Analisi della specificazione del modello

y_t variabile dipendente
x_t regressore

$$E[y_t|x_t] = \beta' x_t \qquad \text{oppure} \qquad y_t = x_t'\beta + \varepsilon_t, \quad E[\varepsilon_t|x_t]=0$$

[ML classico] $V[y_t|x_t] = \sigma^2$ $\qquad V[\varepsilon_t|x_t]=0$

Inoltre (incorrelazione):

$$C[y_t, y_{t'}|X]=0 \quad t \neq t' \qquad C[\varepsilon_t, \varepsilon_{t'}|X]=0$$

Specificare un modello di questo tipo implica:

- ipotesi sulla media condizionale di y_t
- " " componente stocastica
- " " incorrelazione tra x_t e ε_t

{ x_t contiene tutti i regressori rilevanti
{ linearità
{ omosch. incorrelaz. normalità

Consideriamo il problema della scelta dei regressori.

Caso di modello sottoparametrizzato

DGP $\quad y = X_1\beta_1 + X_2\beta_2 + \varepsilon \qquad \varepsilon | X_1, X_2 \sim (0, \sigma^2 I_n)$

modello stimato: $\quad y = X_1\beta_1 + \varepsilon$

Proprietà di $\hat\beta_1 = (X_1'X_1)^{-1}X_1'y$

$$\hat\beta_1 = (X_1'X_1)^{-1}X_1'(X_1\beta_1 + X_2\beta_2 + \varepsilon) =$$

$$= \cancel{(X_1'X_1)^{-1}X_1'X_1}\beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 + (X_1'X_1)^{-1}X_1'\varepsilon$$

$$= \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 + (X_1'X_1)^{-1}X_1'\varepsilon$$

$$E[\hat\beta_1 | X_1, X_2] = \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 \qquad \text{in quanto } E[X_1'\varepsilon | X_1, X_2] = 0$$

In generale, $E[\hat{\beta}_1 | X_1, X_2] \neq \beta_1$ - Lo stimatore NON è corretto.

Facendo l'analisi asintotica, $\hat{\beta}_1 - \beta_1 \not\xrightarrow{P} 0$ (inconsistente)

Gli unici casi in cui $E[\hat{\beta} | X_1, X_2] = \beta$ sono:

(i) $\beta_2 = 0$ [X_2 non entra nel modello]

(ii) $X_1' X_2 = 0$ [X_1 e X_2 sono ortogonali]

Cosa succede alla stima di σ^2 ?

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-K_1}, \quad \hat{\varepsilon} := y - X_1\hat{\beta}_1 = M_{X_1} y$$

(K_1 è il rango di X_1)

$$\hat{\sigma}^2 = \frac{y' M_{X_1} y}{n-K_1} = \frac{(\cancel{X_1\beta_1} + X_2\beta_2 + \varepsilon)' M_{X_1} (\cancel{X_1\beta_1} + X_2\beta_2 + \varepsilon)}{n-K_1}$$

$$= \frac{\beta_2' X_2' M_{X_1} X_2 \beta_2}{n-K_1} + \frac{\beta_2' X_2' M_{X_1} \varepsilon}{n-K_1} + \frac{\varepsilon' M_{X_1} X_2 \beta_2}{n-K_1} + \frac{\varepsilon' M_{X_1} \varepsilon}{n-K_1}$$

$$E[\hat{\sigma}^2 | X_1, X_2] = \frac{\beta_2' X_2' M_{X_1} X_2 \beta_2}{n-K_1} + 0 + 0 + \underbrace{\frac{E[\varepsilon' M_{X_1} \varepsilon | X_1, X_2]}{n-K_1}}_{\sigma^2} > \sigma^2$$

La varianza di ε viene sovrastimata

Caso di modello sovraparametrizzato

DGP: $y = X_1 \beta_1 + \varepsilon$

modello stimato $y = X_1 \beta_1 + X_2 \beta_2 + \varepsilon = X\beta + \varepsilon$, $\quad \underset{n \times k}{X} = [\underset{n \times K_1}{X_1} \vdots \underset{n \times K_2}{X_2}]$, $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$

$k = (K_1 + K_2)$

Proprietà di $\hat{\beta} = (X'X)^{-1} X' y$

$$\hat{\beta} = (X'X)^{-1} X' (X_1 \beta_1 + \varepsilon) = \underbrace{(X'X)^{-1} X' X_1}_{} \beta_1 + (X'X)^{-1} X' \varepsilon$$

$\rightarrow$ coefficiente (stimato) della regressione di X_1 su $[X_1 \vdots X_2] = X$

da cui

$$(X'X)^{-1}X'X_1 = \begin{bmatrix} I_{k_1} \\ 0 \end{bmatrix}$$
$$\underbrace{\underset{k\times k}{\ } \quad \underset{k\times n}{\ } \quad \underset{n\times k_1}{\ }}_{k\times k_1}$$

Quindi:

$$\hat\beta = \begin{bmatrix} \hat\beta_1 \\ \hat\beta_2 \end{bmatrix} = \begin{bmatrix} I_{k_1} \\ 0 \end{bmatrix}\beta_1 + (X'X)^{-1}X'\varepsilon$$
$$= \begin{bmatrix} \beta_1 \\ 0 \end{bmatrix} + (X'X)^{-1}X'\varepsilon$$

$$E[\hat\beta|X] = \begin{bmatrix} \beta_1 \\ 0 \end{bmatrix} + \underbrace{E[(X'X)^{-1}X'\varepsilon|X]}_{0} = \begin{bmatrix} \beta_1 \\ 0 \end{bmatrix}$$

Lo stimatore è corretto.

E' corretto anche $\hat\sigma^2$: $E[\hat\sigma^2|X] = E\left[\frac{\varepsilon' M_X \varepsilon}{n-k}\Big|X\right] = \sigma^2$

Quello che cambia è la varianza di $\hat\beta$:

$$V(\hat\beta|X) = \sigma^2(X'X)^{-1} = \sigma^2\begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1}$$

Ci interessa $V(\hat\beta_1|X)$.

Proprietà: Se $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, allora $M^{-1} = \left(\begin{array}{c|c} (A - BD^{-1}C)^{-1} & \dots\dots \\ \hline \dots\dots & \dots\dots \end{array}\right)$

Osserviamo:

$$V(\hat{\beta}_1 | X) = \sigma^2 \left[X_1' X_1 - X_1' X_2 (X_2' X_2)^{-1} X_2' X_1 \right]^{-1}$$

Notiamo che è maggiore della varianza dello stimatore efficiente:

$$\sigma^2 (X_1' X_1)^{-1}$$

Lo stimatore è quindi inefficiente.

