МИНИCTEPCTBO НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

**Кафедра прикладной информатики**

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ № 7

ДИСЦИПЛИНЫ «Информационно-коммуникационные технологии»

НА ТЕМУ:

**Регрессионный анализ**

**Выполнил:**

студент группы ПИН-б-о-21-1

Стороежнко Артем Владимирович

Проверил: ассистент кафедры прикладной информатики

Мартыновская А.С.

Ставрополь, 2023

**Цель работы:** научиться строить регрессионные модели, проверять их адекватность и значимость их коэффициентов.

**Ход выполнения работы**

1. Сначала, необходимо записать исходные данные (значения X и значения Y) в текстовые файлы, и затем извлечем из них значения с помощью команды READPRN. Так как нам дан один столбец значений X и 5 столбцов со значениями Y, рассчитаем средние значения для каждого столбца Y и сформируем из них новый массив (рис. 1).

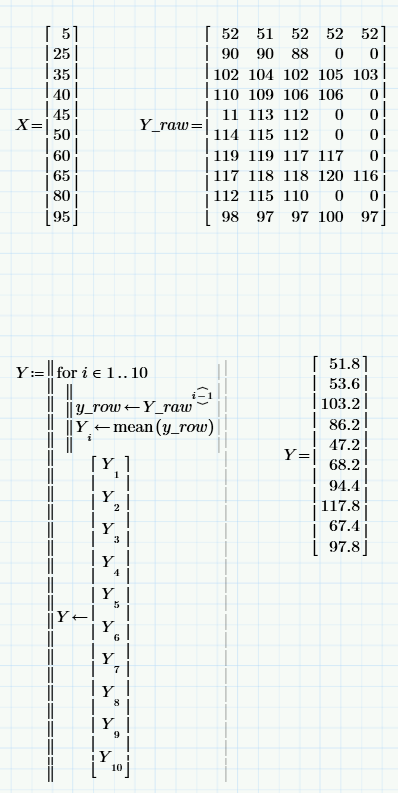


Рисунок 1 – Извлечение исходных данных из текстовых файлов

1. Далее, с помощью функций slope(X, Y) и intercept(X, Y) найдем линейное уравнение регрессии. Функция slope(X, Y) возвращает значение коэффициента X, а intercept – коэффициент свободного члена. Результат выполнения обоих функций показан на рисунке 2



Рисунок 2 – Нахождение линейного уравнения регрессии

Исходя из результатов функции получаем линейное уравнение регрессии: Y = 0.435X + 59.024

1. Далее, найдем квадратичное и кубическое уравнения для исходных данных. Это действие можно выполнить с помощью функции regress(X, Y, n), где n – степень уравнения. Результат регрессионного анализа показан на рисунке 3.

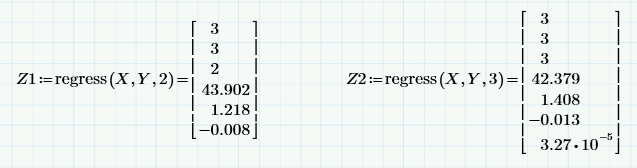


Рисунок 3 – Поиск квадратичного и кубического уравнения для исходных данных

На основе результатов функций regress() построим квадратичное и кубическое уравнения: y = -0.008x^2 + 1.218x + 43.902; y = -0.0000327x^3 – 0.0013x^2 + 1.408x + 45.355.

1. Теперь, найдем экспоненциальное уравнение регрессии. Для этого воспользуемся функцией genfit(X, Y,vb,F), где X – вектор значений X, Y – вектор значений Y, vb – начальное приближение, F – приближающая функция. Функция genfit возвращает коэффициенты регрессионой модели, которые затем можно подставить в приближающую функцию для получения экспоненциального уравнения регрессии. Используя полученную экспоненциальную функцию, построим график, на котором также отметим экспериментальные данные. Процесс поиска экспоненциального уравнения регресии и построенный график представлены на рисунке 4.

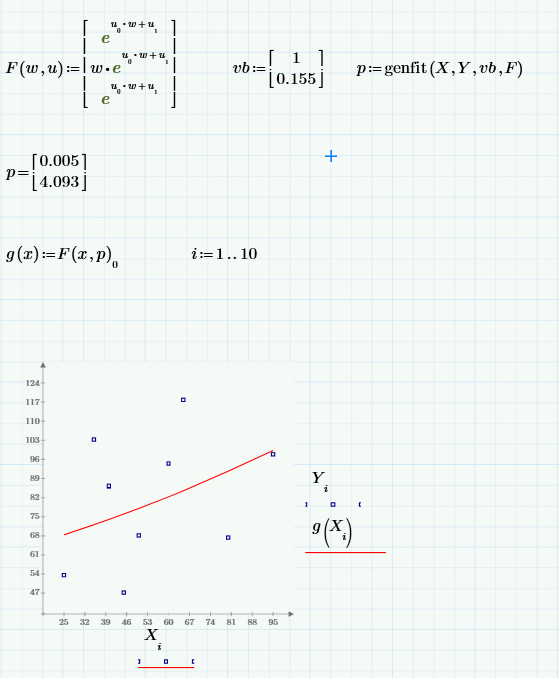


Рисунок 4 – Поиск экспоненциального уравнения регрессии

1. Построим график для остатков экспоненциальной регрессионной модели (рис. 5).

