

ГЛАВА 8.

ПЕРЕВІРКА РЕГРЕСІЙНОЇ МОДЕЛІ

8.1. Статистичні властивості класичної регресійної моделі

В результаті використання методу найменших квадратів визначають коефіцієнти регресії. Нижче розглядається питання оцінки властивостей отриманих коефіцієнтів. Функцією оцінки (оцінник) зветься

Статистичні властивості МНК – оцінника $\hat{\beta}_j$ в лінійній класичній регресійній моделі (тут розглядається як вектор випадкових змінних) є такими:

1. **Лінійність** по y або лінійна залежність (середнього значення) математичного очікування оцінки $\hat{\beta}_j$ від функції мети y :

$$\mathbb{E}(\hat{\beta}_j) = (X^T X)^{-1} X^T y.$$

2. **Незміщеність**: оцінник $\hat{\beta}_j$ дійсного значення β_j є незміщеним, тобто його математичне очікування дорівнює дійсному значенню, тобто $\mathbb{E}(\hat{\beta}_j) = \beta_j$ при будь-якому об'ємі виборки T . Зміщена оцінка має густину розподілу, максимум якої може знаходитись ліворуч або праворуч від математичного очікування значення коефіцієнта регресії.

3. **Покомпонентна мінімізація дисперсій**. Математично доведено, що МНК-оцінник дає мінімальне значення дисперсії, або мінімальну середню квадратичну похибку (теорема Гаусса-Маркова).

Якщо зміщена оцінка має меншу дисперсію порівняно з незміщеною, то перевага надається змішаній оцінці з меншою дисперсією.

4. Асимптотична незміщеність оцінника: оцінка вважається асимптотично незмінною, якщо виконується умова

$$\lim_{T \rightarrow \infty} E(\hat{v}_j) = v_j,$$

де T визначає об'єм виборки (довжину ряду даних).

При збільшенні T (збільшенні довжини ряду даних) наближається до істинної величини параметра β .

Користувача цікавить, починаючи з якого об'єму виборки має місце асимптотична незміщеність, тобто починаючи з якої довжини ряду даних T зміщеність буде незначною. При цьому не розглядається дисперсія: можливо, що зі зростанням об'єму вибірки T величина зміщення зменшиться, але дисперсія зростає.

5. Стохастична збіжність та обґрунтування ІМНК-оцінник
Вважається, що послідовність функцій оцінювання для \hat{v}_{jT} стохастично збігається з числом v_j та є обґрунтованою, якщо для будь-якого досить малого $\xi > 0$ ймовірність P виконання умови при об'ємі виборки $T \rightarrow \infty$ дорівнює "1":

$$\lim_{T \rightarrow \infty} P(|\hat{v}_{jT} - v_j| \leq \xi) = 1,$$

де P – ймовірність події $T \rightarrow \infty$ – об'єм вибірки (довжина ряду);
– абсолютне значення.

Еквівалентним може бути запис

$$\lim_{T \rightarrow \infty} P(|\hat{v}_{jT} - v_j| > \xi) = 0.$$

6. Регресійна модель справедлива для середніх точок, тобто з рівняння

$$y_t = \sum_{j=1}^K \beta_j x_{jt} + u_t$$

можна написати

$$\bar{y} = \sum_{j=1}^K \beta_j \bar{x}_j,$$

звідки

$$\beta_1 = \bar{y} - \sum_{j=2}^K \beta_j \bar{x}_j.$$

7. Середнє значення функції мети (регресанда, залежної змінної) дорівнює середньому значенню виміряних вихідних фактичних значень

$$\begin{aligned} \bar{y} &= \frac{1}{T} \sum_{t=1}^T (y_t + u_t) = \frac{1}{T} \sum_{t=1}^T \sum_{j=1}^K \beta_j x_{jt} + \frac{1}{T} \sum_{t=1}^T u_t = \\ &= \frac{1}{T} \sum_{t=1}^T (\beta_1 + \sum_{j=2}^K \beta_j x_{jt}) = \\ &= \frac{1}{T} \sum_{t=1}^T (\bar{y} - \sum_{j=2}^K \beta_j \bar{x}_j + \sum_{j=2}^K \beta_j \bar{x}) = \frac{1}{T} \sum_{t=1}^T \bar{y} = \bar{y}. \end{aligned}$$

Тут ми врахували, що

$$\frac{1}{T} \sum_{t=1}^T u_t = 0, \quad y_t = \sum_{j=1}^K \beta_j x_{jt}, \quad \beta_1 = \bar{y} - \sum_{j=2}^K \beta_j \bar{x}_j.$$

8. Сума похибок дорівнює нулю:

$$\sum_{t=1}^T u_t = 0.$$

Похибки є некорельовані з $x_{1t}, x_{2t}, \dots, x_{Kt}$ та u_t :

$$\sum_{t=1}^T u_t x_{1t} = \sum_{t=1}^T u_t x_{2t} = \sum_{t=1}^T u_t x_{3t} = \dots = \sum_{t=1}^T u_t x_{Kt} = 0;$$

$$\sum_{t=1}^T u_t y_t = 0.$$

Таким чином, оцінник $\beta = f(x, y)$, отриманий на базі 1МНК, є

ідеальним методом оцінки класичної регресійної моделі. Але слід також вказати, що в економічних та соціальних процесах такі умови до регресорів x , як некорельованість, не витримуються. Тому використання 1МНК для аналізу є досить формальним. Але, з іншого боку, 1МНК та інші методи, які базуються на ньому, в даний час є основою для розвитку узагальнених моделей економетрії, соціометрії та ін. Тут треба лише пам'ятати, що не статистичні та вимірні дані SFQ у $\sum_{t=1}^T y_t^2$ для ілюстрації особливостей 1МНК, а навпаки: 1МНК треба грамотно використовувати та пристосовувати до спостережених даних.