Nota: modello di regressione partizionato in blocchi

$$y = X\beta + \varepsilon = X_1\beta_1 + X_2\beta_2 + \varepsilon \qquad X = [X_1, X_2], \quad \beta = \begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}$$

Lo stimatore OLS è

$$\hat\beta = \begin{pmatrix}\hat\beta_1\\ \hat\beta_2\end{pmatrix} = (X'X)^{-1}X'y = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2\end{bmatrix}^{-1}\begin{bmatrix}X_1'y\\ X_2'y\end{bmatrix}$$

Supponiamo di essere interessati a $\hat\beta_2$ (e non a $\hat\beta_1$). Sia $M_{X_1} := I_n - X_1(X_1'X_1)^{-1}X_1'$. Si dimostra (teorema Frisch–Waugh) che vale la seguente espressione

$$\hat\beta_2 = (X_2'M_{X_1}X_2)^{-1}X_2'M_{X_1}y$$

che corrisponde alla stima OLS della regressione di $\tilde y := M_{X_1}y$ su $\tilde X_2 = M_{X_1}X_2$.

Quindi, la stima di β_2 si ottiene regredendo tra loro i residui delle proiezioni di y e X_2 su X_1. Questo risultato semplifica il calcolo della varianza di $\hat\beta_2$:

$$\begin{aligned}
Var(\hat\beta_2 | X) &= Var\left((X_2'M_{X_1}X_2)^{-1}X_2'M_{X_1}y \,|\, X\right)\\
&= (X_2'M_{X_1}X_2)^{-1}X_2'M_{X_1}V(y|X)M_{X_1}X_2(X_2'M_{X_1}X_2)^{-1}\\
&= \sigma^2(X_2'M_{X_1}X_2)^{-1}X_2'\cancel{M_{X_1}X_2}(\cancel{X_2'M_{X_1}}X_2)^{-1} = \sigma^2(X_2'M_{X_1}X_2)^{-1}\\
&= \sigma^2(X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2)^{-1}
\end{aligned}$$

Errata specificazione della forma funzionale

$$E[y_t | x_t] = x_t' \beta \qquad \text{(ipotesi di linearità)}$$

DGP $\quad E[y_t | x_t] = f(x_t, \beta) \neq x_t' \beta$

esempio: stimato: $\quad y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$

DGP $\quad y_t = \beta_0 + \beta_1 \log(x_t) + \varepsilon_t$

Gli effetti di questo tipo di errore sono analoghi al caso di modello sottoparametrizzato (abbiamo escluso il regressore rilevante $\log(x_t)$):

$$E[X\hat{\beta} | X] \neq E[y | X]$$

Per verificare la correttezza della specificazione della forma funzionale si usa il test RESET.

Test per l'esclusione di regressori

modello stimato: $\quad y = X_1 \underset{K_1\times 1}{\beta_1} + X_2 \underset{K_2\times 2}{\beta_2} + \varepsilon$

ipotesi nulla di esclusione di X_2: $\quad H_0: \beta_2 = \underset{K_2\times 2}{0} \quad$ (K_2 vincoli)

Forma $H_0: R'\beta - r = 0$:

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \qquad \underset{K_2\times(K_1+K_2)}{\begin{bmatrix} 0 & I_{K_2} \end{bmatrix}} \underset{(K_1+K_2)\times 1}{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}} = \underset{K_2\times 1}{0} \qquad \underset{K\times K_2}{R} = \begin{bmatrix} 0 \\ I_{K_2} \end{bmatrix} \qquad \underset{K_2\times 1}{r = 0}$$

Si usa un test F, basato su $\quad F_n := \frac{1}{K_2}(R'\hat\beta - r)'\left[\hat\sigma^2 R'(X'X)^{-1}R\right]^{-1}(R'\hat\beta - r)$

Poiché: - $R'\hat\beta - r = \hat\beta_2$

- $R'(X'X)^{-1}R$ è il blocco in basso a destra di $(X'X)^{-1}$ che, usando le proprietà scritte precedentemente, è uguale a

$$(X_2'X_2)^{-1} - X_2'X_1(X_1'X_1)^{-1}X_1'X_2$$

- $\hat\sigma^2 R'(X'X)^{-1}R$ corrisponde a $\widehat{Var}(\hat\beta_2|X)$

la statistica test si può anche scrivere come

$$F_n := \hat{\beta}_2' \, \widehat{\mathrm{Var}}(\hat{\beta}_2 | X)^{-1} \hat{\beta}^2 / k_2 =$$

$$= y' M_{X_1} X_2 (X_2' M_{X_1} X_2)^{-1} \left(\hat{\sigma}^2 \cancel{(X_2' M_{X_1} X_2)^{-1}} \right)^{-1} \cancel{(X_2' M_{X_1} X_2)^{-1}} X_2' M_{X_1} y \, / k_2$$

$$= y' M_{X_1} X_2 (X_2' M_{X_1} X_2)^{-1} X_2' M_{X_1} y \Big/ \hat{\sigma}^2 k_2$$

Test di inclusione di regressori

modello stimato: $y = X_1 \beta_1 + \varepsilon$

Vogliamo testare se il modello vero è $y = X_1 \beta_1 + X_2 \beta_2 + \varepsilon$

Possibile strategia: stimare il modello allargato, $y = X_1 \beta_1 + X_2 \beta_2 + \varepsilon$, e testare (test F) se $\beta_2 = 0$

Strategia alternativa:
- consideriamo i residui del modello stimato, $\hat{\varepsilon} = y - X_1 \hat{\beta}_1$
- regrediamo $\hat{\varepsilon}$ su X_2 e testiamo (F) la significatività di questa regressione

Nei modelli lineari, le due strategie sono equivalenti

Test RESET - regression specification error test

$$y_t = g(x_t, \beta) + \varepsilon_t \quad , \quad E(\varepsilon_t | X) = 0, \quad V(\varepsilon | X) = \sigma^2 I_n$$

esempio: $g(x, \beta) = e^{x'\beta}$

$$y_t = e^{x_t'\beta} + \varepsilon_t$$

$H_0: g(x, \beta) = x'\beta$ ipotesi di linearità

Passi del test:

1. stima sotto $H_0 \Rightarrow \hat{\beta}$, $\hat{y}_t = x_t'\hat{\beta}$ (valori stimati)
2. stimare la regressione:

$$y_t = x_t'\beta + \gamma_2 \hat{y}_t^2 + \gamma_3 \hat{y}_t^3 + \dots + \gamma_m \hat{y}_t^m + \text{errore}_t$$

per un m selezionato (di solito $m = 2$ 3)

3. test F per $H_0: \gamma_2 = \gamma_3 = \dots = \gamma_m = 0$

Effetti della violazione dell'ipotesi di sfericità di ε

$$E[\varepsilon_t^2 | X] = \sigma_t^2 \quad \text{(eteroschedasticità)}$$

$$E[\varepsilon_t \varepsilon_{t'} | X] \neq 0 \quad t \neq t' \quad \text{(autocorrelazione)}.$$

Abbiamo visto che, se viene usato OLS sotto le ipotesi del modello generalizzato:

DGP: $y = X\beta + \varepsilon$, $E[\varepsilon | X] = 0$, $V[\varepsilon | X] = \Sigma \neq \sigma^2 I_n$

allora

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

$$V[\hat{\beta} | X] = (X'X)^{-1} X' \Sigma X (X'X)^{-1}$$

Conseguentemente, lo stimatore OLS di $V[\hat{\beta} | X]$,

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

non è né corretto, né consistente.

