Лабораторна робота №1 з ММС

Некряч Владислав, ТК-31

**Завдання до роботи**

Задано спостереження за деякою системою у вигляді таблиці 1 (х1, х2, х3 – вхідні (незалежні) змінні, у – вихідна (залежна змінна).

Перевірити вхідні змінні на наявність мультиколінеарності за алгоритмом Фаррара-Глобера.

Використовуючи МНК провести ідентифікацію об’єкта моделювання, тобто визначивши невідомі параметри b0, b1, b2, b3 у моделі

Таблицю 1 наповнити даними самостійно. Функції f1(x1, x2, x3), f2(x1, x2, x3), f3(x1, x2, x3) вибрати самостійно. Для отриманої апроксимаційної функції y(x1, x2, x3) (моделі об’єкта спостереження) виконати кореляційно-регресійний аналіз. Проаналізувати результати.

Дослідити отриману модель на гетероскедастичність (тест обираєте самі) та на автокореляцію (тест Дарбіна-Уотсона).

| № | *x*1 | *x*2 | *x*3 | *y* |
| --- | --- | --- | --- | --- |
| 1 | 1 | 2 | 5 | 10 |
| 2 | 3 | 7 | 11 | 29 |
| 3 | 5 | 9 | 13 | 27 |
| 4 | 9 | 4 | 10 | 25 |
| 5 | 6 | 11 | 13 | 28 |
| 6 | 9 | 12 | 17 | 34 |
| 7 | 19 | 22 | 17 | 37 |
| 8 | 23 | 25 | 19 | 39 |
| 9 | 24 | 26 | 27 | 47 |
| 10 | 26 | 30 | 29 | 51 |
| 11 | 27 | 32 | 33 | 56 |
| 12 | 29 | 36 | 34 | 56 |
| 13 | 32 | 36 | 40 | 63 |
| 14 | 32 | 39 | 44 | 67 |
| 15 | 35 | 41 | 46 | 69 |
| 16 | 37 | 42 | 47 | 65 |
| 17 | 52 | 52 | 37 | 63 |
| 18 | 54 | 55 | 39 | 63 |
| 19 | 55 | 56 | 40 | 64 |
| 20 | 59 | 57 | 44 | 77 |

**Хід роботи**

1) Перевірка вхідних змінних на мультиколінеарність за алгоритмом Фаррара-Глобера

1. Проводимо стандартизацію (нормалізацію) вхідних змінних за формулою:

– дисперсія змінної ,

– середнє значення змінної ,

– кількість вхідних змінних,

– кількість спостережень.

2. Знаходимо кореляційну матрицю: , яка складається з коефіцієнтів кореляції між та , взятих попарно; – стандартизована матриця, отримана у пункті 1, на її діагоналі – одиниці.

3. Визначаємо наявність мультиколінеарності за критерієм Пірсона :

серед змінних існує мультиколінеарність.

4. Визначимо матрицю .

5. Визначимо наявність мультиколінеарності між кожною зі змінних та рештою змінних за критерієм Фішера:

– діагональний елемент матриці з пункту 4.

кожна зі змінних мультиколінеарна з рештою.

6. Знаходимо частинні коефіцієнти кореляції для кожної пари змінних :

– елементи матриці з пункту 4.

7. Перевіряємо наявність мультиколінеарності між парами змінних за критерієм Стьюдента:

змінні попарно мультиколінеарні.

2) Формування гіпотези про вид функціональної залежності

Критерій МНК:

Необхідні умови екстремуму критерію:

Цю систему нормальних рівнянь зручно подати у матричній формі:

Тоді розв'язком цієї системи відносно невідомих буде:

Маємо:

(*x1+x3*)

Значення відгуку моделі та значення відгуку моделі, розраховані за функцією:

| No | Y | Y\* |
| --- | --- | --- |
| 1 | 10 | 9.86995864252882 |
| 2 | 29 | 21.4383704927416 |
| 3 | 27 | 26.4056783930146 |
| 4 | 25 | 33.3507039283343 |
| 5 | 28 | 32.1719037159606 |
| 6 | 34 | 34.564328848296 |
| 7 | 37 | 41.2402742054158 |
| 8 | 39 | 44.0477815855065 |
| 9 | 47 | 49.8297885104965 |
| 10 | 51 | 53.2745240860203 |
| 11 | 56 | 51.964096067027 |
| 12 | 56 | 56.2488612997419 |
| 13 | 63 | 59.4291189709739 |
| 14 | 67 | 60.5855931694383 |
| 15 | 69 | 58.1462020470552 |
| 16 | 65 | 62.7591874821717 |
| 17 | 63 | 63.4474154605848 |
| 18 | 63 | 64.6885678848611 |
| 19 | 64 | 61.8477685552202 |
| 20 | 77 | 84.6898766546111 |

3) Значення критерію найменших квадратів та кореляційно-регресійний аналіз

Стандартні похибки параметрів :

Коефіцієнт детермінації: – близький до 1, побудована залежність є високоякісною.

