Inferenza nel modello lineare (classico)

Ipotizziamo che $\Sigma = \sigma^2 I_n$ (o, in alternativa, che il nostro modello sia stato trasformato in un modello con matrice di varianze e covarianze degli errori sferica).

$$\underset{n\times 1}{y} = \underset{n\times k}{X}\,\underset{k\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \qquad E[\varepsilon | X] = 0 \;,\; V[\varepsilon | X] = \sigma^2 I_n$$

Stimatore OLS:

$$\hat{\beta} = (X'X)^{-1} X'y \qquad (\text{rango}(X) = \max = k)$$

Proprietà (Gauss-Markov):

$$E[\hat{\beta} | X] = \beta$$

$$V[\hat{\beta} | X] = \sigma^2 (X'X)^{-1} = \frac{\sigma^2}{n} S_{XX}^{-1} \;, \quad S_{XX} := \frac{1}{n}\sum_{i=1}^{n} x_i x_i' = \frac{1}{n} X'X$$

<u>nota</u> il vettore di parametri del modello è

$$\theta = \begin{pmatrix} \beta \\ \sigma^2 \end{pmatrix}$$

Stima di σ^2:

Ricordiamo che $\sigma^2 := E[\varepsilon_i^2 | X]$, $i = 1, 2, \dots, n$

Se ε_i fosse osservabile, potremmo considerare lo stimatore:

$$\tilde{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\varepsilon_i - \bar{\varepsilon}_n)^2 \qquad \bar{\varepsilon}_n := \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i$$

$$\tilde{\tilde{\sigma}}^2 = \frac{1}{n} \sum \varepsilon_i^2$$

Purtroppo, $\varepsilon_1 \dots \varepsilon_n$ non sono osservabili.

Possiamo, in alternativa, considerare i residui:

$$\hat{\varepsilon}_i = y_i - x_i' \hat{\beta} \qquad \left(\hat{\beta} = (X'X)^{-1} X'y \right)$$

La relazione tra $\hat{\varepsilon}_i$ e ε_i è:

$$\hat{\varepsilon}_i = x_i' \beta + \varepsilon_i - x_i' \hat{\beta}$$

$$= \varepsilon_i - x_i' (\hat{\beta} - \beta)$$

Consideriamo la somma dei quadrati dei residui.

$$\underbrace{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}_{1\times 1} = \hat{\varepsilon}'\hat{\varepsilon}$$

$$\hat{\varepsilon} = y - X\hat{\beta} = M_X \cdot \varepsilon$$

$$M_X = I_n - X(X'X)^{-1}X' \quad \text{simmetrica idempot.}$$

$$= (M_X \varepsilon)'(M_X \varepsilon)$$

$$= \varepsilon' M_X' M_X \varepsilon$$

$$= \underbrace{\underset{1\times n}{\varepsilon'} \underset{n\times n}{M_X} \underset{n\times 1}{\varepsilon}}_{1\times 1}$$

in quanto $M_X' = M_X$ e $M_X \cdot M_X = M_X$

$$tr(a) = a$$

$$tr(AB) = tr(BA)$$

$$= tr\{\varepsilon' M_X \varepsilon\}$$

$$= tr\{\varepsilon\varepsilon' M_X\}$$

Calcoliamo il valore atteso (condizionato)

$$E[tr(X)] = tr\,E(X)$$

$$E\left[\sum_{i=1}^{n} \hat{\varepsilon}_i^2 \,\middle|\, X\right] = E\left[tr\{\varepsilon\varepsilon' M_X\} \,|\, X\right]$$

$$= tr\{E[\varepsilon\varepsilon' M_X \,|\, X]\} = tr\{E[\varepsilon\varepsilon' \,|\, X] \cdot M_X\}$$

$$= tr\{\sigma^2 I_n M_X\} = \sigma^2\, tr\, M_X = \sigma^2(n-K)$$

in quanto $\operatorname{tr} M_X = \operatorname{rango} M_X = n-K$. Infatti:

$\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$

$$\begin{aligned}
\operatorname{tr}\{M_X\} &= \operatorname{tr}\{I_n - X(X'X)^{-1}X'\} = \\
&= \operatorname{tr}\{I_n\} - \operatorname{tr}\{X(X'X)^{-1}X'\} \\
&= n - \operatorname{tr}\{(X'X)^{-1}X'X\} \\
&= n - \operatorname{tr}\{I_K\} = n-K
\end{aligned}$$

Riassumendo:

$$E\left[\sum_{i=1}^{n} \hat{\varepsilon}_i^2 \mid X\right] = E\left[\hat{\varepsilon}'\hat{\varepsilon} \mid X\right] = \sigma^2(n-K)$$

Quindi, possiamo considerare lo stimatore

$$\hat{\sigma}^2 := \frac{1}{n-K}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-K}$$

che risulta corretto:

$$E[\hat{\sigma}^2 \mid X] = E\left[\frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-K} \mid X\right] = \frac{1}{n-K} E[\hat{\varepsilon}'\hat{\varepsilon} \mid X] = \frac{1}{\cancel{n-K}}\sigma^2\cancel{(n-K)} = \sigma^2$$

nota: una stima di $V(\hat\beta|X) = \sigma^2(X'X)^{-1}$ è

$$\hat V(\hat\beta|X) := \hat\sigma^2(X'X)^{-1}$$

Questa stima è corretta:

$$E[\hat V(\hat\beta|X)\,|X] = E[\hat\sigma^2(X'X)^{-1}|X] = E[\hat\sigma^2|X]\cdot(X'X)^{-1}$$
$$= \sigma^2(X'X)^{-1} = V[\hat\beta|X]$$

nota: lo stimatore $\hat V[\hat\beta|X] = \hat\sigma^2(X'X)^{-1}$ viene riportato da tutti i software. In particolare, se poniamo

a_{ij} generico elemento di A, $A := (X'X)^{-1}$

allora $\hat V[\hat\beta_i|X]$, ovvero la stima OLS della varianza di $\hat\beta_i$, è

$$\hat V[\hat\beta_i|X] = \hat\sigma^2 a_{ii}.$$

La quantità:

$$s.e.(\hat\beta_i) := \sqrt{\hat V[\hat\beta_i|X]} = \hat\sigma\sqrt{a_{ii}}$$

è detto standard error di $\hat\beta_i$ $(i = 1, 2, \ldots, k)$.