Test di omoschedasticità

- White
- Breusch-Pagan

Test di incorrelazione seriale

- Durbin-Watson
- Portmanteau
- Breusch-Godfrey.

Test di White

$$y = X\beta + \varepsilon \quad , \quad E[\varepsilon | X] = 0$$

$$V[\varepsilon | X] = \mathrm{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2\}$$

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

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

$$\hat{V}_W(\hat{\beta}) = (X'X)^{-1} X'\hat{\Sigma} X (X'X)^{-1} \quad \text{dove} \quad X'X = \sum_{t=1}^{n} x_t x_t'$$

$$X'\hat{\Sigma} X = \sum_{t=1}^{n} \hat{\varepsilon}_t^2 x_t x_t'$$

Sotto H_0, $\hat{V}(\hat{\beta})$ e $\hat{V}_W(\hat{\beta})$ devono essere vicine tra loro.

$$\underline{\hat{V}(\hat{\beta})} - \underline{\hat{V}_W(\hat{\beta})} = (X'X)^{-1} \hat{\sigma}^2 (X'X)(X'X)^{-1} - (X'X)^{-1} X'\hat{\Sigma} X (X'X)^{-1}$$

$$= (X'X)^{-1} \underbrace{\left[\hat{\sigma}^2 (X'X) - X'\hat{\Sigma} X \right]}_{\hat{A}_n} (X'X)^{-1}$$

dove

$$\hat{A}_n = \hat{\sigma}^2 \sum_{t=1}^{n} x_t x_t' - \sum_{t=1}^{n} \hat{\varepsilon}_t^2 x_t x_t' = -\sum_{t=1}^{n} \left(\hat{\varepsilon}_t^2 - \hat{\sigma}^2 \right) x_t x_t'$$

Sotto H_0, $\hat{A}_n \frac{1}{n} \xrightarrow{p} 0$, $\frac{1}{\sqrt{n}} \hat{A}_n \xrightarrow{d} N$ multivariata.

White dimostra che, per calcolare la statistica test, possiamo procedere nel seguente modo:

① Stima OLS della regressione di y su X, e calcolo dei residui:

$$\hat{\varepsilon} = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_n)'$$

② Regressione di $\hat{\varepsilon}_t^2$ su $x_t = (x_{1t}, \dots, x_{Kt})'$ e

$$x_{1t}^2, x_{2t}^2, \dots, x_{Kt}^2, \; x_{1t}x_{2t}, x_{1t}x_{3t}, \dots, x_{K-1t}x_{Kt}$$

③ Calcolare un test F per la significatività dei regressori

Alternativamente, test χ^2 calcolato come:

$$LM = n \cdot R^2$$

dove R^2 è l'ind. det. lineare della regressione ② nota: la regressione al passo ② contiene $K + \frac{K(K+1)}{2}$ regressori. A volte sono troppi.

Una possibile soluzione è escludere i prodotti incrociati, e quindi sostituire il passo ② con il seguente:

②' Regressione di $\hat{\varepsilon}_t^2$ su $x_t = (x_{1t}, \dots, x_{Kt})'$ e $x_{1t}^2, x_{2t}^2, \dots, x_{Kt}^2$

Questa regressione ha $2K$ regressori.

Nota: Se $x_{1t} = 1$ (il modello ha una costante), non dobbiamo includere (ai passi ② o ②') x_{1t}^2 e $x_{1t} x_{jt}$ $(j \neq 1)$

note

Al passo ③ abbiamo menzionato la statistica $LM = nR^2$. Perché?

Supponiamo di avere la regressione

$$y_t = \beta_1 + \beta_2 x_{2t} + \dots + \beta_k x_{kt} + \varepsilon_t$$

Vogliamo testare

$$H_0: \beta_2 = \beta_3 = \dots = \beta_k = 0 \qquad (k-1 \text{ vincoli})$$

Abbiamo visto come calcolare la statistica test F
$\left(F \underset{H_0}{\sim} F(k-1, n-k)\right)$

Si può dimostrare che:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

dove R^2 è l'indice di determinazione lineare.

Notiamo che, sotto H_0, $R^2 \simeq 0$. Quindi:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = \frac{(n-k)}{(k-1)} \frac{R^2}{1-R^2}$$

$$\cong \frac{(n-k)R^2}{(k-1)} \quad \text{che possiamo scrivere come:}$$

$$(n-k)R^2 \cong (k-1)\cdot F(k-1, n-k) \qquad (\text{sotto } H_0)$$

$$\boxed{q\,F(q,n) \xrightarrow[n\to\infty]{} \chi^2(q)}$$

$$nR^2 \cong \underbrace{(k-1)\,F(k-1, n-k)}_{\approx\ \chi^2(k-1) \text{ per } n \text{ grande}}$$

Riassumendo, sotto H_0,

$$LM = nR^2 \xrightarrow{d} \chi^2(k-1)$$

quando $n \to \infty$.

Test di Breusch-Pagan

$$y_t = \beta' x_t + \varepsilon_t \quad , \quad E[\varepsilon_t | x] = 0 \quad , \quad V[\varepsilon_t | x] = \sigma_t^2$$

$$\sigma_t^2 = h(\alpha_0 + \alpha_1' z_t) > 0$$

dove z_t non coincide necessariamente con x_t.
L'ipotesi di omoschedasticità corrisponde alla:

$$H_0 : \alpha_1 = 0$$

Infatti, in questo caso, $\sigma_t^2 = h(\alpha_0) =$ costante.
Il test si sviluppa nei seguenti passi:

1. stima del modello sotto H_0 (regressione di y su X) e calcolo dei residui $\hat{\varepsilon}_t$
2. regressione di $\hat{\varepsilon}_t^2$ su $(1, \underset{m \times 1}{z_t'})'$
3. Rifiuto H_0 se la statistica F per la significatività della regressione (2) supera il valore critico ($F(m, n-m-1)$)

Alternativamente:

$$LM = n \cdot R^2 \xrightarrow[H_0]{d} \chi^2(m)$$

(R^2 ottenuto in 2)

nota: GRETL pone $(1, z_t')' = x_t'$.

Test di incorrelazione seriale.

