

# Економетрія

*Кафедра інформаційної економіки,  
підприємництва та фінансів*

*Лектор:*

*ГЛУЩЕВСЬКИЙ В'ячеслав Валентинович*

# ЗМІСТОВИЙ МОДУЛЬ 5

## ОСОБЛИВОСТІ ЗАСТОСУВАННЯ МНК ДЛЯ БАГАТОФАКТОРНИХ МОДЕЛЕЙ РЕГРЕСІЇ

### Теми 10-11

1. Мультиколінеарність у багаторефакторних моделях регресії.
2. Автокореляція у багаторефакторних моделях регресії.

# МУЛЬТИКОЛІНЕАРНІСТЬ У ЛІНІЙНИХ БАГАТОФАКТОРНИХ МОДЕЛЯХ РЕГРЕСІЇ:

## Мультиколінеарність -

наявність лінійних зв'язків між незалежними змінними  $X_i$  багаторфакторної моделі регресії

Суть мультиколінеарності полягає в тому, що в багаторфакторній регресійній моделі дві або більше незалежних змінних пов'язані між собою лінійною залежністю або, іншими словами, мають високий ступінь кореляції, тобто коли

для пари деяких змінних  $X_i$  і  $X_k$  парний коефіцієнт кореляції  $r_{ik} \rightarrow 1, i \neq k$

$$r_{ik} = r[X_i, X_k] = \frac{\sum_{j=1}^n [(x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)]}{\sqrt{\sum_{j=1}^n (x_{ji} - \bar{x}_i)^2} \cdot \sqrt{\sum_{j=1}^n (x_{jk} - \bar{x}_k)^2}}$$

# МУЛЬТИКОЛІНЕАРНІСТЬ У ЛІНІЙНИХ БАГАТОФАКТОРНИХ МОДЕЛЯХ РЕГРЕСІЇ:

$$\vec{a} = \left[ [X]^T [X] \right]^{-1} [X]^T \vec{Y}$$

оцінки невідомих параметрів  
множинної лінійної регресії за МНК

$$\left[ [X]^T [X] \right]$$

матриця спостережень

якщо

$$r_{ik} = r[X_i, X_k] \rightarrow 1$$

то визначник матриці спостережень  $\rightarrow 0$   
і відбувається фактичне ділення на нуль!

$$\left[ [X]^T [X] \right]^{-1} - \text{не існує}$$

оцінки невідомих параметрів регресії

мають суттєву похибку  
та є взагалі беззмистовними

ВИСНОВКИ, ЯКІ ОТРИМАНО З ВИКОРИСТАННЯМ  
ПОБУДОВАНОЇ МОДЕЛІ Є ХИБНИМИ!



# ТЕСТУВАННЯ НА ЯВНОСТІ МУЛЬТИКОЛІНЕАРНОСТІ

ЄДИНОГО СПОСОБУ ВИЗНАЧЕННЯ  
МУЛЬТИКОЛІНЕАРНОСТІ, НАЖАЛЬ, НЕ МАЄ!



**НАЙПОВНІШЕ ДОСЛІДИТИ МУЛЬТИКОЛІНЕАРНІСТЬ**  
дає змогу

**алгоритм ФАРРАРА-ГЛОБЕРА**



поєднує три види статистичних критеріїв:

1. Виявлення мультиколінеарності  
всього масиву незалежних змінних

 критерій

$\chi^2$

2. Виявлення мультиколінеарності  
кожної незалежної змінної з усіма іншими

 критерій  
Фішера

$F$

3. Виявлення колінеарності  
кожної пари незалежних змінних

 критерій  
Стюдента

$t$

# АЛГОРИТМ ФАРРАРА-ГЛОБЕРА

1. Побудувати кореляційну матрицю на основі статистичних даних для незалежних змінних:

$$R = [r_{ik}] = \begin{bmatrix} r_{11} & \cdots & r_{1k} & \cdots & r_{1m} \\ r_{21} & \cdots & r_{2k} & \cdots & r_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{i1} & \cdots & r_{ik} & \cdots & r_{im} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{m1} & \cdots & r_{mk} & \cdots & r_{mm} \end{bmatrix}$$

$$r_{ik} = r[X_i, X_k], r_{ii} = 1$$

— парні коефіцієнти кореляції

2. Знайти визначник кореляційної матриці  $|R|$  та обчислити критерій  $\chi^2_{\text{розрах}}$

$$\chi^2_{\text{розрах}} = - \left[ n - 1 - \frac{1}{6} (2m + 5) \cdot \ln |R| \right]$$

та порівняти його з табличним значенням: якщо  $\chi^2_{\text{розрах}} > \chi^2_{\text{табл}}$

то в масиві незалежних змінних існує мультиколінеарність.

# АЛГОРИТМ ФАРРАРА-ГЛОБЕРА

3. Перевірити, які само з незалежних змінних  $X_k$  мультиколінеарна з іншими:

3.1. Побудувати матрицю похибок

$$Z = [R]^{-1} = [z_{ik}] = \begin{bmatrix} z_{11} & \cdots & z_{1k} & \cdots & z_{1m} \\ z_{21} & \cdots & z_{2k} & \cdots & z_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z_{i1} & \cdots & z_{ik} & \cdots & z_{im} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z_{m1} & \cdots & z_{mk} & \cdots & z_{mm} \end{bmatrix}$$

незалежна змінна  $X_k$   
мультиколінеарна з іншими

3.2. Провести  $F$ -тест

$$F_k = \frac{(z_{kk} - 1)(n - m)}{(m - 1)}$$

$$F_k \geq F_{табл}$$

так      ні

незалежна змінна  $X_k$   
не мультиколінеарна з іншими

4. Перевірити, які само з незалежних змінних  $X_i$  попарно колінеарні з  $X_k$ :

4.1. Знайти часткові коефіцієнти кореляції:

$$r_{ki} = \frac{-z_{ki}}{\sqrt{z_{kk} z_{ii}}}$$

$$k, i = \overline{1, m}, k \neq i$$

4.2. Провести  $t$ -тест

$$t_{ki} = |r_{ki}| \cdot \frac{\sqrt{n - m}}{\sqrt{1 - r_{ki}^2}}$$

$$t_{ki} \geq t_{табл}$$

так      ні

між  $X_k$  та  $X_i$  існує мультиколінеарність

мультиколінеарність  
між  $X_k$  та  $X_i$  відсутня

УСУНЕННЯ МУЛЬТИКОЛІНЕАРНОСТІ

# СПОСОБИ УСУНЕННЯ МУЛЬТИКОЛІНЕАРНОСТІ:

1. Використання додаткової або первинної інформації.
2. Об'єднання інформації.
3. Збільшення кількості спостережень.
4. Відкидання змінної  $X_k$  з високою кореляцією.
5. Перетворення даних (використання перших різниць):  $X_k^* = X_k - X_i$ .

Примітка. Якщо способами 1-5 не вдається усунути мультиколінеарність, то для оцінювання параметрів множинної лінійної регресії доцільно застосувати *метод головних компонент*.