ECONOMETRIA APPLICATA PER L'IMPRESA

Eteroschedasticità nel modello di regressione

Davide Raggi

davide.raggi@unibo.it

24 febbraio 2010

L'ipotesi di Eteroschedasticità

- Per dimostrare il teorema di **Gauss-Markov** risulta necessario assumere, tra le altre ipotesi, che le varianze condizionali dei termini d'errore ϵ_i sono costanti, nel senso che non dipendono da qualche generico regressore, infatti

$$\text{Var}(\epsilon_i | X_{1i}, \ldots, X_{ki}) = \sigma^2 \qquad \forall i$$

- Sotto questa ipotesi si dimostra che gli stimatori dei minimi quadrati sono BLUE, cioé i *migliori* nel senso che hanno varianza minima tra quelli lineari e non distorti.
- L'ipotesi di omoschedasticità tuttavia non viene utilizzata per verificare se uno stimatore è non distorto e consistente. ***Ma allora perchè porsi il problema?***

Si ricordi che...

L'ipotesi di omoschedasticità è importante quando calcolano gli standard error della stima. Nel caso del modello di regressione semplice quindi si tratta di calcolare $\hat{\sigma}_{\hat{\beta}_1}$ e $\hat{\sigma}_{\hat{\beta}_0}$ che sono espressi da:

$$\hat{\sigma}_{\hat{\beta}_1} = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}$$

$$\hat{\sigma}_{\hat{\beta}_0} = \sqrt{\frac{\hat{\sigma}^2\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right)}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}$$

$\mathrm{Var}(\epsilon_i | X_{1i}, \ldots, X_{ki}) = \sigma_i^2$: Conseguenze...

- Lo stimatore delle varianze richiamato precedentemente risulta distorto senza l'assunzione di omoschedasticità. La stima associata allo stimatore, ipotizzando omoschedasticità, fornisce una cattiva approssimazione della vera varianza di stima.
- Si ricordi che **la statistica t dipende dallo stimatore di $\sigma_{\hat{\beta}_i}$**. Quindi se lo stimatore della varianza di stima è distorto, allora anche la verifica di ipotesi basata su una *varianza sbagliata* assumerà valori errati. Analogo ragionamento si può fare con i test di tipo F.
- Sembra evidente allora che sia la verifica di ipotesi che il calcolo degli intervalli di confidenza risenta pesantemente dell'ipotesi di eteroschedasticità.

Esempio

Si consideri il modello che regredisce il prezzo delle case rispetto ad alcuni regressori quali la metratura, il lotto ed il numero di stanze, usando OLS standard (Modello 1) e correggendo gli errori standard opportunamente (Modello 2).

Modello 1 *(non corretto per l'eteroschedasticità)*:

	coefficiente	errore std.	rapporto t	p-value	
const	-21.7703	29.4750	-0.7386	0.4622	
bdrms	13.8525	9.01015	1.537	0.1279	
lotsize	0.00206771	0.000642126	3.220	0.0018	<=
sqrft	0.122778	0.0132374	9.275	1.66E-014	

Modello 2 *(corretto per l'eteroschedasticità)*:

	coefficiente	errore std.	rapporto t	p-value	
const	-21.7703	37.1382	-0.5862	0.5593	
bdrms	13.8525	8.47862	1.634	0.1060	
lotsize	0.00206771	0.00125142	1.652	0.1022	<=
sqrft	0.122778	0.0177253	6.927	8.10E-010	

Esempio (cont.)

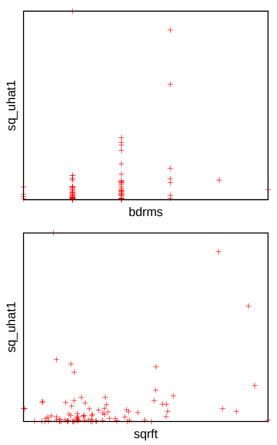

Figura 1: Residui al quadrato regrediti sui regressori. C'è o no dipendenza lineare tra $\hat{\epsilon}_i^2$ e, rispettivamente, *lotsize*, *bdrms* e *sqrft*?

Due Possibili Soluzioni

In letteratura sono stati proposti almeno due metodi per risolvere il problema della stima degli standard error degli stimatori di β_i. Il primo, che forse è meno utile in pratica ma più intuitivo si basa su una opportuna standardizzazione del termine di errore. Si parla allora di *Weighted Least Squares (WLS)*. Nel secondo caso si sfruttano le proprietà di consistenza degli stimatori OLS per ottenere delle *stime robuste* rispetto all'eteroschedasticità.

