**Харківський національний економічний університет**

**імені Семена Кузнеця**

**ЗВІТ**

**З ВИКОНАННЯ ЛАБОРАТОРНОЇ РОБОТИ № 4**

**за дисципліною: *“*Основи математичного моделювання**”

**на тему: “** **Побудова багатофакторної регресійної моделі та перевірка її на адекватність”**

**Варіант № 4**

**Виконав: студент факультету Інформаційних технологій**

**3 курсу, спец. Кібербезпека,**

**групи 6.04.125.010.21.2**

**Бойко Вадим Віталійович**

**Перевірила: Шаповалова Олена Олександрівна**

**ХНЕУ ім. С. Кузнеця**

**2023**

**Мета роботи –** навчитися будувати багатофакторні регресійні моделі, перевіряти їх на адекватність, визначати значущість окремих коефіцієнтів регресії.

**Після виконання роботи студент повинен:**

ЗНАТИ визначення лінійної багатофакторної моделі, алгоритм перевірки адекватності багатофакторної моделі, алгоритм розрахунку невідомих параметрів моделі за методом найменших квадратів, алгоритм визначення значущості коефіцієнтів регресії.

УМІТИ будувати багатофакторні моделі за методом найменших квадратів, розраховувати коефіцієнт детермінації, перевіряти значущість множинного коефіцієнта кореляції, перевіряти модель на адекватність, оцінювати вплив кожного регресора на якість моделі.

МАТИ УЯВЛЕННЯ про властивості оцінок параметрів, які знайдені за методом найменших квадратів; якими методами можна побудувати моделі множинної регресії.

**Завдання:**

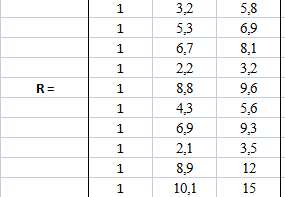
1. Знайти параметри багатофакторної моделі методом найменших квадратів (за варіантом).
2. Обчислити коефіцієнт детермінації, скорегований коефіцієнт детермінації, вибірковий коефіцієнт детермінації.
3. Перевірити модель на адекватність за критерієм Фішера з довірчою ймовірністю (n+60)/100.
4. Визначити статистичну значущість коефіцієнта кореляції за критерієм Стьюдента.
5. Знайти значущість окремих коефіцієнтів регресії за t – статистикою з q=0,01\*n.
6. Розрахувати часткові коефіцієнти детермінації.
7. Для статистичної вибірки виконати обчислення пп.1-6 та зробити висновки щодо якості отриманої моделі.

**Початкові дані:**

|  |  |  |  |
| --- | --- | --- | --- |
| N | X1 | X2 | Y9 |
| 1 | 3,20 | 5,80 | 30,00 |
| 2 | 5,30 | 6,90 | 43,00 |
| 3 | 6,70 | 8,10 | 51,00 |
| 4 | 2,20 | 3,20 | 18,00 |
| 5 | 8,80 | 9,60 | 62,00 |
| 6 | 4,30 | 5,60 | 33,00 |
| 7 | 6,90 | 9,30 | 55,00 |
| 8 | 2,10 | 3,50 | 19,00 |
| 9 | 8,90 | 12,00 | 72,00 |
| 10 | 10,10 | 15,00 | 88,00 |

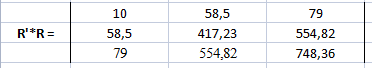
**Хід роботи (1)**

1. Знайдемо оцінні значення параметрів моделі , , . Побудуємо|спорудити| спочатку матрицю *R*, складену з|із| одиничного|поодинокий| стовпця і значень *хi*

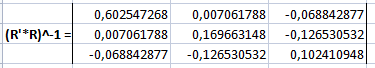


транспонуємо її.

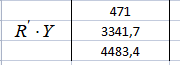
 помножимо транспоновану матрицю  на початкову|вихідний| *R* .



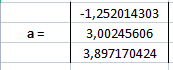
Визначник даної матриці  дорівнює **7323,924**. Знайдемо матрицю, обернену до отриманої|одержаної|.



Визначимо вектор :



Тоді вектор оцінок параметрів моделі матиме вигляд:



Отримано оцінні значення параметрів моделі: = **-1,25,** = **3,002**, = **3,897**.

