

18. Проверка гипотезы о линейном ограничении общего вида $H_0 : H\beta = r$.

Линейное ограничение общего вида $H_0 : H\beta = r$. Пусть H — $q \times k$ матрица, β — $k \times 1$ вектор коэффициентов, r — $q \times 1$ вектор.

Число ограничений не превосходит числа параметров и ограничения линейно независимы, т.е. $q \leq k$ и матрица H имеет максимальный ранг: $\text{rank}(H) = q$.

Пример. В качестве примера рассмотрим следующие матрицы H, r для $k = 3, q = 2$:

$$H\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = r.$$

Это условие соответствует системе двух линейных ограничений

$$\begin{cases} \beta_1 = 2 \\ \beta_2 - \beta_3 = 0 \end{cases}$$

$$F = \frac{(H\hat{\beta} - r)'(H(X'X)^{-1}H')^{-1}(H\hat{\beta} - r)/q}{e'e/(n-k)} \sim F(q, n-k).$$

Если справедлива гипотеза $H_0 : H\beta - r = 0$, то статистика F не должна принимать слишком больших значений, а именно, с 95%-вероятностью $F < F_c(q, n-k)$, где $F_c(q, n-k)$ есть 95%-квантиль распределения Фишера $F(q, n-k)$.

Другой вид F -статистики:

$$F = \frac{(\hat{\beta} - \beta)'H'(H(X'X)^{-1}H')^{-1}H(\hat{\beta} - \beta)/q}{e'e/(n-k)} \sim F(q, n-k)$$

Условие $F < F_c(q, n-k)$ задает 95%-доверительную область для коэффициентов β (выпуклую).

В случае $H = I$ статистика F выглядит следующим образом:

$$F = \frac{(\hat{\beta} - \beta)'(X'X)(\hat{\beta} - \beta)/k}{e'e/(n-k)} \sim F(k, n-k).$$

$H_0 : \beta_{k-q+1} = \beta_{k-q+2} = \dots = \beta_k = 0$. Гипотеза является, конечно, частным случаем общей линейной гипотезы $H\beta = r$.

$$F = \frac{(e'e - e'e)/q}{e'e/(n-k)} = \frac{(\text{ESS}_R - \text{ESS}_{UR})/q}{\text{ESS}_{UR}/(n-k)} \sim F(q, n-k).$$

Здесь ESS_R — сумма квадратов остатков "короткой" (*restricted*) регрессии; ESS_{UR} — сумма квадратов остатков "длинной" (*unrestricted*) регрессии.

F -статистику можно выразить через коэффициенты детерминации R^2 для "короткой" и "длинной" регрессий:

$$F = \frac{(R_{UR}^2 - R_R^2)/q}{(1 - R_{UR}^2)/(n-k)} \sim F(q, n-k)$$

При $F < F_c(q, n-k) \Rightarrow H_0$ о "короткой" регрессии принимается, т.е. использование q дополнительных регрессоров ничего не добавляет к качеству оценивания модели.