8.2. Коефіцієнти детермінації для оцінки адекватності моделі

Оцінити адекватність різних моделей (конкуруючих варіантів рішень) можна за допомогою суми квадратів помилок

$y_{\text{рег}}^2 = \frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^2$
але верхня припустима межа цього рівняння відсутня (нижча межа відома – вона дорівнює нулю). Цей недолік SFQ усувається використанням коефіцієнта детермінації R^2 , значення якого вказує, наскільки значення регресанта \hat{y}_t , отримані за допомогою 1МНК, відхиляються від реальних спостережень y_t . Коефіцієнт детермінації можна записати у вигляді відношення

$$R^2 = y_{\text{рег}}^2 / y_{\text{заг}}^2, \quad 0 \leq R_1^2 \leq 1;$$

де $y_{\text{рег}}^2$ – дисперсія, яка пояснюється регресією;

$y_{\text{заг}}^2$ – загальна дисперсія.

Нижче наведені три рівноцінні коефіцієнти детермінації R^2 , які часто використовуються для оцінки адекватності:

$$R_1^2 = \frac{[\sum_{t=1}^T (y_t - \bar{y})(\hat{y}_t - \bar{y})]^2}{\sum_{t=1}^T (y_t - \bar{y})^2 \sum_{t=1}^T (\hat{y}_t - \bar{y})^2} \leq 1,$$

$$R_2^2 = \frac{\sum_{t=1}^T (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^T (y_t - \bar{y})^2};$$

$$R_3^2 = 1 - \frac{\sum_{t=1}^T \hat{u}_t^2}{\sum_{t=1}^T (y_t - \bar{y})^2};$$

де y_t – теоретичні значення регресанда (спостережений ряд y_t); \hat{y}_t – розраховані (оцінені) значення регресанда; \bar{y} – середнє арифметичне значення ряду y_t ; \hat{u}_t – похибка t -го спостереження.

При $R^2 = 1$ отримуємо випадок повної адекватності: всі спостережені значення знаходяться на регресійній гіперповерхні. При $R^2 = 0$ функція регресії нічого не показує.

Недоліком коефіцієнта детермінації R^2 є його збільшення разом із збільшенням числа регресорів. Вважається, що зі збільшенням числа регресорів модель більш адекватно оцінює процес. Але з кожним додатковим регресором втрачається один ступінь свободи (при числі спостережень $T = \text{const}$) і збільшується дисперсія. Цей недолік не враховує при обчисленні R^2 . Кількість ступенів свободи $FG = T - K$. При збільшенні K на “1” (введення додаткового регресора) FG зменшується на “1” і зростає дисперсія.

При використанні t -та F -тестів, а також при побудові довірчих і прогнозних інтервалів бажано мати (при рівних інших умовах) як можна більше ступенів свободи FG .

При цьому довірчі та прогнозні інтервали будуть тим меншими, чим більше FG . Тому в статистичному відношенні наявність додаткового регресанда може бути не завжди бажаною. Скориговані коефіцієнти детермінації враховують зменшення ступенів свободи при збільшенні числа регресорів:

- скоригований коефіцієнт детермінації по Тейлу (Theil):

$$R_T^2 = 1 - \left(1 - R^2\right) \frac{T-1}{T-K};$$

- скоригований коефіцієнт детермінації по Амемії (Amemiya):

Тут коефіцієнт “підштовхує” до обрання рівняння з меншим числом регресорів.

Якщо показники не узгоджуються між собою, то кінцевий вибір рівняння регресії повинен виконуватись з урахуванням інших статистичних показників (t -, F -статистик та ін.).

Примітка: якщо в регресійному рівнянні змінюється кількість регресорів, то змінюються оцінки всіх коефіцієнтів регресії а іноді – знаки перед ними. Змінюються також оцінки інших показників моделі (вектори скаляри).

8.3. Дисперсно-коваріаційна матриця для коефіцієнтів регресії

Коваріаційна матриця (дисперсно-коваріаційна матриця) використовується для: $V = (X^T X)^{-1} X^T Y_t$

- МНК-оцінки дисперсії коефіцієнтів
- МНК-оцінки коваріації (зв'язку) між коефіцієнтами шляхом використання коефіцієнта кореляції $Y_t = X\beta + U_t$;

• t -тестування і розрахунків довірчих інтервалів.

Раніше ми вказували, що коефіцієнти не є стохастичними, бо єдиною стохастичною величиною у рівнянні регресії вважали помилку U_t . В дійсності отримане нами раніше матричне рівняння

$$(8.3.1)$$

вміщує спостережену функцію мети, яку можна розглядати як

$$V_\beta = (X^T X)^{-1} X^T U_t \tag{8.3.2}$$

де β – матриця дійсних коефіцієнтів, які оцінюються матрицею оцінок коефіцієнтів U_t – вектор помилок при визначенні \hat{y}_t .