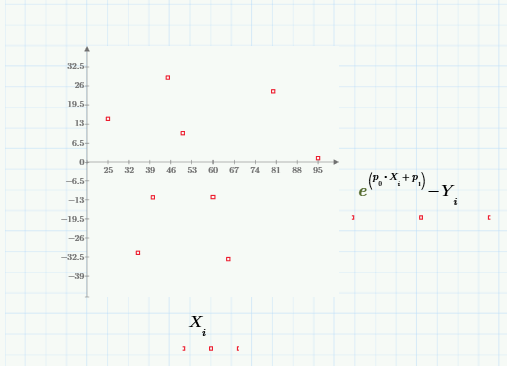


Рисунок 5 – График остатков экспоненциальной регрессионной модели

Теперь, создадим программу в Mathcad для построения полноценных линейной и квадратичной регрессионных моделей с учетом весов, проверки адекватности построенных моделей и значимости всех коэффициентов.

1. Получим из файлов значения X и Y и вычислим среднее значение для каждой строки массива Y (рис. 6).

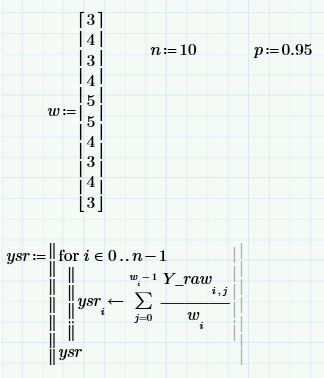


Рисунок 6 – Получение данных и рассчет среднего арифметического для каждой строки Y

1. Создадим программу для построения линейной регрессионной модели с учетом весов. Для этого зададим вектор весов w, затем посчитаем среднее каждой строки вектора значений Y. Далее, напишем формулы для получения коэффициентов a и b. Для этого создадим две системы уравнений – первую – для коэффициентов x и вторую – для свободных членов уравнения. Системы уравнений решим с помощью метода lsolve(a, b). Процесс создания линейной регрессионной модели показан на рисунке 7.

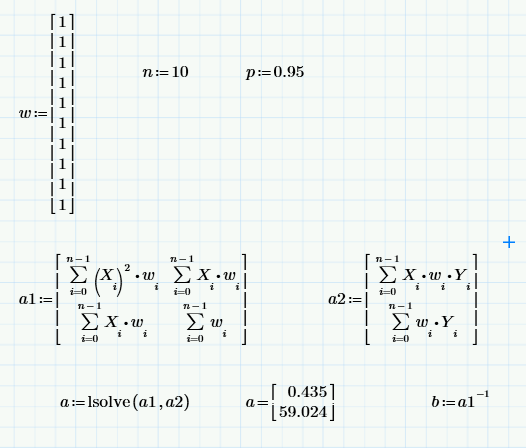


Рисунок 7 – Программа для создания линейной регрессионной модели

Полученные коэффициенты совпадают с таковыми, полученными при использовании функций slope и intercept. Полученное уравнение линейной регрессии для исходных данных: y = 0.435x + 59.024.

1. Создадим программу для построения квадратичной модели регрессии. Суть программы будет той же – мы решим системы уравнений для коэффициентов x и свободных членов с помощью метода lsolve(). Программа для построения квадратичной модели регрессии представлена на рисунке 8.

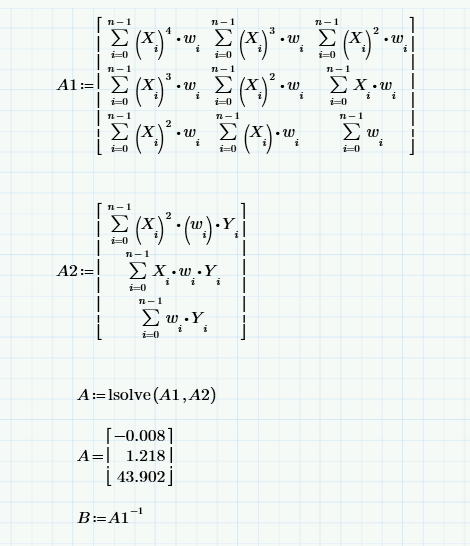


Рисунок 8 – Программа для пострения квадратичной модели регрессии

Коэффициенты уравнения совпадают с таковыми, полученными при помощи метода regress(). Полученное уравнение квадратичной регрессии выглядит следующим образом: y = -0.008x^2 + 1.218x + 43.902.

1. Теперь, напишем программу, проводящую проверку полученных моделей на адекватность. Код программы показан на рисунках 9 и 10.