- **WLS:** $\text{Var}(\epsilon_i|X) = \sigma_i^2 = (\sigma\tilde{\sigma}_i)^2 \Rightarrow$ divido ϵ_i per $\tilde{\sigma}_i$.
- **Stima robusta:** Utilizzo $\hat{\epsilon}_i$ per stimare la *vera* varianza di $\hat{\beta}_i$.

Weighted Least Squares

Sia data la solita regressione multivariata

$$Y_i = \beta_0 + \beta_1 X_{i1} + \ldots + \beta_k X_{ik} + \epsilon_i$$

in cui si ipotizza che la varianza condizionale $\text{Var}(\epsilon_i|X) = \sigma_i^2$ sia esprimibile in una forma nota ma dipendente da alcuni parametri ignoti.

Ad esempio potrebbe essere sensato ipotizzare che

$$\sigma_i^2 = \sigma^2 \exp\{\delta_0 + \delta_1 Z_1 + \ldots + \delta_m Z_m\} = (\sigma\tilde{\sigma}_i)^2$$

In particolare, $Z_1, \ldots, Z_m$ sono opportuni regressori che potrebbero coincidere con $X_1, \ldots, X_k$.

Se i parametri $\delta_i,\ i = 0, \ldots, m$ fossero noti, allora σ_i sarebbe a sua volta noto. Si potrebbe quindi costruire la seguente regressione *ausiliaria*

$$\frac{Y_i}{\tilde{\sigma}_i} = \beta_0 \frac{1}{\tilde{\sigma}_i} + \beta_1 \frac{X_{i1}}{\tilde{\sigma}_i} + \ldots + \beta_k \frac{X_{ik}}{\tilde{\sigma}_i} + \frac{\epsilon_i}{\tilde{\sigma}_i}$$

da cui si ricava la formulazione equivalente del modello

$$Y_i^* = \beta_0 X_{i0}^* + \beta_1 X_{i1}^* + \ldots + \beta_k X_{ik}^* + \epsilon_i^*$$

In questa versione Var$(\epsilon_i^* | X) = \sigma^2$ e quindi l'uso dei Minimi Quadrati Ordinari per questa regressione porterebbe ad ottenere uno stimatore BLUE (poichè ci si è ricondotti alle ipotesi del teorema di **Gauss-Markov**).

Minimi Quadrati Generalizzati Calcolabili

- Lo stimatore OLS del modello su Y^* produce risultati differenti rispetto al metodo OLS *classico*. Gli stimatori cosí ottenuti sono un esempio dei Minimi Quadrati Generalizzati o GLS di cui WLS ne è un caso particolare.
- In generale, tuttavia, non siamo in grado di conoscere i parametri δ_j e quindi non possiamo *standardizzare* i regressori e la variabile dipendente.
- Per questo motivo occorre stimare la funzione che descrive la varianza condizionale. Si parla allora di *Minimi Quadrati Generalizzati Calcolabili (Feasible Generalized Least Squares)*. La varianza condizionale viene stimata tramite il metodo dei minimi quadrati.

Si noti che...

- Il modello di regressione ipotizza che la media condizionale delle Y sia
$$\mathrm{E}(Y|X) = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k$$
- E se volessi costruire un modello che descriva la varianza condizionale $\mathrm{Var}(Y|X)$?
- Potrei modellare la varianza condizionale come
$$\mathrm{Var}(Y|X) = \mathrm{E}[\epsilon^2|X] = g(Z_1, \ldots, Z_m) \qquad g(\cdot) > 0$$
- Questo equivale ad ipotizzare **ad esempio** un modello del tipo
$$\epsilon^2 = \sigma^2 \exp\{\delta_0 + \delta_1 Z_1 + \ldots + \delta_m Z_m\} \nu$$

Procedura FGLS

$$Y_i = \beta_0 + \beta_1 X_{i1} + \ldots, X_{ik} + \epsilon_i$$

$$\epsilon^2 = \sigma^2 \exp\{\delta_0 + \delta_1 Z_1 + \ldots + \delta_m Z_m\} \nu$$

- Si regredisca Y su $X_1, \ldots, X_k$ e si stimi $\hat{\epsilon}$;
- Si costruisca $\log(\hat{\epsilon}^2)$
- Si regredisca $\log(\hat{\epsilon}^2)$ su $Z_1, \ldots, Z_m$;
- Si calcolino i valori predetti dall'ultima regressione, $\hat{g} = \widehat{\log(\hat{\epsilon}^2)}$ e si calcoli $\hat{\sigma}_i = \exp\left\{\frac{\hat{g}_i}{2}\right\}$
- Si usi lo stimatore WLS dividendo variabili dipendente e regressori per $\hat{\sigma}_i$.