Підставимо значення (8.3.2) в (8.3.1) і виконаємо перетворення:

$$V = (X^T X)^{-1} X^T (X\beta + U_t) = (X^T X)^{-1} X^T X\beta + (X^T X)^{-1} X^T U_t =$$

$$= \frac{1}{T-K} \left[(X^T X)^{-1} X^T U \right] * \left[(X^T X)^{-1} X^T U \right] =$$

$$= \beta + (X^T X)^{-1} X^T U = \beta + V_\beta,$$

де V_β – вектор оцінених помилок

Тоді дисперсно-коваріаційна матриця V_β , яка враховує дисперсію при розрахунках β може розглядатись так $V_\beta = \sigma_u^2 (X^T X)^{-1}$ – читається: “Велика сигма з кутом, унизу мала бета з кутом”):

$$= \frac{1}{T-K} \left[(X^T X)^{-1} X^T U \right] * U^T \left[(X^T X)^{-1} X^T \right]^T =$$

$$= \frac{1}{T-K} \left[(X^T X)^{-1} X^T U \right] * U^T X \left[(X^T X)^{-1} \right]^T =$$

$$\frac{1}{T-K} U^T U = \sigma_u^2 = \frac{1}{T-K} \sum_{i=1}^T u_i^2; I = (X^T X)^{-1} X^T X; y_u^2$$

$$= \frac{1}{T-K} \left[(X^T X)^{-1} X^T U \right] * U^T X \left[(X^T X)^{-1} \right]^T =$$

$$\left[(X^T X)^{-1} \right]^T = (X^T X)^{-1}$$

$$= \frac{1}{T-K} (X^T X)^{-1} X^T [U_i U_i^T] X (X^T X)^{-1} =$$

$$\frac{1}{T-K} \left(\frac{1}{T-K} U_i U_i^T \right) (X^T X)^{-1} X^T X * (X^T X)^{-1} = \sum_{i=1}^T \frac{1}{T-K} \frac{1}{T-K} U_i U_i^T$$

$$y_{\beta_i \beta_j} = y_u^2 (X^T X)^{-1} = \begin{bmatrix} y_{\beta_1 \beta_1}^2 & y_{\beta_1 \beta_2} & \dots & y_{\beta_1 \beta_K} \\ y_{\beta_2 \beta_1} & y_{\beta_2 \beta_2}^2 & \dots & y_{\beta_2 \beta_K} \\ \dots & \dots & \dots & \dots \\ y_{\beta_K \beta_1} & y_{\beta_K \beta_2} & \dots & y_{\beta_K \beta_K}^2 \end{bmatrix},$$

де $-1 \leq r_{\beta_i \beta_j} \leq 1$.

значення $y_{\beta_i \beta_j}$, що показує середню суму квадратів відхилень членів ряду y_i від його середніх значень (математичного очікування) \bar{y}_i ; I – одинична квадратна матриця з одиницями в діагоналі та нульовими іншими елементами. Цю одиничну матрицю з β_j рівняння можна вилучити (замінити на скаляр “1”).

Тут прийнято, що $r_{\beta_i \beta_i} = +1$, – дисперсія ви-
симетричність матриці $r_{\beta_i \beta_j}$.

При розрахунку дисперсно-коваріаційної матриці використані відомі перетворення при транспонуванні матриць:

$$(A * B)^T = B^T A^T; \quad (A^T)^T = A.$$

На головній діагоналі дисперсно-коваріаційної матриці елементи є 1МНК-оцінником дисперсії регресійного коефіцієнта β_i . Елементи (за головної діагоналю) є 1МНК-оцінником $\beta_i \beta_j$ (зв’язку між коефіцієнтами β_i та β_j), $y_i = \sum_{j=1}^k \beta_j x_{ij}$.

За даними цієї матриці може бути розрахований коефіцієнт кореляції між двома коефіцієнтами регресії β_i та β_j :

$$-1 \leq r_{\beta_i \beta_j} = \frac{y_{\beta_i \beta_j}}{y_{\beta_i \beta_i} y_{\beta_j \beta_j}} \leq +1, \tag{8.4.1}$$

де

Якщо коефіцієнт кореляції $r_{\beta_i \beta_j}$ то ми отримали систему з дуже позитивним зв’язком. При коефіцієнті кореляції $r_{\beta_i \beta_j} < 0$ ми отримали систему з негативним (зімним) зв’язком між коефіцієнтами β_i та β_j . При взаємному зв’язку та впливу у коефіцієнтів регресії немає.

Доведено, що в класичній моделі з нормальним розподілом збурень U_t коефіцієнти розподілені нормально, а коефіцієнт детермінації дорівнює квадрату коефіцієнта кореляції $R^2 = r^2$.

8.4. t -тест

За допомогою t -тесту можна відповісти на запитання: чи вірно визначено коефіцієнти регресії, чи не дорівнюють вони, наприклад, нулю?

Для цього при t -тестуванні розглядають дві гіпотези:

1. Нульова гіпотеза: H_0 Звичайно

2. Альтернативна гіпотеза:

Тут β_j – один з коефіцієнтів регресії, а t_j – будь-яка визначена дослідником величина; вона може бути довільною або може бути визначеною за якимись правилами.

