ГУАП

КАФЕДРА № 43

ОТЧЁТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| к.т.н., доцент |  |  |  | В. В. Мышко |
| должность, уч. Степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ОТЧЁТ О ЛАБОРАТОРНОЙ РАБОТЕ №5 | | | | | |
| МНОГОФАКТОРНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ | | | | | |
| по дисциплине: ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ | | | | | |
|  | | | | | |
| РАБОТУ ВЫПОЛНИЛ | | | | | |
| СТУДЕНТ ГР. | 4931 |  |  |  | А.А. Кинько |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург 2023

**Текст задания**

Согласно варианту №4931-12:

На основе заданного массива данных:

* построить уравнение регрессии в виде *линейного алгебраического полинома от двух переменных*;
* проверить адекватность уравнения регрессии;
* проверить значимость факторов регрессии.

Расчеты произвести в матричной форме.

Порядок выполнения задания:

1. Выполнить центрирование факторов (массив экспериментальных данных, таблица 1);
2. Составить матричное уравнение с вектором неизвестных оценок коэффициентов регрессии;
3. Найти оценки коэффициентов регрессии посредством решения матричного уравнения;
4. Проверить адекватность построенного уравнения регрессии экспериментальным данным по критерию Фишера при уровне значимости α = 0,05;
5. Выполнить селекцию факторов по критерию Стьюдента при таком же уровне значимости;
6. Повторно проверить адекватность уравнения регрессии после исключения незначимых факторов.

*Таблица №1. Экспериментальные данные варианта №4931-12*

| **Массив экспериментальных данных** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | -2 | -1 | 0 | 1 | 2 | -0,5 |
|  | 1 | 2 | 3 | 5 | 6 | 7 |
|  | -20 | -11 | -3 | 8 | 14 | -3 |

**Расчетные формулы**

1. Поиск функции регрессии.

В рассматриваемой задаче функция регрессии представляет собой линейный алгебраический полином от двух независимых переменных:

Составляем таблицу экспериментальных данных с центрированными значениями факторов, вычислив средние значения факторов:

Производим центрирование факторов, используя формулу :

*Таблица №2. Расчетная таблица для центрирования факторов*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | -1.917 | -0.917 | 0.083 | 1.083 | 2.083 | -0,417 |
|  | -3 | -2 | 1 | 1 | 2 | 3 |
|  | -20 | -11 | -3 | 8 | 14 | -3 |

Таким образом, функция регрессии принимает вид:

Система нормальных уравнений представляется следующим образом:

Представим и решим систему выше в матричной форме. Матричное уравнение, эквивалентное данной системе, принимает вид:

где

Выражение для вычисления оценок коэффициентов регрессии представляется равенством:

Вычисляем матрицу

Обратная матрица

Далее находим матрицу в правой части уравнения:

Вычисляем оценку вектора коэффициентов регрессии:

Таким образом, получаем следующее уравнение регрессии:

2. Проверка адекватности уравнения

Вычисляем оценки дисперсий, для чего составим расчетную таблицу:

*Таблица №3. Расчетная таблица для проверки адекватности*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| -1.917 | -3 | -20 | -17.5 | 306.25 | -19.65 | -0.35 | 0.123 |
| -0.917 | -2 | -11 | 8.5 | 72.25 | -11.22 | 0.22 | 0.048 |
| 0.083 | 1 | -3 | -0.5 | 0.25 | -2.789 | 0.211 | 0.044 |
| 1.083 | 1 | 8 | 10.5 | 110.25 | 6.554 | 1.446 | 2.091 |
| 2.083 | 2 | 14 | 16.5 | 275.25 | 14.984 | -0.984 | 0.969 |
| -0,417 | 3 | -3 | -0.5 | 0.25 | -2.898 | -0.102 | 0.01 |
|  | | | | |  | | |

Вычисляем оценки средних квадратичных отклонений:

Наблюдаемое значение показателя согласованности:

Для отыскания критического значения показателя согласованности при уровне значимости и степенях свободы , , используя приложение №5, получаем

Поскольку неравенство выполняется, нулевая гипотеза об адекватности функции регрессии экспериментальным данным принимается.

