценки, их свойства

Carlos Buitrago

Table of contents

1	Вероятностно-статистическая модель	1
1.1	Параметрическая статистическая модель	1
2	Примеры статистик	2
3	Свойства точечных оценок	4
3.1	Наследование свойств	5
4	Важные результаты	6
4.1	Всегда ли существует несмещённая оценка?	6
4.2	Всегда ли несмещённые оценки полезны?	6
5	Полезные результаты для решения задач	7
5.0.1	Минимум и максимум равномерной выборки	7
5.0.2	Предельное распределение максимума	7

1 Вероятностно-статистическая модель

Пусть X — случайная величина или случайный вектор со значениями в измеримом пространстве $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. Предполагается, что распределение P случайной величины X неизвестно, но известно, что

$$P \in \mathcal{P},$$

где $\mathcal{P}$ — заданное семейство вероятностных распределений. Тройка $(\mathcal{X}, \mathcal{B}(\mathcal{X}), \mathcal{P})$ называется **статистической моделью**.

1.1 Параметрическая статистическая модель

Если семейство распределений можно записать в виде

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\},$$

где Θ — параметрическое пространство, то модель называется **параметрической**.

Параметр $\theta \in \Theta$ считается неизвестным.

Задача статистического оценивания состоит в восстановлении самого параметра θ или некоторой его функции $\tau(\theta)$.

Выборка

Пусть $X_1, \dots, X_n$ — независимые одинаково распределённые случайные величины с распределением P_θ . Вектор

$$X = (X_1, \dots, X_n)$$

называется **выборкой размера n** из распределения P_θ .

Кратко пишут $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} P_\theta$. Если каждый элемент выборки принимает значения в $\mathcal{X}$, то вся выборка принимает значения в $\mathcal{X}^n$.

При истинном значении параметра θ совместное распределение выборки равно

$$P_\theta^n = \underbrace{P_\theta \otimes \dots \otimes P_\theta}_{n \text{ раз}}.$$

Статистика

Пусть $(\mathcal{Y}, \mathcal{E})$ — измеримое пространство. Любое измеримое отображение

$$S : \mathcal{X}^n \rightarrow \mathcal{Y}$$

называется **статистикой**.

Таким образом, $S = S(X_1, \dots, X_n)$ — это любая измеримая функция от выборки, не зависящая от неизвестного параметра.

Оценка

Пусть параметрическая модель задаётся семейством распределений $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, и требуется оценить параметрическую функцию

$$\tau : \Theta \rightarrow \mathcal{T}.$$

Оценкой величины $\tau(\theta)$ называется статистика

$$\hat{\tau}_n = \hat{\tau}_n(X_1, \dots, X_n),$$

принимаяющая значения в пространстве $\mathcal{T}$ значений параметрической функции τ .

Иными словами, оценка является статистикой, значения которой лежат в том же пространстве, что и оцениваемая величина $\tau(\theta)$.

2 Примеры статистик

Пусть $g : \mathcal{X} \rightarrow \mathbb{R}$ — борелевская функция.

Статистика

$$\overline{g(X)} = \frac{1}{n} \sum_{i=1}^n g(X_i)$$

называется **выборочной характеристикой**, соответствующей функции g .

Заметим, что...

Если $\mathbb{E}_\theta |g(X_1)| < \infty$, то по закону больших чисел

$$\overline{g(X)} \xrightarrow{\mathbb{P}_\theta} \mathbb{E}_\theta g(X_1).$$

Выборочное среднее

При $g(x) = x$ получаем **выборочное среднее**

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

Выборочные моменты

Для $k \in \mathbb{N}$ величина

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

называется **выборочным моментом порядка k** .

Соответствующий теоретический момент обозначим $m_k(\theta) = \mathbb{E}_\theta X_1^k$.

Заметим, что...

Если $\mathbb{E}_\theta |X_1|^k < \infty$, то

$$\overline{X^k} \xrightarrow{\mathbb{P}_\theta} m_k(\theta).$$

Если $\mathbb{E}_\theta |X_1|^{2k} < \infty$, то центральная предельная теорема даёт

$$\sqrt{n} \left(\overline{X^k} - \mathbb{E}_\theta X_1^k \right) \xrightarrow{d_\theta} N(0, \text{Var}_\theta(X_1^k)).$$

Кроме того,

$$\mathbb{E}_\theta \overline{X^k} = \mathbb{E}_\theta X_1^k.$$

Следовательно, выборочный момент является одновременно:

- несмещённой оценкой теоретического момента;
- состоятельной оценкой;
- при наличии момента порядка $2k$ — асимптотически нормальной оценкой.

