

Seminar 2 in Applied Statistics

Carlos Buitrago

Table of contents

0.1	а)	1
0.2	б)	1
0.3	в)	1
0.4	г)	1
0.5	1.	2
0.6	2.	2
0.7	3.	2
0.8	4.	3

💡 Задача 1

Пусть $X_1, \dots, X_n$ — независимая случайная выборка из распределения с плотностью

$$f_X(x) = \frac{2\theta^2}{x^3} \mathbf{1}_{[\theta, \infty)}(x), \quad \theta \geq 1.$$

0.1 а)

Найдите **оценку максимального правдоподобия (ОМП)** параметра θ .

0.2 б)

Найдите распределение ОМП-оценки и, используя его, постройте точный доверительный интервал уровня $100(1 - \alpha)\%$ для параметра θ .

0.3 в)

Найдите оценку θ методом моментов, используя соотношения

$$\mathbb{E}[\ln X_1] = \ln \theta + \frac{1}{2},$$

$$\mathbb{E}[(\ln X_1)^2] = (\ln \theta)^2 + \ln \theta + \frac{1}{2}.$$

Используя полученную оценку, постройте асимптотический доверительный интервал для θ .

0.4 г)

С помощью статистического моделирования постройте 95%-ные бутстрэп-доверительные интервалы для θ , используя:

- ОМП-оценку из пункта **б)**;
- оценку методом моментов из пункта **в)**.

Сравните полученные бутстрэп-интервалы:

1. с точным доверительным интервалом из пункта 6);
2. с асимптотическим доверительным интервалом, основанным на $\hat{\theta}_{\text{ММ}}$.

Сравните ширины полученных доверительных интервалов.

С помощью многократного моделирования оцените **фактический уровень покрытия** каждого интервала и сравните его с заявленным уровнем 95%.

💡 Задача 2

Пусть $X_1, \dots, X_n$ — независимая случайная выборка из равномерного распределения

$$X_i \sim U(0, \theta),$$

где $\theta > 0$ — неизвестный параметр.

0.5 1.

Рассмотрим статистику

$$T_n = n(X_{(n)} - \theta),$$

где

$$X_{(n)} = \max\{X_1, \dots, X_n\}$$

— максимальный элемент выборки.

Покажите, что функция распределения T_n имеет вид

$$F_{T_n}(t) = \left(1 + \frac{t}{n\theta}\right)^n, \quad t \in [-n\theta, 0].$$

Найдите предельное распределение T_n при $n \rightarrow \infty$.

Является ли $X_{(n)}$ состоятельной оценкой параметра θ ?

0.6 2.

Предложите доверительные интервалы для θ уровня $1 - \alpha$, используя:

1. точное распределение, найденное в пункте 1;
2. предельное распределение, найденное в пункте 1.

Сравните полученные доверительные интервалы.

0.7 3.

Пусть $X_1^*, \dots, X_n^*$ — непараметрическая бутстрэп-выборка, полученная выбором с возвращением из исходной выборки

$$X_1, \dots, X_n.$$

Обозначим её максимум через

$$X_{(n)}^* = \max\{X_1^*, \dots, X_n^*\}.$$

Рассмотрим бутстрэп-аналог статистики T_n :

$$T_n^* = n(X_{(n)}^* - X_{(n)}).$$

Покажите, что

$$\mathbb{P}^*(T_n^* = 0) = 1 - \left(1 - \frac{1}{n}\right)^n.$$

Найдите предел этой вероятности при $n \rightarrow \infty$.

0.8 4.

Покажите, что условное бутстрэп-распределение статистики T_n^* **не сходится** к предельному распределению статистики T_n при $n \rightarrow \infty$.

Объясните причину несостоятельности обычного непараметрического бутстрэпа в данной задаче.

Что можно сказать о доверительном бутстрэп-интервале для θ , построенном **методом Холла** на основе распределения T_n^* ?