

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение

высшего образования

**«Дальневосточный федеральный университет»**

**(ДВФУ)**

|  |  |
| --- | --- |
|  | **ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ** |

**Лабораторная работа №9**

**Линейная регрессия**

Дисциплина «Теория вероятностей и математическая статистика»

Студент группы Б9123-01.03.02ии

Моттуева Уруйдана Михайловна

г. Владивосток

2025

Целью данной лабораторной работы является построение модели линейной регрессии.

**Уравнение регрессии**

Имеется признаков и зависящей от них целевой признак .

Уравнением регрессии на называется уравнение

где — случайный остаток(ошибка).

**Линейная регрессия**

Если условное математическое ожидание является линейной функцией:

То *уравнение линейной регрессии* выглядит как:

Модель определяется параметрами которые оцениваются с помощью выборки при условии

Имеется выборка , из . Тогда, подставляя значения в уравнение линейной регрессии, получаем

Оценкой является

Оптимальные значения , минимизирующие сумму квадратов отклонений, находятся как:

Оценки параметров являются несмещёнными, состоятельными и с наименьшей дисперсией при соблюдении условий *теоремы Гаусса-Маркова.*

**Теорема Гаусса-Маркова** утверждает, что оценка МНК (метода наименьших квадратов) является наилучшей линейной несмещенной оценкой (BLUE), если:

;

Данное условие проверяется с помощью одновыборочного t-теста для гипотезы:

;

.

;

Для проверки автокорреляции использовался тест Дарбина–Уотсона, который позволяет выявить наличие линейной автокорреляции первого порядка. Интерпретация статистики:

— автокорреляции нет,

— есть положительная автокорреляция,

— есть отрицательная автокорреляция.

3. Ранг матрицы X равен числу факторов: ;

4. Ошибки имеют постоянную дисперсию:

Для её проверки использовался тест Бреуша–Пагана:

дисперсия остатков постоянна

: дисперсия остатков изменяется

5. Математическое ожидание ошибок равно нулю: .

Примечание: нормальность остатков **не требуется** для теоремы Гаусса-Маркова, но необходима для проверки значимости коэффициентов (t-тесты, F-тесты).

**Реализация**

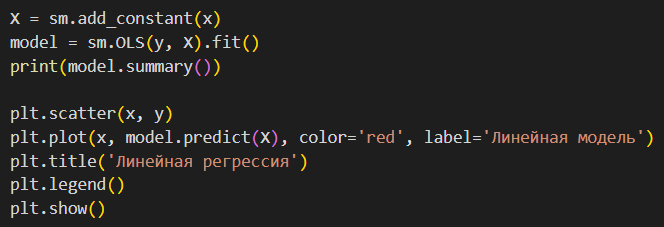
Линейная модель:

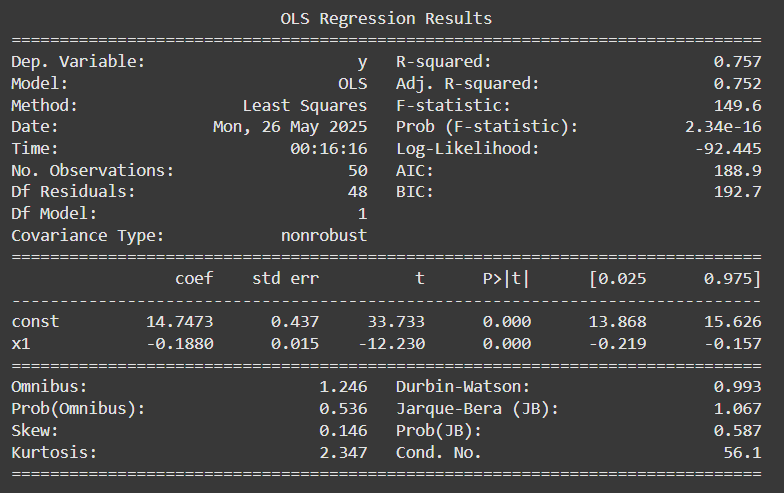
1. Построить диаграмму рассеяния

Изображение выглядит как текст, снимок экрана, диаграмма, линия

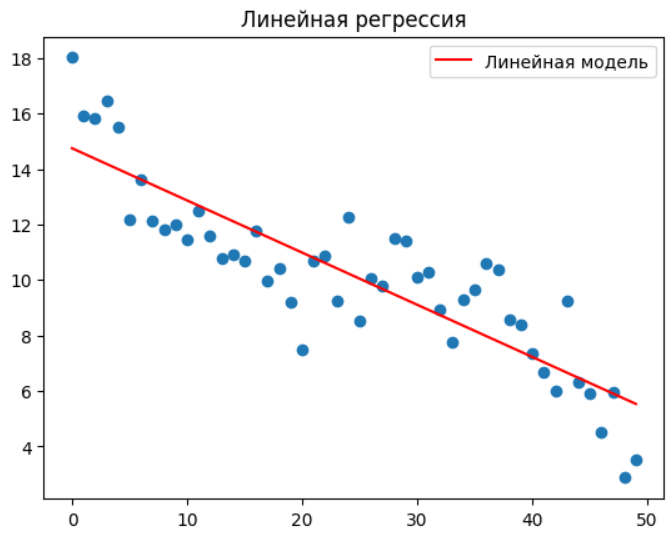
Контент, сгенерированный ИИ, может содержать ошибки.

1. Построить модель линейной регрессии с помощью statsmodels.api.OLS взяв в качестве условного матожидания линейную функцию.

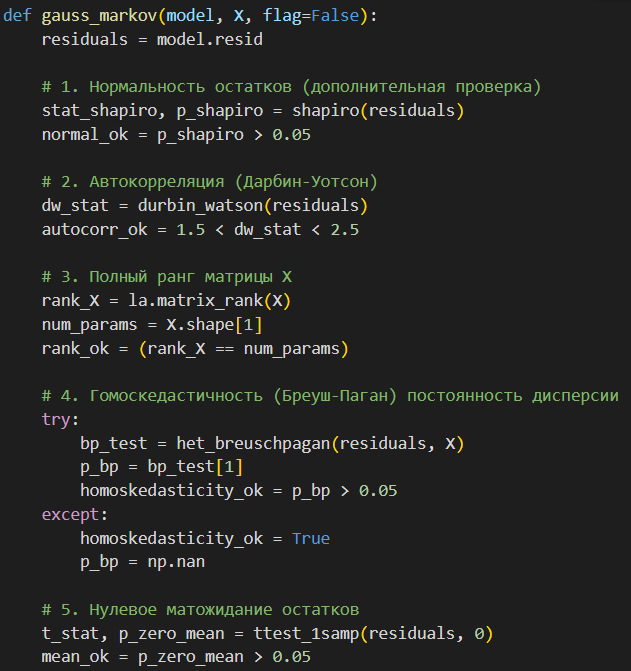




, коэффициенты значимы

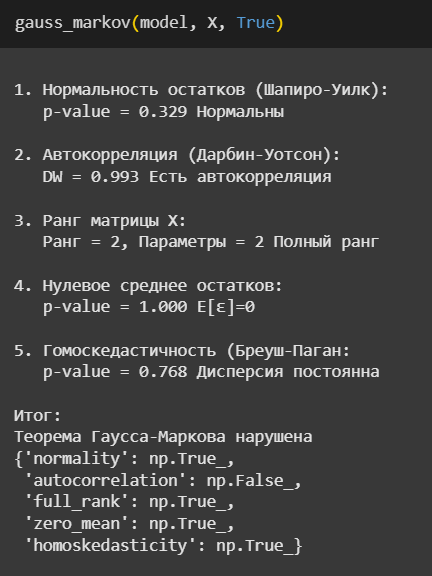


1. Оценить значимость коэффициентов. Проверить положения теоремы Гаусса-Маркова.



Изображение выглядит как текст, снимок экрана, Шрифт, меню

Контент, сгенерированный ИИ, может содержать ошибки.



Квадратичная модель:

Изображение выглядит как текст, снимок экрана, Шрифт

Контент, сгенерированный ИИ, может содержать ошибки.

Изображение выглядит как текст, снимок экрана, число

Контент, сгенерированный ИИ, может содержать ошибки.

8, не все коэффициенты значимы

Изображение выглядит как линия, текст, График, диаграмма

Контент, сгенерированный ИИ, может содержать ошибки.