Stima Robusta

- WLS (FGLS) sembrerebbe un modo *efficace* e relativamente semplice per risolvere il problema dell'eteroschedasticità.
- **Problema:** Quando si utilizzano procedure di tipo GLS, **si stanno facendo ipotesi esplicite sulla forma funzionale di σ_i^2**. L'efficienza degli stimatori dipende dal fatto che il modello specificato per la varianza condizionale sia *appropriato*. **Questo in generale non lo sappiamo a priori**, tranne in casi molto particolari. In sostanza, l'efficienza di WLS-GLS dipende dalla *corretta specificazione* del modello per la varianza condizionale. Se questo non fosse vero, non è possibile dire se gli stimatori sono efficienti oppure no.
- In molti casi è allora più utile costruire delle procedure alternative, che non fanno ipotesi esplicite sulla struttura della varianza condizionale $\Rightarrow$ **Stimatori Robusti all'Eteroschedasticità**.

Stima Robusta (cont.)

- Si ricordi OLS produce stimatori consistenti ($\hat{\beta}_j \xrightarrow{p} \beta_j$) e non distorti (**indipendentemente dalle ipotesi fatte sulle varianze condizionali**).
- Si prenda per esempio il caso della regressione semplice. Si dimostra che

$$\begin{aligned} \hat{\epsilon} = Y - \hat{Y} &= \underbrace{\beta_0 + \beta_1 X + \epsilon}_{Y} - \underbrace{(\hat{\beta}_0 + \hat{\beta}_1 X)}_{\hat{Y}} \\ &= (\underbrace{\beta_0 - \hat{\beta}_0}_{=0 \text{ se } n \to +\infty}) + (\underbrace{\beta_1 - \hat{\beta}_1}_{=0 \text{ se } n \to +\infty})X + \epsilon \end{aligned}$$

- Ne consegue che $\hat{\epsilon} \xrightarrow{p} \epsilon$, quindi **è uno stimatore consistente del termine d'errore**. L'idea è allora sfruttare questa proprietà per stimare la varianza di $\epsilon|X$.

Procedura di White (-Huber-Heicker)

- Consideriamo per semplicità il modello di regressione semplice

$$Y = \beta_0 + \beta_1 X + \epsilon.$$

Come abbiamo già notato nelle lezioni precedenti, un po' di algebra ci consente di esprimere lo stimatore del parametro in funzione del vero valore del parametro β_1 e dell'errore ϵ

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n}(X_i - \bar{X})\epsilon_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

- Nel caso generale con k regressori si ottiene che

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \underbrace{\boldsymbol{Y}}_{=\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{\epsilon}} \\ &= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\epsilon} \end{aligned}$$

- Siamo interessati a trovare uno stimatore per $\mathrm{Var}(\hat{\beta}_1)$. La ***vera*** varianza di $\hat{\beta}_1$ è data da

$$\mathrm{Var}(\hat{\beta}_1) = \frac{\sum_{i=1}^{n} \overbrace{\sigma_i^2}^{\mathrm{Var}(\epsilon_i|X)} (X_i - \bar{X})^2}{\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right]^2}$$

- L'idea di White (1980) è *relativamente semplice*. **Siccome $\hat{\epsilon}_i$ è *consistente* per ϵ_i**, se lo sono gli stimatori OLS $\hat{\beta}_i$, allora $\hat{\epsilon}_i^2$ **potrebbe essere uno stimatore sensato di $E\left(\epsilon_i^2\right) = \sigma_i^2$**, nel senso che è anche lui consistente (Teo. di Slutsky). Di conseguenza propone

$$\widehat{\mathrm{Var}}(\hat{\beta}_1) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2 \ \hat{\epsilon}_i^2}{\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right]^2}$$

- Si può dimostrare che $\widehat{\mathrm{Var}}(\hat{\beta}_1)$ **è consistente** per $\mathrm{Var}(\hat{\beta}_1)$.