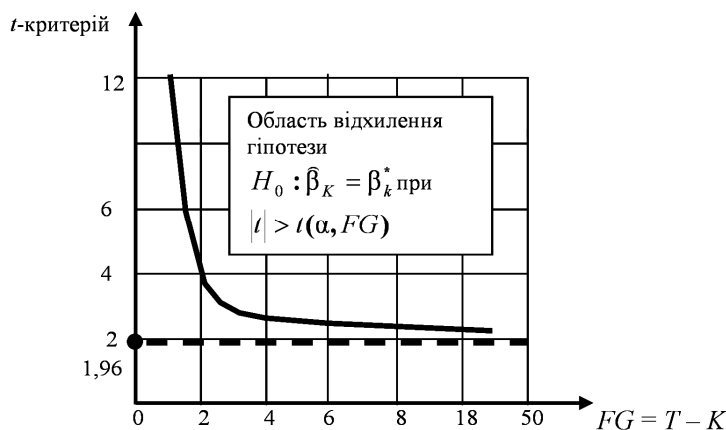
Для прийняття гіпотези розраховується t -статистика (розрахунок дає нам реалізацію центральної t -розподіленої випадкової величини):

$$|t_j| > t(\alpha, FG) \quad t_j = \frac{\hat{\beta}_j - \beta_j^*}{\sqrt{\hat{\sigma}^2 \cdot \frac{1}{B_j}}}$$

де $\hat{\beta}_j$ беруть як корінь квадратний від діагонального елемента дисперсно-коваріаційної матриці $\hat{\sigma}^2$ – оцінка коефіцієнта регресії, отримана методом ІМНК для рівняння регресії $y = \beta_j x + t$ – випадкова змінна, яка підпорядковується центральному t -розподілу з кількістю ступенів свободи $FG = T - K$.

Отримана t -статистика може бути додатково порівняна з критичними значеннями (бо коефіцієнт β_j може бути додатковим). Таблиця t -критерія Стьюдента ($FG = T - K$) регресії

FG	α			
	0,5	0,05	0,01	0,001
1	1,000	12,706	63,657	636,619
2	0,816	4,303	9,925	31,598
3	0,765	3,182	5,841	12,941
4	0,741	2,776	4,604	8,610
10	0,700	2,228	3,169	4,587
20	0,687	2,086	2,845	3,850
30	0,683	2,042	2,750	3,646
200	0,676	1,972	2,601	3,340
-	0,675	1,960	2,576	3,291

Рис. 8.4.1. Область відхилення гіпотези H_0 при $\alpha = 0,05$

Практичне правило: при зростанні числа ступенів свободи $FG = T - K > 4$ та при розрахованій величині t -статистики $|t| > 2$, ($\alpha = 0,05$) нульову гіпотезу H_0 можна відхилити.

Іноді потрібно перевірити твердження: чи перевищує коефіцієнт β_j (це невідома величина) визначену величину β_j^* . В цьому випадку формують **односторонні нульові та альтернативні гіпотези**.

Всі ці гіпотези можуть бути відображені в табл. 8.4.2.

Таблиця 8.4.2

Односторонні нульові та альтернативні гіпотези

№	Нульова гіпотеза H_0	Альтернативна гіпотеза H_A	Тест	Область відхилення гіпотези H_0
1	$H_0 : \beta_j = \beta_j^*$	$H_A : \beta_j \neq \beta_j^*$	Двосторонній	$ t > t(\alpha, FG)$
2	$H_0 : \beta_j \leq \beta_j^*$	$H_A : \beta_j > \beta_j^*$	Односторонній	$t > t(2\alpha, FG)$
3	$H_0 : \beta_j \geq \beta_j^*$	$H_A : \beta_j < \beta_j^*$	Односторонній	$t < -t(2\alpha, FG)$

Односторонній тест проводиться подібно до двостороннього, але область прийняття рішення тепер інша, бо виконуються інші дії з таблицею t -критеріїв:

- для двостороннього тесту з рівнем значущості α обрати $t(\alpha, FG)$;
- для одностороннього тесту з рівнем значущості α обрати $t(2\alpha, FG)$.

8.5. Точкові та інтервальні прогнози регресанда

Раніше ми отримали ІМНК-оцінник регресанда

$$\hat{y}_t = \sum_{j=1}^K \hat{\beta}_j x_{jt}$$

Якщо ми отримали часові спостереження за $t = 1, 2, \dots, T$ минулих періодів, то ми можемо отримати точкові прогнози: в минулому; в базовому періоді; в майбутньому. Прогнози для минулого (в тому числі і для базового періоду) не є дійсними прогнозами в точному значенні цього слова: вони можуть використовуватись, наприклад, для інтерполяції – отримання відсутніх даних для деякого діапазону минулого часу.

Справжні прогнози у часі знаходяться за межами періоду спостережень (він також зветься базовим періодом) при $t > T$.

Якщо визначити вектор факторів для одного спостереження (тут $x_{m1t} = 1 = \text{const}$)

$$X_{Mt} = (x_{M1t}, x_{M2t}, \dots, x_{Mjt}, \dots, x_{MKt})$$

для майбутнього періоду, то ми отримуємо точковий прогноз (його математичне очікування) за межами базового періоду (ex-post-прогнози для минулого, ex-ante-прогнози для майбутнього).

