

Практическое занятие №3

Тема: Нелинейная регрессия. Спецификация модели

Цель работы:

Освоить методы построения и анализа нелинейных регрессионных моделей, научиться выбирать оптимальную функциональную форму зависимости, оценивать и интерпретировать параметры нелинейных моделей.

Алгоритм выполнения работы:

1. Сначала выполните расчеты в Python для получения точных результатов.
2. Затем повторите ключевые расчеты в Excel для углубленного понимания.
3. Сравните результаты из обеих сред.
4. Сделайте выводы о качестве моделей и интерпретируйте результаты.

Правило выбора варианта:

- **Вариант 1:** если сумма двух последних цифр пароля $\in [0, 6]$
- **Вариант 2:** если сумма двух последних цифр пароля $\in [7, 13]$
- **Вариант 3:** если сумма двух последних цифр пароля $\in [14, 18]$

Вариант 1: Анализ производственной функции

Исходные данные:

Зависимость объема выпуска продукции (Y , тыс. ед.) от количества работников (L , чел.) и капиталовложений (K , тыс. руб.) для 20 предприятий:

Предприятие	Y	L	K
1	120	10	50
2	180	15	80
3	250	20	120
4	310	25	150
5	380	30	200
6	440	35	240
7	500	40	280
8	560	45	320
9	620	50	360
10	680	55	400
11	740	60	440
12	800	65	480
13	860	70	520

14	920	75	560
15	980	80	600
16	1040	85	640
17	1100	90	680
18	1160	95	720
19	1220	100	760
20	1280	105	800

Вариант 2: Анализ себестоимости продукции

Исходные данные:

Зависимость себестоимости единицы продукции (Y, руб.) от накопленного объема производства (X, тыс. ед.) для 18 периодов:

Период	Y	X
1	120	1
2	115	2
3	110	3
4	105	5
5	100	8
6	95	12
7	90	18
8	85	25
9	80	35
10	75	50
11	70	70
12	65	95
13	60	130
14	55	180
15	50	250
16	45	350
17	40	500
18	35	700

Вариант 3: Анализ кривой Филлипса

Исходные данные:

Зависимость между уровнем безработицы (X, %) и инфляцией (Y, %) для 16 стран:

Страна	Y	X
1	8.5	3.0
2	7.2	3.5
3	6.1	4.0
4	5.3	4.5
5	4.8	5.0
6	4.5	5.5

7	4.3	6.0
8	4.2	6.5
9	4.1	7.0
10	4.0	7.5
11	3.9	8.0
12	3.8	8.5
13	3.7	9.0
14	3.6	9.5
15	3.5	10.0
16	3.4	10.5

Краткие теоретические сведения

1. Понятие нелинейной регрессии

Нелинейная регрессия используется для моделирования зависимостей, которые не могут быть адекватно описаны линейной функцией. Различают два типа нелинейности:

- **Нелинейность по переменным, линейность по параметрам:** Модели, которые могут быть линеаризованы с помощью преобразований переменных
- **Нелинейность по параметрам:** Модели, которые не могут быть линеаризованы (требуют нелинейного МНК)

2. Основные типы линеаризуемых моделей

2.1. Степенная модель (модель Кобба-Дугласа)

- Форма: $Y = aX^b$
- Линеаризация: $\ln Y = \ln a + b \ln X$
- Интерпретация: Коэффициент b представляет эластичность Y по X

2.2. Экспоненциальная модель

- Форма: $Y = ae^{bX}$
- Линеаризация: $\ln Y = \ln a + bX$
- Интерпретация: Коэффициент b показывает **относительное изменение** Y при абсолютном изменении X

2.3. Логарифмическая модель

- Форма: $Y = a + b \ln X$

- **Линеаризация:** Не требует преобразования (уже линейна по параметрам)
- **Интерпретация:** Коэффициент b показывает **абсолютное изменение** Y при относительном изменении X

