Econometria

Obiettivo: Fornire le competenze di base per l'analisi dei fenomeni economici attraverso la costruzione di modelli econometrici

Applicazioni:

- interpretazione del fenomeno
- testare le teorie economiche
- analisi di policy (valutazione delle politiche)
- previsione

Ambiti fenomenici

- macro/micro economico
- finanziario
- aziendale

Dati

- individuali
- aggregati

- cross section
- time series
- panel

Cos'è l'econometria?

1930: Fondazione dell'Econometric Society

"The Econometric Society is an international society for the advancement of economic theory in its relation to statistics and mathematics"

1933: Pubblicazione di Econometrica

Nuova disciplina con l'obiettivo di rendere più rigorosa la scienza economica unificando studi teorici ed empirici

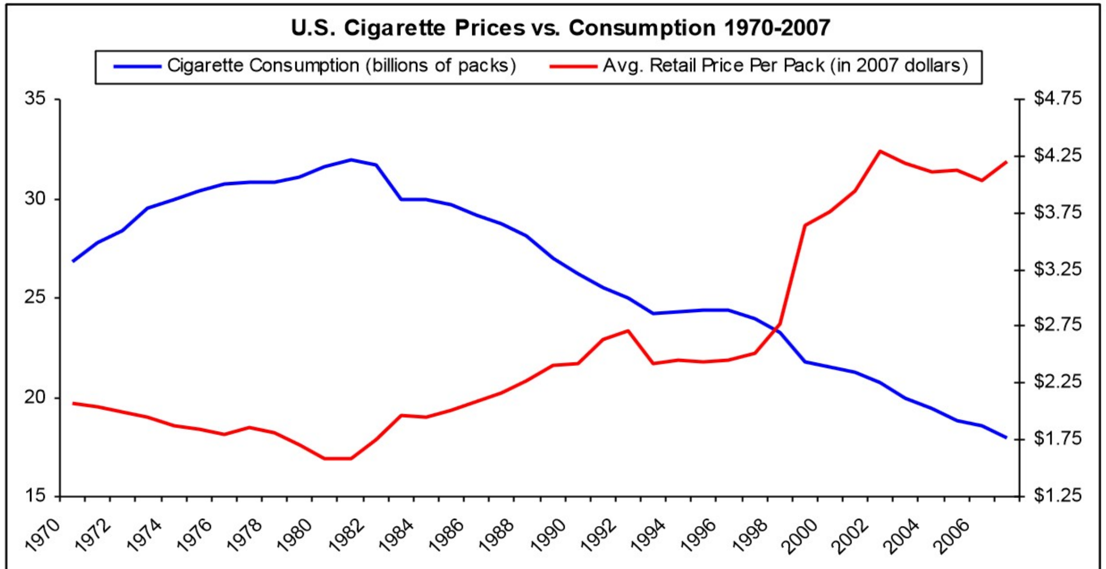
Frish (1933):

" But there are several aspects of the quantitative approach to economics, and no single one of these aspects, taken by itself, should be confounded with econometrics. "

"Econometrics is by no means economic statistics. Nor it is identical to ... economic theory. Nor it is a synonomous of the application of mathematics to economics"

"La teoria economica si deve ispirare in misura crescente all'analisi empirica analizzata con appropriate tecniche statistiche"

Prezzi e consumo di sigarette (US)

