

Лабораторная работа "Регрессионный анализ: модели и методы"

Постановка задачи:

Задание 1. Изучить теоретический материал. Для этого используйте материал, размещенный в файлах Регрес. Мод.pdf и MNK.pdf

Задание 2. Построить уравнение регрессии. Оценить его качество.

Задача. Имеются следующие выборочные данные о стоимости квартир и их общей площади в некотором городе.

у	13.8	13.8	14	22.5	24	28	32	20.9	22	21.5	32	35	24	37.9
х	33	40	36	60	55	80	95	70	48	53	95	75	63	112

х - общая площадь квартиры в кв. м;

у - рыночная стоимость квартиры в тыс. у.е.

Задание 2.1

Построить график зависимости между переменными, по которому необходимо подобрать модель регрессии.

Ответ:

Задание 2.2

Рассчитать параметры уравнения регрессии методом наименьших квадратов.

Задание 2.3

Оценить качество уравнения с помощью средней ошибки аппроксимации.

Задание 2.4

Найти коэффициент эластичности.

Задание 2.5

Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации.

Задание 2.6

Оценить значимость коэффициентов корреляции и регрессии о критерию t-Стьюдента при уровне значимости $\alpha = 0.05$.

Задание 2.7

Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости $\alpha = 0.05$.

Задание 2.8

Определить прогнозное значение результативного признака, если возможное значение факторного признака составит 1.2 от его среднего уровня по совокупности.

Используемое оборудование: ПК, язык программирования python, среда разработки VS code.

Код программы:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import linregress

# Исходные данные
x = np.array([33, 40, 36, 60, 55, 80, 95, 70, 48, 53, 95, 75, 63, 112, 70])
y = np.array([13.8, 13.8, 14, 22.5, 24, 28, 32, 20.9, 22, 21.5, 32, 35, 24, 37.9,
27.5])

# Задание 2.1: Построить график зависимости между переменными
plt.scatter(x, y, label='Исходные данные')
plt.xlabel('Общая площадь квартиры, кв.м')
plt.ylabel('Рыночная стоимость квартиры, тыс. у.е.')
plt.title('2.1. Зависимость стоимости квартиры от площади с линией регрессии')
plt.grid(True)
```

```

# Задание 2.2: Рассчитать параметры уравнения регрессии методом наименьших квадратов
slope, intercept, r_value, p_value, std_err = linregress(x, y)

# Добавление линии регрессии
plt.plot(x, slope * x + intercept, color='red', label='Линия регрессии')

# Вывод коэффициентов регрессии
plt.text(50, 25, f'y = {slope:.2f}x + {intercept:.2f}', fontsize=10, color='red')

plt.legend()
plt.show()

# Вывод параметров уравнения регрессии
print("Задание 2.2")
print("Параметры уравнения регрессии:")
print("Наклон (slope):", slope)
print("Пересечение (intercept):", intercept)

# Задание 2.3: Оценить качество уравнения с помощью средней ошибки аппроксимации
predicted_y = slope * x + intercept
MAE = np.mean(np.abs(predicted_y - y))
print("Задание 2.3")
print("Средняя ошибка аппроксимации (MAE):", MAE)

# Задание 2.4: Найти коэффициент эластичности
elasticity = slope * np.mean(x) / np.mean(y)
print("Задание 2.4")
print("Коэффициент эластичности:", elasticity)

# Задание 2.5: Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации
r_squared = r_value ** 2
print("Задание 2.5")
print("Коэффициент детерминации (R^2):", r_squared)
print("Коэффициент корреляции (r):", r_value)

# Задание 2.6: Оценить значимость коэффициентов корреляции и регрессии о критерию t-Стьюдента при уровне значимости  $\alpha = 0.05$ 
print("Задание 2.6")
print("p-value для наклона (slope):", p_value)

