****

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

Условие задачи

Получить навыки построения и оценивания модели множественной регрессии.

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 |

Порядок выполнения

1 Построение уравнения множественной регрессии

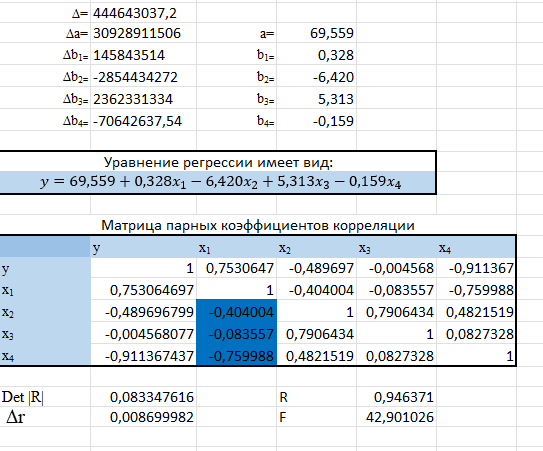
Уравнение множественной регрессии по 4 коэффициентам построено с помощью вычислений в MS Excel. Результаты представлены на рисунке 1.

Рисунок 1 – Уравнение множественной регрессии (по 4 коэффициентам)

Уравнение регрессии со статистически значимыми факторами

Упрощенное уравнение регрессии строится путем исключения наименее значимого фактора из уравнения регрессии. Для определения значимости факторов построена матрица парных коэффициентов корреляции (см. таблицу 2).

Таблица 2 – Матрица парных коэффициентов корреляции

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | y | x1 | x2 | x3 | x4 |
| y | 1 | 0,753064697 | -0,489696799 | -0,004568077 | -0,911367437 |
| x1 | 0,753064697 | 1 | -0,404003861 | -0,083556615 | -0,759988055 |
| x2 | -0,489696799 | -0,404003861 | 1 | 0,790643383 | 0,482151869 |
| x3 | -0,004568077 | -0,083556615 | 0,790643383 | 1 | 0,082732766 |
| x4 | -0,911367437 | -0,759988055 | 0,482151869 | 0,082732766 | 1 |

Коэффициенты интеркорреляции (т.е. корреляции между объясняющими переменными) позволяют исключать из модели дублирующие факторы. Считается, что две переменные явно коллинеарны, т.е. находятся между собой в линейной зависимости, если . Фактор х1 имеет наибольшую тесноту связи, значит остаются факторы х2, х3,х4.

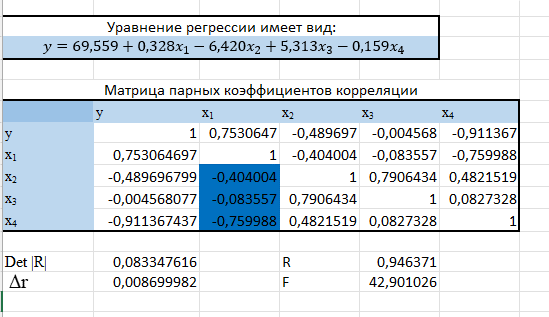
Уравнение регрессии со статистически значимыми факторами имеет вид: ŷ = a + b2x2 + b3x3 + b4x4. Параметры статистически значимого уравнения регрессии представлены на рисунке 2.

Рисунок 2 – Уравнение регрессии со статистически значимыми факторами

Данные матрицы парных коэффициентов корреляции для уравнения регрессии со статистически значимыми факторами представлены в таблице 3.

Таблица 3 – Матрица парных коэффициентов корреляции для уравнения регрессии со статистически значимыми факторами

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | y | x2 | x3 | x4 |
| y | 1 | -0,48970 | -0,00457 | -0,91137 |
| x2 | -0,48970 | 1 | 0,79064 | 0,48215 |
| x3 | -0,00457 | 0,79064 | 1 | 0,08273 |
| x4 | -0,91137 | 0,48215 | 0,08273 | 1 |

В матрице слабая теснота связи факторов, что доказывает статистическую значимость выбранных коэффициентов.

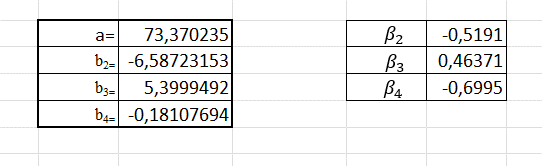
Произведен расчет стандартизированных коэффициентов регрессии. Результаты представлены на рисунке 3.

Рисунок 3 – Стандартизированные коэффициенты регрессии

По полученным стандартизированным коэффициентам регрессии видно, что наибольшее влияние имеет фактор B3.

Значимость уравнения множественной регрессии в целом определяется с помощью коэффициента Фишера.

При сравнении значимости полученных коэффициентов Фишера для полного уравнения регрессии и уравнения регрессии со статистически значимыми факторами получаем 42,90102621<56,36830533 (значимость коэффициента Фишера для статистически значимого уравнения больше значимости коэффициента Фишера для полного уравнения), из чего следует вывод: статистически значимые факторы определены правильно.

Независимость остатков от факторов, включенных в регрессию: проверяется с помощью графика остатков для факторов X2, X3, X4. Результаты представлены на рисунках 4–6.

Рисунок 4 – Остатки по Х2

Рисунок 5 – Остатки по Х3

Рисунок 6 – Остатки по Х4

По всем анализируемым показателям был сделан вывод, что уравнение показательной регрессии лучше всего подходит к исходному набору данных. После этого была проведена оценка уравнения согласно предпосылкам Гаусса-Маркова:

* Гомоскедастичность выполняется: остатки для Y могут быть размещены между двумя параллельными прямыми (см. Рисунок 7).
* Оценки параметров несмещенные

Рисунок 7 – Остатки по У

Полученные результаты

Навыки построения и оценивания модели множественной регрессии получены. Лабораторная работа выполнена, цель достигнута.