Ознакою точкового прогнозу є те, що це – одне число. При прогнозуванні індекс t означає майбутній період або додатковий елемент $t > T$, у той час як базовий період має

Ця оцінена похибка прогнозу використовується при побудові інтервального прогнозу для $E(y_t)$ (інтервалу довіри).

Інтервальний прогноз математичного очікування регресанда (інтервал довіри для математичного сподівання залежної змінної y_t) розраховується за формулою

$$\hat{y}_{Mt,2,3} = \hat{y}_{Mt} \pm t(\alpha; FG) * \hat{y}_{Me}, \quad (8.5.1)$$

$$\hat{y}_{Me} = \sqrt{\hat{y}_u^2 X_{Mt} (X^T X)^{-1} X_{Mt}^T}; \quad \hat{y}_u^2 = \frac{1}{T-K} \sum_{t=1}^T u_t^2; \quad \hat{y}_{Mt} = \sum_{j=1}^K \hat{\beta}_j x_{Mjt};$$

$$X_{Mt} = (x_{M1t}, x_{M2t}, \dots, x_{Mjt}, \dots, x_{MKt})$$

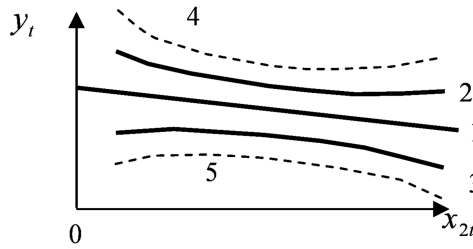
\hat{y}_{Me} – вектор регресорів для майбутнього періоду для однієї точки спостереження; X, X^T – матриця спостережених значень регресорів для базового періоду; \hat{y}_u^2 – дисперсія базового періоду; $t(\alpha; FG)$ – береться із таблиці t -критерія для рівня значущості α (наприклад, $\alpha = 0,05$) і числа ступенів

свободи $FG = T - K$.

$$\hat{y}_t = \hat{b}_1 * 1 + \hat{b}_2 x_{t2},$$

Якщо розглядати рівняння першого порядку то геометрична інтерпретація прогнозних інтервалів наведена на рис. 8.5.1.

У формулі (8.5.1) знак «+» відноситься до верхньої межі 2 прогнозного інтервалу математичного очікування регресанда а знак «-» - до нижньої межі 3 цього інтервалу (рис. 8.5.1). На рис. 8.5.1 додатково позначено: 1 - математичне очікування



4, 5 - межі прогнозного інтервалу

індивідуального значення регресанда.

Інтервальный прогноз індивідуального значення регресанда 4 та 5 (рис. 8.5.1) (інтервал довіри для індивідуального значення залежної змінної y_t при крапковому прогнозі окремої реалізації) рахується за аналогічною формулою

$$\hat{y}_{Mei} = \sqrt{\hat{y}_u^2 + \hat{y}_{Me}^2} = \sqrt{\hat{y}_u^2 + \hat{y}_{Me}^2 + X_{Mt}^T (B^{-1} - K G^{-1} K^T) X_{Mt}} \hat{y}_u,$$

$$\hat{y}_u^2 > 0,$$

де

Оскільки то при інших рівних умовах інтервал індивідуального значення регресанда (криві 4 та 5) буде завжди більше величини математичного очікування регресанда (криві 2, 3).

Для інтервального прогнозу \hat{y}_{Mei} індивідуального значення y_{Mt} регресанда y_{Mt} (тобто при оцінці реалізації регресанда y_{Mt}) похибка прогнозу e_{Mei}

Тут y_{Mt} є реалізацією випадкової змінної y_{Mt} , яка в ІМНК має таку ж дисперсію як і змінна похибок U_t . Але компонента для математичного очікування має дисперсію. Оскільки та y_{Mt} не корелюються, то дисперсія похибки прогнозу e_{Mei} реалізації дорівнює:

$$\hat{y}_{Mei}^2 = \hat{y}_u^2 + \hat{y}_{Me}^2 = \hat{y}_u^2 + \hat{y}_u^2 X_{Mt}^T (X^T X)^{-1} X_{Mt}^T.$$

На основі того, що в класичній лінійній моделі нормальної

регресії похибка математичного очікування \hat{y}_{Mt} ,

$$e_{Mt} = E(y_{Mt}) - \hat{y}_{Mt} = E(y_{Mt}) - X_{Mt} \hat{\mathbf{b}}$$

є лінійною функцією нормально розподіленої змінної y_{Mt} то можна вважати, що ця похибка теж розподілена за нормальним законом і має відповідну дисперсію

$$e_{Mt} \sim N \left[E(y_{Mt}) - X_{Mt} \hat{\mathbf{b}}; X_{Mt} \sum_{b_j} X_{Mt}^T \right].$$

Тут істинна дисперсійно-коваріаційна матриця $\Sigma_{\mathbf{b}}$ невідома і її змінюють на оцінку

$$\hat{\Sigma}_{b_j} = \hat{y}_u^2 (X^T X)^{-1}.$$

Тоді оцінена дисперсія похибки прогнозу для $E(Y_i)$

$$\hat{y}_{Me}^2 = X_{Mt} \hat{\Sigma}_{b_j} X_{Mt}^T \hat{y}_u^2 (X^T X)^{-1} X_{Mt}^T = \hat{y}_u^2 X_{Mt} (X^T X)^{-1} X_{Mt}^T,$$

тому що скаляр \hat{y}_u^2 може бути винесений за межі матриць.