2.4. Гиперболическая модель

- **Форма:** $Y = a + \frac{b}{x}$
- **Линеаризация:** не требует преобразования
- **Интерпретация:** используется для моделирования убывающей с насыщением зависимости

3. Критерии выбора функциональной формы

- **Теоретическая обоснованность:** Соответствие экономической теории
- **Статистические критерии:**
 - Коэффициент детерминации R^2 и скорректированный R^2
 - Информационные критерии (AIC, BIC)
 - Стандартная ошибка регрессии
- **Графический анализ:** Визуальное соответствие линии регрессии данным
- **Анализ остатков:** Проверка на гомоскедастичность и отсутствие автокорреляции

4. Особенности интерпретации коэффициентов

В нелинейных моделях интерпретация коэффициентов отличается от линейного случая:

- **Эластичность:** Процентное изменение Y при изменении X на 1%
- **Полуэластичность:** Процентное изменение Y при изменении X на единицу
- **Предельный эффект:** Изменение Y при малом изменении X (рассчитывается как производная)

5. Производственная функция Кобба-Дугласа

Специальный случай степенной модели для анализа производства:

$$Y = AL^{\alpha}K^{\beta},$$

где:

- Y - объем выпуска
- L - труд
- K - капитал
- α, β - эластичности выпуска по труду и капиталу
- A - коэффициент общей производительности

Методические указания

ЧАСТЬ 1: ВЫПОЛНЕНИЕ В PYTHON

```

"""
ЛАБОРАТОРНАЯ РАБОТА №3: НЕЛИНЕЙНАЯ РЕГРЕССИЯ
Этот код сравнивает различные функциональные формы регрессии
"""

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import statsmodels.api as sm

# БЛОК 1: Подготовка данных
# Пояснение: Замените данные на свои в соответствии с вариантом
X = np.array([1, 2, 3, 5, 8, 12, 18, 25, 35, 50, 70, 95, 130, 180, 250, 350, 500, 700])
Y = np.array([120, 115, 110, 105, 100, 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 45, 40, 35])

df = pd.DataFrame({'X': X, 'Y': Y})

# БЛОК 2: Сравнение различных моделей
# Пояснение: Оцениваем параметры разных функциональных форм

models_results = {}

# Линейная модель
X_linear = sm.add_constant(df['X'])
model_linear = sm.OLS(df['Y'], X_linear).fit()
models_results['Линейная'] = {
    'r2': model_linear.rsquared,
    'r2_adj': model_linear.rsquared_adj,
    'aic': model_linear.aic
}

```

```

}

# Логарифмическая модель
df['lnX'] = np.log(df['X'])
X_log = sm.add_constant(df['lnX'])
model_log = sm.OLS(df['Y'], X_log).fit()
models_results['Логарифмическая'] = {
    'r2': model_log.rsquared,
    'r2_adj': model_log.rsquared_adj,
    'aic': model_log.aic
}

# Степенная модель (логарифмируем обе переменные)
df['lnY'] = np.log(df['Y'])
model_power = sm.OLS(df['lnY'], X_log).fit()
models_results['Степенная'] = {
    'r2': model_power.rsquared,
    'r2_adj': model_power.rsquared_adj,
    'aic': model_power.aic
}

# БЛОК 3: Сравнительный анализ моделей
print("=== СПРАВНЕНИЕ МОДЕЛЕЙ ===")
comparison_df = pd.DataFrame(models_results).T
print(comparison_df.round(4))

# БЛОК 4: Визуализация моделей
plt.figure(figsize=(12, 8))

plt.subplot(2, 2, 1)
plt.scatter(df['X'], df['Y'], alpha=0.7)
plt.plot(df['X'], model_linear.predict(X_linear), 'r-',
label='Линейная')
plt.title('Линейная модель')
plt.legend()

plt.subplot(2, 2, 2)
plt.scatter(df['X'], df['Y'], alpha=0.7)
plt.plot(df['X'], model_log.predict(X_log), 'g-',
label='Логарифмическая')
plt.title('Логарифмическая модель')
plt.legend()

plt.subplot(2, 2, 3)
plt.scatter(df['X'], df['Y'], alpha=0.7)
plt.plot(df['X'], np.exp(model_power.predict(X_log)), 'b-',
label='Степенная')
plt.title('Степенная модель')
plt.legend()

plt.subplot(2, 2, 4)
plt.scatter(df['X'], df['Y'], alpha=0.7, label='Данные')