Порядковые статистики

Упорядочим элементы выборки:

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}.$$

Статистики $X_{(1)}, \dots, X_{(n)}$ называются **порядковыми статистиками**, а вся совокупность называется **вариационным рядом**.

В частности,

$$X_{(1)} = \min_{1 \leq i \leq n} X_i,$$

$$X_{(n)} = \max_{1 \leq i \leq n} X_i.$$

Заметим, что...

$$F_{X_{(n)}}(x) = F(x)^n.$$

$$F_{X_{(1)}}(x) = 1 - (1 - F(x))^n.$$

Выборочная дисперсия

Статистика

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

называется **выборочной дисперсией**.

Используя тождество $\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$, получаем

$$s_n^2 = \overline{X^2} - \bar{X}^2.$$

3 Свойства точечных оценок

Несмещенность

Оценка $\hat{\tau}_n$ называется **несмещённой оценкой** параметрической функции $\tau(\theta)$, если

$$\mathbb{E}_\theta \hat{\tau}_n = \tau(\theta) \quad \forall \theta \in \Theta.$$

В частности, $\hat{\theta}_n$ является несмещённой оценкой параметра θ , если $\mathbb{E}_\theta \hat{\theta}_n = \theta$.

Смещением оценки $\hat{\tau}_n$ называется величина

$$\text{Bias}_\theta(\hat{\tau}_n) = \mathbb{E}_\theta \hat{\tau}_n - \tau(\theta).$$

Состоятельность

Последовательность оценок $\hat{\tau}_n$ называется **состоятельной оценкой** $\tau(\theta)$, если для любого $\theta \in \Theta$

$$\hat{\tau}_n \xrightarrow{\mathbb{P}_\theta} \tau(\theta), \quad n \rightarrow \infty.$$

Сильно состоятельная оценка

Последовательность оценок $\hat{\tau}_n$ называется **сильно состоятельной**, если для любого $\theta \in \Theta$

$$\hat{\tau}_n \xrightarrow{\text{п.н.}} \tau(\theta).$$

Асимптотическая нормальность

Оценка $\hat{\tau}_n$ называется **асимптотически нормальной** оценкой $\tau(\theta)$ с асимптотической дисперсией $\sigma^2(\theta)$, если

$$\sqrt{n} (\hat{\tau}_n - \tau(\theta)) \xrightarrow{d_\theta} N(0, \sigma^2(\theta)).$$

Функция $\sigma^2(\theta)$ называется **асимптотической дисперсией** оценки.

3.1 Наследование свойств

Наследование состоятельности

Пусть

$$\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta$$

и функция τ непрерывна в точке θ . Тогда по теореме о непрерывном отображении

$$\tau(\hat{\theta}_n) \xrightarrow{\mathbb{P}} \tau(\theta).$$

Наследование сильной состоятельности

Если

$$\hat{\theta}_n \xrightarrow{\text{п.н.}} \theta$$

и τ непрерывна в точке θ , то

$$\tau(\hat{\theta}_n) \xrightarrow{\text{п.н.}} \tau(\theta).$$

Наследование асимптотической нормальности

Пусть

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta)),$$

а функция τ дифференцируема в точке θ . Тогда по дельта-методу

$$\sqrt{n}(\tau(\hat{\theta}_n) - \tau(\theta)) \xrightarrow{d} N(0, \sigma^2(\theta)[\tau'(\theta)]^2).$$

