# ПЕРЕВІРКА ВАЖНОСТІ ПОЯСНЮВАЛЬНИХ ЗМІННИХ

## ЗНАЧЕННЯ ЄДИНОЇ ПОЯСНЮВАЛЬНОЇ ЗМІННОЇ
## критерій t-Стьдента

При дослідженні значущості впливу змін значення i-ї пояснювальної змінної на зміни значення пояснюваної змінної найчастіше використовують тест статистичної значущості.:

I. Формуємо гіпотези:

$H_0 : \alpha_i = 0$ (змінна $X_i$ не має значення для моделі, що розглядається)

$H_1 : \alpha_i \neq 0$ (змінна $X_i$ має статистично значущий вплив на залежну змінну)

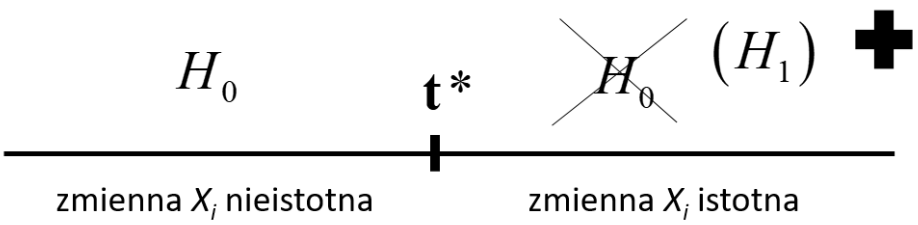
II. Розраховуємо статистику: $t = \dfrac{|a_i|}{S(a_i)}$ , який має розподіл з n- (k + 1) ступенями свободи.

III. З таблиць t-розподілу Стьюдента читаємо критичне значення $t^* = t_{\alpha,r}$ для заданого рівня значущості $\alpha$ (зазвичай $\alpha = 0,05$) і для $r = n - (k+1)$ ступенів свободи.

IV. Порівнюємо статистику $t$ і $t^*$ робимо висновок:

- Якщо $t > t^*$ необхідно відхилити нульову гіпотезу $H_0$ на користь альтернативної гіпотези $H_1$. Це означає, що змінна $X_i$ (з параметром $a_i$) має статистично значущий вплив на залежну змінну $Y$ (іншими словами: змінна .. $X_i$ .. є статистично значущою). Ми оцінюємо модель позитивно.
Ймовірність зробити помилку, яка полягає в прийнятті неправильного рішення щодо перевірки $\alpha$ .

- Якщо $t < t^*$ немає підстав відхиляти нульову гіпотезу $H_0$. Це означає, що змінна $X_i$ має статистично незначний вплив на залежну змінну $Y$, тобто вона незначна для моделі, що розглядається. Цю змінну слід видалити з моделі.
Ми оцінюємо модель негативно.
Ймовірність зробити помилку, яка полягає в прийнятті неправильного рішення щодо перевірки $\alpha$ .




Joanna Grochowska eTrapez
www.etrapez.pl
tel. 534 864 486

# ВИВЧЕННЯ ЗНАЧЕННЯ ПІДМНОЖИНИ ПОЯСНЮВАЛЬНИХ ЗМІННИХ узагальнений тест Вальда

Коли є необхідність перевірити гіпотезу про одночасну значущість вибраної підмножини пояснювальних змінних, застосовується наступний тест:

I. Формуємо гіпотези:

$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_k = 0$ (всі змінні статистично несуттєві)

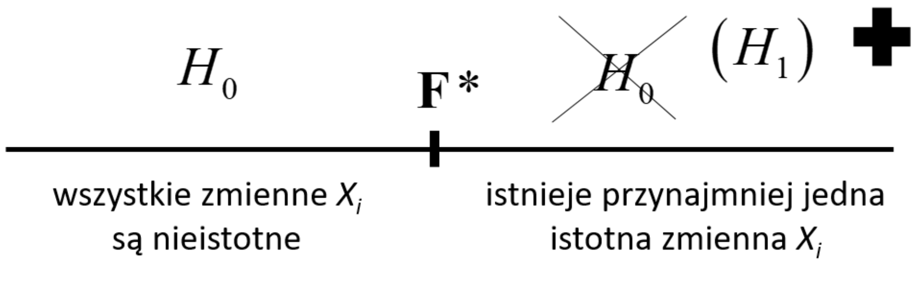
$H_1 : co\ \ najmniej\ \ jeden\ \ z\ \ parametrów\ \ \alpha_i \neq 0$ (є якась важлива змінна *Xi*)