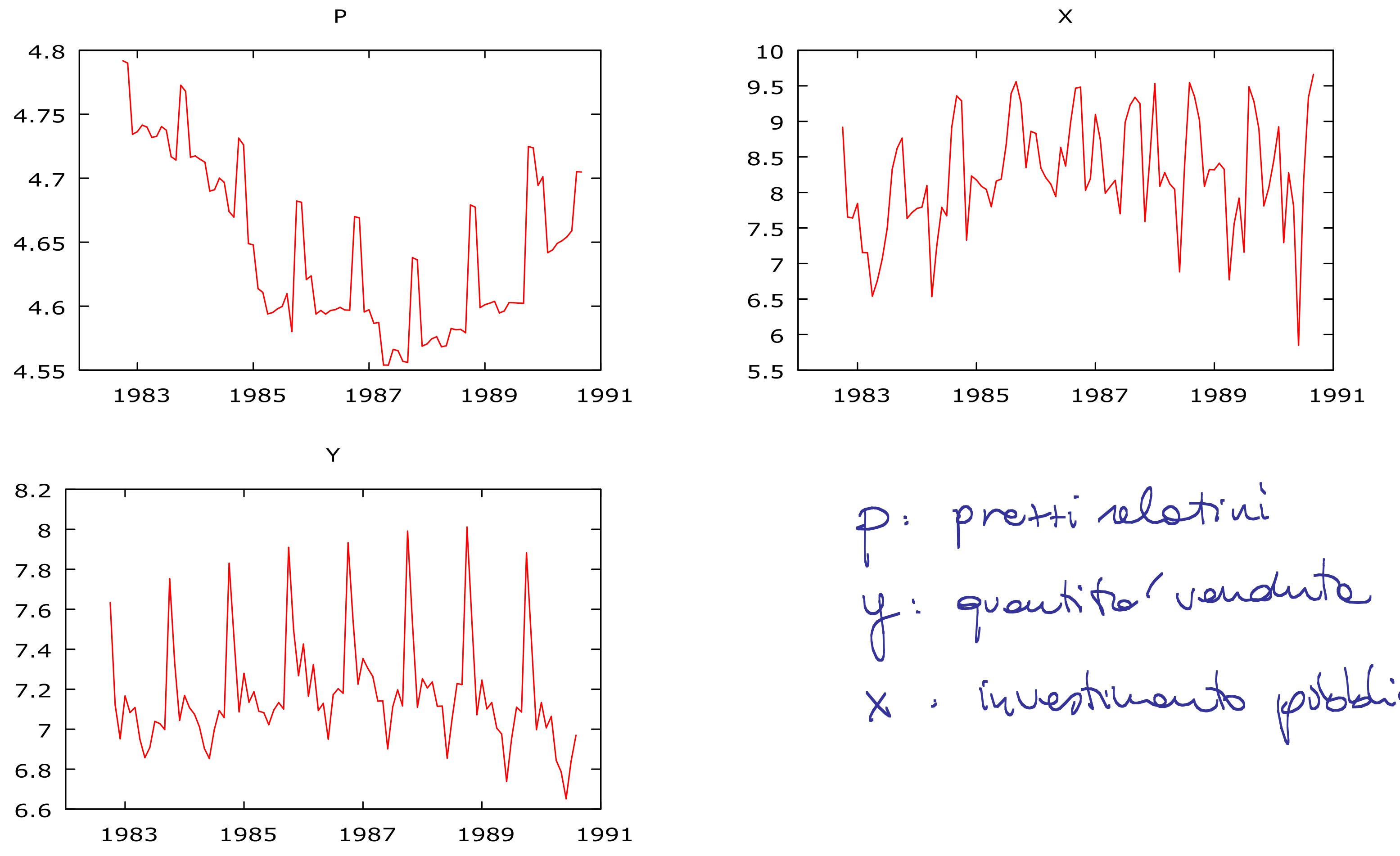


Sources: *The Tax Burden on Tobacco*, 2007; USDA Economic Research Service; U.S. Bureau of Labor Statistics.

Table 5a: UK: Prices, cigarette consumption and smoking intensity (if smoker)

	OLS			Heckman Selection Model		
	Log Cig	Log Cot	Log Inten	Log Cig	Log Cot	Log Inten
Panel A: Average Price Effects						
Average price effect	-0.81**	0.22	0.43	-0.85**	0.05	0.26
	(0.40)	(0.71)	(0.56)	(0.40)	(0.71)	(0.56)

Consumo di Whisky in Italia



p: prezzi relativi

y: quantità venduta

x: investimento pubblicitario

Figure 4. Oil Prices and CPI Inflation

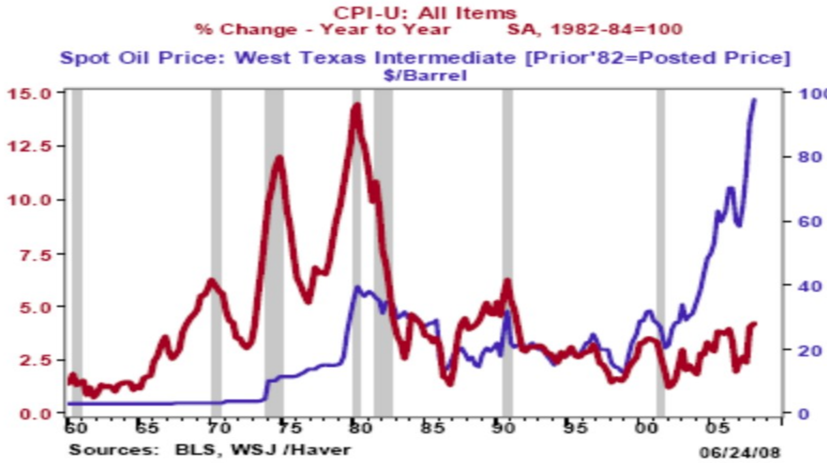


Figure 5. Oil Prices and Real GDP Growth

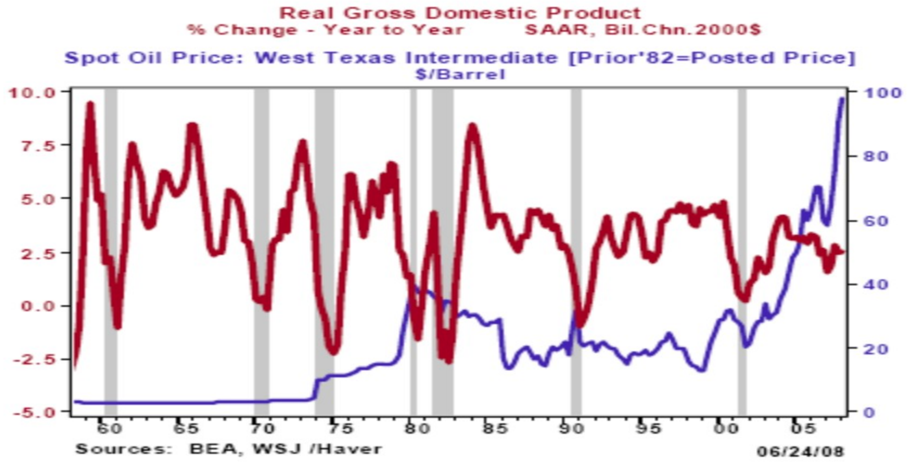
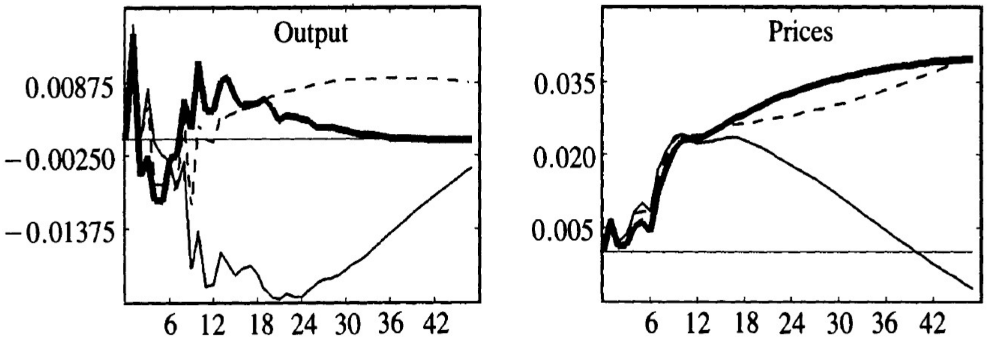