Estensione al caso multiplo

La stessa idea si può applicare al caso della regressione multipla

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \epsilon$$

nel qual caso si dimostra che

$$\widehat{\mathrm{Var}}(\hat{\beta}_j) = \frac{\sum_{i=1}^n \hat{r}_{ij}\ \hat{\epsilon}_i^2}{\left[\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2\right]^2} \qquad j = 1, \ldots, k$$

$\hat{r}_{ij}$ è la stima dell'i-esimo residuo ottenuto dalla regressione di X_j su tutti gli altri regressori, cioè

$$X_j = \alpha_0 + \alpha_1 X_1 + \alpha_{j-1} X_{j-1} + \alpha_{j+1} X_{j+1} + \ldots + \alpha_k X_k + u$$

Estensione al caso multiplo (cont.)

Si consideri la matrice di varianze e covarianze (diagonale) $\boldsymbol{\Sigma}$ definita da

$$\mathrm{E}\left[\boldsymbol{\epsilon}\boldsymbol{\epsilon}'\right] = \boldsymbol{\Sigma}_{n\times n} = \begin{pmatrix} \sigma_1^2 & 0 & \ldots & 0 \\ 0 & \sigma_2^2 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & \ldots & \sigma_n^2 \end{pmatrix}$$

Si dimostra facilmente che

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\mathrm{E}\left[\boldsymbol{\epsilon}\boldsymbol{\epsilon}'\right]\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

In generale White dimostra che uno stimatore consistente di $\frac{\boldsymbol{X}'\boldsymbol{\Sigma}\boldsymbol{X}}{n}$ è dato da $\frac{\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}\boldsymbol{X}}{n}$ in cui gli elementi sulla diagonale di $\hat{\boldsymbol{\Sigma}}$ sono stimati con $\hat{\sigma}_i^2 = \hat{\epsilon}_i^2$

Riassumendo

- Lo stimatore $\hat{\sigma}_{\hat{\beta}_j}$ calcolato con **OLS** ipotizzando omoschedasticità quando questa non è verificata, **può non essere appropriato**.
- Tuttavia, gli stimatori dei parametri stimati con OLS sono comunque consistenti e non distorti. Di conseguenza $\hat{\epsilon}$ risulta consistente per ϵ. L'idea di White-(Huber-Heicker) è di utilizzare questa proprietà per correggere gli standard error.
- Il grande vantaggio di questa procedura è:
  - Nessuna ipotesi sul tipo di eteroschedasticità (flessibilità);
  - La procedura OLS rimane valida (semplicità).
- Per contro lo stimatore dei β_j non è più BLUE.

Test Robusti all'eteroschedasticità

- Il **Test t**
  - La statistica test t, costruita in maniera robusta rispetto all'eteroschedasticità, è sostanzialmente identica al caso omoschedastico. In questo caso, però, il denominatore deve essere quello calcolato nella sua versione robusta. La statistica t robusta rispetto all'ipotesi di eteroschedasticità risulta quindi

$$t = \frac{\text{stimatore - valore ipotizzato}}{\text{standard error robusto}}$$

- Il **Test F**.
  - Calcolare la statistica F calcolata in maniera robusta risulta più complicato. Tuttavia molti pacchetti applicativi, tra cui GRETL, consentono di calcolarla in maniera automatica.

Come verificare l'ipotesi di eteroschedasticità?

Abbiamo visto che l'utilizzo di metodi robusti consente di calcolare opportunamente i test t ed F indipendentemente dal fatto che ci sia o no eteroschedasticità. La ragione principale deriva dal fatto che OLS sotto ipotesi di eteroschedasticità non risulta efficiente rispetto a GLS.

Qualche volta allora può essere interessante il voler costruire delle procedure che ci consentano di decidere se è più o meno plausibile parlare di omo oppure eteroschedasticità...

IMPORTANTE: Tutti i test che seguono **non permettono di accettare l'ipotesi nulla** che è l'omoschedasticità, ma ci **consentono al più di rifiutare l'alternativa** (eteroschedasticità) .

