

# Домашнее Задание 5

Дедлайн – 5 февраля

## Задание 1 (Python).

Создайте две переменные  $X_1 \sim \mathcal{N}(3, 1)$  (1 – дисперсия) и  $X_2 \sim \text{Poisson}(4)$  с 1000 наблюдений для каждой переменной. Предположим, что  $Y$ ,  $X_1$  и  $X_2$  связаны следующим уравнением для генеральной совокупности:

$$Y = 5 + 3X_1 - 2X_2 + u$$

Предположения Гаусса-Маркова верны. Ответьте на следующие вопросы:

I. Сгенерируйте 1000 случайных выборок переменной  $Y$  в соответствии с уравнением выше. Ошибки  $u$  следуют нормальному распределению со средним 0 и дисперсией 4. Можно использовать генератор случайных чисел из библиотеки `numpy`.

II. Для каждой выборки оцените МНК-регрессию, используя  $X_1$  и  $X_2$  в качестве независимых переменных. Используйте OLS из библиотеки `statsmodels`, примеры кода есть в ноутбуке из семинара `Quasi-experiments`.

III. Для каждой выборки сохраните получившиеся параметры модели (интерсепт плюс два коэффициента наклона). Для каждого параметра модели посчитайте 95-процентные доверительные интервалы и проверьте, включают ли (покрывают ли) данные интервалы параметры генеральной совокупности.

IV. Посчитайте средние арифметические по каждому из параметров. Близки ли получившиеся значения параметрам модели в генеральной совокупности? Какое свойство МНК-оценок иллюстрируют данные результаты?

V. Посчитайте по каждому параметру модели процент доверительных интервалов, покрывших реальные параметры модели в генеральной совокупности. Вы должны получить числа, близкие к 95 %.

VI. Посчитайте дисперсию по каждому из параметров модели (дисперсию 1000 получившихся значений по каждому из параметров модели). Как будет меняться данная дисперсия с ростом числа наблюдений? К какому значению она будет асимптотически (количество наблюдений  $\rightarrow \infty$ ) приближаться? О каком свойстве МНК-оценок здесь идёт речь?

## Задание 2 (Python).

Выполните задание C11 из учебника Вулдриджа, Глава 4. В дополнение к заданию, представьте результаты регрессий из пунктов (i), (iii) и (iv) в профессиональной

регрессионной таблице (библиотека stargazer Вам поможет).

**Задание 3.**

Как можно обосновать предположение Гаусса-Маркова о нормальном распределении ошибок?