****

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования

**«Дальневосточный федеральный университет»**

(ДВФУ)

|  |
| --- |
| **ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ**  **Департамент информационных и компьютерных систем** |

**ОТЧЕТ**

по лабораторной работе №2

**по дисциплине «Эконометрика»**

**на тему «Множественная регрессия и корреляция»**

вариант 2

**направление подготовки**

**09.03.03 Прикладная информатика**

**Прикладная информатика в экономике**

|  |  |
| --- | --- |
| Выполнил: | студент  гр. Б9121-09.03.03пиэ |
|  | Туровец В. Ю. |
| Проверил: | старший преподаватель |
|  | Шувалова Е. И. |
|  | |
| (оценка) | |

г. Владивосток

2024

**1. Условие задачи**

Цель: получить навыки построения и оценивания модели множественной регрессии.

Используя данные о стоимости автомобиля, производителе, объеме двигателя, мощности двигателя и пробеге:

1. Постройте уравнение множественной регрессии в линейной форме с полным набором факторов.
2. Установить какие факторы коллинеарны и удалить зависимые факторы
3. Постройте уравнение регрессии со статистически значимыми факторами.
4. Оцените статистическую значимость уравнений регрессии с помощью критерия Фишера.

**2. Ход работы**

Изучается зависимость средней ожидаемой продолжительности жизни от нескольких факторов по данным за 1995 г. Исходные данные представлены в таблице 1.

Таблица 1 – Исходные данные

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Страна | у | X1 | X2 | X3 | X4 |
| Мозамбик | 47 | 3 | 2,6 | 2,4 | 113 |
| Бурунди | 49 | 2,3 | 2,6 | 2,7 | 98 |
| Чад | 48 | 2,6 | 2,5 | 2,5 | 117 |
| Непал | 55 | 4,3 | 2,5 | 2,4 | 91 |
| Буркина-Фасо | 49 | 2,9 | 2,8 | 2,1 | 99 |
| Мадагаскар | 52 | 2,4 | 3,1 | 3,1 | 89 |
| Бангладеш | 58 | 5,1 | 1,6 | 2,1 | 79 |
| Гаити | 57 | 3,4 | 2 | 1,7 | 72 |
| Мали | 50 | 2 | 2,9 | 2,7 | 123 |
| Нигерия | 53 | 4,5 | 2,9 | 2,8 | 80 |
| Кения | 58 | 5,1 | 2,7 | 2,7 | 58 |
| Того | 56 | 4,2 | 3 | 2,8 | 88 |
| Индия | 62 | 5,2 | 1,8 | 2 | 68 |
| Бенин | 50 | 6,5 | 2,9 | 2,5 | 95 |
| Никарагуа | 68 | 7,4 | 3,1 | 4 | 46 |
| Гана | 59 | 7,4 | 2,8 | 2,7 | 73 |
| Ангола | 47 | 4,9 | 3,1 | 2,8 | 124 |
| Пакистан | 60 | 8,3 | 2,9 | 3,3 | 90 |
| Мавритания | 51 | 5,7 | 2,5 | 2,7 | 96 |
| Зимбабве | 57 | 7,5 | 2,4 | 2,2 | 55 |
| Гондурас | 67 | 7 | 3 | 3,8 | 45 |
| Китай | 69 | 10,8 | 1,1 | 1,1 | 34 |
| Камерун | 57 | 7,8 | 2,9 | 3,1 | 56 |
| Конго | 51 | 7,6 | 2,9 | 2,6 | 90 |
| Шри-Ланка | 72 | 12,1 | 1,3 | 2 | 16 |

**2.1 Уравнение в явном виде**

Для проведения дальнейшего анализа, используя пакет анализа MS Excel, вычислены параметры множественной регрессии в линейной форме с полным набором факторов. Параметры представлены на рисунке 1.

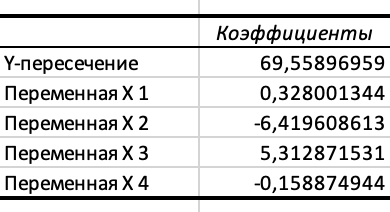


Рисунок 1 – Коэффициенты уравнения

Полученное уравнение в явном виде:

Характеристика поведения параметров при неизменных других:

При увеличении ВВП в паритетах покупательной способности на 1 единицу средняя ожидаемая продолжительность жизни увеличиться на 0,328 года.