3. Селекция факторов.

Для селекции найдем элементы главной диагонали корреляционной матрицы и умножим каждый из элементов на , получив:

Оценки средних квадратических отклонений коэффициентов принимают значения:

Для каждого фактора находим наблюдаемое значение показателя согласованности:

Для числа степеней свободы и уровня значимости критическое значение показателя согласованности .

Следовательно, , ,

Для данного варианта все факторы значимы, значит пересчет не нужен. В любом другом случае в отношении факторов принимается конкурирующая гипотеза о их незначимости. Тогда в правой части выражения регрессии соответствующие слагаемые необходимо приравнять к нулю и пересчитать это выражение.

**Результат работы**

В ходе выполнения данной лабораторной работы была написана программа на языке Python 3.10, решающая задачу в общем виде для всех вариантов – находит факторы регрессии уравнения в виде линейного алгебраического полинома от двух переменных, проверяет их адекватность и значимость, а также производит перерасчет. Так, для варианта №4931-12 были получены следующие результаты:

Загружаем данные для варианта № 12 ...

Массив x1: [-2.0, -1.0, 0.0, 1.0, 2.0, -0.5]

Массив x2: [1.0, 2.0, 3.0, 5.0, 6.0, 7.0]

Массив y: [-20.0, -11.0, -3.0, 8.0, 14.0, -3.0]

Найдем функцию регрессии в виде алгебраического полинома: y = b0 + b1 \* x1 + b2 \* x2

Среднее значение фактора x1: -0.083

Среднее значение фактора x2: 4.0

Производим центрирование данных:

Массив x1\*: [-1.917, -0.917, 0.083, 1.083, 2.083, -0.417]

Массив x2\*: [-3.0, -2.0, -1.0, 1.0, 2.0, 3.0]

Массив y: [-20.0, -11.0, -3.0, 8.0, 14.0, -3.0]

Для вычисления оценок коэффициентов регрессии решим уравнение B ̃ = (Ẋ^T \* Ẋ)^(-1) (Ẋ^T \* Y)

Матрица (Ẋ^T \* Ẋ):

[[ 6. -0. 0. ]

[-0. 10.21 11.5 ]

[ 0. 11.5 28. ]]

Обратная для этой матрица:

[[ 0.1667 0. 0. ]

[ 0. 0.1823 -0.0749]

[-0. -0.0749 0.0665]]

Матрица (Ẋ^T \* Y):

[[-15. ]

[ 87.255]

[112. ]]

Оценка вектора коэффициентов регрессии:

[[-2.5005 ]

[ 7.5177865]

[ 0.9126005]]

Получаем уравнение: -2.5 + 7.518 ẋ1 + 0.913 ẋ2

---- Проверка адекватности уравнения

Составляем расчетную таблицу:

x1\_i [-1.917, -0.917, 0.083, 1.083, 2.083, -0.417]

x2\_i [-3.0, -2.0, -1.0, 1.0, 2.0, 3.0]

y\_i [-20.0, -11.0, -3.0, 8.0, 14.0, -3.0]

y\_i - y^- [-17.5, -8.5, -0.5, 10.5, 16.5, -0.5]

(y\_i - y^-)^2 [306.25, 72.25, 0.25, 110.25, 272.25, 0.25] - сумма:761.5

y~\_i [-19.65, -11.22, -2.789, 6.554, 14.984, -2.898]

y\_i - y~\_i [-0.35, 0.22, -0.211, 1.446, -0.984, -0.102]

(y\_i - y~\_i)^2 [0.123, 0.048, 0.044, 2.091, 0.969, 0.01] - сумма:3.2849999999999997

Оценка отклонения sigma^2: 152.3

Оценка отклонения sigma1^2: 0.8212499999999999

Значение показателя согласованности: 185.44901065449014

Критическое значение показателя согласованности: 6.26

Так как значение показателя согласованности больше значения показателя F, нулевая гипотеза об адекватности функции регрессии, принимается