Durbin-Watson

$$y = X\beta + \varepsilon \qquad y_t = x_t'\beta + \varepsilon_t \qquad t = 1, \dots, n$$

$$\varepsilon_t = \alpha\, \varepsilon_{t-1} + u_t, \qquad u_t \sim iid(0, \sigma^2)$$

Per $\alpha \neq 0$, ε_t è autocorrelato.

$H_0: \alpha = 0$ corrisponde all'ipotesi di assenza di autocorrelazione.

Il test si sviluppa nei seguenti passi:

① stima OLS della regressione di y_t su x_t e calcolo dei residui $\hat{\varepsilon}_t$

② Calcolo della statistica

$$DW = \frac{\sum_t (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum \hat{\varepsilon}_{t-1}^2} \cong 2(1-\hat{\alpha})$$

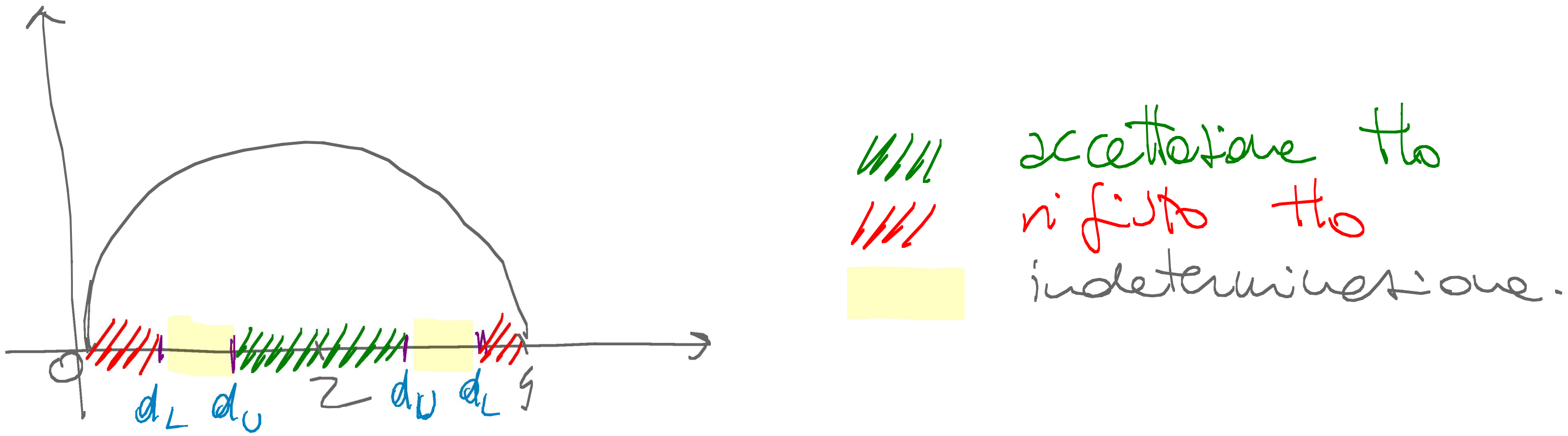
dove $\hat{\alpha} = \frac{\sum \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum \hat{\varepsilon}_{t-1}^2}$ (autocorrelazione di $\hat{\varepsilon}_t$ al ritardo 1)

I valori critici non sono standard. In particolare

Sotto $H_0: \alpha = 0$, $\hat{\alpha} \simeq 0$. In questo caso $DW \cong 2$

Se $\hat{\alpha} = 1$ (massima autocorrelazione positiva), allora $DW \cong 0$

Se $\hat{\alpha} = -1$ (" " negativa), allora $DW \cong 4$



Note – Alcuni software (GRETL) presentano il p-value.

– il test richiede che vi sia una costante nel modello

– il test non funziona se, tra i regressori del modello, compare la dipendente ritardata.

Test portmanteau

Si basa sul seguente risultato:

$$\varepsilon_t \sim iid(0, \sigma^2)$$

$\hat{\rho}_K$ = coefficiente di autocorrelaz. campionario, ritardo K.

$$\hat{\rho}_k = \frac{\sum \varepsilon_t \varepsilon_{t-k}}{\sum \varepsilon_t^2}$$

Costruiamo il vettore

$$\sqrt{n} \underbrace{\begin{pmatrix} \hat{\rho}_1 \\ \hat{\rho}_2 \\ \vdots \\ \hat{\rho}_m \end{pmatrix}}_{m \times 1} \xrightarrow{d} N(\underbrace{0}_{m \times 1}, \underbrace{I_m}_{m \times m})$$

da cui

$$Q_m := n \sum_{k=1}^{m} \hat{\rho}_k^2 \xrightarrow{d} \chi^2(m)$$

(la somma di m coeff. di autocorr al quadrato, moltiplicato per n, converge al $\chi^2(m)$).

Il test Portmanteau corrisponde a Q_m, calcolato sui residui della regressione OLS di y su X.

nota: non funziona se c'è la dipendente ritardata tra i regressori

Test di Breusch-Godfrey

$$y_t = \beta' x_t + \varepsilon_t$$

AR(m): $\varepsilon_t = \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \dots + \alpha_m \varepsilon_{t-m} + u_t$, $u_t \sim iid(0, \sigma^2)$

$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$ (no autocorrelazione)

Passi del test:

1. regressione di y_t su x_t e calcolo residui, $\hat{\varepsilon}_t$
2. regressione di $\hat{\varepsilon}_t$ su $x_t, \hat{\varepsilon}_{t-1}, \hat{\varepsilon}_{t-2}, \dots, \hat{\varepsilon}_{t-m}$
3. calcolo del test F per la significatività della regressione 2

oppure, alternativamente:

$$LM = nR^2$$

dove R^2 è l'indice di det. lineare della regressione ② . Sotto H_0,

$$LM \xrightarrow{d} \chi^2(m)$$

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

$$H_0: \varepsilon | X \sim N(0, \sigma^2 I_n)$$

Test di normalità (Jarque-Bera)

La teoria asintotica NON richiede la normalità.
I risultati per campioni finiti (t, F) SI.

Passi del test

1. regressione di y su X e calcolo residui $\hat{\varepsilon}_t$
2. calcolo coefficienti di asimmetria e curtosi di $\hat{\varepsilon}_t$

$$\hat{S} = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{\hat{\varepsilon}_t - \bar{\varepsilon}_n}{\hat{S}_n}\right)^3 \qquad \bar{\varepsilon}_n = \frac{1}{n}\sum \hat{\varepsilon}_t, \quad \hat{S}_n = \sqrt{\frac{1}{n}\sum(\hat{\varepsilon}_t - \bar{\varepsilon}_n)^2}$$

$$\hat{K} = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{\hat{\varepsilon}_t - \bar{\varepsilon}_n}{\hat{S}_n}\right)^4$$

(ricordiamo che, se $X \sim N(\mu, \sigma^2)$, asimm $= 0$, curtosi $= 3$)

③ calcolo della statistica test

$$BJ := \frac{n-k+1}{6}\left(\hat{S}^2 + \frac{1}{4}(\hat{K}-3)^2\right)$$

④ Sotto H_0: $y|X \sim N(0, \sigma^2 I_n)$,

$$BJ \xrightarrow{a} \chi^2(2)$$

Test di cambiamento strutturale

$$y_t = \beta' x_t + \varepsilon_t \qquad t = 1, 2, \dots, n$$

Implicitamente, si assume che β sia invariante nel campione.