При увеличении темпов прироста населения по сравнению с предыдущим годом на 1% единицу средняя ожидаемая продолжительность жизни уменьшиться на 6,42 лет.

При увеличении темпов прироста рабочей силы по сравнению с предыдущим годом на 1% единицу средняя ожидаемая продолжительность жизни увеличиться на 5,313 лет.

При увеличении коэффициента младенческой смертности на 1% единицу средняя ожидаемая продолжительность жизни уменьшиться на 0,159 года.

Критерий Фишера (F-критерий), который является показателем надежности и значимости уравнения в целом равен: 42,9, табличное значение критерия Фишера: 2,64.

Так как , уравнение надежно, значимо, показатель тесноты связи надежен и отражает устойчивую зависимость средней ожидаемой продолжительности жизни от нескольких факторов.

**2.2 Анализ факторов**

Для анализа факторов была построена матрица парных корреляций, представленная на рисунке 2.

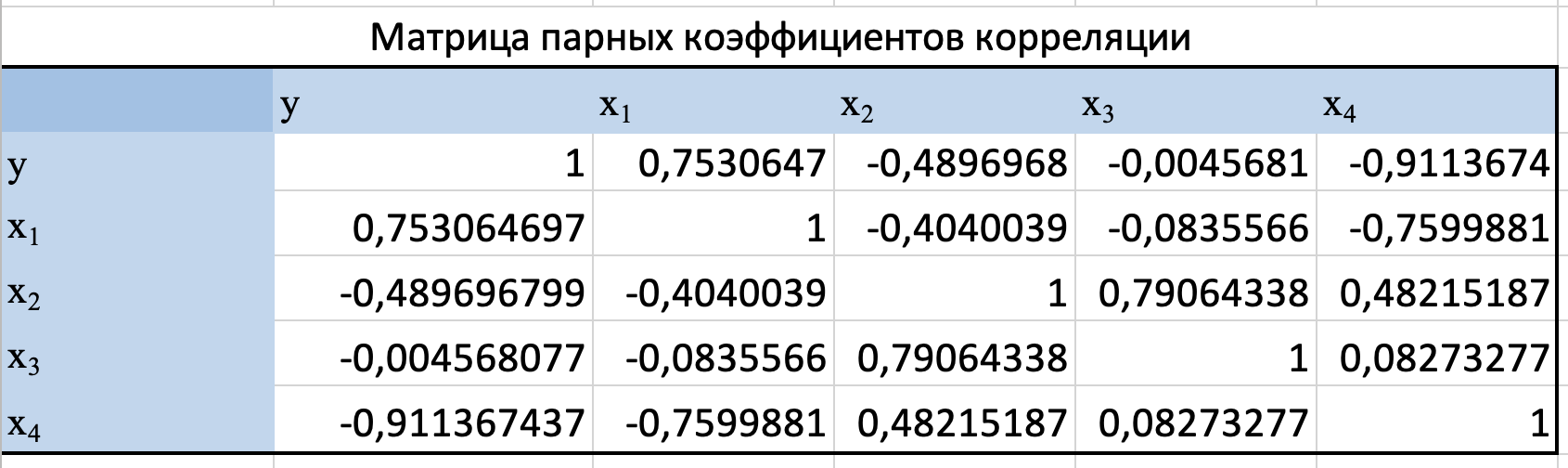


Рисунок 2 – Матрица парных корреляций

Исходя из теоретического анализа принято решение о подробном рассмотрении x1.

Коэффициенты корреляции переменной x1: 0,75; -0,4; -0,08; -0,75;

В результате сравнения корреляции признака x1 с другими было принято решении об исключении признака x1, так как признак x1 имеет меньшую тесноту связи с другими факторами.

**2.3 Уравнение без коррелирующих признаков**

Вычисленные коэффициенты множественной регрессии в линейной форме с новым набором факторов представлены на рисунке 3.

