**Лабораторна робота №4**

**Побудова багатофакторної регресійної моделі та перевірка її на адекватність**

**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:**

1Знайти параметри багатофакторної моделі методом найменших квадратів (за варіантом).

2Обчислити 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.

**Хід роботи**:

1 Знайдемо вектор параметрів багатофакторної моделі за формулою|в

.

Вектор  дозволяє знайти оцінки параметрів моделі та отримати|одержати| наступний|слідуючий| вираз|вираження|, який оцінює багатофакторну регресійну модель|фактор|:

.

 дозволяє отримати прогнозні значення за лінійною багатофакторною регресійною моделлю|фактор|, які можуть, в загальному випадку, і не співпадати|збігатися| з|із| .

2Коефіцієнт детермінації визначає загальний вплив факторів на залежну змінну і розраховується за формулою

,

де *еі* – залишки регресії.

Скорегований коефіцієнт детермінації обчислюється за формулою



Вибірковий коефіцієнт детермінації показує щільність зв’язку всіх незалежних факторів і залежної змінної:



3 Перевіримо модель на адекватність за *F*–статистикою:

 або 

Розрахункове значення *Fр* порівнюємо з табличним *Fтабл*(додаток В), використовуючи наступні аргументи:

– рівень значущості *q*;

– кількість ступенів свободи регресії, яка дорівнює кількості незалежних факторів моделі *k*;

– кількість ступенів свободи залишків, що дорівнює *(n–k–1).*

Якщо *Fр>Fтабл*, то коефіцієнти регресії є значущими, якщо ж *Fр<Fтабл* , то не є значущими з довірчою ймовірністю α.

1. Визначимо значущість коефіцієнта кореляції за критерієм Стьюдента за формулою:

.

Розрахункове значення *tр* порівнюємо з табличним, використовуючи наступні аргументи (додаток Б):

* заданий рівень значущості *q*;
* кількість ступенів свободи залишків, яке дорівнює *(n–k–1).*

Якщо *tр>tтабл* , то можна зробити висновок про *(1–q)\*100%* значущість коефіцієнта кореляції, який характеризує щільність зв’язків між залежною та незалежними змінними.

5 Перевіримо значущість окремих коефіцієнтів регресії. *tj* – статистика для кожного коефіцієнта регресії знаходиться за формулою:

 (),

.

де .

*cjj* – діагональні елементи матриці .

Розрахункове значення t–критерію порівнюємо з табличним з рівнем значущості *q* та кількістю ступенів свободи *(n–k–1*) (додаток В). Якщо *| tр |> tтабл*, то оцінки є значущими, у протилежному випадку – *|tр|<tтабл* – незначущими.

6 Знайдемо часткові коефіцієнти детермінації для кожного з факторів моделі. Вони показують, в якій мірі зменшиться коефіцієнт детермінації за умови виключення відповідного фактора з моделі:



Найбільше впливає на залежну змінну той фактор, який має найбільше значення .

**Приклад 4**

Побудувати|спорудити| багатофакторну|численний| лінійну регресійну модель залежності роздрібного товарообігу Y, (тис.шт.), від кількості підприємств роздрібної торгівлі X1, (шт.), і обсягу укладених угод на біржах X2, (шт.), за даними, наведеними|уявленим| в табл. 4.1 і перевірити її на адекватність з рівнем значущості 5%.

Таблиця 4.1 – Вихідні дані до прикладу 4

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Y | 10 | 12 | 17 | 13 | 15 | 10 | 14 | 12 | 16 | 18 |
| X1 | 2 | 2 | 8 | 2 | 6 | 3 | 5 | 3 | 9 | 10 |
| X2 | 1 | 2 | 10 | 4 | 8 | 4 | 7 | 3 | 10 | 11 |

**Розв’язання:**

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 7160. Знайдемо матрицю, обернену до отриманої|одержаної|.

.

Визначимо вектор :

.

