

Дедлайн: 2022-12-04, 21:00. Если вычислительные возможности позволяют, берите  $n_{boot} = 10000$  бутстрэп-выборок. Если нет, то сокращайте в соответствии с вычислительными возможностями.

1. Рассмотрим наивный бутстрэп для исходной выборки из  $n$  наблюдений непрерывного распределения. Обозначим исходные наблюдения с помощью  $y_i$ , а бутстрэпированные — с помощью  $y_i^*$ .

- а) Какова вероятность того, что пятое исходное наблюдение ровно три раза попадет в бутстрэп выборку?
- б) Чему равен предел найденной в (а) вероятности при  $n \rightarrow \infty$ ?
- в) Чему равен предел вероятности  $\mathbb{P}(\max y_i > \max y_i^*)$  при  $n \rightarrow \infty$ ?

2. Величины  $y_i$  независимы  $\mathcal{N}(\mu; 1)$ .

- а) Без компьютера явно найдите оценку  $\mu$  методом максимального правдоподобия, если а-приори известно, что  $\mu \geq 42$ .

Винни-Пух использует наивный бутстрэп для построения 95%-го интервала для  $\mu$ .

- б) С помощью симуляций оцените фактическую вероятность накрытия, если истинное  $\mu = 42$ .
- в) С помощью симуляций оцените фактическую вероятность накрытия, если истинное  $\mu = 44$ .

3. Величины  $y_i$  независимы  $\mathcal{N}(0; 1)$ . Кролик не знает истинное ожидание, и для надежности использует очень хитрый способ оценки  $\mu$ . В качестве  $\hat{\mu}$  Кролик выбирает середину самого короткого отрезка, содержащего 50% выборки. Для построения 95%-го интервала для  $\mu$  Кролик использует наивный бутстрэп.

С помощью симуляций оцените фактическую вероятность накрытия для  $n = 100$ .

Следующие вопросы относятся к набору данных [https://github.com/bdemeshev/psmo\\_2022-23/raw/main/ha\\_02/psmo-ha\\_02.csv](https://github.com/bdemeshev/psmo_2022-23/raw/main/ha_02/psmo-ha_02.csv).

4. Постройте 95%-й доверительный интервал для медианы величины  $y$ , используя наивный бутстрэп, бутстрэп  $t$ -статистики и ВСА-бутстрэп.
5. Рассмотрим модель линейной регрессии

$$\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i + \hat{\beta}_3 x_i^2.$$

- а) С помощью наивного бутстрэпа постройте 95%-й доверительный интервал для каждого коэффициента.
- б) С помощью наивного бутстрэпа постройте 95%-й доверительный интервал для вершины параболы.

6. Рассмотрим модель квантильной регрессии<sup>1</sup> для квантилей  $\tau = 0.1, \tau = 0.5, \tau = 0.9$ :

$$\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i + \hat{\beta}_3 x_i^2.$$

- а) Постройте три полученные линии регрессии на одном графике с данными.
- б) Для медианной регрессии с помощью наивного бутстрэпа постройте 95%-й доверительный интервал для каждого коэффициента.
- в) Для медианной регрессии с помощью наивного бутстрэпа постройте 95%-й доверительный интервал для вершины параболы.

<sup>1</sup> Целевая функция  $\sum_{i=1}^n \rho_\tau(y_i - \hat{y}_i)$ , где  $\rho_\tau(u) = u \cdot (\tau - I(u < 0))$ .