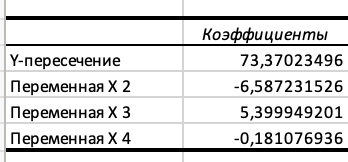


Рисунок 3 – Коэффициенты уравнения

Уравнение без коррелирующих признаков в явном виде:

ŷ = 73,37 - 6,59 x2 + 5,4 x3 – 0,18x4

Характеристика поведения параметров при неизменных других:

При увеличении темпов прироста населения по сравнению с предыдущим годом на 1% единицу средняя ожидаемая продолжительность жизни уменьшиться на 6,59 лет.

При увеличении темпов прироста рабочей силы по сравнению с предыдущим годом на 1% единицу средняя ожидаемая продолжительность жизни увеличиться на 5,4 лет.

При увеличении коэффициента младенческой смертности на 1% единицу средняя ожидаемая продолжительность жизни уменьшиться на 0,18 года.

Так как , уравнение надежно, значимо, показатель тесноты связи надежен и отражает устойчивую зависимость средней ожидаемой продолжительности жизни от нескольких факторов.

Критерий Фишера (F-критерий), который является показателем надежности и значимости уравнения в целом равен 56,37, табличное значение критерия Фишера - 2,85.

**2.4 Стандартизация уравнения**

Рассчитанные стандартизированные параметры β представлены на рисунке 4.

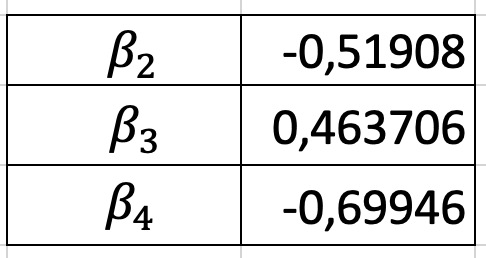


Рисунок 4 – Стандартизированные параметры

Уравнение в стандартизированном виде:

Влияние факторов:

1. наибольшее влияние имеет x4;
2. среднее влияние оказывает x2;
3. наименьшее влияние имеет x3.

**2.6 Теорема Гаусса-Маркова**

Исследование случайных отклонений.

На рисунке 5 представлена зависимость остатков от теоретических значений y. Точки распределены равномерно, что показывает, что остатки являются случайными величинами, и использование метода наименьших квадратов (МНК) оправдано. Теоретические значения ŷx хорошо соответствуют фактическим значениям y. Дисперсия случайной ошибки не меняется в зависимости от номера наблюдения, поэтому условие гомоскедастичности выполняется.

Рисунок 5 - График зависимости остатков от значений y

Построен график зависимости случайных остатков от фактора x1. Остатки на графике представлены в виде случайных величин, они не зависят от значений x2, что свидетельствует о несмещённости оценок коэффициентов регрессии. Данный график представлен на рисунке 6.

Рисунок 6 - График зависимости остатков от переменной x1

Построен график зависимости случайных остатков от фактора x3. Остатки на графике представлены в виде случайных величин, они не зависят от значений x3, что свидетельствует о несмещённости оценок коэффициентов регрессии. Данный график представлен на рисунке 7.

Рисунок 7 - График зависимости остатков от переменной x3

Построен график зависимости случайных остатков от фактора x4. Остатки на графике представлены в виде случайных величин, они не зависят от значений x4, что свидетельствует о несмещённости оценок коэффициентов регрессии. Данный график представлен на рисунке 8.

Рисунок 8 - График зависимости остатков от переменной x4

График нормального распределения отображает стандартный вид и соответствует ожидаемому распределению. Случайные ошибки на нем распределены нормально. График представлен на рисунке 9.

Рисунок 9 – График нормального распределения

Условия Гаусса-Маркова выполняются, оценки параметров регрессии являются несмещёнными, состоятельными и эффективными.

Полученные результаты

В результате выполнения данной лабораторной работы:

1. Построено уравнение множественной регрессии в линейной форме с полным набором факторов.
2. Установлено какие факторы коллинеарны и удален зависимый фактор.
3. Построено уравнение регрессии со статистически значимыми факторами.
4. Произведена оценка статистической значимости уравнений регрессии с помощью критерия Фишера.
5. Произведено исследование случайных отклонений используя теорему Гаусса-Маркова.