---- Селекция факторов

Элементы главной диагонали корреляционной матрицы: [0.13690237499999997, 0.14971387499999997, 0.054613125]

Оценки средних квадратичных отклонений: [0.37000320944553977, 0.3869287725150457, 0.2336945121306874]

Соответствующие показатели согласованности: [6.758049487589768, 19.429380893889636, 3.905100259648602]

Критическое значение показателя согласованности при уровне значимости alpha=0,05 степени свободы f = 3: 3.18

Фактор ẋ0 является значимым

Фактор ẋ1 является значимым

Фактор ẋ2 является значимым

---- Пересчет регрессионного выражения

Все факторы значимы, пересчет не требуется.

Также был получен результат для тестового варианта (64 в файле data.txt):

*Таблица №4. Экспериментальные данные тестового варианта*

| **Массив экспериментальных данных** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | -0.5 | 0 | 0.8 | 0.4 | 0.5 | -0,6 |
|  | -3 | -1 | 2 | 0.5 | 1.5 | 6 |
|  | -15.1 | -1 | -19.9 | 9.5 | 16.5 | 47.9 |

Загружаем данные для варианта № 64 ...

Массив x1: [-0.5, 0.0, 0.8, 0.4, 0.5, 0.6]

Массив x2: [-3.0, -1.0, 2.0, 0.5, 1.5, 6.0]

Массив y: [-15.1, -1.0, 19.9, 9.5, 16.5, 47.9]

Найдем функцию регрессии в виде алгебраического полинома: y = b0 + b1 \* x1 + b2 \* x2

Среднее значение фактора x1: 0.3

Среднее значение фактора x2: 1.0

Производим центрирование данных:

Массив x1\*: [-0.8, -0.3, 0.5, 0.1, 0.2, 0.3]

Массив x2\*: [-4.0, -2.0, 1.0, -0.5, 0.5, 5.0]

Массив y: [-15.1, -1.0, 19.9, 9.5, 16.5, 47.9]

Для вычисления оценок коэффициентов регрессии решим уравнение B ̃ = (Ẋ^T \* Ẋ)^(-1) (Ẋ^T \* Y)

Матрица (Ẋ^T \* Ẋ):

[[ 6. -0. 0. ]

[-0. 1.12 5.85]

[ 0. 5.85 46.5 ]]

Обратная для этой матрица:

[[ 0.1667 0. 0. ]

[ 0. 2.6039 -0.3276]

[-0. -0.3276 0.0627]]

Матрица (Ẋ^T \* Y):

[[ 77.7 ]

[ 40.95]

[325.3 ]]

Оценка вектора коэффициентов регрессии:

[[12.95259 ]

[ 0.061425]

[ 6.98109 ]]

Получаем уравнение: 12.953 + 0.061 ẋ1 + 6.981 ẋ2

---- Проверка адекватности уравнения

Составляем расчетную таблицу:

x1\_i [-0.8, -0.3, 0.5, 0.1, 0.2, 0.3]

x2\_i [-4.0, -2.0, 1.0, -0.5, 0.5, 5.0]

y\_i [-15.1, -1.0, 19.9, 9.5, 16.5, 47.9]

y\_i - y^- [-28.05, -13.95, 6.95, -3.45, 3.55, 34.95]

(y\_i - y^-)^2 [786.802, 194.602, 48.302, 11.902, 12.603, 1221.503] - сумма:2275.714

y~\_i [-15.021, -1.028, 19.964, 9.468, 16.455, 47.876]

y\_i - y~\_i [-0.079, 0.028, -0.064, 0.032, 0.045, 0.024]