Test di Breush-Pagan

- Consideriamo il solito modello di regressione multipla

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \alpha_k X_k + \epsilon$$

- Vogliamo costruire un test per cui l'ipotesi nulla è

$$H_0: \ \mathrm{Var}(\epsilon | X_1, \ldots, X_k) = E[\epsilon^2 | X_1, \ldots, X_k] = \sigma^2$$

- Per le varianze condizionali si potrebbe ipotizzare una relazione del tipo

$$\mathrm{E}[\epsilon^2 | \boldsymbol{X}] = \delta_0 + \delta_1 X_1 + \ldots + \delta_k X_k$$

  e di conseguenza si potrebbe pensare al modello

$$\epsilon^2 = \delta_0 + \delta_1 X_1 + \ldots + \delta_k X_k + u$$

- Dato questo modello **l'ipotesi nulla di omoschedasticità** potrebbe essere verificata attraverso il sistema di ipotesi per cui $H_0 : \text{Var}(\epsilon|X) = \sigma^2 \Leftrightarrow H_0 : \delta_1 = \delta_2 = \ldots = \delta_k = 0$ mentre
H_1 : Almeno una delle uguaglianze sotto H_0 non è verificata
- Questa ipotesi può essere facilmente verificata usando una statistica test di tipo F. Ovviamente ϵ^2 non è osservabile direttamente. Sappiamo però che una sua stima consistente è data da $\hat{\epsilon}^2$
- Il modello ridotto (*quello cioè vero sotto* H_0) diventa

$$\epsilon^2 = \delta_0 + err$$

nel qual caso è facile dimostrare che l'R_v^2 associato è nullo.

- Di conseguenza il test di eteroschedasticità basato sulla statistica F risulta

$$F_{k,n-k-1} = \frac{R^2_{\hat{\epsilon}^2}/k}{(1 - R^2_{\hat{\epsilon}^2})/(n-k-1)}$$

- Questo test può essere derivato anche in modo leggermente diverso, e prende il nome di test di Breusch-Pagan per l'eteroschedasticità che si calcola come

$$T_{BP} = nR^2_{\hat{\epsilon}^2}$$

- L'unica differenza rispetto al test F è che la statistica T_{BP} si distribuisce asintoticamente come un χ^2_k.

Test di eteroschedasticità di White

Un altro test piuttosto popolare per verificare l'ipotesi nulla di omoschedasticità si deve a White ma in sostanza l'idea di base è la stessa del test di Breush-Pagan. Per cercare di capire se i residui al quadrato dipendono o no in qualche modo dai regressori si consideri la seguente regressione *in cui per semplicità consideriamo* $k = 3$

$$\begin{aligned} \hat{\epsilon}^2 &= \delta_0 + \delta_1 X_1 + \delta_2 X_2 + \delta_3 X_3 + \delta_4 X_1^2 + \delta_5 X_2^2 \\ &\quad + \delta_6 X_3^2 + \delta_7 X_1 X_2 + \delta_8 X_1 X_3 + \delta_9 X_2 X_3 + errore \end{aligned}$$

Una volta fatta inferenza per questo modello ausiliario, è possibile costruire un test F in cui $H_0: \; \delta_1 = \ldots = \delta_9 = 0$.

Così come nel caso del test di Breush-Pagan esiste una versione alternativa del test basata sull'R^2 della regressione in cui la variabile dipendente è la stima del residuo al quadrato. Il test di White (che è un test di tipo Moltiplicatori di Lagrange) assume la seguente forma

$$T_W = n R^2_{\hat{\epsilon}^2} \sim \chi^2_{\sharp \text{ restrizioni}}$$

in cui, come prima che $R^2_{\hat{\epsilon}^2}$ è il coefficiente di bontà di adattamento del modello per $\hat{\epsilon}^2$.

Nota: Da un punto di vista pratico il test F su tutti i coefficienti ed il test T_W sono equivalenti

```
Test di White per l'eteroschedasticita'. Var. dipendente: uhat^2
  VARIABILE   COEFFICIENTE       ERRORE STD    STAT T   P-VALUE
  const         -3.84112E+06      6.11300E+06  -0.628   0.53049
  roe       669134           499522             1.340   0.18192
  ros       -17459.7          46409.7          -0.376   0.70716
  sales         22.9821         429.575         0.053   0.95739
  sq_roe     -8969.42          8057.41         -1.113   0.26697
  roe_ros    -1188.52          2001.93         -0.594   0.55340
  roe_sales    -13.8571          25.8617       -0.536   0.59268
  sq_ros        74.1444         133.731         0.554   0.57991
  ros_sales      1.52295          2.74052       0.556   0.57903
  sq_sales       0.00171829       0.00367704    0.467   0.64079

  R-quadro = 0.0130225         Numerosita' campionaria = 209

Statistica test: TR^2 = 2.721706, p-value = P(Chi-quadro(9) >
2.721706) = 0.974315
```