Figure 1: Responses to a Hamilton oil price shock



Notes: Thin line—standard VAR; dotted line—fixed FF rate, no expectations; thick line—fixed FF rate with consistent expectations.

Source: Bernanke *et al*. (1997, Figure 4, p. 117), reproduced with permission from the Brookings Institution.

ELSEVIER Journal of Monetary Economics 51 (2004) 1007–1032
www.elsevier.com/locate/econbase

The macroeconomic consequences of terrorism ☆

S. Brock Blomberg[a,*], Gregory D. Hess[a,b], Athanasios Orphanides[c]

[a] Department of Economics, Claremont McKenna College, Claremont, CA 91711, USA
[b] Department of Economics, Claremont McKenna College, Claremont, CA 91711, USA and CESifo
[c] Board of Governors of the Federal Reserve System, Washington, DC 20551, USA

Received 21 November 2003; received in revised form 9 March 2004; accepted 28 April 2004

Correlations with terrorism across sub-samples

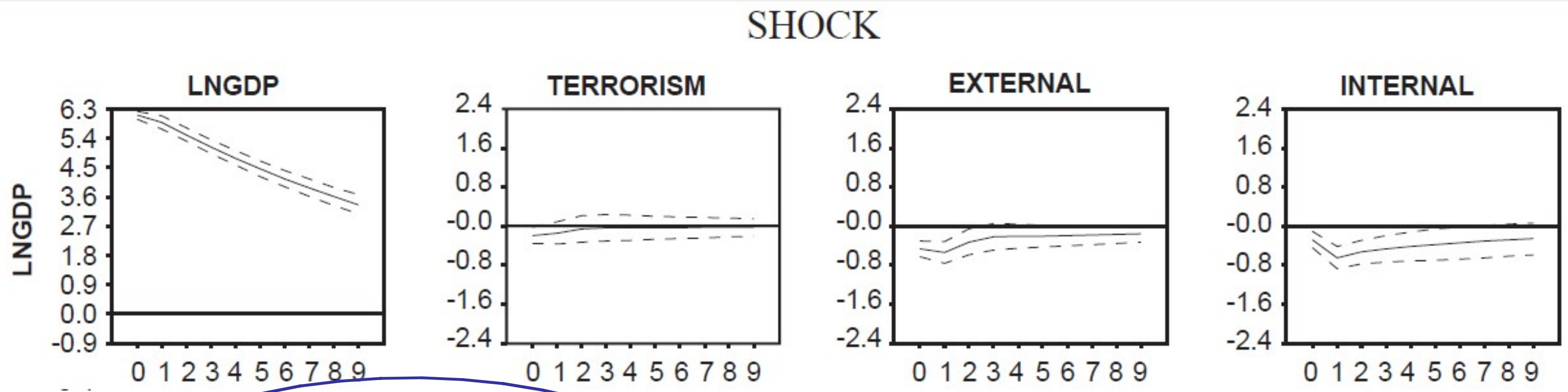
	ALL	NONDEMO	OECD	AFRICA	MIDEAST	ASIA
Correlation with terrorism (T)						
H	0.037	0.045	0.000	0.031	0.084	0.050
A	0.056	0.089	0.062	0.127	0.055	0.084
I	0.151	0.077	0.078	0.225	0.117	0.194
y	0.197	0.060	−0.076	0.113	0.325	0.021
Δy	0.000	−0.043	−0.043	−0.021	0.055	−0.124
OPEN	−0.164	−0.269	−0.207	−0.154	0.220	−0.177
GINI	−0.018	0.056	0.187	−0.079	−0.178	0.090

$$\Delta y_i = \beta_0 + \beta_1\,\mathrm{COM}_i + \beta_2\,\mathrm{AFRICA}_i + \beta_3 \ln y_{0i} + \beta_4 I/Y_i + \beta_5 T_i + \beta_6 I_i + \beta_7 E_i + \epsilon_i,$$