L'estremo opposto è il caso in cui:

$$y_t = \beta_t' X_t + \varepsilon_t$$

dove $\beta = \beta_t$ dipende da t.

Il modello senza cambiamenti strutturali corrisponde ad

$$H_0: \beta_1 = \beta_2 = \dots = \beta_N = \beta$$

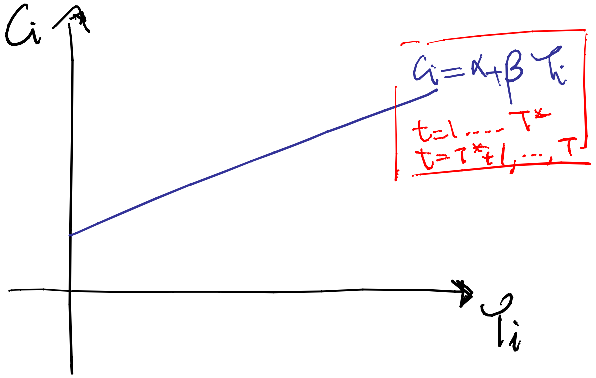
(β non varia in $t = 1, 2, \dots, n$)

Test che richiedono una ipotesi sulla data (sulle date) del possibile break (dei possibili break)

- Test di Chow

Test che NON richiedono tale ipotesi

- test basati sui residui ricorsivi (CUSUM, CUSUM Q)

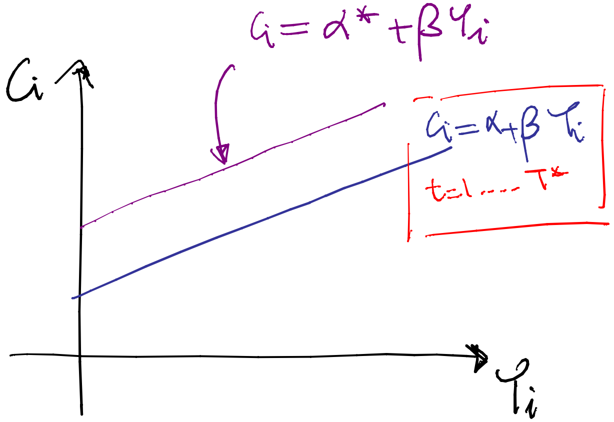


$t = 1, 2, \ldots\ldots, T$

$t = 1, 2, \ldots, T^*$

$t = T^*+1, \ldots, T$

Nessun cambiamento struttura

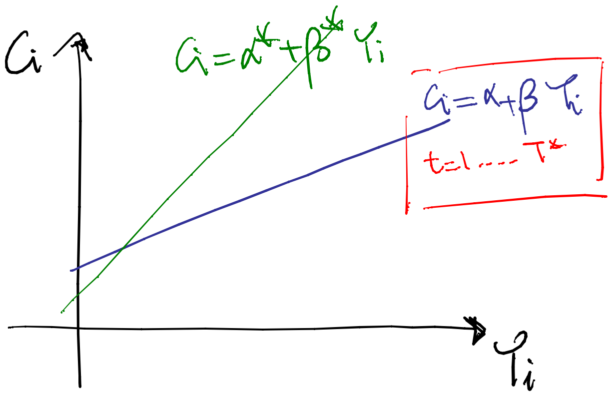


$t = 1, 2, \ldots\ldots, T$

$t = 1, 2, \ldots, T^*$

$t = T^*+1, \ldots, T$

cambiamento strutturale nell'intercetta



$t = 1, 2, \ldots\ldots, T$

$t = 1, 2, \ldots, T^*$

$t = T^*+1, \ldots, T$

cambiamento strutturale nell'intera regressione

Test di Chow.

$$y_t = \beta' x_t + \varepsilon_t \quad , \quad t = 1, 2, \dots, n$$

Ipotesi: cambiamento strutturale in β al tempo n^*+1

$$y_t = \beta_A' x_t + \varepsilon_t \quad , \quad t = 1, 2, \dots, n^*$$

$$y_t = \beta_B' x_t + \varepsilon_t \quad , \quad t = n^*+1, \dots, n$$

n^* è noto a priori.

L'ipotesi nulla di assenza di cambiamento strutturale è

$$H_0: \beta_A = \beta_B$$

o, analogamente, $H_0: \beta_A - \beta_B = 0$.

Per costruire un test, usiamo le variabili dummy.

$$d_{At} = \begin{cases} 1 & \text{se } t \leq n^* \\ 0 & \text{altrimenti} \end{cases}$$

(t appartiene al primo sottocampione

$$\boxed{\begin{cases} y_t = \beta_A' x_t + \varepsilon_t, & t = 1, \dots, n^* \\ y_t = \beta_B' x_t + \varepsilon_t, & t = n^*+1, \dots, n \end{cases}}$$

$$d_{Bt} = \begin{cases} 1 & \text{se } t \geq n^*+1 \\ 0 & \text{altrimenti.} \end{cases}$$

Notiamo che $d_{At} + d_{Bt} = 1$ $\quad (d_{Bt} = 1 - d_{At})$.
Possiamo scrivere:

$$y_t = \beta_A' x_t \cdot d_{At} + \beta_B' x_t \cdot d_{Bt} + \varepsilon_t$$

Infatti, se ad esempio $t \leq n^*$, allora $d_{At} = 1$, $d_{Bt} = 0$ e l'equazione diventa

$$y_t = \beta_A' x_t + \varepsilon_t$$

Possiamo definire: $\quad x_{At} := x_t \cdot d_{At}$, $\quad x_{Bt} := x_t \cdot d_{Bt}$ $\qquad \tilde{\beta} := \begin{pmatrix} \beta_A \\ \beta_B \end{pmatrix}$

e scrivere:

$$y_t = (\beta_A' \ \ \beta_B') \begin{pmatrix} x_{At} \\ x_{Bt} \end{pmatrix} + \varepsilon_t$$

$$(*) \quad y_t = \tilde{\beta}' \tilde{x}_t + \varepsilon_t, \qquad \tilde{x}_t = \begin{pmatrix} x_{At} \\ x_{Bt} \end{pmatrix}$$

Per testare H_0 è sufficiente stimare con OLS il modello (*) e sottoporre a verifica l'ipotesi che $\beta_A = \beta_B$. Questa ipotesi si può scrivere come:

$$H_0 : R'\tilde{\beta} - r = 0$$

dove

$$\underbrace{\begin{bmatrix} I_K & -I_K \end{bmatrix}}_{R'} \underbrace{\begin{bmatrix} \beta_A \\ \beta_B \end{bmatrix}}_{\tilde{\beta}} = \beta_A - \beta_B = 0$$

Usiamo un test F, χ^2.