II. Розраховуємо статистику: $F = \frac{R^2}{1-R^2} \cdot \frac{n-(k+1)}{k}$ .

III. Критичне значення читаємо з таблиць Фішера-Снедекора $F^* = F_{\alpha, r_1, r_2}$ для заданий рівень значущості $\alpha$ (звичайно $\alpha = 0,05$) та для $r_1 = k$ і $r_2 = n-(k+1)$ ступені свободи.

IV. Порівнюємо статистику $F$ і $F^*$ робимо висновок:

- Якщо $F > F^*$ необхідно відхилити нульову гіпотезу *H0* на користь альтернативної гіпотези *H1*. Це означає, що існує така пояснювальна змінна $X_i$, яка має статистично значущий вплив на залежну змінну *Y*.
Ймовірність зробити помилку, яка полягає в прийнятті неправильного рішення щодо перевірки $\alpha$ .
- Якщо $F < F^*$ немає підстав відхиляти нульову гіпотезу *H0*. Це означає, що не існує такої змінної $X_i$, яка матиме статистично значущий вплив на залежну змінну *Y*, тобто всі змінні є нерелевантними для моделі, що розглядається. Ймовірність зробити помилку, яка полягає в прийнятті неправильного рішення щодо перевірки $\alpha$ .




Joanna Grochowska eTrapez
www.etrapez.pl
tel. 534 864 486

# ДОВІРЕВІ ІНТЕРВАЛИ ДЛЯ СТРУКТУРНИХ ПАРАМЕТРІВ

Щоб побудувати довірчий інтервал для параметра $\alpha_i$, i = 1,2, ..., k, з коефіцієнтом довіри $(1-\alpha)$, слід використовувати таке співвідношення:

$$P\left(a_i - t_{\alpha,n-(k+1)} \cdot S\left(a_i\right) \le \alpha_i \le a_i + t_{\alpha,n-(k+1)} \cdot S\left(a_i\right)\right) = 1-\alpha$$

Отже, довірчий інтервал для параметра $\alpha_i$ має вигляд:

$$\left(a_i - t_{\alpha,n-(k+1)} \cdot S\left(a_i\right)\ ,\ a_i + t_{\alpha,n-(k+1)} \cdot S\left(a_i\right)\right)$$

**ІНТЕРПРЕТАЦІЯ (ВІДПОВІДЬ):**

З імовірністю ..$(1-\alpha)$.. (наприклад, 0,95) (з упевненістю, наприклад, 0,95)
діапазон $\left(a_0 - t_{\alpha,n-(k+1)} \cdot S\left(a_0\right)\ ,\ a_0 + t_{\alpha,n-(k+1)} \cdot S\left(a_0\right)\right)$ містить невідоме значення параметр $\alpha_0$,
діапазон $\left(a_1 - t_{\alpha,n-(k+1)} \cdot S\left(a_1\right)\ ,\ a_1 + t_{\alpha,n-(k+1)} \cdot S\left(a_1\right)\right)$ містить невідоме значення параметр $\alpha_1$,
діапазон $\left(a_2 - t_{\alpha,n-(k+1)} \cdot S\left(a_2\right)\ ,\ a_2 + t_{\alpha,n-(k+1)} \cdot S\left(a_2\right)\right)$ містить невідоме значення параметр $\alpha_2$,
(послідовні інтервали аналогічно).

Довжина довірчого інтервалу залежить від рівня значущості $(1-\alpha)$, кількості ступенів свободи і розмір стандартних помилок в оцінці параметрів.
Для них це ВУЖЕ:

- значення рівня значущості вище $\alpha$,
- чим більша кількість ступенів свободи (і, отже, більша вибірка),
- нижнє значення стандартної помилки оцінки параметра *S(a$_i$)*.

Joanna Grochowska eTrapez
www.etrapez.pl
tel. 534 864 486