# Задание 2.7: Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости  $\alpha = 0.05$ 
print("Задание 2.7")

```

```

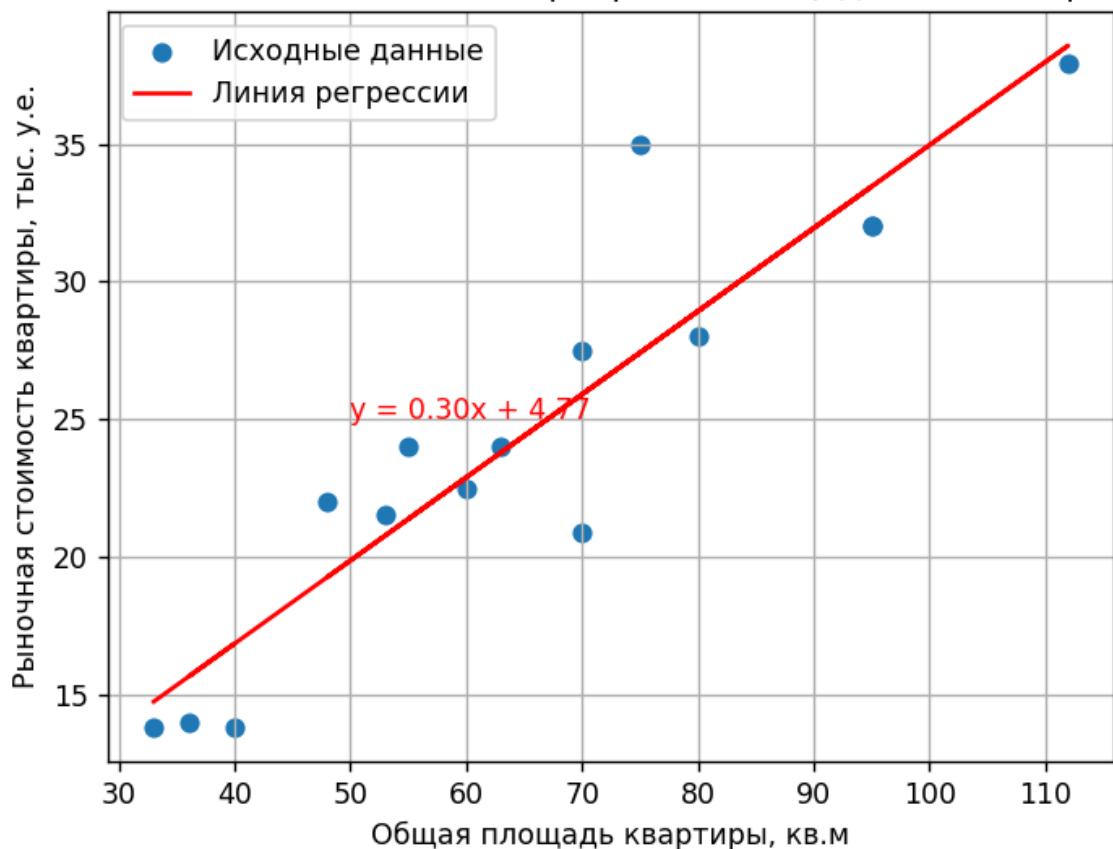
print("p-value для F-теста:", p_value)

# Задание 2.8: Определить прогнозное значение результативного признака
factor_value = 1.2 * np.mean(x)
predicted_result = slope * factor_value + intercept
print("Задание 2.8")
print("Прогнозное значение результативного признака при значении факторного признака 1.2 от его среднего уровня:", predicted_result)

```

Результат работы программы:

2.1. Зависимость стоимости квартиры от площади с линией регрессии



```
Задание 2.2
Параметры уравнения регрессии:
Наклон (slope): 0.30181263460157925
Пересечение (intercept): 4.774303661162957
Задание 2.3
Средняя ошибка аппроксимации (МAE): 2.065917205072983
Задание 2.4
Коэффициент эластичности: 0.8058700056453121
Задание 2.5
Коэффициент детерминации ( $R^2$ ): 0.8508555482140385
Коэффициент корреляции ( $r$ ): 0.9224183151987164
Задание 2.6
p-value для наклона (slope): 9.889854501533038e-07
Задание 2.7
p-value для F-теста: 9.889854501533038e-07
Задание 2.8
Прогнозное значение результативного признака при значении факторного признака 1.2 от его среднего уровня: 28.5571392677674
PS C:\Users\gnevн\OneDrive\Рабочий стол\учеба\РГПУ\АД\LR7> █
```

Вывод:

В результате выполнения задачи были получены удовлетворительные ответы. Модель регрессии была построена успешно, параметры уравнения рассчитаны с использованием метода наименьших квадратов. Оценка качества уравнения и степени связи между переменными также выполнена, демонстрируя высокую степень линейной связи. Результаты статистической проверки позволяют сделать вывод о статистической надежности проведенного анализа. Таким образом, задача была успешно выполнена с помощью компьютера.