Nota: dobbiamo avere un numero di osservazioni sufficientemente elevato da poter stimare la regressione su entrambi i sottocampioni.

Chow ha costruito un test applicabile quando $n - n^* < K$.

Nota:

$$y_t = \beta_A' \underbrace{x_t d_{At}} + \beta_B' \underbrace{x_t d_{Bt}} + \varepsilon_t$$
$$= \beta_A' x_{At} + \beta_B' x_{Bt} + \varepsilon_t$$

Test di chow: ① regressione di y_t su x_{At} e x_{Bt}

② Test F per H_0: $\beta_A - \beta_B = 0$
(solo K vincoli)

Nota: alternativamente:

$$y_t = \beta' x_t + \gamma' x_t d_{2t} + \varepsilon_t \,, \quad d_{2t} = \begin{cases} 1 & \text{se } t = n^*+1, \dots, n \\ 0 & \text{altrimenti} \end{cases}$$

per $t = 1, \dots, n^*$: $y_t = \beta' x_t + \varepsilon_t$

per $t = n^*+1, \dots, n$ $y_t = \beta' x_t + \gamma' x_t + \varepsilon_t$

$$y_t = (\beta + \gamma)' x_t + \varepsilon_t$$

γ misura la variazione nel coefficiente β passando dal primo campione al secondo

L'ipotesi nulla di assenza di break: $H_0: \gamma = 0$ $\quad (\gamma = \beta_B - \beta_A)$

Nota: ipotesi di cambiamento strutturale nell'intercetta

$$y_t = \begin{cases} \alpha_A + \beta' x_t + \varepsilon_t & \text{se} \quad t = 1, \dots, n^* \\ \alpha_B + \beta' x_t + \varepsilon_t & \text{se} \quad t = n^*+1, \dots, n \end{cases}$$

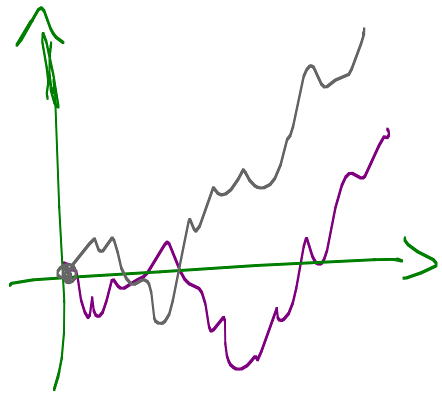
Con:

$$d_{2t} := \begin{cases} 1 & \text{se } t = n^*+1, \dots, n \\ 0 & \text{altrimenti} \end{cases}$$

$$y_t = \alpha + \gamma d_{2t} + \beta' x_t + \varepsilon_t$$

Se $t = 1, \dots, n^*$, allora $d_{2t} = 0$ e $y_t = \alpha + \beta' x_t + \varepsilon_t$
Se $t = n^*+1, \dots, n$, allora $d_{2t} = 1$ e $y_t = (\alpha + \gamma) + \beta' x_t + \varepsilon_t$

Test CUSUM/CUSUMQ



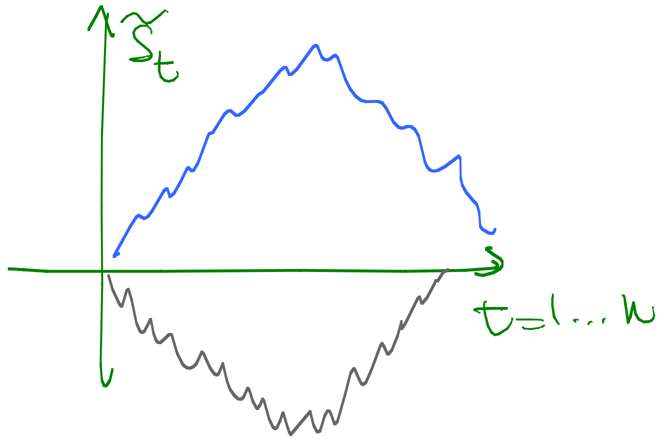
$$y_t = x_t'\beta + \varepsilon_t \quad , \quad t=1,\dots,n$$

$$S_t := \sum_{i=1}^{t} \varepsilon_i \qquad \text{(somme parziale di } \varepsilon_t\text{)}$$

Poiché ε_t non è osservabile, posso considerare i residui

$$\hat{\varepsilon}_t = y_t - x_t'\hat{\beta} \quad \text{e} \quad \tilde{S}_t := \sum_{i=1}^{t} \tilde{\varepsilon}_i$$

Se ci sono dei cambiamenti strutturali, dovrebbero influenzare la dinamica di $\tilde{S}_t$



I residui sono tuttavia correlati:

$$\hat{\varepsilon} = M_X y = M_X \varepsilon, \quad \text{da cui}$$

$$V[\hat{\varepsilon}|X] = E[\hat{\varepsilon}\hat{\varepsilon}'|X] = E[M_X \varepsilon \varepsilon' M_X | X]$$

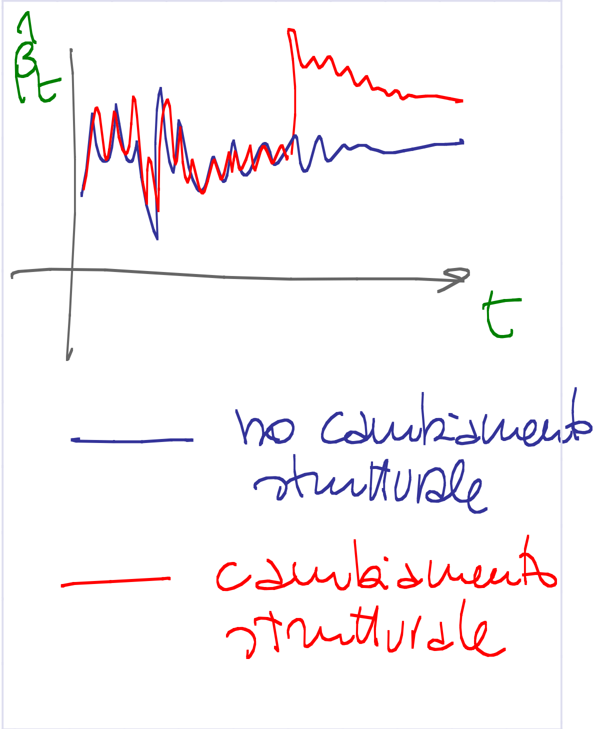
$$= \sigma^2 M_X \neq \sigma^2 I_n$$

Un tipo di residuo che risulta correlato è il residuo ricorsivo.