(y\_i - y~\_i)^2 [0.006, 0.001, 0.004, 0.001, 0.002, 0.001] - сумма:0.015

Оценка отклонения sigma^2: 455.14279999999997

Оценка отклонения sigma1^2: 0.00375

Значение показателя согласованности: 121371.41333333333

Критическое значение показателя согласованности: 6.26

Так как значение показателя согласованности больше значения показателя F, нулевая гипотеза об адекватности функции регрессии, принимается

---- Селекция факторов

Элементы главной диагонали корреляционной матрицы: [0.0006251249999999999, 0.009764624999999999, 0.000235125]

Оценки средних квадратичных отклонений: [0.025002499875012497, 0.09881611710647205, 0.015333786225195654]

Соответствующие показатели согласованности: [518.0517974102589, 0.6216091240846149, 455.2750310636944]

Критическое значение показателя согласованности при уровне значимости alpha=0,05 степени свободы f = 3: 3.18

Фактор ẋ0 является значимым

Фактор ẋ1 является не значимым

Фактор ẋ2 является значимым

---- Пересчет регрессионного выражения

Пересчет

Найдем функцию регрессии в виде алгебраического полинома: y = b0 + b2 \* x2

Среднее значение фактора x2: 0.0

Производим центрирование данных:

Массив x2\*: [-4.0, -2.0, 1.0, -0.5, 0.5, 5.0]

Массив y: [-15.1, -1.0, 19.9, 9.5, 16.5, 47.9]

Для вычисления оценок коэффициентов регрессии решим уравнение B ̃ = (Ẋ^T \* Ẋ)^(-1) (Ẋ^T \* Y)

Матрица (Ẋ^T \* Ẋ):

[[ 6. 0. ]

[ 0. 46.5]]

Обратная для этой матрица:

[[0.1667 0. ]

[0. 0.0215]]

Матрица (Ẋ^T \* Y):

[[ 77.7]

[325.3]]

Оценка вектора коэффициентов регрессии:

[[12.95259]

[ 6.99395]]

Получаем уравнение: 12.953 + 6.994 ẋ2

---- Проверка адекватности уравнения

Составляем расчетную таблицу:

x1\_i [-0.8, -0.3, 0.5, 0.1, 0.2, 0.3]

x2\_i [-4.0, -2.0, 1.0, -0.5, 0.5, 5.0]

y\_i [-15.1, -1.0, 19.9, 9.5, 16.5, 47.9]

y\_i - y^- [-28.05, -13.95, 6.95, -3.45, 3.55, 34.95]

(y\_i - y^-)^2 [786.802, 194.602, 48.302, 11.902, 12.603, 1221.503] - сумма:2275.714

y~\_i [-15.023, -1.035, 19.947, 9.456, 16.45, 47.922]

y\_i - y~\_i [-0.077, 0.035, -0.047, 0.044, 0.05, -0.022]

(y\_i - y~\_i)^2 [0.006, 0.001, 0.002, 0.002, 0.003, 0.0] - сумма:0.014000000000000002

Оценка отклонения sigma^2: 455.14279999999997

Оценка отклонения sigma1^2: 0.0035000000000000005

Значение показателя согласованности: 130040.79999999997

Критическое значение показателя согласованности: 6.26

Так как значение показателя согласованности больше значения показателя F, нулевая гипотеза об адекватности функции регрессии, принимается

---- Селекция факторов

Элементы главной диагонали корреляционной матрицы: [0.00058345, 7.525e-05]

Оценки средних квадратичных отклонений: [0.0241547096856907, 0.008674675786448736]

Соответствующие показатели согласованности: [536.2345550223334, 806.2491523805069]

Критическое значение показателя согласованности при уровне значимости alpha=0,05 степени свободы f = 3: 3.18

Фактор ẋ0 является значимым

Фактор ẋ2 является значимым

---- Пересчет регрессионного выражения

Все факторы значимы, пересчет не требуется

**Выводы**

В ходе данной лабораторной работы были получены навыки получения уравнения регрессии в виде линейного алгебраического полинома от двух переменных, а также проверка его адекватности и значимости факторов. Так, для варианта №4931-12 было получено уравнение:

**Приложение А. Листинг программы**