Vogliamo ora usare questi risultati per effettuare
test di ipotesi.

esempi $\quad \beta = (\beta_1, \beta_2, \dots, \beta_K)'$

① $H_0: \beta_i = 0$ (nullità i-esimo coeff. di regressione)

② $H_0: \beta_i = \bar{\beta}$

③ $H_0: \beta_i = \beta_j$ $\quad (H_0: \beta_i - \beta_j = 0)$

④ $H_0: \beta_i + \beta_j = 1$ $\quad (H_0: \beta_i + \beta_j - 1 = 0)$

- ①-④ sono ipotesi su uno o più coefficienti di regressione
- ciascuna di queste Hp. esprime <u>un</u> solo vincolo

⑤ $H_0: \beta_1 = 0, \beta_2 = 0, \dots, \beta_K = 0$

⑥ $H_0: \beta_2 = 0, \beta_3 = 0, \dots, \beta_K = 0$ (non-significatività dell'intera regressione — esclusa l'intercetta)

$$\underset{q\times 1}{g}(\underset{k\times 1}{\beta}) = 0_{q\times 1}$$

(7) $H_0: \beta_i = 0, \; \beta_j = 0$

(8) $H_0: \beta_1 = 0, \; \beta_2 + \beta_3 - 1 = 0$

le ipotesi (5)-(8) identificano $q>1$ vincoli sui k coefficienti

(9) $H_0: \beta_i \cdot \beta_j = 1$

la (9), al contrario delle (1)-(8) è non lineare.

Consideriamo ipotesi lineari — ipotesi su particolari combinazioni lineari dei coefficienti

Proprietà: un vincolo lineare sui coefficienti si può esprimere nella forma

$$\underset{1\times k}{R'}\,\underset{k\times 1}{\beta} - \underset{1\times 1}{r} = \underset{1\times 1}{0}$$

dove

R è un vettore $k\times 1$ e r uno scalare

Nel caso dei precedenti esempi:

1. $H_0: \beta_i = 0 \Rightarrow \underbrace{\begin{array}{c} \begin{bmatrix} 0 \ldots 0 \ldots 0 & 1 & 0 \ldots 0 \end{bmatrix} \\ \uparrow \text{ i-esima posizione} \end{array}}_{R'} \underbrace{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_k \end{bmatrix}}_{\beta} = \underbrace{0}_{r}$

o, analogamente, $R'\beta - r = 0$

2. $H_0: \beta_i = \bar{\beta} \Rightarrow R' = [0 \ldots 0 \; 1 \; 0 \ldots 0]$, $r = \bar{\beta}$

elemento i-esimo

3. $H_0: \beta_i - \beta_j = 0 \Rightarrow R' = [0 \ldots 0 \; 1 \; 0 \ldots 0 \; -1 \; 0 \ldots 0]$, $r = 0$

$(H_0: \beta_i = \beta_j)$

i-esimo j-esimo

infatti $R'\beta = \beta_i - \beta_j$.

④ $H_0: \beta_1+\beta_2 = 1 \implies R' = [1\ 1\ 0 \ldots\ 0],\ r = 1$

infatti $R'\beta - r = \beta_1+\beta_2-1$.

Proprietà: q vincoli lineari sui coefficienti si possono esprimere nella forma

$$\underset{q\times k}{R'}\ \underset{k\times 1}{\beta} - \underset{q\times 1}{r} = \underset{q\times 1}{0}$$

dove

R è una matrice $k\times q$ e r un vettore $q\times 1$.

esempio ⑧ $H_0: \underbrace{\beta_1 = 0}, \underbrace{\beta_2+\beta_3-1=0}$ $q=2$ vincoli

$$\underbrace{[1\ 0 \ldots\ 0]}_{R_1'}\cdot\beta = \underbrace{0}_{r_1} \qquad \underbrace{[0\ 1\ 1\ 0 \ldots\ 0]}_{R_2'}\beta = \underbrace{1}_{r_2}$$

Il vettore β deve soddisfare $q=2$ vincoli:

$$R_1'\beta - r_1 = 0$$
$$R_2'\beta - r_2 = 0$$

ovvero

$$\underbrace{\begin{bmatrix} R_1' \\ R_2' \end{bmatrix}}_{R'} \beta - \underbrace{\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}}_{r} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\underset{q\times K}{R'} \; \underset{K\times 1}{\beta} - \underset{q\times 1}{r} = \underset{q\times 1}{0} \qquad (q=2)$$

dove

$$\underset{K\times q}{R} = \begin{bmatrix} R_1 & R_2 \end{bmatrix} \qquad \underset{q\times 1}{r} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

Esercizio: scrivere nella forma $R'\beta - r = 0$ i vincoli (6)-(7)-(8).

Vogliamo testare un'ipotesi nella forma $R'\beta - r = 0$.

R è $k\times q$ e r è $q\times 1$, dove q è il numero di vincoli.