1. Розрахуємо коефіцієнт детермінації, спираючись на дані кореляційної таблиці:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***n*** | ***yi*** | ***xi1*** | ***xi2*** | ***y mod*** | ***ymod-y*** | ***(ymod-y)^2*** | ***ymod-yср*** | ***(ymod-ycp)^2*** | ***yi-ycp*** | ***(yi-ycp)^2*** |
| 1 | 30 | 3,2 | 5,8 | 30,95943355 | 0,95943355 | 0,920512736 | -16,14056645 | 260,5178853 | -17,1 | 292,41 |
| 2 | 43 | 5,3 | 6,9 | 41,55147874 | -1,448521257 | 2,098213833 | -5,548521257 | 30,78608814 | -4,1 | 16,81 |
| 3 | 51 | 6,7 | 8,1 | 50,43152174 | -0,568478264 | 0,323167536 | 3,331521736 | 11,09903708 | 3,9 | 15,21 |
| 4 | 18 | 2,2 | 3,2 | 17,82433439 | -0,175665613 | 0,030858408 | -29,27566561 | 857,0645971 | -29,1 | 846,81 |
| 5 | 62 | 8,8 | 9,6 | 62,5824351 | 0,582435099 | 0,339230645 | 15,4824351 | 239,7057966 | 14,9 | 222,01 |
| 6 | 33 | 4,3 | 5,6 | 33,48270113 | 0,482701131 | 0,233000382 | -13,61729887 | 185,4308285 | -14,1 | 198,81 |
| 7 | 55 | 6,9 | 9,3 | 55,70861746 | 0,708617457 | 0,5021387 | 8,608617457 | 74,10829452 | 7,9 | 62,41 |
| 8 | 19 | 2,1 | 3,5 | 18,69323991 | -0,306760092 | 0,094101754 | -28,40676009 | 806,9440189 | -28,1 | 789,61 |
| 9 | 72 | 8,9 | 12 | 72,23588972 | 0,235889723 | 0,055643961 | 25,13588972 | 631,8129521 | 24,9 | 620,01 |
| 10 | 88 | 10,1 | 15 | 87,53034827 | -0,469651733 | 0,22057275 | 40,43034827 | 1634,613061 | 40,9 | 1672,81 |
| Сума | 471 | 58,5 | 79 | 471 | 2,17071E-12 | 4,817440705 | 2,14584E-12 | 4732,082559 | 0 | 4736,9 |
| Ср. зн. | 47,1 |  |  |  |  | e |  |  |  |  |

= **0,9989**

Коефіцієнт детермінації наближається до одиниці, тому можна зробити висновок, що варіація залежної змінної значною мірою залежить від варіації незалежних змінних.

На основі знайденого коефіцієнта детермінації розрахуємо скорегований

**= 0,997**

та вибірковий коефіцієнт множинної кореляції:

**= 0,99949**

Значення коефіцієнта кореляції наближається до 1, тому можна припустити, що існує суттєвий зв'язок між всіма незалежними факторами і залежною змінною.

1. Перевіримо модель на адекватність за *F*–статистикою з довірчою ймовірністю (n+60)/100=84/100=**0.84**:

= **3929,13**

Розрахункове значення порівняємо з|із| табличним при 2–х і 7–ми ступенях свободи і рівнем значущості ***q*=16%**; **F табл. =** **0,178.**

Так як то можна сказати, що дана модель з довірчою ймовірністю 84% є адекватною.

1. За критерієм Стьюдента визначимо значущість коефіцієнта кореляції:

**= 82,92**

Знайдемо відповідне табличне значення t–розподілу Стьюдента з 7-ма ступенями свободи та рівнем значущості *q*=16%;  *tтабл(16%, 7)*= **1,415**.

Оскільки *|tр|>tтабл*, можна зробити висновок щодо значущості коефіцієнта кореляції з ймовірністю 84%.

1. Перевіримо значущість окремих коефіцієнтів регресії:

**t0 p = -1,94**

**t1 p = 8,79**

**t2 p = 14,68**

Розрахункові значення порівняємо з табличним з рівнем значущості q=0,01\*n =0,24= **24%** та 7-ма ступенями свободи *tтабл(24%,7)*=**0,711**. Оскільки **|t0,1,2 р| > t табл**, то оцінка α0, α1, α2 є значущими.

1. Знайдемо часткові коефіцієнти детермінації для кожного з факторів моделі:



R0^2 = 0,000549

R1^2 = 0,011217

R2^2 = 0,031308

На залежну змінну найбільше впливає фактор, який має найбільше значення = **R2^2 = 0,031.**

**Хід роботи (2)**

1. Тепер, для статистичної вибірки виконаємо обчислення. Для прикладу, візьмемо вибірку «Список країн світу за чисельністю населення».
2. Таблиця з розрахунками наведена у файлі Ексель, лист-2.

**Висновок**

В даній лабораторній роботі, навчилися будувати багатофакторні регресійні моделі, перевіряти їх на адекватність, визначили значущість окремих коефіцієнтів регресії.