ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»

Университет ИТМО

Отчёт по лабораторной работе № 3

«Бутстрап-оценки»

Выполнили работу:

Кащеев Максим Николаевич,

гр. J3111, ИСУ 466147

Косарев Илья Андреевич,

группа J3110, ИСУ 466304

Капустина Юлия Ильинична,

группа J3110, ИСУ 466110

Санкт-Петербург 2025

1. Цели и задачи

**Цель работы:** познакомиться с методом бутстрапа и применить его на тестовой выборке.

**Задачи:**

* Сгенерировать выборку по заранее известному распределению и рассчитать точечные оценки (среднее, медиану, дисперсию, интерквартильный размах) на выборке и генеральной выборке.
* Реализовать алгоритм бутстрапа для оценок.
* Построить доверительные интервалы для оценки среднего и медианы.
* Выявить влияние размера выборки и итераций алгоритма бутстрапа на ширину доверительных интервалов.
* Проверить точность построения доверительного интервала эмпирически.
* Визуализировать данные и сделать выводы.

2. Теоретическая часть

Бутстрап — это статистический метод, который используется для оценки распределения интересующей статистики (например, среднего, медианы, стандартной ошибки) на основе одной выборки данных. Этот метод основывается на идее повторного случайного (с заменой) извлечения подвыборок из исходной выборки и вычислении статистики для каждой из этих подвыборок.

Идея бутстрапа состоит в том, что исходная выборка считается приближённым представлением генеральной совокупности. Поскольку собрать дополнительные данные может быть затруднительно или дорого, бутстрап позволяет имитировать процесс повторного эксперимента. Именно это помогает оценить неопределённость, стандартную ошибку и построить доверительный интервал для интересующей статистики, без необходимости прибегать к строгим предположениям о виде распределения данных.

KDE (Kernel Density Estimation), или ядерная оценка плотности, — это непараметрический метод оценки функции плотности вероятности случайной величины на основе конечной выборки данных.

Доверительный интервал — это диапазон значений, вычисленный на основе выборки данных, который, с заданной степенью уверенности (например, 95%), предполагается охватывать истинное значение параметра генеральной совокупности (например, среднего, медианы, пропорции и т.д.). То есть, если бы мы повторяли эксперимент много раз, то выбранная процедура построения доверительного интервала приводила бы к тому, что, например, 95% полученных интервалов содержали бы истинное значение параметра.

3. Практическая часть

В качестве непрерывного распределения было взято нормальное распределение с центром в 0 и дисперсионным коэффициентом 5. На его основе была взята выборка из 500 исходов и на ее основе рассчитаны параметры:

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 1: точечные оценки и их сравнение с генеральной выборкой

После этого было построено примерное распределение выборки:

Изображение выглядит как текст, График, линия, диаграмма

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 2: распределение выборки в виде гистограммы и его сравнение с KDE

Далее был реализован алгоритм бутстрапа (множественное взятие подвыборок с повторным включением элементов) и вычислены распределения основных точечных оценок, один из примеров:

Изображение выглядит как текст, снимок экрана, линия, диаграмма

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 3: распределение оценки среднего по бутстрапу и сравнение с истинными значениями

Далее зададим доверительный интервал для произвольной подвыборки какой-либо оценки (на примере 95%): . Вычислим его для данного набора результатов бутстрапа:

Изображение выглядит как текст, снимок экрана, диаграмма, линия

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 4: распределение оценок среднего по подвыборкам бутстрапа с обозначенными доверительными интервалами разной уверенности

После этого были взяты несколько значений размеров и построены соответствующие выборки, после чего на основе бутстрапа вычислены доверительные интервалы и вычислены их длины:

Изображение выглядит как текст, линия, График, диаграмма

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 5: зависимость ширины доверительного интервала от размера выборки

Как можно заметить, при увеличении размера выборки ширина интервала также уменьшается, что свидетельствует о прямой корреляции. Скорее всего, результаты статистики оказываются менее подверженными случайным колебаниям, поэтому распределение оценки сужается вокруг истинного значения, что сужает доверительный интервал.

Также был построен график зависимости от кол-ва итераций в бутстрапе. В силу, однако, случайности метода, явной закономерности выявить не удалось:

Изображение выглядит как текст, линия, диаграмма, График

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 6: зависимость ширины интервала от кол-ва итераций бутстрапа

Можно сказать, что при достаточно большом B ширина интервала будет сходится к некоторому определенному значению, так как данные при многократном выборе будут давать приближенные к центру оценки; эмпирическое распределение становится всё более стабильно и приближается к «истинному» распределению оценок, которые могли бы получиться при бесконечном числе повторений.

Наконец, было проведено эмпирическое вычисление истинного уровня доверенности на случайных выборках в зависимости от N и B. В качестве распределения было взято нормальное стандартное:

Изображение выглядит как текст, снимок экрана, диаграмма, Красочность

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 7: тепловая карта истинных значений доверенности

Как видно, хоть и наблюдаются выбросы, большинство значений доверенности действительно близко к теоретически заданному (0.95).

4. Выводы

В ходе выполнения лабораторной работы мы познакомились с понятием бутстрап-метода и применили его на практике. Мы также познакомились с вспомогательными инструментами оценки распределений метрик бутстрап подвыборок, реализовали их и выявили закономерности между ними и параметрами задания изначальной выборки и итерационного алгоритма.

Практические задачи позволили удостовериться в универсальности алгоритма и его способности оценивать выборки неопределенного распределения.