Lavoriamo sotto le ipotesi di Gauss-Markov, che implicano

$$E[\hat{\beta}|X] = \beta, \quad V[\hat{\beta}|X] = \sigma^2 (X'X)^{-1}$$

$$E[\hat{\sigma}^2|X] = \sigma^2$$

In aggiunta introduciamo l'ipotesi di normalità:

$$y|X \sim N \qquad [\text{o, equivalentemente, } \varepsilon|X \sim N]$$

Poiché, sotto le ipotesi del mod. lineare classico, si ha:

$$E[y|X] = X\beta$$

$$V[y|X] = \sigma^2 I_n$$

e poiché la distribuzione normale dipende solo da valore atteso e varianza, possiamo scrivere

$$\underset{n\times 1}{y}|X \sim N(X\beta, \sigma^2 I_n)$$

o, in termini di ε,

$$\underset{n\times 1}{\varepsilon}|X \sim N(\underset{n\times 1}{0}, \underset{n\times n}{\sigma^2 I_n})$$

Proprietà di $\hat\beta$ e $\hat\sigma^2$ sotto l'ipotesi di normalità:

$$\hat\beta = (X'X)^{-1}X'y \qquad\qquad y|X \sim N(X\beta, \sigma^2 I_n)$$

$\hat\beta$, essendo (condizionatamente a X) una combinazione lineare di y, che è normale, è a sua volta normale

1 $$\boxed{\hat\beta|X \sim N(\beta, \sigma^2(X'X)^{-1})}$$ ovvero

$$\boxed{\hat\beta-\beta|X \sim N(0, \sigma^2(X'X)^{-1})}$$

$\boxed{\begin{array}{l}\underline{\text{esercizio}}: \text{ partire da } y|X \sim N(X\beta, \sigma^2 I) \text{ e ottenere la} \\ \text{distribuzione di } \hat{\beta}|X \text{ usando il fatto che} \\ \hat{\beta} \text{ è una comb. lineare di } y.\end{array}}$

Consideriamo $\hat{\sigma}^2 := \hat{\varepsilon}'\hat{\varepsilon}/(n-k)$

Abbiamo visto che

$$\hat{\sigma}^2 = \hat{\varepsilon}'\hat{\varepsilon}/(n-k) = \varepsilon' M_X \varepsilon / (n-k),$$

dove $M_X = I_n - X(X'X)^{-1}X'$, simmetrica e idempotente.

Usiamo il seguente risultato:

$\boxed{\begin{array}{l}\text{Se } \underset{m\times 1}{z} \sim N(0, I_m) \text{ e se } A \text{ idempotente con rango } q, \\ \text{allora } \quad z'Az \sim \chi^2(q)\end{array}}$

da cui:

$$\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma^2} = \frac{1}{\sigma^2}\varepsilon' M_X \varepsilon = \left(\frac{1}{\sigma}\varepsilon\right)' M_X \left(\frac{1}{\sigma}\varepsilon\right) \sim \chi^2(n-k)$$

in quanto $\varepsilon \sim N(0, \sigma^2 I_n)$ (da cui $\frac{1}{\sigma}\varepsilon \sim N(0, I_n)$)
e M_X idempotente di rango $n-K$.

Quindi

$$\hat{\sigma}^2 = \frac{\varepsilon' M_X \varepsilon}{n-K} = \frac{\sigma^2}{n-K} \cdot \frac{\varepsilon' M_X \varepsilon}{\sigma^2} = \frac{\sigma^2}{n-K} \cdot \chi^2(n-K)$$

ovvero

② $$\boxed{(n-K)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-K)}$$

$$\hat{\sigma}^2 = \sigma^2 \cdot \left\{ \frac{\chi^2(n-K)}{n-K} \right\}$$

Un terzo importante risultato riguarda l'indipendenza (condizionata a X) tra $\hat{\beta}$ e $\hat{\sigma}^2$. Per dimostrarla, possiamo usare il seguente risultato:

Se $\underset{m \times 1}{z} \sim N(0, V)$, allora $R'z$ e $z'Az$ sono stocasticamente indipendenti se e solo se $R'A = 0$

Poiché $\hat{\beta} = (X'X)^{-1}X'y = (X'X)^{-1}X'[X\beta + \varepsilon] = \beta + (X'X)^{-1}X'\varepsilon$
abbiamo che

$$\hat{\beta} - \beta = \underline{(X'X)^{-1}X'}\,\varepsilon$$

Inoltre $\hat{\varepsilon}'\hat{\varepsilon} = \varepsilon' M_X \varepsilon$ è una forma quadratica in ε.
Poiché

$$\hat{\beta} - \beta = [(X'X)^{-1}X']\,\varepsilon$$

e $[(X'X)^{-1}X'] \cdot M_X = 0$, allora

(3) $\boxed{\hat{\beta} - \beta \text{ e } \hat{\varepsilon}'\hat{\varepsilon} \text{ sono (dato } X) \text{ stocasticamente indipendenti}}$

Usiamo questi tre risultati per testare ipotesi su β

Caso più semplice: $H_0: \beta_i = \bar{\beta}$ ($\bar{\beta}$ fissato)

Consideriamo $\hat{\beta}_i$ - Qual è la sua distribuzione?

Poiché $\hat{\beta}|X \sim N(\beta, \sigma^2(X'X)^{-1})$ allora

$\hat{\beta}_i|X \sim N(\beta_i, \sigma^2 a_{ii})$, dove a_{ii} è l'elemento i,i di $A := (X'X)^{-1}$

Standardizziamo:

$$\frac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{a_{ii}}} \sim N(0,1) \quad \text{(condizionatamente a } X\text{)}$$

Se è vera H_0, allora

$$z_n := \frac{\hat{\beta}_i - \bar{\beta}}{\sigma\sqrt{a_{ii}}} \sim N(0,1)$$

Se σ fosse noto, potremmo usare la statistica z_n

per testare H_0. In alternativa, possiamo sostituire σ con un suo stimatore, ad esempio $\hat{\sigma}$. In questo caso otteniamo:

$$t_n = \frac{\hat{\beta}_i - \bar{\beta}}{\hat{\sigma}\sqrt{a_{ii}}} \qquad \text{(statistica } t\text{)}$$

Si può dimostrare il seguente risultato:

Sotto le ipotesi fatte in precedenza ($y|X \sim N(X\beta, \sigma^2(X'X)^{-1})$, rango$(X)$ = max), allora

$$t_n := \frac{\hat{\beta}_i - \bar{\beta}}{\hat{\sigma}\sqrt{a_{ii}}} \sim t(n-k)$$

sotto l'ipotesi nulla $H_0: \beta_i = \bar{\beta}$ e condizionatamente a X.