Тоді вектор оцінок параметрів моделі матиме вигляд:

.

Отримано оцінні значення параметрів моделі: =9,4, =0,128, =0,612. Якщо операції з|із| матрицями проводяться|виробляються,справляються| за допомогою табличного процесора, рекомендується використовувати стандартні вбудовані функції MDETERM (МОПРЕД), TRANSPOSE (ТРАНСП), MMULT (МУМНОЖ), MINVERSE (МОБР). Під час роботи з|із| матрицями введення даних завершується комбінацією клавіш Shift+Ctrl+Enter.

2 Розрахуємо коефіцієнт детермінації, спираючись на дані кореляційної таблиці (табл. 4.2):

Таблиця 4.2 – Кореляційна таблиця до прикладу 4

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *n* | *yi* | *х1* | *х2* |  |  | ()2 |  | ()2 |  |  |
| 1 | 10 | 2 | 1 | 10,25 | –0,25 | 0,07 | –3,44 | 11,86 | 3,44 | 11,86 |
| 2 | 12 | 2 | 2 | 10,86 | 1,13 | 1,28 | –2,83 | 8,02 | 2,83 | 8,02 |
| 3 | 17 | 8 | 10 | 16,53 | 0,46 | 0,22 | 2,83 | 8,02 | –2,83 | 8,02 |
| 4 | 13 | 2 | 4 | 12,09 | 0,90 | 0,83 | –1,61 | 2,59 | 1,61 | 2,59 |
| 5 | 15 | 6 | 8 | 15,05 | –0,05 | 0,00 | 1,35 | 1,83 | –1,35 | 1,83 |
| 6 | 10 | 3 | 4 | 12,21 | –2,21 | 4,93 | –1,48 | 2,19 | 1,48 | 2,19 |
| 7 | 14 | 5 | 7 | 14,31 | –0,31 | 0,10 | 0,61 | 0,37 | –0,61 | 0,37 |
| 8 | 12 | 3 | 3 | 11,60 | 0,39 | 0,15 | –2,09 | 4,38 | 2,09 | 4,38 |
| 9 | 16 | 9 | 10 | 16,66 | –0,66 | 0,44 | 2,96 | 8,77 | –2,96 | 8,77 |
| 10 | 18 | 10 | 11 | 17,40 | 0,59 | 0,36 | 3,70 | 13,70 | –3,70 | 13,70 |
| ∑ | 137 | 50 | 60 | 137 | 0 | 8,368 | 0 | 61,73 | 0,00 | 61,73 |
| Середнє значення | 13,7 |  |  |  |  |  |  |  |  |  |



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

,

та вибірковий коефіцієнт множинної кореляції:

.

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

3 Перевіримо модель на адекватність за *F*–статистикою:

.

Розрахункове значення порівняємо з|із| табличним при 2–х і 7–ми ступенях свободи і рівнем значущості *q*=5%; .

Так як , то можна сказати, що дана модель з довірчою ймовірністю 95% є адекватною.

4 За критерієм Стьюдента визначимо значущість коефіцієнта кореляції:



Знайдемо відповідне табличне значення t–розподілу Стьюдента з 7-ма ступенями свободи та рівнем значущості *q*=5%;  *tтабл(5%, 7)*= 2,365.

Оскільки *|tр|>tтабл*, можна зробити висновок щодо значущості коефіцієнта кореляції з ймовірністю 95%.

5 Перевіримо значущість окремих коефіцієнтів регресії:







Розрахункові значення порівняємо з табличним з рівнем значущості 5% та 7-ма ступенями свободи *tтабл(5%, 7)*=1,895. Оскільки *|t0р|>tтабл*, то оцінка *α0* є значущою. Так як *|t1р|<tтабл* та *|t2р|<tтабл ,* оцінки *α1, α2* єнезначущими.

6Знайдемо часткові коефіцієнти детермінації для кожного з факторів моделі:







На залежну змінну найбільше впливає фактор, який має найбільше значення .