Table 3
Cross-sectional regression: terrorism and growth

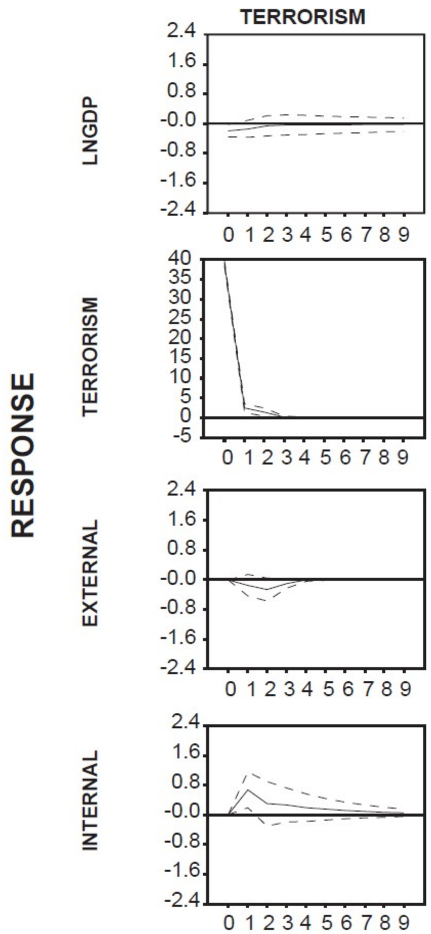
Model	1	2	3	4	5	6	7	8	9	10
specification	Base	& T	& I	& E	& T, I, E	Base	& T	& I	& E	& T, I, E
estimation	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV	IV
COM	−1.116***	−1.200***	−1.186***	−1.123***	−1.220***	−1.155***	−1.248***	−1.246***	−1.165***	−1.278***
	[0.326]	[0.334]	[0.325]	[0.325]	[0.335]	[0.339]	[0.340]	[0.334]	[0.337]	[0.339]
AFRICA	−0.963***	−1.358***	−1.109***	−0.987***	−1.357***	−1.234***	−1.650***	−1.404***	−1.268***	−1.636***
	[0.366]	[0.400]	[0.366]	[0.373]	[0.397]	[0.378]	[0.408]	[0.374]	[0.383]	[0.405]
$\ln y_0$	−0.715***	−0.641***	−0.782***	−0.712***	−0.692***	−0.498***	−0.440***	−0.606***	−0.492***	−0.526***
	[0.170]	[0.166]	[0.174]	[0.171]	[0.179]	[0.167]	[0.166]	[0.171]	[0.167]	[0.174]
I/Y	0.147***	0.142***	0.135***	0.146***	0.137***	0.094***	0.094***	0.083***	0.092***	0.087***
	[0.022]	[0.022]	[0.022]	[0.022]	[0.022]	[0.026]	[0.025]	[0.026]	[0.025]	[0.026]
T		−1.587***			−1.291*		−1.774***			−1.257*
		[0.569]			[0.683]		[0.579]			[0.708]
I			−0.663**		−0.356			−0.877***		−0.565*
			[0.279]		[0.331]			[0.266]		[0.329]
E				−2.837	0.696				−3.838	0.243
				[3.326]	[3.009]				[3.230]	[2.888]
Obs	115	115	115	115	115	113	113	113	113	113
R^2	0.53	0.56	0.55	0.53	0.57	0.50	0.54	0.52	0.50	0.54

Notes: Robust standard errors are presented in square brackets. *, ** and *** represent statistical significance at the 0.10, 0.05 and 0.01 levels, respectively. Models (1)–(10) are different specifications of cross country growth regressions. Models (1)–(5) are the basic OLS model adding separately the different forms of conflict, i.e. terrorism (T), internal conflict (I), home (H) and away (A) wars and their sum, external wars (E). Models (6)–(10) repeat the exercises but estimate the model as IV/GMM with initial investment as a percent of GDP (I/Y) as the instrument. Included in each regression is a dummy for non-oil exporting commodity countries (com), Africa (afr), initial GDP per capita ($\ln y_0$) and average investment as a percent of GDP (I/Y). The R^2 measure excludes the contribution from the individual fixed effects.