Dimostrazione: per il risultato 1, $\hat{\beta}_i - \beta_i \mid X \sim N(0, \sigma^2 a_{ii})$,
da cui

$$\frac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{a_{ii}}} \sim N(0,1)$$

Inoltre

$$\frac{\hat{\beta}_i - \beta_i}{\hat{\sigma}\sqrt{a_{ii}}} = \frac{\sigma}{\hat{\sigma}} \cdot \frac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{a_{ii}}}$$

$$= \frac{\boxed{\dfrac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{a_{ii}}}}}{\sqrt{\boxed{\hat{\sigma}^2/\sigma^2}}}$$

$\xrightarrow{1} N(0,1)$

3 indipendenti

$\xrightarrow{2} \dfrac{\chi^2(n-k)}{n-k}$

Abbiamo quindi scritto t_n come il rapporto tra una $N(0,1)$ e la radice quadrata di un $\chi^2(n-k)$ diviso per i suoi gradi di libertà. Poiché $N(0,1)$ e $\chi^2(n-k)$ sono indipendenti

$$\frac{\hat{\beta}_i - \beta_i}{\hat{\sigma}\sqrt{a_{ii}}} \sim t(n-k)$$

Infine, sotto H_0, $\beta_i = \bar{\beta}$, per cui:

$$t_u := \frac{\hat{\beta}_i - \bar{\beta}}{\hat{\sigma}\sqrt{a_{ii}}} \underset{H_0}{\sim} t(n-k)$$

condizionatamente a X. $\square$

Nota: cosa succede se H_0 è falsa? Ovvero $\beta_i \neq \bar{\beta}$?

$$t_u = \frac{\hat{\beta}_i - \bar{\beta}}{\hat{\sigma}\sqrt{a_{ii}}} = \frac{\hat{\beta}_i - \beta_i + \beta_i - \bar{\beta}}{\hat{\sigma}\sqrt{a_{ii}}} = \underbrace{\frac{\hat{\beta}_i - \beta_i}{\hat{\sigma}\sqrt{a_{ii}}}}_{t(n-k)} + \frac{\beta_i - \bar{\beta}}{\hat{\sigma}\sqrt{a_{ii}}}$$

Il termine residuo, $(\beta_i - \bar{\beta})/\hat{\sigma}\sqrt{a_{ii}}$ è nullo sotto H_0

Se invece $\beta_i > \bar{\beta}$, assumerà un valore positivo, tanto più grande quanto più piccolo è lo standard error di $\hat{\beta}_i$, ovvero $\hat{\sigma}\sqrt{a_{ii}}$ -

Vincolo su più coefficienti

Esempio Cobb-Douglas (funzione di produzione)

$Y_t = A L_t^{\alpha} K_t^{\beta}$ L: lavoro K: capitale Y: prodotto

$\alpha + \beta = 1$ rendimenti di scala costanti

Passando ai logaritmi:

$y = a + \alpha \ell_t + \beta k_t$, $\ell_t = \log(L_t)$, $k_t = \log(K_t)$, $a = \log A$

Denominando $\theta := (a, \alpha, \beta)'$ il vettore dei parametri, l'ipotesi $\alpha + \beta = 1$ corrisponde a:

$$R'\theta - r = 0$$

dove $R' = [0 \quad 1 \quad 1]$ e $r = 1$

Consideriamo il generico caso di un modello lineare

$$y = X\beta + \varepsilon$$

con vincolo (da testare) $H_0 : \underset{1\times k}{R'}\,\underset{k\times 1}{\beta} - \underset{1\times 1}{r} = \underset{1\times 1}{0}$

Vogliamo derivare una statistica test per H_0.

Sappiamo che $\hat{\beta} - \beta \sim N\left(\underset{k\times 1}{0},\ \sigma^2 (X'X)^{-1}\right)$ (dato X)

da cui:

$$\underset{1\times k}{R'}\,(\underset{k\times 1}{\hat{\beta} - \beta}) \sim N\left(\underbrace{\underset{1\times k}{R'}\,\underset{k\times 1}{0}}_{1\times 1},\ \sigma^2 \underbrace{\underset{1\times k}{R'}\,\underset{k\times k}{(X'X)^{-1}}\,\underset{k\times 1}{R}}_{1\times 1}\right)$$

$\underset{q\times k}{R'}\,\underset{k\times 1}{X} =: \underset{q\times 1}{Y}$

$E[Y] = R'\,E[X]$

$V[Y] = R'\,V[X]\,R$

Se H_0 è vero, allora $R'\beta = r$, e quindi:

$$R'(\hat{\beta} - \beta) \overset{H_0}{=} R'\hat{\beta} - r \sim N\left(0,\ \sigma^2 R'(X'X)^{-1}R\right)$$

da cui

$$Z_n = \frac{R'\hat{\beta} - r}{\sqrt{\sigma^2 R'(X'X)^{-1}R}} \overset{H_0}{\sim} N(0,1) \qquad \text{(dato } X\text{)}$$

Sostituendo σ con $\hat{\sigma}$:

$$t_n := \frac{R'\hat{\beta} - r}{\hat{\sigma}\sqrt{R'(X'X)^{-1}R}}$$

Sotto H_0,

$$t_n \sim t(n-k)$$

Dimostrazione: per esercizio.

<u>Nota</u>: t_n è il rapporto tra $N(0,1)$ e $\sqrt{\chi^2(n-k)/(n-k)}$, indipendenti tra loro.