Відповідно оцінена стандартна похибка прогнозу математичного очікування

Контрольні завдання

Скласти таблицю спостережень регресанда та регресорів з $R_1^2; R_2^2; R_3^2; R_T^2; R_A^2$

Таблиця 8.5.1

Дані для розрахунку коефіцієнтів детермінації

t	y_t	x_{2t}	\hat{y}_t
1	0,6*N	0	0,5*N
2	1,4*N	2	1,0*N
3	1,2*N	4	1,5*N
4	1,8*N	6	2,0*N
5	2,4*N	8	2,5*N
6	3,1*N	10	3,0*N

Таблиця 8.5.2
Дані для розрахунку коефіцієнтів детермінації

t	y_t	x_{2t}	\hat{y}_t
1	2,6°N	0	2,5°N
2	3,6°N	1	4,0°N
3	5,7°N	2	5,5°N
4	7,2°N	3	7,0°N
5	8,8°N	4	8,5°N
6	9,6°N	5	10,0°N

2. Виконати t -тестування коефіцієнтів регресійного рівняння з попереднього індивідуального завдання.

3. Отримати точкові та інтервальні прогнози регресанда.

4. Розрахувати дисперсійно-коваріаційну матрицю для одного з отриманих регресандів. Визначити коефіцієнти кореляції між коефіцієнтами B_{11} , B_{12} .

5. Виконати t -тест для отриманого регресанда.

8.6. F -тест гіпотез для груп регресійних коефіцієнтів та лінійних комбінацій

8.6.1. Загальні відомості про F -тест (тест Фішера)

Якщо дві змінні мають незалежні x -квадрат розподіли x_1^2 та x_2^2 з V_1 та V_2 ступенями вільності, то F -статистика

$$F = \frac{x_1^2 / V_1}{x_2^2 / V_2} = t^2$$

має F -розподіл Фішера з V_1 та V_2 ступенями вільності (тут t^2 – це квадрат t -статистики Стьюдента).

Іноді F -статистику називають відношенням дисперсій. F -статистику найчастіше використовують для тестування рівності цих незалежних оцінок дисперсій y_1 та y_2 з V_1 та V_2 ступенями вільності за

формулою $F = \frac{y_1^2}{y_2^2}$ – відношення дисперсій.

Якщо дві оцінки дисперсії \hat{y}_1^2 та \hat{y}_2^2 близькі між собою, то їх відношення наближується до "1". Чим більше різниця між двома дисперсіями, тим більше значення F -відношення відрізняється від "1". У загальному випадку значення F -відношення є завжди додатним і знаходиться у межах $0 \leq F \leq \infty$). Для перевірки регресійної моделі (однофакторної та багатфакторної, у яких $x_{1t} = 1 = const$)

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_K x_{Kt} + u_t$$

використовують такий алгоритм $H_0: \beta_1 = \beta_2 = \dots = \beta_K = 0$ тестування.

1. Формується нуль-гіпотеза H_0 на нуль належить розглядати не лише як тест на значення коефіцієнтів регресії, але й як відповідь на питання: чи впливають обрані змінні x_{it} на значення регресанта y_t .

Альтернативна гіпотеза H_A стверджує: якщо нульова гіпотеза H_0 відхиляється (не приймається), то хоча б одне значення (і відповідне значення фактора x_{it}) статистично значущо впливає на величину регресанда y_t . Тобто у цьому випадку не всі коефіцієнти приблизно дорівнюють нулю.

2. Задається рівень значущості $\alpha * 100\%$ (наприклад, 5%), де α вказує ймовірність помилки в оцінці значущості коефіцієнтів.

3. Розраховується F -статистика Фішера з $(K - 1)$ та $(T - K)$ ступенями вільності

$$F_{K-1, T-K} = \frac{\sum_{t=1}^T (\hat{y}_t - \bar{y})^2 / (K-1)}{\sum_{t=1}^T (y_t - \hat{y}_t)^2 / (T-K)} = \frac{(T-K) * SSR}{(K-1) * SSE} = \frac{R^2 / (K-1)}{(1-R^2) / (T-K)},$$

де $SSE = \sum_{t=1}^T (y_t - \hat{y}_t)^2 = \sum_{t=1}^T u_t^2$;

де $(K - 1), (T - K)$ – ступені вільності; K – кількість

коефіцієнтів регресії які увійшли у модель; T – загальна кількість спостережень; R^2 – коефіцієнт детермінації.

4. З таблиці розподілу Фішера знаходимо F -критерій при заданому рівні значущості (або помилки) та з $(K - 1), (T - K)$ ступнями вільності $F_{0,95}(K - 1, T - K)$.

5. Нульова гіпотеза відхиляється з $\alpha * 100\%$ -им ризиком помилитися, якщо $F_{K-1, T-K} > F_{0,95}(K - 1, T - K)$.

F-тест має недолік у порівнянні з t -тестом: він перевіряє лише двосторонні гіпотези, у той час як t -тест перевіряє односторонні та двосторонні гіпотези.

