

Задача №5.

В этом учебном году вам не довелось писать контрольную работу. Решив задачу №5, вы рассчитаете оценку, которая у вас могла бы быть.

Файл «Данные к задаче 5.ods» содержит сведения о студентах, обучавшихся на втором курсе Программной Инженерии в 2015-2016 учебном году:

x — итоговая оценка за курс «Теория вероятностей и математическая статистика»;

d — пол студента (0 — юноша, 1 — девушка);

y — оценка за контрольную работу по нашему курсу.

К названиям переменных добавлен номер варианта (**x8, d8, y8** — данные восьмого варианта и т. п.). Каждый вариант данных содержит наблюдения за 65 студентами.

Оцениваемые модели.

«Длинная» регрессия: $y_i = \beta_1 + \beta_2 x_i + \beta_3 d_i + \epsilon_i$.

«Короткая» регрессия: $y_i = \gamma_1 + \gamma_2 x_i + v_i$.

Пункты задания.

1. Оцените длинную и короткую регрессии по наблюдениям 1-50 вашего варианта. Выпишите оценённые уравнения, дайте интерпретацию полученным коэффициентам.

2. Рассчитайте значения R^2 для обеих моделей.

3. По каждой модели рассчитайте прогноз оценки за контрольную работу для наблюдений 51–65

вашего варианта $\hat{y}_{51}, \dots, \hat{y}_{65}$ и среднюю абсолютную ошибку прогноза $MAE = \frac{1}{15} \sum_{i=51}^{65} |y_i - \hat{y}_i|$

(MAE — Mean Absolute Error).

4. Выберите модель с меньшей ошибкой и с её помощью спрогнозируйте собственную оценку за контрольную работу по курсу «Прикладные методы математической статистики».

Требования к оформлению.

Отчёт должен быть представлен в текстовом документе (.doc, .docx, .pdf) и представлять собой связный текст.

Срок сдачи: до 11 июня включительно.