Consideriamo $t_n^2 =: F_n$ - Qual è la sua distribuzione?

$$t_n = \frac{N(0,1)}{\sqrt{\chi^2(n-k)/(n-k)}} \Rightarrow t_n^2 = \frac{N(0,1)^2}{\chi^2(n-k)/(n-k)} = \frac{\chi^2(1)}{\chi^2(n-k)/(n-k)}$$

$$=: F(1, n-k)$$

Nel nostro caso, $$F_n := \frac{(R'\hat{\beta} - r)^2}{\hat{\sigma}^2 R'(X'X)^{-1}R}$$

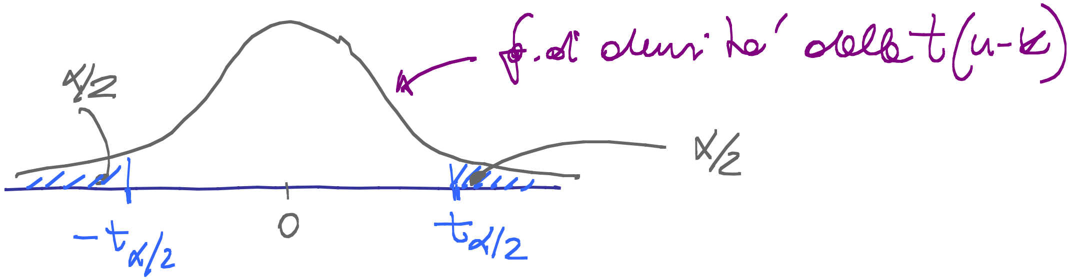
nota: Ipotesi uni/bi-direzionali

$H_0: \beta_i = 0 \qquad H_1: \beta_i > 0$ (one-sided)

$H_0: \beta_i = 0 \qquad H_1: \beta_i \neq 0$ (two sided)

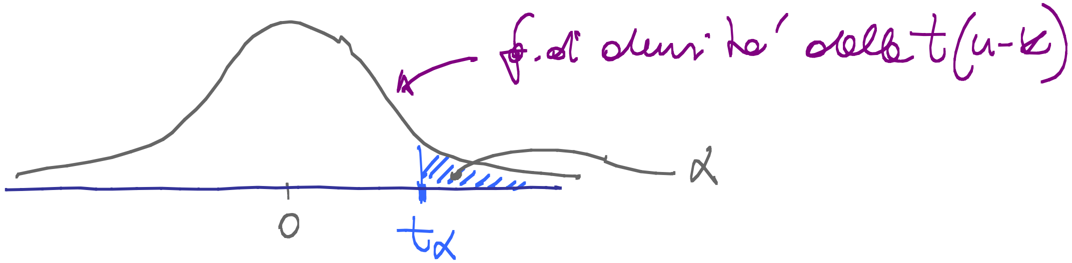
Nel caso di ipotesi bidirezionali, i test t_n ed F_n sono equivalenti.
Nel caso di ipotesi unidirezionali, il test t_n è più potente del test F_n

Test bidirezionale



$$P(|t_n| > t_{\alpha/2}) = \alpha \quad \text{(livello di significatività)}$$

Test unidirezionale



$H_0: \beta = 0$ vs $H_1: \beta > 0$

$$P(t_n > t_\alpha) = \alpha$$

Poiché $t_\alpha < t_{\alpha/2}$, se la statistica test t_m tende ad assumere valori positivi (grandi) sotto l'ipotesi alternativa, allora il test t unidirezionale ha potenza maggiore del test bidirezionale.

Caso di più vincoli lineari

Esempio : $\alpha = 1/2$, $\beta = 1/2$

Il generico insieme di q vincoli lineari si può scrivere come

$$\underset{q\times K}{R'}\,\underset{K\times 1}{\beta} - \underset{q\times 1}{r} = \underset{q\times 1}{0}$$

Costruiamo un test per $H_0 : R'\beta - r = 0$.

<u>nota</u> : una condizione che useremo è rango$(R) = q = $ max

Escludiamo quindi vincoli ridondanti.

Esempio :

$$\alpha = \frac{1}{2},\ \beta = \frac{1}{2},\ \alpha + \beta = 1$$

In questo caso: $R = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ $r = \begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$ $\theta = \begin{pmatrix} a \\ \alpha \\ \beta \end{pmatrix}$

e rango$(R) < q = 3$

Partiamo da $\hat{\beta} - \beta \sim N(0, \sigma^2 (X'X)^{-1})$ (dato X)

Premoltiplicando per R':

$$\underset{q \times K}{R'} \underset{K \times 1}{(\hat{\beta} - \beta)} \sim N\Big(\underbrace{\underset{q \times K}{R'} \underset{K \times 1}{0}}_{q \times 1}, \; \sigma^2 \underbrace{\underset{q \times K}{R'} \underset{K \times K}{(X'X)^{-1}} \underset{K \times q}{R}}_{q \times q}\Big)$$

Sotto H_0, $R'\beta = r$, da cui

$$\underset{q \times 1}{R'(\hat{\beta} - \beta)} \overset{H_0}{=} \underset{q \times 1}{R'\hat{\beta} - r} \sim N\big(\underset{q \times 1}{0}, \; \underset{q \times q}{\sigma^2 R'(X'X)^{-1}R}\big) \quad (\text{dato } X)$$