Але **F-тест має і переваги** перед t -тестом: він статистично перевіряє гіпотези про значення **кількох коефіцієнтів** або **кількох лінійних комбінацій**, а також сполучень того та іншого. Між тим t -тест завжди відноситься до одного параметра однієї лінійної комбінації.

Нижче ми опишемо три двосторонні гіпотези (три пари гіпотез):

1. Двостороння гіпотеза про значення одного, двох або кількох регресійних коефіцієнтів (t -тест перевіряє лише один регресійний коефіцієнт).
2. Двостороння гіпотеза про значення однієї, двох або кількох лінійних комбінацій регресійних коефіцієнтів (t -тест перевіряє лише одну лінійну комбінацію).
3. Сукупність гіпотез про значення регресійних коефіцієнтів та їх лінійних комбінацій (t -тест таку задачу не розв'язує).

Всі ці вищезазвані гіпотези є трьома версіями загальної лінійної гіпотези.

8.6.2. Загальна лінійна гіпотеза для F-тесту.

Загальна лінійна гіпотеза F -тесту (складається з нульової гіпотези H_0 та альтернативної гіпотези H_A), яка перевіряється за допомогою F -тесту у класичній лінійній моделі нормальної регресії, має вигляд:

$$H_0: C\beta = C^*; \quad (8.6.2.1)$$

$$H_A: C\beta \neq C^*; \quad (8.6.2.2)$$

де C^* – вектор-колонка з елементами (v_1, v_2, \dots, v_k) . Значення елементів задаються на основі професійно-теоретичних міркувань; C – матриця розмірності $m \times k$, яка складається з одиниць $C_{j1}v_1 + \dots + C_{jk}v_k + \dots + C_{jK}v_K = C_{j*}v^*$, де v^* – вектор-колонка регресійних коефіцієнтів, які потрібно оцінювати і які статистично перевіряються гіпотезами; C – матриця розмірності $m \times k$, яка складається з одиниць $C_{j1}v_1 + \dots + C_{jk}v_k + \dots + C_{jK}v_K = C_{j*}v^*$, де v^* – вектор-колонка регресійних коефіцієнтів, які потрібно оцінювати і які статистично перевіряються гіпотезами; m – кількість рядків у матриці C , або кількість лінійних рівнянь (лінійних гіпотез), які сумісно створюють групу H_0 та H_A – гіпотез F -тесту.

У матричному рівнянні (8.6.2.1) окреме j -те рівняння має вигляд

$$(8.6.2.3)$$

Несуперечливість всіх цих m лінійних рівнянь (лінійних гіпотез) забезпечується лише у тому випадку, якщо ранг матриці C дорівнює

m , тобто кількість рядків даної матриці $r(C) = m$. В подальшому вважаємо, що ця умова виконується.

Приклад 8.6.2.1. В певних C та C^* для перевірених t -тестом гіпотез ($H_0: \beta_2 = 0$) та ($H_A: \beta_2 \neq 0$) будуть мати такі значення:

$$C = (010); C^* = 0.$$

Приклад 8.6.2.2. Якщо взяти $C = \begin{bmatrix} 010 \\ 110 \end{bmatrix}$; $C^* = 1$, то в цьому випадку відповідає перевіреним t -тестуванням гіпотеза про лінійну комбінацію $\beta_1 + \beta_2 = 1$.

Приклад 8.6.2.3. $C =$ $C^* =$ Ця лінійна гіпотеза з'єднує

(комбінує) наведені у попередніх прикладах гіпотези про $\beta_2 = 0$ та $\beta_1 + \beta_2 = 1$, що не може бути перевірено t -тестом.

Якщо $m = 1$ (як в обох вищезгаданих випадках, то F -тест приводить до результатів, аналогічних t -тесту.

Виникає питання: чи не можна при $m \geq 2$ замість F -тесту використати двосторонній t -тест для перевірки m часткових гіпотез (тобто виконати m тестів і отримати ті ж результати, що і з F -тестом)? Відповідь: ні, не можна, бо F -тест перевіряє групу (пакет) часткових гіпотез як єдине ціле, як сукупність. Якщо у прикладі 8.6.2.3 провести F -тест на відхилення нульової гіпотези, яка складається з двох лінійних гіпотез, то отримали б рішення, що гіпотези прикладів 8.6.2.1, 8.6.2.2 згідно з F - та t -тестами не можуть бути відхилені.

8.6.3. F -статистика і правила використання F -тесту

Для перевірки гіпотез розділа 8.6.2 потрібно розрахувати F -статистику за формулою

$$(8.6.3.1)$$

В умовах класичної лінійної моделі нормальної регресії і справедливості гіпотези $H_0: C\beta = C^*$ випадкова змінна F (формула (8.6.3.1)) має F -розподіл з $(m; n)$ ступенями вільності, де $n = T - K$.