Коефіцієнт кореляції: – згідно зі шкалою Чеддока, маємо дуже сильний зв'язок .

-статистика:

значення коефіцієнта кореляції істотне, модель адекватна.

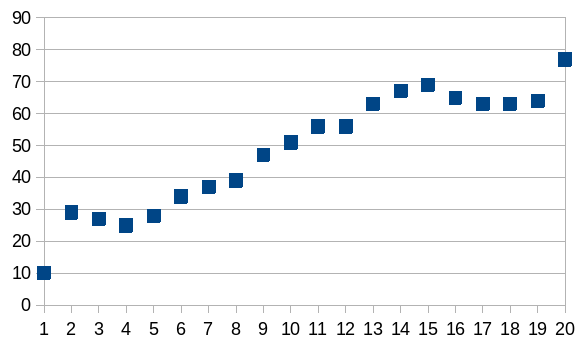
Статистичну значущість коефіцієнтів перевіримо за критерієм Стьюдента:

, статистично незначущі, статистично значущі.

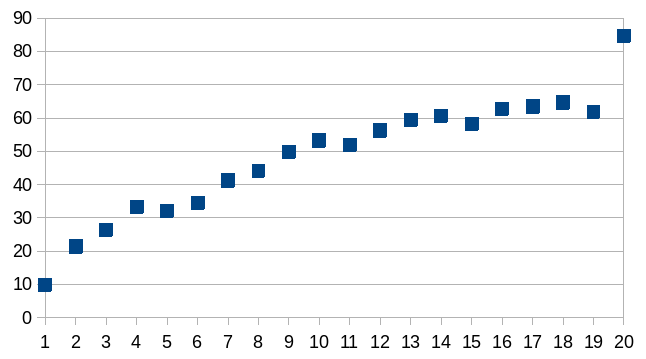
Довірчий інтервал для коефіцієнта :

Графіки функцій:

*Зафіксовані дані:*

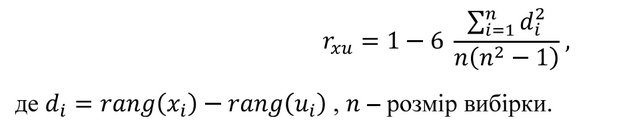


*Апроксимація:*

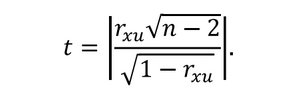


Графік функції за апроксимацією доволі близький до графіку функції за спостережуваними даними.

3) Гетероскедатичність (тест кореляції Спірмена)

Розрахуємо коефіціенти рангової кореляції Спірмена між значеннями незалежних змінних та залишками , використовуючи наступну формулу:  
  


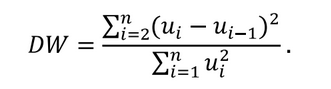
| Corr X1 ↔ y | Corr X2 ↔ y | Corr X3 ↔ y |
| --- | --- | --- |
| 0.109857066464056 | 0.0518992138996467 | 0.295822544605529 |

Знаходимо для кожної змінної відповідне значення t-критерію за наступною формулою:  
 

| 0.494 | 0.2261 | 1.4956 |
| --- | --- | --- |

Критичне значення: = 2.445  
Гетероскедатичність відсутня для кожної з змінних, т.я. значення критерію менше за критичне значення -> гіпотеза про наявність гетероскедатичності відкидається.

4) Автокореляція залишків (тест Дарбіна-Уотсона)

Вираховується критерій Дарбіна-Уотсона, який тісно пов’язаний з автокореляцією: DW~2(1- r ( - ) ).  
  


В нашому випадку DW = 1.314.  
  
Табличні межі критерію для 3 змінних та 20 спостережень для рівня значущості 0.05: DL = 1, DU = 1.68

Т.я. DL < DW < DU, то висновків про наявність або відсутність автокореляції зробити не можна.

**Висновок**

Незважаючи на те що серед змінних наявна мультиколінеарність, побудована модель є адекватною та має високі показники якості зв'язку, більшість її коефіцієнтів є статистично значущими, тому підсумковий графік функції, побудований за апроксимацією, є близьким до графіку функції, побудованого за спостережуваними даними. Гетероскедастичність відсутня, що є добре, а про наявність автокореляції висновків зробити не можна.