Usiamo la seguente proprietà:

$$\boxed{\begin{array}{l} \text{Se } Y \sim N(0, V),\ V \text{ rango pieno, allora } Y'V^{-1}Y \sim \chi^2(m) \\ \text{dove } m \text{ è la dimensione del vettore } Y \end{array}}$$

$$Y \sim N(0, \underline{\underline{I_m}}) \qquad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_m \end{pmatrix}, \quad Y_i \sim N(0,1) \text{ indipendenti}$$

$$Y' V^{-1} Y = Y'Y = \sum_{i=1}^{m} \underbrace{Y_i}_{\to N(0,1)}{}^2 = \sum_{i=1}^{m} \{N(0,1)\}^2 =: \chi^2(m)$$

$$Y \sim N(0, V) \qquad V^{1/2} : \quad V^{1/2} \cdot V^{1/2} = V$$

$$\tilde{Y} := V^{-1/2} Y \sim N(0, \underbrace{V^{-1/2} \cdot V \cdot V^{-1/2}}_{I_m}) \equiv N(0, I_m)$$

$$Y' V^{-1} Y = Y' V^{-1/2} V^{-1/2} Y = (V^{-1/2} Y)' (V^{-1/2} Y)$$

$$= \underbrace{\tilde{Y}' \tilde{Y}}_{\to N(0, I_m)} =: \chi^2(m).$$

Torniamo a $R'\hat{\beta} - r$:

$$R'\hat{\beta} - r \overset{H_0}{\sim} N(\underset{q\times 1}{0}, \sigma^2 [R'(X'X)^{-1} R]) \qquad \text{dato } X$$

Poiché, se rango$(X) = K$ e rango$(R) = q$, allora $R'(X'X)^{-1}R$ ha rango pieno, possiamo applicare la precedente proprietà:

$$W_n := (R'\hat{\beta} - r)'\left(\sigma^2 [R'(X'X)^{-1}R]\right)^{-1}(R'\hat{\beta} - r) \overset{H_0}{\sim} \chi^2(q) \quad , \text{dato } X$$

Se conoscessimo σ^2, potremmo usare W_n per testare H_0 (test χ^2).

Possiamo sostituire σ^2 con $\hat{\sigma}^2$:

$$\tilde{W}_n := (R'\hat{\beta} - r)'\left(\hat{\sigma}^2 [R'(X'X)^{-1}R]\right)^{-1}(R'\hat{\beta} - r)$$

Potremmo confrontare $\tilde{W}_n$ con i valori critici della $\chi^2(q)$.
In questo caso il test non è "esatto".

In alternativa si può considerare:

$$F_n := \tilde{W}_n / q = \frac{1}{\hat{\sigma}^2}(R'\hat{\beta} - r)'(R'(X'X)^{-1}R)^{-1}(R'\hat{\beta} - r)/q$$

Si dimostra che

$$F_n \overset{H_0}{\sim} F(q, n-k) \quad , \text{dato } X.$$

Dimostrazione

$$F_n = \frac{1}{\hat{\sigma}^2/\sigma^2} \cdot \underbrace{(R'\hat{\beta}-r)'\left(\sigma^2 R'(X'X)^{-1}R\right)^{-1}(R'\hat{\beta}-r)/q}_{\tilde{W}_n \overset{H_0}{\sim} \chi^2(q) \text{ sotto } H_0}$$

$$(n-K)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-K)$$

$$\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-K)/(n-K)$$

numeratore $\left((R'\hat{\beta}-r)'(\ldots\ldots \right)$ e denominatore $\left(\hat{\sigma}^2/\sigma^2\right)$ sono indipendenti. Quindi:

$$F_n \overset{H_0}{=} \frac{\chi^2(q)/q}{\chi^2(n-K)/(n-K)} =: F(q, n-K)$$

in quanto i due χ^2 sono tra loro indipendenti.
Si dimostra usando la proprietà ③ vista in precedenza.

Alternativamente, l'indipendenza si dimostra usando il seguente risultato.

Se $Y \sim N(0, V)$, rango(V) = max, allora le forme quadratiche

$$Y'MY \quad e \quad Y'NY$$

sono indipendenti se e solo se $M \cdot N = 0$

Nel nostro caso, $\varepsilon \sim N(0, \sigma^2 I_n)$, e

$$(n-k)\frac{\hat{\sigma}^2}{\sigma^2} = \varepsilon' M_X \varepsilon \quad , \quad M_X := I_n - X(X'X)^{-1}X'$$

Analogamente,

$$\hat{\beta} - \beta = (X'X)^{-1}X'\varepsilon$$

$$R'\hat{\beta} - r \overset{H_0}{=} R(X'X)^{-1}X'\varepsilon$$

e

$$(R'\hat{\beta} - r)'[R'(X'X)^{-1}R]^{-1}(R'\hat{\beta} - r) = \varepsilon' N \varepsilon$$

$$N := X(X'X)^{-1}R[R'(X'X)^{-1}R]^{-1}R(X'X)^{-1}X'$$

che soddisfa

$$M_X N = 0$$

nota: non tutte le ipotesi sono lineari:

$$\beta_1 \cdot \beta_2 = 1$$

$$\beta_1 + \beta_2^2 = 0$$

Vincoli non lineari: $\underset{q\times 1}{g}(\beta) = \underset{q\times 1}{0}$ $\quad g: \mathbb{R}^K \to \mathbb{R}^q$

Test nel modello lineare generalizzato.

Abbiamo ipotizzato $\varepsilon | X \sim N(0, \sigma^2 I_n)$.

Questa ipotesi è cruciale per derivare la distribuzione di F_n / t_n sotto H_0.

Cosa succede se invece $\varepsilon | X \sim N(0, \Sigma)$, $\Sigma \neq \sigma^2 I_n$?

In questo caso siamo nell'ambito del mod. lineare generalizzato

Se usiamo OLS:

$$\hat{\beta} - \beta = (X'X)^{-1}X'\varepsilon \sim N(\underset{K\times 1}{0}, V), \quad V := (X'X)^{-1}X'\Sigma X(X'X)^{-1}$$

$$R'\hat{\beta} - r \overset{H_0}{=} N(\underset{q\times 1}{0}, R'VR)$$

La trasformazione $N \to \chi^2$ è:

$$\underbrace{(R'\hat{\beta} - r)'[R'VR]^{-1}(R'\hat{\beta} - r)}\overset{H_0}{\sim} \chi^2(q)$$

$$(R'\hat{\beta} - r)'[R'(X'X)^{-1}X'\Sigma X(X'X)^{-1}R]^{-1}(R'\hat{\beta} - r) \overset{H_0}{\sim} \chi^2(q)$$

In pratica, potrei usare questo risultato per ottenere una statistica test nel caso in cui sia possibile stimare

$$X'\Sigma X$$

Non sempre si può fare; un caso in cui è possibile è il modello lineare con errori eteroschedastici, usando lo stimatore di White di $X'\Sigma X$.