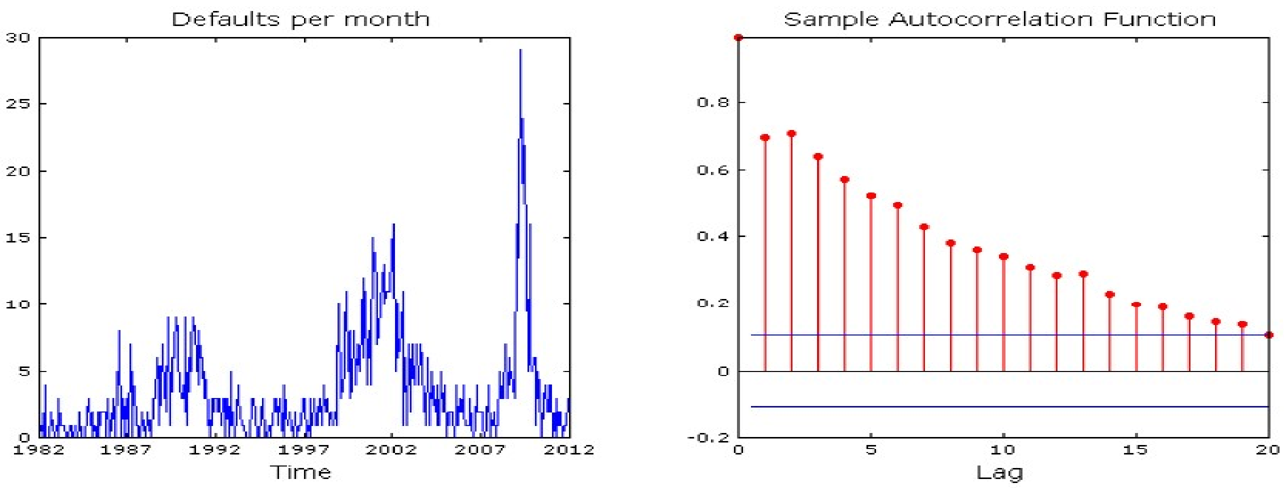


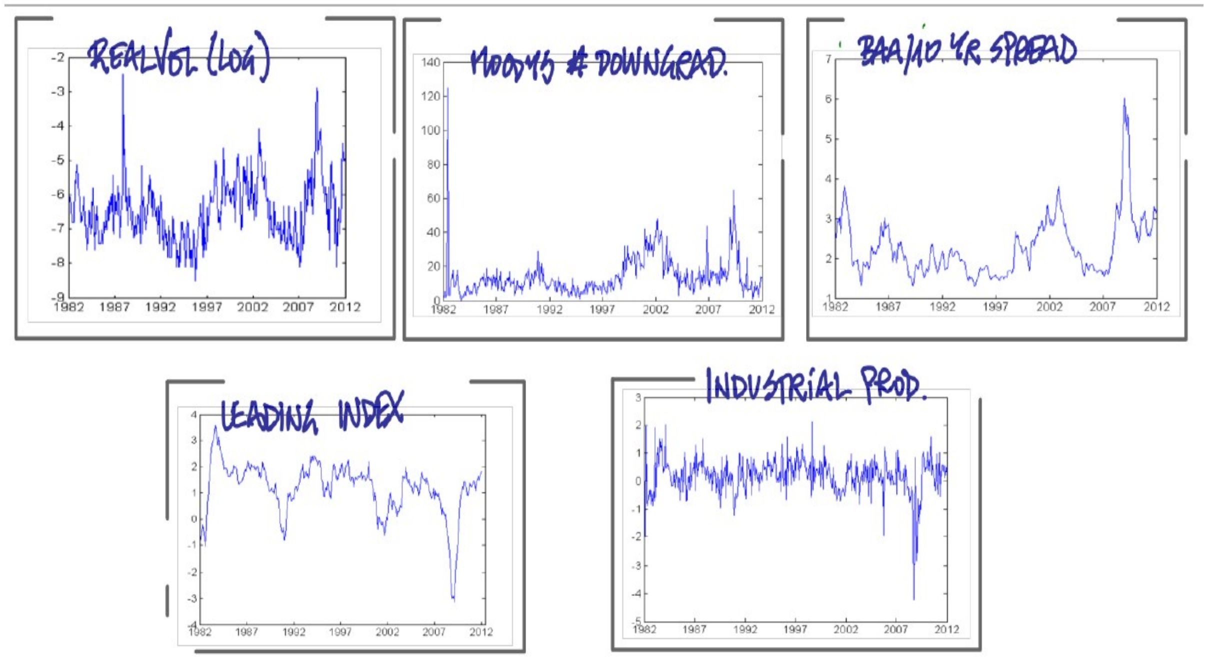
Impatto dei vari shock sulla crescita

impatto del terrorismo sulle variabili considerate



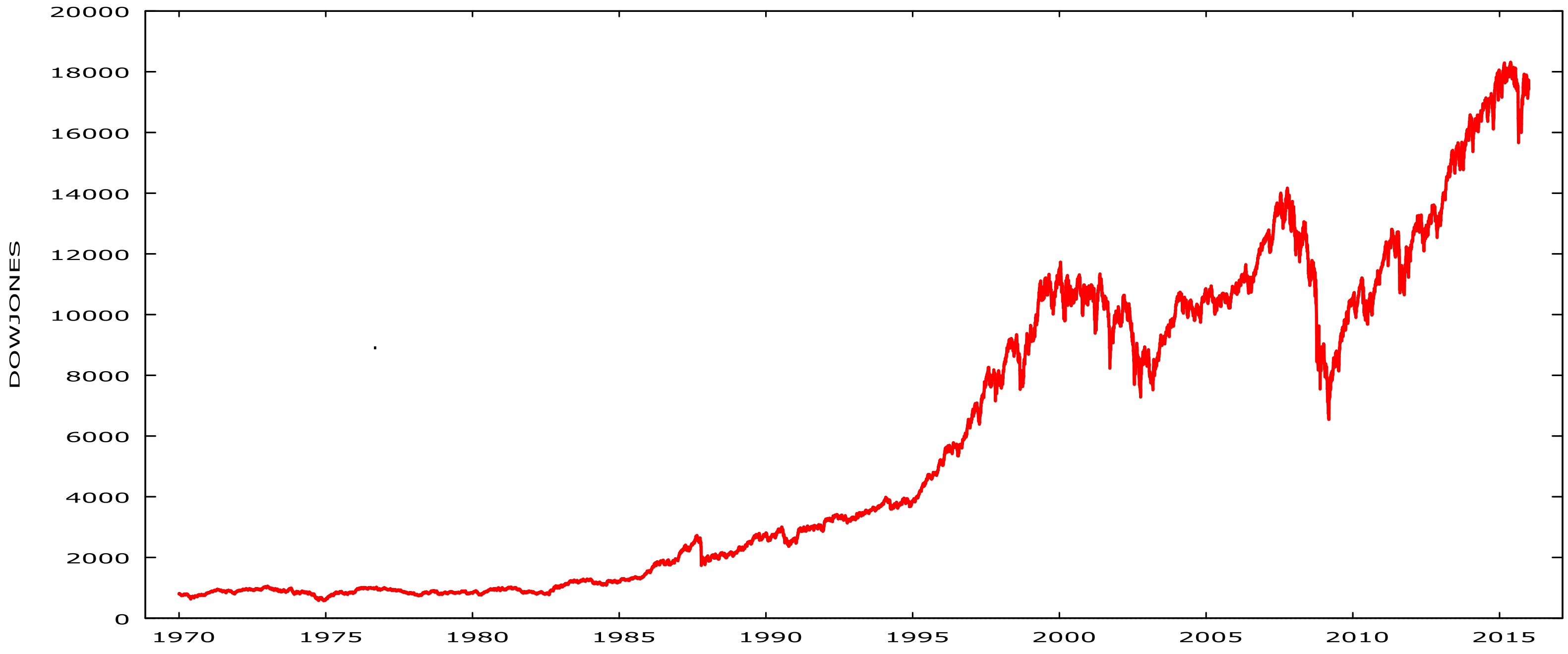
US corporate default counts and their autocorrelations