```

```

plt.plot(df['X'], model_linear.predict(X_linear), 'r-',
label='Линейная')
plt.plot(df['X'], model_log.predict(X_log), 'g-',
label='Логарифмическая')
plt.plot(df['X'], np.exp(model_power.predict(X_log)), 'b-',
label='Степенная')
plt.title('Все модели')
plt.legend()

plt.tight_layout()
plt.show()

# БЛОК 5: Интерпретация коэффициентов
print("\n=== ИНТЕРПРЕТАЦИЯ КОЭФФИЦИЕНТОВ ===")
print("Линейная модель:")
print(f"Y = {model_linear.params[0]:.2f} + {model_linear.params[1]:.4f} * X")

print("Логарифмическая модель:")
print(f"Y = {model_log.params[0]:.2f} + {model_log.params[1]:.4f} * ln(X)")

print("Степенная модель:")
print(f"ln(Y) = {model_power.params[0]:.4f} + {model_power.params[1]:.4f} * ln(X)")
print(f"Эластичность: {model_power.params[1]:.4f}")

```

ЧАСТЬ 2: ВЫПОЛНЕНИЕ В EXCEL

Шаг 1: Подготовка преобразованных переменных

A (X)	B (Y)	C (lnX)	D (lnY)	E (X²)	F (Предсказания)

Шаг 2: Построение различных моделей

Линейная модель:

G1: =НАКЛОН(B:B,A:A) # Коэффициент b1

G2: =ОТСЕЧЕННАЯ(B:B,A:A) # Константа b0

F2: =\$G\$2+\$G\$1*A2 # Предсказания

Логарифмическая модель:

H1: =НАКЛОН(B:B,C:C) # Коэффициент

H2: =ОТСЕЧЕННАЯ(B:B,C:C) # Константа

Степенная модель:

I1: =НАКЛОН(D:D,C:C) # Коэффициент (эластичность)

I2: =ОТСЕЧЕННАЯ(D:D,C:C) # Константа

Шаг 3: Расчет показателей качества

Для каждой модели рассчитываем:

J1: =КВПИРСОН(B:B,A:A) # R² для линейной

J2: =КВПИРСОН(B:B,C:C) # R² для логарифмической

J3: =КВПИРСОН(D:D,C:C) # R² для степенной

ЧАСТЬ 3: СРАВНЕНИЕ РЕЗУЛЬТАТОВ

Модель	Python R ²	Excel R ²	Лучшая модель
Линейная			
Логарифмическая			
Степенная			

Требования к отчету:

1. **Титульный лист** (название работы, вариант, ФИО, группа)
2. **Постановка задачи** (цель анализа, описание данных)
3. **Исходные данные** (таблица)
4. **Результаты расчетов:**
 - Основные статистики.
 - Таблица сравнения результатов, полученных с применением Python и Excel.

- Графики.

5. **Выводы** (интерпретация результатов, какая модель оказалась лучше и почему?)

Критерии оценки для получения зачета:

Зачет:

- полностью выполнены все расчеты, построены все графики, дана грамотная интерпретация, ответы на контрольные вопросы
- выполнены основные расчеты, есть незначительные ошибки в интерпретации или оформлении графиков
- выполнены не все расчеты, имеются ошибки в методике анализа, возможно потребуется доработка

Незачет: работа не выполнена или выполнена неверно