Stime ricorsive di β:

$$\hat{\beta}_t = \left(\sum_{i=1}^{t} x_i x_i'\right)^{-1} \sum_{i=1}^{t} x_i y_i$$

$$= (X_t' X_t)^{-1} X_t' Y_t \,, \qquad t = K, \dots, n$$



dove

$$X_t = \begin{bmatrix} x_1' \\ \vdots \\ x_t' \end{bmatrix} \qquad Y_t = \begin{bmatrix} y_1 \\ \vdots \\ y_t \end{bmatrix}$$

Residuo ricorsivo:

$$\tilde{\varepsilon}_t := y_t - \hat{y}_{t|t-1} \,, \qquad \hat{y}_{t|t-1} := \hat{\beta}_{t-1}' x_t$$

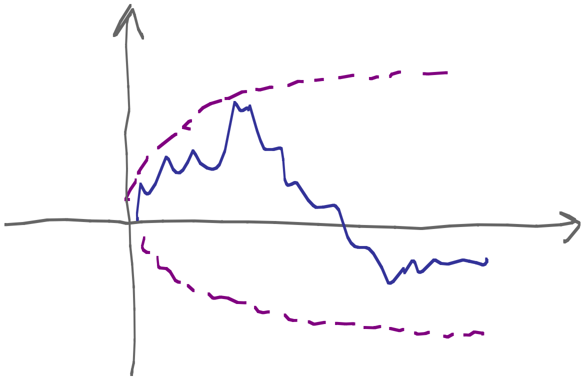
(previsione di y_t basata sulle osservazioni fino al tempo $t-1$)

Proprietà:
$$E[\tilde{\varepsilon}_t | X] = 0$$
$$V[\tilde{\varepsilon}_t | X] = \sigma^2 [1 + x_t'(X_t'X_t)^{-1}x_t]$$
$$C[\tilde{\varepsilon}_t, \tilde{\varepsilon}_{t'} | X] = 0 \text{ per } t \neq t'$$

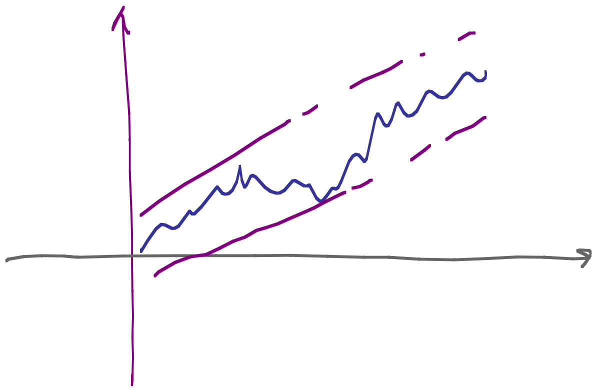
Da questa proprietà segue che
$$w_t := \frac{\tilde{\varepsilon}_t}{\sqrt{1 + x_t'(X_t'X_t)^{-1}x_t}}$$

è incorrelato e omoschedastico. Usando w_t definiamo
$$CUSUM_t := \frac{1}{\hat{\sigma}} \sum_{i=k+1}^{t} w_i \qquad t = k+1, \dots, n$$
$$CUSUMQ_t := \left(\sum_{i=k+1}^{t} w_i^2\right) \Big/ \left(\sum_{i=k+1}^{n} w_i^2\right) \qquad t = k+1, \dots, n$$

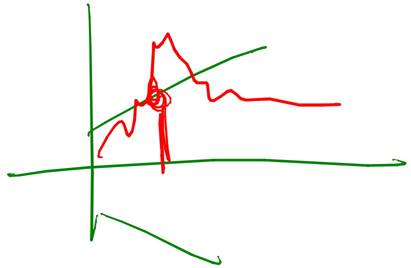
Queste due sequenze hanno una distribuzione asintotica nota sotto l'ipotesi di assenza di cambiamenti strutturali



— CUSUM $\sum \hat{\varepsilon}_i$



— CUSUMQ $\sum \hat{\varepsilon}_i^2$



Note generali sui test di ipotesi

$\underset{k\times 1}{\theta}$ vettore di parametri $\qquad \ell(\theta)$ log-verosimiglianza

$H_0: g(\theta) = \underset{q\times 1}{0}$ insieme di vincoli non lineari

nota: $g(\theta) = R'\theta - r = 0$ è il caso lineare

Esistono tre "classi" di test per H_0.

① Test di Wald.

- Stima di θ senza imporre H_0:

$$\hat{\theta} := \underset{\theta \in \Theta}{\arg\max}\ \ell(\theta)$$

Θ spazio dei parametri

- Calcolo di $g(\hat{\theta})$ - Se H_0 è vero, ci aspettiamo che

$g(\hat{\theta}) \approx 0.$

- Calcolo della statistica test:

$$W_n := g(\hat{\theta})' \, \hat{V}[g(\hat{\theta})]^{-1} g(\hat{\theta})$$

Sotto condizioni di regolarità (stazionarietà + ergodicità),

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, V_\theta)$$

$$\sqrt{n}\, g(\hat{\theta}) \xrightarrow[H_0]{d} N(0, V_g) \qquad \hat{V}[g(\hat{\theta})] \xrightarrow{P} V_g$$

da cui:

$$W_n \xrightarrow[H_0]{d} \chi^2(q)$$

② Test dei moltiplicatori di Lagrange

- Stima del modello imponendo il vincolo:

$$\tilde{\theta} := \underset{\{\theta \in \Theta : g(\theta) = 0\}}{\operatorname{argmax}} \ \ell(\theta)$$

Abbiamo visto che un problema di massimo vincolato si risolve risolvendo il problema:

$$\max \ m(\theta, \lambda) = \ell(\theta) - \lambda' g(\theta)$$

rispetto a θ e λ:

$$(\tilde{\theta}, \tilde{\lambda}) = \operatorname{argmax} \ m(\theta, \lambda)$$

Ovviamente $g(\tilde{\theta}) = 0$. Inoltre $g(\hat{\theta}) = 0$ se e solo se $\tilde{\lambda} = 0$

- Il test LM è basato su $\tilde{\lambda}$. In particolare, si dimostra che

$$LM = \tilde{\lambda}' \left[\hat{V}(\tilde{\lambda}) \right]^{-1} \tilde{\lambda} \xrightarrow[H_0]{d} \chi^2(q)$$

Esiste una versione equivalente basata sullo "score" valutato in $\tilde{\theta}$.