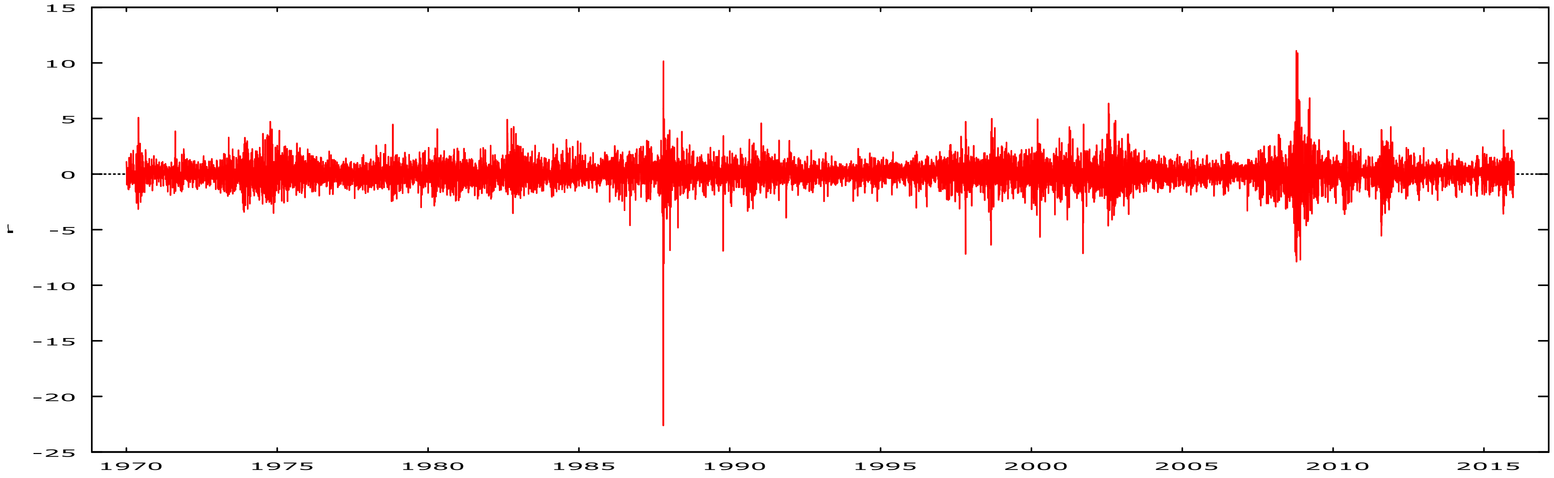


$$y_t | \mathcal{F}_{t-1} = \text{Poisson}\,(\lambda_t)$$

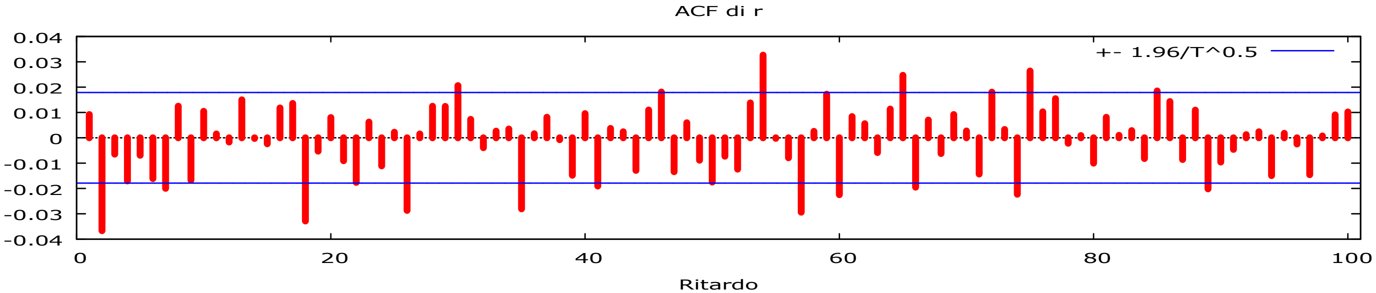
DOW JONES INDUSTRIAL AVERAGE, 1970-2015

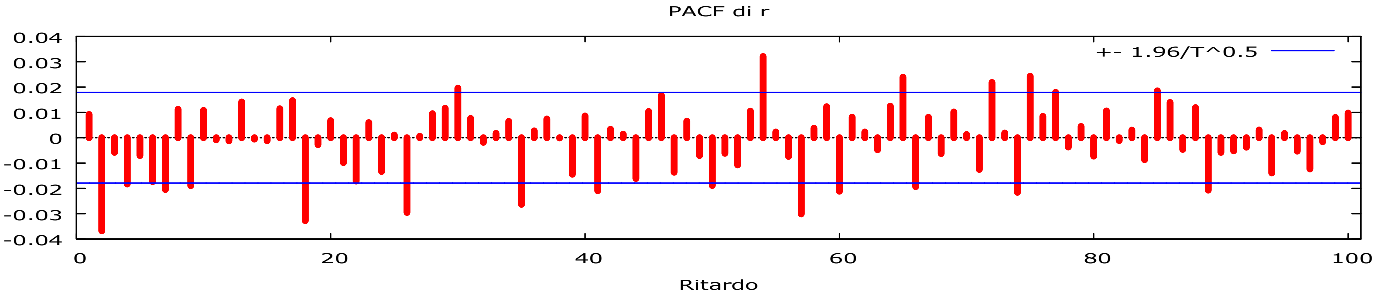


P_t

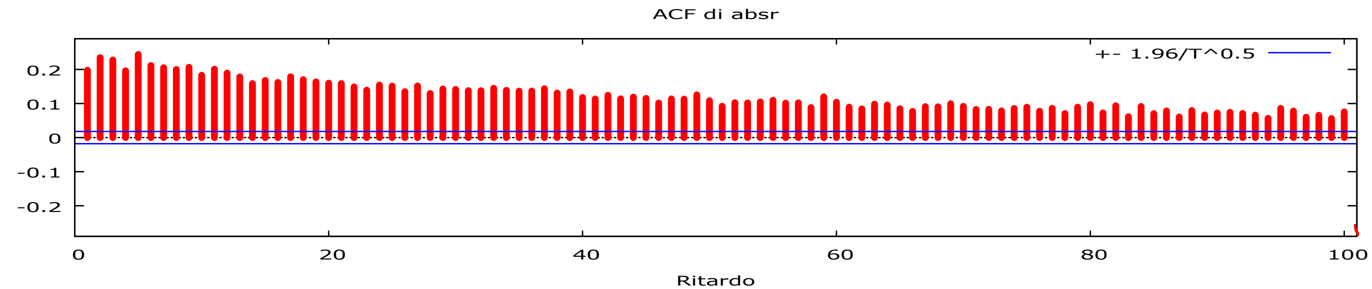


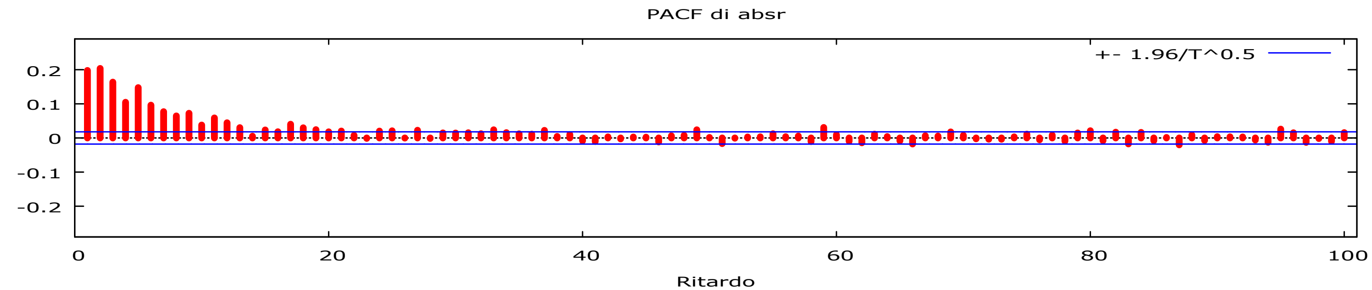
$$r_t := \frac{P_t - P_{t-1}}{P_{t-1}} \times 100$$





$\widehat{Corr}(r_t, r_{t-k})$

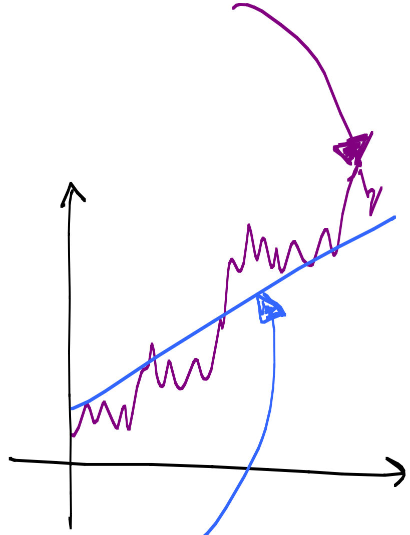




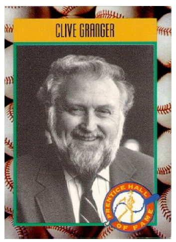
$\widehat{Corr}(|r_t|, |r_{t-k}|)$

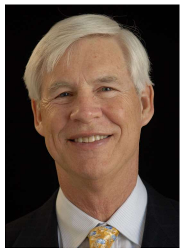
Sviluppo dell'econometria

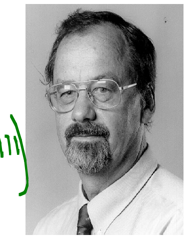
- nascita delle agenzie governative di statistica ufficiale (dati)
- Paradigma Keynesiano (contabilità nazionale)
- Cowles commission
  - nuove metodologie statistiche
  - macro modelli
    - interpretazione
    - simulazioni di policy
    - previsione



- Rational expectations (anni '70) e critica di Lucas
- Metodologia VAR e LSE (1980)
- Non-stazionarietà dei dati economici e modelli per la volatilità
  - Co-integrazione
  - ARCH
- Financial crisis

(i) 

 (ii)

(iii) 

(i) Clive Granger
(ii) Rob. Engle
(iii) Søren Johansen

Programma

PARTE 1

- Introduzione all'econometria
- La costruzione del modello econometrico
  - modelli econometrici uniequazionali
  - condizionamento ed esogeneità
  - Stime [OLS, GLS, ML]
  - Ipotesi sui coefficienti
  - Modelli vincolati e non vincolati
- L'analisi della specificazione
  - confronto tra modelli: [Wald, LM, LR] "nested" e valutazione della bontà dell'adattamento
  - test delle ipotesi del modello
  - (test sui residui)
    break strutturali e stabilità dei parametri
- Modelli dinamici e teoria asintotica
  - stazionarietà ed ergodicità, leggi dei grandi numeri e teoremi centrali del limite
  - il modello lineare dinamico: stima e test

PARTE 2

- Modelli multiequazionali
  - endogeneità, strutture e forma ridotta
  - inconsistenza di OLS ; variabili strumentali [IV]
- Modelli per la varianza (volatilità)
  - caratteristiche dei dati finanziari
  - modelli GARCH
- Ulteriori sviluppi
  - il problema della non-stazionarietà
  - ....
- Laboratorio : analisi di casi di studio (economia / finanza)

  Software: GRETL

Testo di riferimento: GARDINI et al. (2001) Econometria, vol I.