La verifica di ipotesi nel modello generalizzato (come anche la stima), è complicata dal fatto che Σ non è nota.

Nei casi in cui Σ sia nota, si può trasformare il modello generalizzato in un modello lineare ordinario, e poi fare inferenza in quest'ultimo.

Stima di modelli vincolati

esempio Cobb-Douglas $\quad y = a + \alpha \ell + \beta k \qquad H_0 : \alpha + \beta = 1$

$$\Updownarrow$$

$$\beta = 1 - \alpha$$

Possiamo imporre l'ipotesi nulla:

$$y = a + \alpha \ell + (1-\alpha) k$$

$$y = a + \alpha \ell + k - \alpha k$$

$$\underbrace{y - k}_{\tilde{y}} = a + \alpha \underbrace{(\ell - k)}_{x}$$

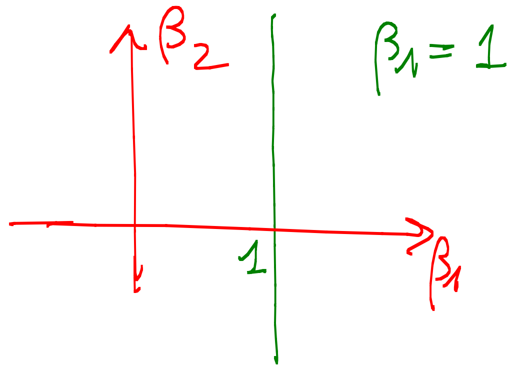
$$\tilde{y} = a + \alpha x \quad \Longrightarrow \quad \begin{cases} \hat{\alpha} \\ \hat{\beta} = 1 - \hat{\alpha} \end{cases}$$

Come possiamo stimare β imponendo il vincolo $R'\beta - r = 0$?

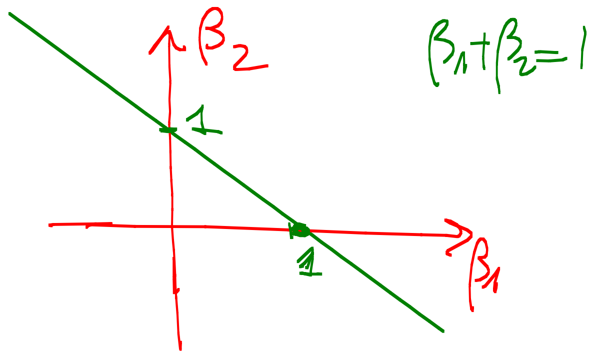
$$\hat{\beta} := \arg\min_{\tilde{\beta} \in \mathbb{R}^K} Q(\tilde{\beta}) \quad , \quad Q(\tilde{\beta}) = (y - X\tilde{\beta})'(y - X\tilde{\beta})$$

Stimatore vincolato:

$$\hat{\beta}_{RLS} := \arg\min_{\tilde{\beta} \in \{x \in \mathbb{R}^K \,:\, R'x - r = 0\}} Q(\tilde{\beta})$$



RLS = Restricted least squares.



$f(x)$ da minimizzare sotto il vincolo $g(x) = 0$

$$L(x, \lambda) = f(x) - \lambda' g(x) \qquad \text{LAGRANGIANA}$$

λ → vettore dei moltiplicatori di Lagrange.

Si dimostra che il punto di minimo si può trovare risolvendo il problema

$$\begin{cases} \dfrac{\partial L(x,\lambda)}{\partial x} = 0 \\ \dfrac{\partial L(x,\lambda)}{\partial \lambda} = 0 \end{cases}$$

Nel nostro caso:

$$Q(\tilde\beta) = (y - X\tilde\beta)'(y - X\tilde\beta) = y'y - 2\tilde\beta' X' y + \tilde\beta' X' X \tilde\beta$$

vincolo: $R'\beta - r = 0$

Costruiamo la lagrangiana:

$$L(\tilde\beta, \lambda) = \underset{1\times 1}{Q(\tilde\beta)} - 2\underset{1\times q}{\lambda'}(\underbrace{\underset{q\times k\, k\times 1}{R'\tilde\beta}}_{q\times 1} - \underset{q\times 1}{r})$$

$$\begin{cases} \dfrac{\partial L(\tilde\beta,\lambda)}{\partial \tilde\beta} = 0 \\ \dfrac{\partial L(\tilde\beta,\lambda)}{\partial \lambda} = 0 \end{cases}$$

$$\boxed{\frac{\partial}{\partial x} x'Ax = 2Ax}$$

$$\boxed{\frac{\partial}{\partial x} x'A = A}$$

$$\underbrace{\lambda' R' \beta}_{\beta' R \lambda} - \lambda' r$$

$$\frac{\partial}{\partial \beta} \beta' R \lambda = R\lambda$$

$$
\begin{cases}
\dfrac{\partial L(\tilde\beta,\lambda)}{\partial \tilde\beta} = -2X'y + 2X'X\tilde\beta - 2R\lambda = \underset{K\times 1}{0} \\[2ex]
\dfrac{\partial L(\tilde\beta,\lambda)}{\partial \lambda} = -2(R'\tilde\beta - r) = \underset{q\times 1}{0}
\end{cases}
$$

$$
\begin{cases}
-X'y + X'X\tilde\beta - R\lambda = 0 \\
R'\tilde\beta - r = 0
\end{cases}
$$