$$s(\tilde{\theta})' \left[\quad \right]^{-1} s(\tilde{\theta}) = LM$$

$$y_t = \beta_1 + \beta_2 x_t^{\alpha} + \varepsilon_t \qquad \theta = (\beta_1, \beta_2, \alpha)$$

$$H_0: \alpha = 1$$

In questo esempio LM è più facile da implementare di W e LR

③ Test del rapporto di verosimiglianza

$\hat{\theta}$ stimatore non vincolato $\Rightarrow$ $\ell(\hat{\theta})$

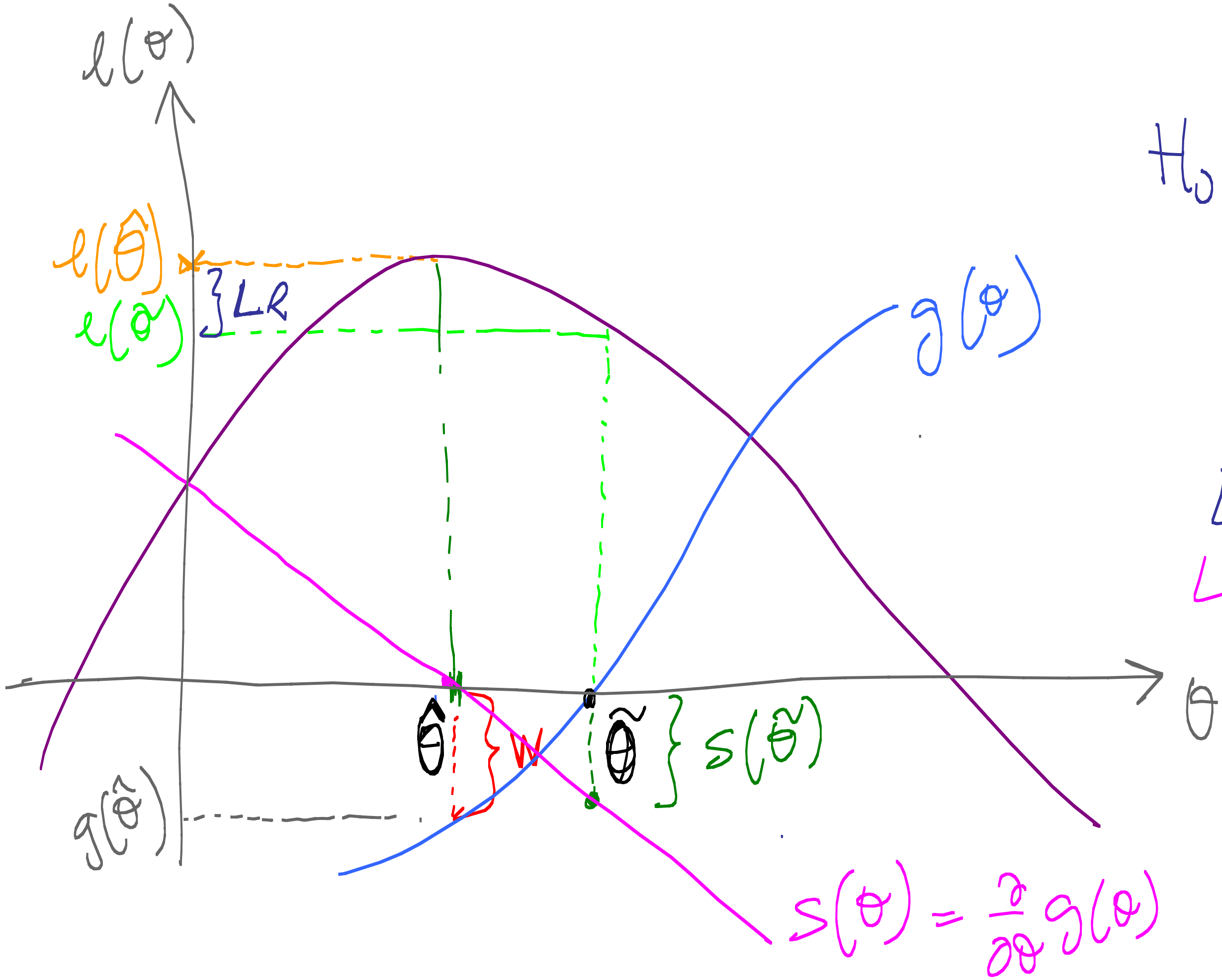
$\tilde{\theta}$ stimatore vincolato $\Rightarrow$ $\ell(\tilde{\theta})$

Se H_0 è vera, $\hat{\theta} - \tilde{\theta}$ non dovrebbe essere significativamente diverso da zero. Analogamente, $\ell(\hat{\theta})$ e $\ell(\tilde{\theta})$ non dovrebbero differire sistematicamente.

Il test LR (likelihood ratio) valuta la significatività della differenza $\ell(\hat{\theta}) - \ell(\tilde{\theta})$. In particolare:

$$LR := 2\left[\ell(\hat{\theta}) - \ell(\tilde{\theta})\right] \xrightarrow[H_0]{d} \chi^2(q)$$

$H_0: g(\theta) = 0$



W based su $g(\hat{\theta})$

LR based on $\ell(\hat{\theta}) - \ell(\tilde{\theta})$

LM based on $s(\tilde{\theta})$

Test di Wald, modello lineare

$$y = X\beta + \varepsilon, \qquad \varepsilon | X \sim N(0, \sigma^2 I_n)$$

$$\theta = (\beta', \sigma^2)' \qquad \ell(\theta) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)$$

$$\Downarrow$$

$$\hat\theta = \begin{pmatrix} \hat\beta \\ \hat\sigma^2 \end{pmatrix} = \begin{pmatrix} (X'X)^{-1}X'y \\ \hat\varepsilon'\hat\varepsilon / n \end{pmatrix}, \qquad \hat\varepsilon = y - X\hat\beta$$

Caso di $H_0: g(\theta) = g(\beta) = R'\beta - r$ (ipotesi su β)

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

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

sotto H_0:

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

- $$g(\hat\beta)' V^{-1} g(\hat\beta) = \frac{1}{\sigma^2}(R'\hat\beta - r)'\left(R'(X'X)^{-1}R\right)^{-1}(R'\hat\beta - r)$$

- Sostituendo σ^2 con la sua stima ML, $\hat{\sigma}^2 = \hat{\varepsilon}'\hat{\varepsilon}/n$,

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

Esercizio :
- riprendere gli appunti sulla stima vincolata ($H_0: R'\beta - r = 0$) e considerare il vettore di moltiplicatori di Lagrange : $\tilde{\lambda}$
- trovare $Var(\tilde{\lambda})$
- trovare $LM_n = \tilde{\lambda}'\left(\widehat{Var}(\tilde{\lambda})\right)^{-1}\tilde{\lambda}$ e mostrare che è χ^2