Вихідні дані для самостійного виконання лабораторної роботи подані в додатку нижче. Номер варіанта (номер Y) обирається за номером студента у журналі.

ДОДАТОК К

Вихідні дані для лабораторної роботи

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | X1 | X2 | Y1 | Y2 | Y3 | Y4 | Y5 | Y6 | Y7 | Y8 | Y9 | Y10 |
| 1 | 3,20 | 5,80 | 9,00 | 11,20 | 37,00 | 22,00 | 17,20 | 17,30 | 33,50 | 10,20 | 30,00 | 4,00 |
| 2 | 5,30 | 6,90 | 15,00 | 12,70 | 37,00 | 23,00 | 10,10 | 26,70 | 42,20 | 13,30 | 43,00 | 8,00 |
| 3 | 6,70 | 8,10 | 20,00 | 12,40 | 40,00 | 26,00 | 9,60 | 31,50 | 52,40 | 16,40 | 51,00 | 8,00 |
| 4 | 2,20 | 3,20 | 8,00 | 14,90 | 28,00 | 16,00 | 13,90 | 4,90 | 13,90 | 9,00 | 18,00 | 6,00 |
| 5 | 8,80 | 9,60 | 25,00 | 12,30 | 41,00 | 26,00 | 1,70 | 42,00 | 65,30 | 20,20 | 62,00 | 9,00 |
| 6 | 4,30 | 5,60 | 13,00 | 13,60 | 34,00 | 20,00 | 9,80 | 18,00 | 33,20 | 12,90 | 33,00 | 7,00 |
| 7 | 6,90 | 9,30 | 20,00 | 11,20 | 44,00 | 32,00 | 9,50 | 38,00 | 61,70 | 15,90 | 55,00 | 6,00 |
| 8 | 2,10 | 3,50 | 9,00 | 14,70 | 30,00 | 14,00 | 14,80 | 6,00 | 16,30 | 10,00 | 19,00 | 5,00 |
| 9 | 8,90 | 12,00 | 24,00 | 8,10 | 50,00 | 38,00 | 6,10 | 53,00 | 81,20 | 18,60 | 72,00 | 5,00 |
| 10 | 10,10 | 15,00 | 26,00 | 3,70 | 60,00 | 44,00 | 3,80 | 65,00 | 102,00 | 20,00 | 88,00 | 4,00 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| N | X1 | X2 | Y11 | Y12 | Y13 | Y14 | Y15 |  |  |  |  |  |
| 1 | 2,30 | 3,50 | 4,80 | 9,00 | 15,00 | 21,00 | 10,00 |  |  |  |  |  |
| 2 | 7,20 | 8,30 | 26,00 | 16,50 | 43,10 | 52,00 | 43,00 |  |  |  |  |  |
| 3 | 2,50 | 4,30 | 6,10 | 7,10 | 16,10 | 27,00 | 15,00 |  |  |  |  |  |
| 4 | 3,70 | 6,90 | 12,50 | 9,10 | 26,10 | 39,00 | 31,00 |  |  |  |  |  |
| 5 | 6,10 | 7,10 | 21,10 | 15,50 | 36,40 | 43,00 | 35,00 |  |  |  |  |  |
| 6 | 2,50 | 5,70 | 7,20 | 8,90 | 19,00 | 33,00 | 21,00 |  |  |  |  |  |
| 7 | 5,30 | 7,20 | 17,40 | 13,50 | 33,00 | 42,00 | 36,00 |  |  |  |  |  |
| 8 | 2,10 | 5,60 | 6,00 | 5,50 | 17,00 | 32,00 | 19,00 |  |  |  |  |  |
| 9 | 4,50 | 7,10 | 15,70 | 11,60 | 30,00 | 42,00 | 32,00 |  |  |  |  |  |
| 10 | 6,20 | 8,60 | 21,20 | 13,90 | 40,00 | 50,00 | 43,00 |  |  |  |  |  |