**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ**

федеральное государственное автономное образовательное учреждение высшего образования

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ**

**ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

ОТЧЕТ

по лабораторной работе

***Регрессионный анализ парных наблюдений***

по курсу

«Математическая статистика»

Выполнил(а): Дзебан А.А. (0В21)

Проверил:

Доцент каф. ВМиМФ: Шинкеев М.Л.

Томск 2017

***Цель работы:***

Подбор и оценка параметров уравнения среднеквадратичной регрессии.

**Задание.**

**Требуется:**

1. Отобразить корреляционное поле наблюдаемых значений величин  и  (построить диаграмму рассеяния).
2. Методом наименьших квадратов найти оценки параметров уравнения линейной , квадратичной , логарифмической  и показательной  регрессии  на . Изобразить полученные зависимости на фоне корреляционного поля.
3. Получить оценки остаточной дисперсии и коэффициента детерминации, а также проверить значимость коэффициентов уравнения регрессии и значимость модели в целом (используя достигнутый уровень значимости) для каждой модели, в предположении, что остатки независимые нормальные случайные величины с одинаковой дисперсией.
4. Сравнить модели, используя значения оценок остаточной дисперсии и коэффициента детерминации, а также полученные (достигнутые) уровни значимости коэффициентов моделей и выбрать одну из них (обосновать выбор).
5. Для выбранной модели определить границы доверительных интервалов для значений линии регрессии (отобразить графически) в предположении, что остатки независимые нормальные случайные величины с одинаковой дисперсией.
6. Проверить качественно свойства остатков модели (нормальность, гомоскедастичность), построив гистограмму остатков и диаграмму рассеяния остатков в зависимости от значений переменной .

**Исходные данные:**

**Ход работы:**

Строим диаграммы рассеяния величин X и Y с нанесенными на них линиями соответствующих регрессий:

Замечаем, что визуально и по значениям квадрата коэффициента детерминации наиболее предпочтительными являются квадратичная и линейниая модель регрессии. Исследуем подробно обе модели регрессии.

***Результаты регрессионного анализа для линейной модели :***

Вектор базисных функций:  , вектор коэффициентов уравнения регрессии: .

Оценку вектора коэффициентов уравнения регрессии метода наименьщих квадратов находим по формуле: , где , :

Таким образом, оценка функции регрессии:

Оценка остаточной дисперсии модели регрессии:

Оценка коэффициента детерминации модели и скорр. оценка коэффициента детерминации:

Значение статистики Фишера:

Уровень значимости:

Следовательно нулевая гипотеза отвергается, значит использование модели статистически обоснованно.

Для линейной модели  значимость коэффициента детерминации равносильна значимости коэффициента уравнения регрессии , поэтому проверка значимости данного коэффициента уравнения регрессии на основе значения статистики Стьюдента избыточна.

***Результаты регрессионного анализа для квадратичной модели***

**:**

Вектор базисных функций:  , вектор коэффициентов уравнения регрессии: .

Оценку вектора коэффициентов уравнения регрессии метода наименьщих квадратов находим по формуле: , где , :

Таким образом, функция регрессии:

Оценка остаточной дисперсии:

Оценка коэффициента детерминации модели и скорр. оценка коэффициента детерминации:

Значение статистики Фишера:

Данное значение статистики является высокозначимым, соответственно R^2 значимо отличается от нуля.

Проверяем значимость коэффициентов регрессионной модели. Для коэффициета  значение статистики Стьюдента:

Уровень значимости определяем как вероятность события , где  случайная величина, имеющая распределение Стьюдента с  степенями свободы:

*.*

Следовательно, можно признать, что значение  не значимо отличается от нуля.

Оценим коэффициент модели :

Соответствующий уровень значимости:

Следовательно, можно признать, что значение  значимо отличается от нуля.

Таким образом, мы получили для квадратичнй модели, что коэффициент  не значим, а коэффициент  значим. Это означает, что мы должны изменить квадратичную модель, исключив из нее соответствующее не значимое слагаемое, то есть, рассмотреть модель вида:  и заново оценить ее коэффициенты.

***Результаты регрессионного анализа для квадратичной модели :***

Вектор базисных функций:  , вектор коэффициентов уравнения регрессии: .

Оценку вектора коэффициентов уравнения регрессии метода наименьщих квадратов находим по формуле: :

Таким образом, оценка функции регрессии:

Оценка коэффициента детерминации модели и скорр. оценка коэффициента детерминации:

Значение статистики Фишера:

Следовательно, данное значение статистики является высокозначимым, то есть коэффициент детерминации модели значимо отличается от нуля.

Для модели  значимость коэффициента детерминации равносильна значимости коэффициента уравнения регрессии , поэтому проверка значимости данного коэффициента уравнения регрессии на основе значения статистики Стьюдента избыточна.