Правило використання F -тесту: $H_0: C\beta = C^*$ відхиляється, якщо

$$F > F(1 - \alpha; m; T - K) \quad (8.6.3.2)$$

Таблиця 8.6.3.1

F -критерій з m (у чисельнику) та n (у знаменнику) ступенями вільності
(дані наведені для $n = 1 \dots 200, \infty$; $m = 1 - 10$)

$n = T - K$	$\alpha = 0,05$					$\alpha = 0,01$				
	m					m				
	1	2	3	5	10	1	2	3	5	10
1	162	200	216	230	242	4052	4999	5403	5764	6056
2	18,5	19,0	19,2	19,3	19,4	98,5	99,0	99,2	99,3	99,4
3	10,1	9,55	9,28	8,94	8,79	34,1	30,8	29,4	28,2	27,2
4	7,71	6,94	6,59	6,16	5,96	21,2	18,0	16,7	15,5	14,5
5	6,61	5,79	5,41	4,95	4,74	16,3	13,3	12,1	11,0	10,1
6	5,99	5,14	4,76	4,28	4,06	13,7	10,9	9,78	8,75	7,87
7	5,59	4,74	4,35	3,87	3,64	12,2	9,55	8,45	7,46	6,62
8	5,32	4,46	4,07	3,58	3,35	11,3	8,65	7,59	6,63	5,81
9	5,12	4,26	3,86	3,37	3,14	10,6	8,02	6,99	6,06	5,26
10	4,96	4,10	3,71	3,22	2,98	10,0	7,56	6,55	5,64	4,85
15	4,54	3,68	3,29	2,90	2,54	8,68	6,36	5,42	4,56	3,80
30	4,17	3,32	2,92	2,53	2,16	7,56	5,39	4,51	3,70	2,98
200	3,88	3,04	2,65	2,26	1,88	6,75	4,71	3,88	3,11	2,41
∞	3,84	3,00	2,60	2,21	1,83	6,63	4,61	3,78	3,02	2,32

8.6.4. Виконання F -тесту та окремі випадки

1. Формулюється пара гіпотез H_0, H_A .
2. Обирається рівень значущості α .
3. Знаходять по таблиці табличне значення F -критерія.
4. Розраховується чисельне значення F -статистики.
5. Порівнюючи розраховану величину F з її табличним значенням, використовують рішення про відповідність правила використання F -тесту.
6. Інтерпретація результату тесту.

F-тест гіпотез для всіх регресійних коефіцієнтів.

Часто перевіряється гіпотеза, що жодний з регресорів не впливає на регресант:

$$H_0: \beta_2 = \beta_3 = \dots = \beta_k = \dots = \beta_K = 0 \quad (8.6.4.1)$$

Ця гіпотеза для економічних досліджень звичайно мало цікава, бо нульова гіпотеза при статистичній перевірці звичайно відхиляється через використання теоретично важливих регресорів.

Значення F -статистики для нульової гіпотези (8.6.4.1) може бути розрахована з використанням коефіцієнта детермінації за спрощеною формулою

$$F = \frac{R^2}{1 - R^2} \cdot \frac{T - K}{K - 1}. \quad (8.6.4.2)$$

Ця формула отримується, якщо в (8.6.3.1) використати окремий випадок (8.6.4.1) загальної лінійної гіпотези (8.6.2.1).

Так, у випадку з трьома регресорами ($k = 3, m = k - 1 = 2$) гіпотеза H_0 з (8.6.4.1) з врахуванням (8.6.2.1) записується як

$$H_0: C\beta = C^*, \quad (8.6.4.3)$$

$$\text{де } C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad C^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Аналогічне спрощення може бути зроблено для будь-якої кількості регресорів $k > 2$.

F-тест гіпотез для кількох регресійних коефіцієнтів.

Звичайно перевіряються група з m -останніх регресорів. Значення F -статистики розраховується за формулою

$$F = \frac{\hat{u}_0^T \hat{u}_0 - \hat{u}^T \hat{u}}{\hat{u}^T \hat{u}} \cdot \frac{T - K}{m}, \quad (8.6.4.4)$$

де $\hat{u}_0^T \hat{u}_0$ – сума квадратів помилок в регресії з k регресорами, $\hat{u}^T \hat{u}$ – сума квадратів помилок в регресії, звідки вилучена група m регресорів, що досліджуються.

Таким чином, розрахунки виконуються по двох регресіях:

- з K регресорами (перевіряється загальна гіпотеза, яка відповідає альтернативній гіпотезі H_A);
- з $(K - m)$ регресорами (перевіряється "обмежена" гіпотеза, яка

відповідає гіпотезі H_0).

Ідею F -тестування на основі формули (8.6.4.4) можна сформулювати так:

1. Якщо нульова гіпотеза вірна (те, що m регресорів, що досліджуються, не впливають на регресанд), то суми квадратів помилок відрізняються незначно.

2. Якщо нульова гіпотеза помилкова, то значення $(\sum_{i=0}^T u_i^2 - \sum_{i=0}^T v_i^2) / m$ збільшиться відносно $(T - K)$, і таким чином, значення F -статистики зростає і нульова гіпотеза може бути відхилена.

F-тест гіпотез для одного регресійного коефіцієнта або однієї лінійної комбінації.

У випадку, коли $m = 1$, F -тест та двосторонній t -тест дають однакові результати. Це дозволяє показати, що у даному випадку F -статистика дорівнює квадрату t -статистики:

$$t^2 = F.$$

Для відповідних табличних значень також є вірним співвідношення

$$[t(\delta, T - K)]^2 = F(1 - \delta; 1; T - K).$$

Контрольні завдання

1. Отримати F -тест гіпотез для всіх регресійних коефіцієнтів.
2. Отримати F -тест гіпотез для одного регресивного коефіцієнта.