K+q equazioni
in K+q incognite

$$X'X\tilde\beta = X'y + R\lambda$$

$$\tilde\beta = (X'X)^{-1}X'y + (X'X)^{-1}R\lambda$$

$$R'\tilde\beta - r = R'(X'X)^{-1}X'y + R'(X'X)^{-1}R\lambda - r = \underset{q\times 1}{0}$$

$$R'\underbrace{(X'X)^{-1}X'y}_{} - r = -\left[R'(X'X)^{-1}R\right]\lambda$$

$$R'\hat{\beta}_{OLS} - r = -[R'(X'X)^{-1}R]\cdot\lambda$$

$\underbrace{R'\hat{\beta}_{OLS} - r}_{}$ "distanza" di $\hat{\beta}_{OLS}$ dal vincolo

$$\hat{\lambda} = -[R'(X'X)^{-1}R]^{-1}(R'\hat{\beta}_{OLS} - r)$$

$$\hat{\beta}_{RLS} = \underbrace{(X'X)^{-1}X'y}_{\hat{\beta}_{OLS}} + (X'X)^{-1}R\hat{\lambda} = \hat{\beta}_{OLS} - (X'X)^{-1}R[R'(X'X)^{-1}R]^{-1}(R'\hat{\beta}_{OLS} - r)$$

nota $\hat{\beta}_{RLS} = \hat{\beta}_{OLS}$ se e solo se $R'\hat{\beta}_{OLS} - r = 0$, ovvero se $\hat{\beta}_{OLS}$ soddisfa il vincolo.

nota $\hat{\beta}_{RLS}$ è corretto (dato X) solo se il vincolo è vero. Infatti:

$$E[\hat{\beta}_{RLS}|X] = E[\hat{\beta}_{OLS}|X] - E\left[(X'X)^{-1}R[R'(X'X)^{-1}R]^{-1}(R'\hat{\beta}_{OLS} - r)|X\right]$$

$$= \beta - (X'X)^{-1}R[R'(X'X)^{-1}R]^{-1}(R'E(\hat{\beta}_{OLS}|X) - r)$$

$$= \beta - (X'X)^{-1}R[R'(X'X)^{-1}R]^{-1}(R'\beta - r)$$

$= \beta$ solo se $R'\beta - r = 0$ (ovvero se il vincolo è vero)

nota: $V[\hat{\beta}_{RLS} | X] \leq V[\hat{\beta}_{OLS} | X]$ sempre (indipendentemente da H_0)

Vincoli in forma esplicita

I vincoli nella forma $R'\beta - r = 0$ si dicono in forma implicita.

Forma esplicita: esprimiamo β in funzione di un "nuovo" vettore di parametri, φ, di dimensione più piccola.

Esempio. $\beta_1 = 0$ $\beta = (\beta_1, \beta_2, \dots, \beta_k)'$

$$\underset{K \times 1}{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}} = \underset{K \times (K-1)}{\begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 1 \end{bmatrix}} \underset{(K-1) \times 1}{\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{k-1} \end{bmatrix}}$$

$$\left[\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} 0 \\ \varphi_1 \\ \vdots \\ \varphi_{k-1} \end{pmatrix} \right]$$

Esempio $\qquad \beta_1 = \beta_2 = \beta_3 \qquad q = 2$ vincoli

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \vdots \\ \beta_K \end{bmatrix} = \begin{bmatrix} \varphi_1 \\ \varphi_1 \\ \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{K-2} \end{bmatrix} \iff \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \vdots \\ \beta_K \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{K-2} \end{bmatrix}$$

$$\underset{K \times 1}{\beta} = \underset{K \times S}{H} \cdot \underset{S}{\varphi}$$

$$S = K - 2 = K - \#\text{ vincoli}$$

In generale, i vincoli in forma esplicita si possono scrivere come:

$$\underset{K\times 1}{\beta} = \underset{K\times S}{H}\,\underset{S\times 1}{\varphi} + \underset{K\times 1}{h}$$

S è il # parametri meno il # di vincoli.

$\boxed{\text{esercizio}: \quad \beta_1 = 1, \ \beta_2 = 2\beta_3}$

E' facile ottenere un modello con vincoli in forma esplicita (non servono i 'minimi quadrati vincolati'):

$$\begin{cases} y = X\beta + \varepsilon & \leftarrow \text{modello} \\ \underset{K\times 1}{\beta} = \underset{K\times S}{H}\,\underset{S\times 1}{\varphi} + \underset{K\times 1}{h} & \leftarrow \text{vincolo} \end{cases}$$

$s = K - q, \quad q = \#\ \text{vincoli}$

Sostituiamo il vincolo nell'equazione del modello:

$$y = X(H\varphi + h) + \varepsilon$$

$$y = XH \cdot \underline{\underline{\varphi}} + Xh + \varepsilon$$

$$\underbrace{y - Xh}_{\tilde{y}} = \underbrace{XH}_{\tilde{X}}\varphi + \varepsilon$$

$$\tilde{y} = \tilde{X}\varphi + \varepsilon$$

Abbiamo una nuova variabile dipendente $\tilde{y}$, e un nuovo insieme di regressori: $\tilde{X}$ – Non abbiamo toccato ε

Usiamo OLS per stimare φ:

$$\hat{\varphi} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y} = (H'X'XH)^{-1} H'X'(y - Xh)$$
$$= (H'X'XH)^{-1}H'X'y - (H'X'XH)^{-1}H'X'Xh$$

Ottenuto $\hat{\varphi}$, ricaviamo $\hat{\beta}$ usando il fatto che $\beta = H\varphi + h$:

$$\hat{\beta} = H\hat{\varphi} + h$$

$$\hat{\beta} = H(H'X'XH)^{-1}H'X'y - H(H'X'XH)^{-1}H'X'Xh + h$$