Билет 18 проверка гипотезы о линейном ограничении общего вида $H_0: H\beta = r$

Источник: Магнус, страница 80

Напоминаю, что такое множественная регрессия:

$y_t = \beta_1 x_{1t} + \dots + \beta_k x_{kt}, t = 1, \dots, n$

Гипотеза $H_0: H\beta = r$

Пусть $H \in R^{q \times k}; \beta \in k \times 1; r \in q \times 1$

$rank(H) = q; q \leq k$

Важные формулы:

(приложение ЛА, п. 10), находим оценку метода наименьших квадратов:

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'y. \quad (3.4)$$

1. Покажем, что МНК-оценка (3.4) является несмещенной оценкой β:

$$\begin{aligned} \mathrm{E}\hat{\beta}_{OLS} &= \mathrm{E}\big((X'X)^{-1}X'y\big) = (X'X)^{-1}X'\mathrm{E}(y) \\ &= (X'X)^{-1}X'\mathrm{E}(X\beta + \varepsilon) \\ &= (X'X)^{-1}X'X\beta + (X'X)^{-1}X'\mathrm{E}\varepsilon = \beta. \end{aligned} \quad (3.7)$$

2. Подсчитаем матрицу ковариаций МНК-оценки:

$$\begin{aligned} \mathrm{V}(\hat{\beta}_{OLS}) &= \mathrm{V}(Ay) = A\mathrm{V}(y)A' = A\sigma^2 IA' \\ &= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} = \sigma^2(X'X)^{-1} \end{aligned} \quad (3.8)$$

(здесь мы использовали симметричность матрицы $X'X$ и свойство матрицы ковариаций (МС.9)).

Далее проверка:

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:

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

Из (3.4), (3.7), (3.8) видно, что вектор $\hat{\beta}_{OLS}$ имеет нормальное распределение со средним β и матрицей ковариаций $\sigma^2(X'X)^{-1}$.

$$\hat{\beta}_{OLS} \sim N\big(\beta, \sigma^2(X'X)^{-1}\big).$$

Отсюда получаем, что $\boldsymbol{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{r} \sim N(\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{r}, \boldsymbol{\Sigma})$, где $\boldsymbol{\Sigma}$ — $q \times q$ матрица и $\boldsymbol{\Sigma} = \mathrm{V}(\boldsymbol{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}) \doteq \mathrm{V}(\boldsymbol{H}\widehat{\boldsymbol{\beta}}) = \boldsymbol{H}\mathrm{V}(\widehat{\boldsymbol{\beta}})\boldsymbol{H}' = \sigma^2\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'$. Итак,

$$\boldsymbol{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{r} \sim N(\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{r}, \sigma^2\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'). \tag{3.37}$$

По лемме (приложение МС, п. 4, N9) из (3.37) при условии справедливости гипотезы H_0: $\boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{r}$, получаем:

$$\frac{1}{\sigma^2}(\boldsymbol{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{r})'\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1}(\boldsymbol{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}) \sim \chi^2(q). \tag{3.38}$$

Из (3.21) и (3.38), используя независимость $\widehat{\boldsymbol{\beta}}$ и $\boldsymbol{e}$, получаем:

$$F = \frac{(\boldsymbol{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{r})'\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1}(\boldsymbol{H}\widehat{\boldsymbol{\beta}} - \boldsymbol{r})/q}{\boldsymbol{e}'\boldsymbol{e}/(n-k)} \sim F(q, n-k). \tag{3.39}$$

Если справедлива гипотеза H_0: $\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{r} = \boldsymbol{0}$, то статистика F в (3.39) не должна принимать слишком больших значений, а именно, с вероятностью $1-\alpha$ имеем $F < F_\alpha(q, n-k)$, где $F_\alpha(q, n-k)$ есть $100\alpha\%$-ная точка распределения Фишера $F(q, n-k)$.

$$F = \frac{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{H}'\left(\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{H}'\right)^{-1}\boldsymbol{H}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})/q}{\boldsymbol{e}'\boldsymbol{e}/(n-k)} \sim F(q, n-k). \tag{3.40}$$

Условие $F < F_\alpha(q, n-k)$ задает $100(1-\alpha)\%$-ную *доверительную область* для коэффициентов $\boldsymbol{\beta}$.

Так как в числителе (3.40) стоит неотрицательно определенная квадратичная форма от β_i, то эта доверительная область является выпуклым множеством.

В случае $\boldsymbol{H} = \boldsymbol{I}$ статистика F в (3.40) выглядит следующим образом:

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

В этом случае доверительная область является эллипсоидом в k-мерном пространстве коэффициентов $\boldsymbol{\beta}$.