4 Важные результаты

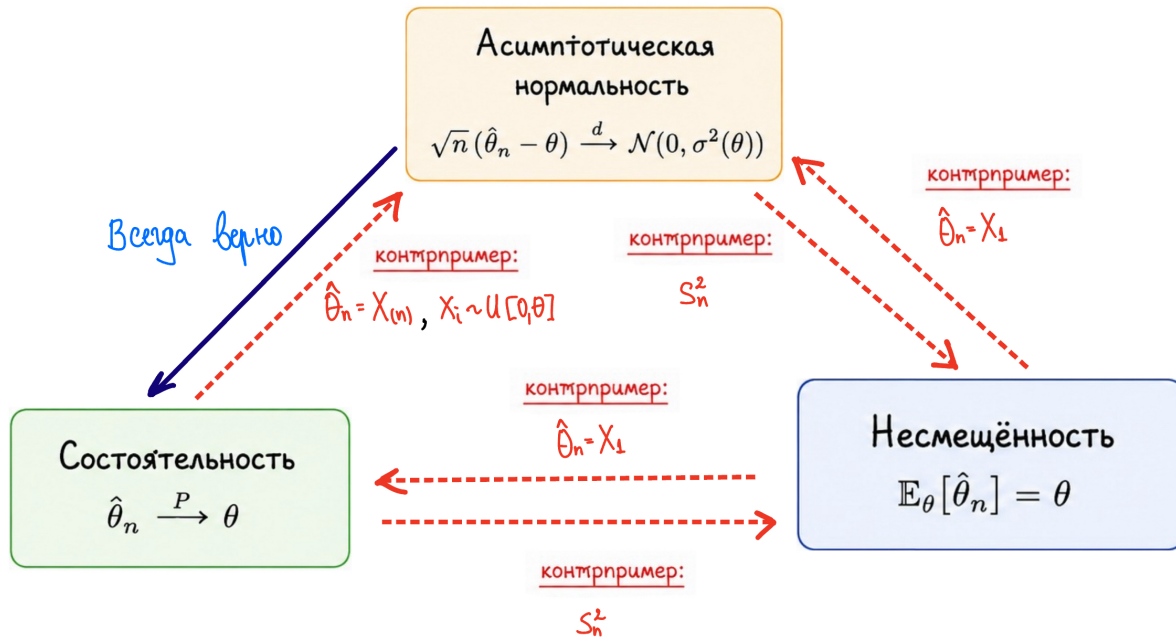


Figure 1: Связь между свойствами статистических оценок

4.1 Всегда ли существует несмещённая оценка?

Несмещённость является естественным и желательным свойством оценки. Однако **несмещённая оценка заданной параметрической функции существует не всегда**.

Рассмотрим простой пример. Пусть

$$X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(\theta), \quad \theta \in (0, 1),$$

и предположим, что требуется оценить параметрическую функцию

$$\tau(\theta) = \frac{1}{\theta}.$$

Можно показать, что несмещённой оценки $\tau(\theta)$ не существует.

4.2 Всегда ли несмещённые оценки полезны?

Даже если несмещённая оценка существует, это ещё не означает, что она обладает хорошими статистическими свойствами.

Рассмотрим следующий пример. Пусть имеется всего одно наблюдение X_1 с распределением

$$\mathbb{P}_\theta(X_1 = k) = \theta^k(1 - \theta), \quad k \in \mathbb{Z}_+, \quad \theta \in (0, 1).$$

Несмещённая оценка имеет вид

$$\hat{\theta} = \mathbf{1}_{\{X_1 \geq 1\}}.$$

Действительно,

$$\mathbb{E}_\theta \hat{\theta} = \mathbb{P}_\theta(X_1 \geq 1) = 1 - \mathbb{P}_\theta(X_1 = 0) = \theta.$$

Однако несмещённость сама по себе ещё ничего не говорит о точности оценки.

5 Полезные результаты для решения задач

Равномерное распределение

Пусть $X \sim U(0, \theta)$.

Тогда для любого $k > -1$

$$\mathbb{E}X^k = \frac{\theta^k}{k+1}.$$

5.0.1 Минимум и максимум равномерной выборки

Пусть $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} U(0, \theta)$.

Для $0 \leq x \leq \theta$, выполнено $\mathbb{P}_\theta(X_{(n)} \leq x) = \left(\frac{x}{\theta}\right)^n$.

Следовательно,

$$\frac{X_{(n)}}{\theta} \sim \text{Beta}(n, 1).$$

Для минимума выполнено $\mathbb{P}_\theta(X_{(1)} > x) = \left(1 - \frac{x}{\theta}\right)^n$.

Следовательно,

$$\frac{X_{(1)}}{\theta} \sim \text{Beta}(1, n).$$

Из свойств бета-распределения следуют формулы

$$\mathbb{E}_\theta X_{(1)} = \frac{\theta}{n+1},$$

$$\mathbb{E}_\theta X_{(n)} = \frac{n}{n+1}\theta.$$

5.0.2 Предельное распределение максимума

Рассмотрим

$$Y_n = \frac{n}{\theta} (\theta - X_{(n)}).$$

Тогда для $x \geq 0$

$$\mathbb{P}(Y_n > x) = \mathbb{P}\left(X_{(n)} < \theta \left(1 - \frac{x}{n}\right)\right) = \left(1 - \frac{x}{n}\right)^n \rightarrow e^{-x}.$$

Следовательно,

$$\frac{n}{\theta} (\theta - X_{(n)}) \xrightarrow{d} \text{Exp}(1).$$

Аналогично для минимума

$$\frac{n}{\theta} X_{(1)} \xrightarrow{d} \text{Exp}(1).$$

Экспоненциальное распределение

Пусть $X \sim \text{Exp}(\theta)$ в параметризации по интенсивности:

$$f_{\theta}(x) = \theta e^{-\theta x}, \quad x \geq 0.$$

Тогда для любого $k \in \mathbb{N}$

$$\mathbb{E} X^k = \frac{k!}{\theta^k}.$$