

## (1) Задача<sup>1</sup>

Имеется статистика расходов на питание ( $Y$ ) и общих расходов ( $X$ ) 55 индийских семей, проживающих в сельской местности (начало 2000 годов). График разброса и выдержка из корреляционной матрицы показателей изображены на Рис. 1.

1. Видна ли на графике линейная взаимосвязь между расходами на питание и общими расходами? Подтверждается ли взаимосвязь коэффициентом корреляции?
2. Построена модель линейной регрессии  $Y = \beta_0 + \beta_1 \cdot X$ . Проанализируйте отчёт Gretl №1. Запишите модель в явном виде. Проверьте значимость параметров. Сколько процентов разброса зависимой переменной объясняет модель? На сколько в среднем отклоняются модельные значения от фактических?

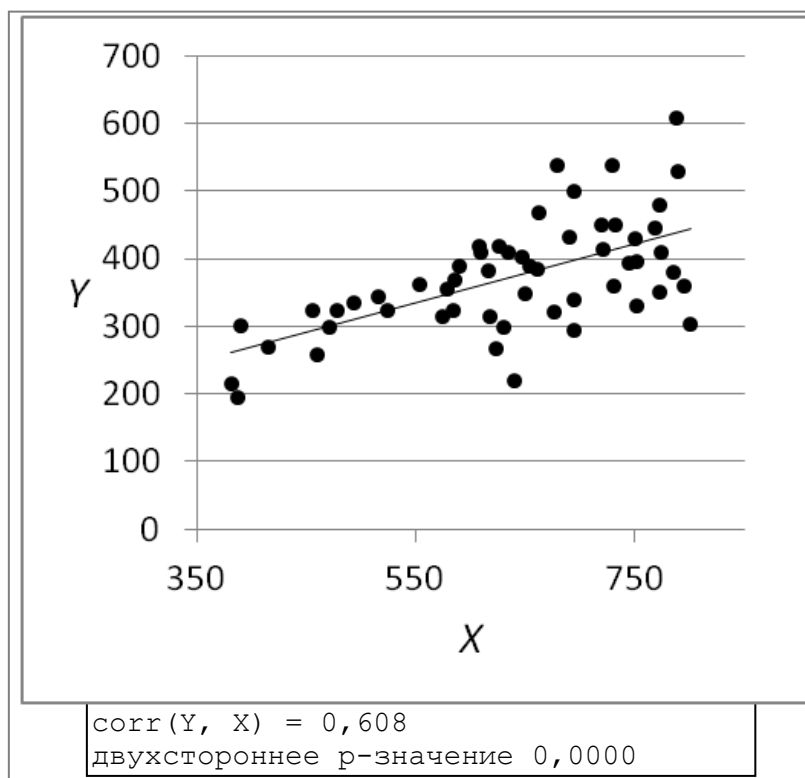


Рис. 1. График разброса  $Y$  и  $X$

### Отчёт 1. Отчёт Gretl по модели парной линейной регрессии

Модель 1: МНК, использованы наблюдения 1–55

Зависимая переменная:  $Y$

	Коэффициент	Ст. ошибка	t-статистика	P-значение
-----	-----	-----	-----	-----
const	94,2088	50,8563	1,852	0,0695
X	0,436809	0,0783226	5,577	8,45e-07
Среднее зав. перемен	373,3455	Ст. откл. зав. перемен	83,43510	
Сумма кв. остатков	236893,6	Ст. ошибка модели	66,85575	
R-квадрат	0,369824	Испр. R-квадрат	0,357934	
F(1, 53)	31,10345	P-значение (F)	8,45e-07	
Лог. правдоподобие	-308,1625	Крит. Акаике	620,3251	

<sup>1</sup> Gujarati: Basic Econometrics, Fourth Edition. The McGraw-Hill Companies, 2004.

## (2) Задача 3.3<sup>2</sup>

Предположим, что у нас есть набор данных с пятью предикторами:  $X_1$  – средний балл за школьные предметы (GPA),  $X_2$  – коэффициент IQ,  $X_3$  – пол (1 для женщин и 0 для мужчин),  $X_4$  – взаимодействие между GPA и IQ, и  $X_5$  – взаимодействие между GPA и полом. Зависимой переменной является начальная заработная плата после окончания университета (тыс. долларов). Представьте, что мы используем метод наименьших квадратов для подгонки модели и получаем  $\hat{\beta}_0 = 50$ ,  $\hat{\beta}_1 = 20$ ,  $\hat{\beta}_2 = 0.07$ ,  $\hat{\beta}_3 = 35$ ,  $\hat{\beta}_4 = 0.01$  и  $\hat{\beta}_5 = -10$ .

- а) Какой из приведённых ниже ответов правильный и почему?
1. При фиксированных значениях IQ и GPA мужчины в среднем зарабатывают больше женщин.
  2. При фиксированных значениях IQ и GPA женщины в среднем зарабатывают больше мужчин.
  3. При фиксированных значениях IQ и GPA мужчины в среднем зарабатывают больше женщин при условии, что GPA достаточно высок.
  4. При фиксированных значениях IQ и GPA женщины в среднем зарабатывают больше мужчин при условии, что GPA достаточно высок.
- б) Дайте прогнозную оценку заработной платы женщины с IQ = 111 и GPA = 4.0.
- в) Верно ли следующее утверждение: поскольку коэффициент для взаимодействия между GPA и IQ очень небольшой, то у нас есть мало оснований сделать вывод о наличии эффекта взаимодействия. Объясните свой ответ.

---

<sup>2</sup> Задачи с нумерацией X.Y – из книги «Введение в статистическое обучение с примерами на языке R». X – номер главы, Y – номер задания из раздела теоретических упражнений.