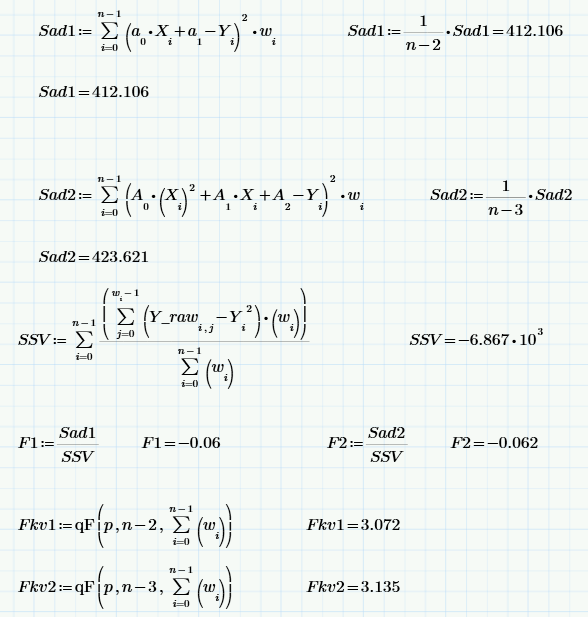


Рисунок 9 – Проверка полученных регрессионных моделей на адекватность, часть 1

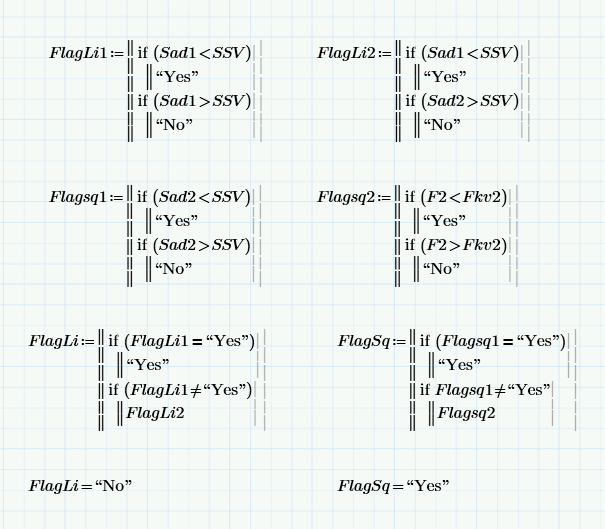


Рисунок 10 – Проверка полученных регрессионных моделей на адекватность, часть 2

Результат проверки показал, что линейная модель регрессии является адекватной, а квадратичная – нет.

1. Проверим значимость коэффициентов обеих моделей. Для этого нам необходимо вычислить стандартные ошибки для всех коэффициентов и рассчитать отношение коэффициентов к ошибкам. Для проверки значимости коэффициентов рассчитывается отношение дисперсии адекватности к свободной дисперсии (Fэ) и квантиль распределения Фишера (Fт = F(p, l – 1, l - n)). Гипотеза принимается, если Fт > Fэ и отвергается в ином случае. Программа для определения значимости коэффициентов представлена на рисунках 11, 12, 13.

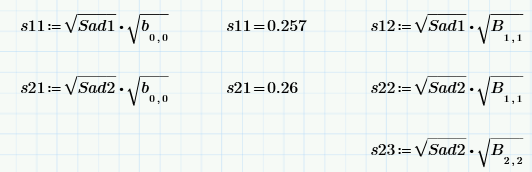


Рисунок 11 – Расчет стандартных ошибок коэффициентов



Рисунок 12 - Расчет отношения коэффициентов к стандартным ошибкам

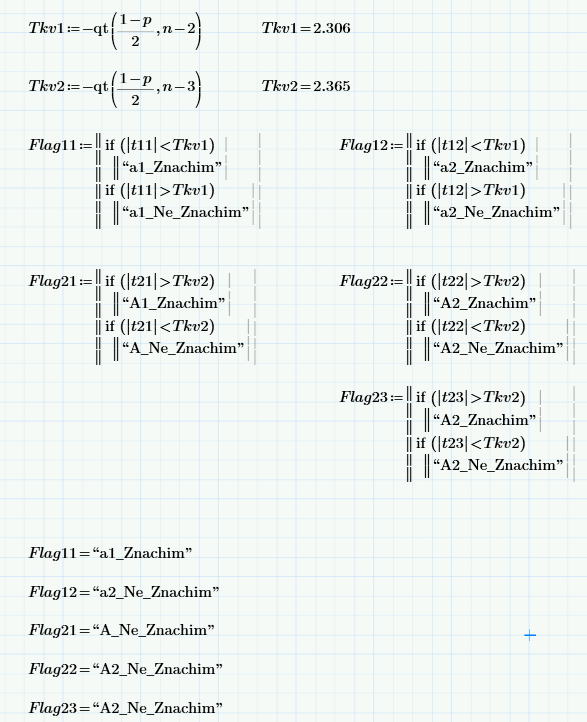


Рисунок 13 – Проверка значимости коэффициентов регрессионных моделей

Результат провреки показал, что значимым является только первый коэффициент.

**Вывод:** выполнив задания лабораторной работы я научился использовать программные средства Mathcad для построения регрессионных моделей, проверки их на адекватность и проверки значимости их коэффициентов.