La costruzione del modello econometrico — introduzione

$$\underset{p\times 1}{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = (y_1, y_2, \dots, y_p)'$$ vettore delle variabili endogene

$$\underset{k\times 1}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = (x_1, x_2, \dots, x_k)'$$ vettore delle variabili esplicative (esogene)

$$\underset{m\times 1}{w} = \begin{pmatrix} y \\ x \end{pmatrix} \qquad m := p+k$$

Il dataset.

- Abbiamo n realizzazioni di $w = \begin{pmatrix} y \\ x \end{pmatrix}$
- Queste n realizzazioni definiscono il campione

$i = 1, 2, \dots, n$ l'indice i identifica l'i-esima realizzazione (l'i-esima unità statistica)

$$\underset{p \times 1}{y_i} = \begin{pmatrix} y_{1,i} \\ y_{2,i} \\ \vdots \\ y_{p,i} \end{pmatrix} \qquad x_i = \begin{pmatrix} x_{1,i} \\ x_{2,i} \\ \vdots \\ x_{k,i} \end{pmatrix} \qquad i = 1, 2, \dots, n$$

forma compatta / matriciale

$$\underset{n \times p}{Y} = \begin{bmatrix} y_{1,1} & y_{2,1} & \cdots & y_{p,1} \\ y_{1,2} & y_{2,2} & \cdots & y_{p,2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1,n} & y_{2,n} & \cdots & y_{p,n} \end{bmatrix}$$

$$= \begin{bmatrix} - & y_1' & - \\ - & y_2' & - \\ & \vdots & \\ - & y_n' & - \end{bmatrix}$$

In riga: tutte le variabili relative ad una particolare unità statistica

In colonna: tutte le osservazioni relative ad una particolare variabile

$$\underset{n \times K}{X} = \begin{bmatrix} x_{1,1} & x_{2,1} & \cdots & x_{K,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{K,2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,n} & x_{2,n} & \cdots & x_{K,n} \end{bmatrix}$$

$$= \begin{bmatrix} - & x_1' & - \\ - & x_2' & - \\ & \vdots & \\ - & x_k' & - \end{bmatrix}$$

In riga: tutte le variabili relative ad una particolare unità statistica

In colonna: tutte le osservazioni relative ad una particolare variabile

La matrice dei dati è

$$\underset{n \times m}{W} = [\underset{n \times p}{Y} \;\vdots\; \underset{n \times K}{X}]$$

$m := p + K$

Vogliamo studiare il meccanismo (non noto) che genera i dati, ovvero il processo generatore dei dati [DGP].

Il modello econometrico è un modello statistico che approssima il DGP.

Se $p=1$ il modello è uniequazionale (univariato)
Se $p>1$ " " " multiequazionale (multivariato)

esempio

$$\underset{1\times 1}{y_i} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \qquad (i=1,\dots,n)$$

uniequazionale

$$\begin{cases} y_{1i} = \beta_{10} + \beta_{11} x_{1i} + \varepsilon_{1i} \\ y_{2i} = \beta_{20} + \beta_{21} x_{2i} + \varepsilon_{2i} \end{cases} \qquad (i=1,\dots,n)$$

multiequazionale

Modelli uniequazionali (single equation models)

La base statistica, in questo caso, è del tipo

$$\underset{n\times m}{W} = \begin{pmatrix} y_1 & x_{11} & \dots & x_{k1} \\ y_2 & x_{12} & \dots & x_{k2} \\ \vdots & \vdots & & \vdots \\ y_n & x_{1n} & \dots & x_{kn} \end{pmatrix} = [\underset{n\times 1}{y} \; \vdots \; \underset{n\times k}{X}] \qquad m = k+1$$

<u>nota</u>: alcune delle variabili esplicative possono essere deterministiche

<u>esempio</u>: - la costante (intercetta) corrisponde all'inclusione, tra i regressori, del vettore

$$x_1 = (1, 1, \dots, 1)'$$

- le componenti stagionali

$$s_i = \begin{cases} 1 & \text{se } i = \text{dicembre} \\ 0 & \text{se } i \neq \text{dicembre} \end{cases}$$

- trend temporali

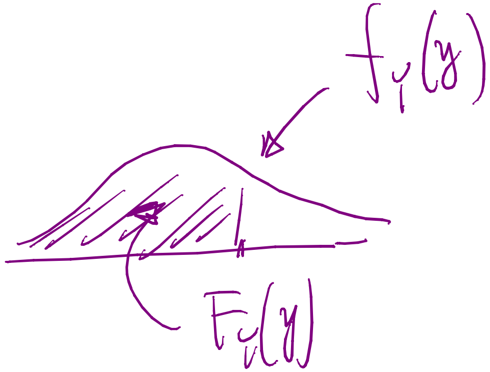
$$s_i = i \qquad (i = 1, \dots, n)$$

De-Tour: distribuzioni congiunte, condizionate e marginali (e valori attesi ad esse associati)

$$w = \begin{pmatrix} Y \\ X \end{pmatrix} \qquad \text{vettore aleatorio (bidimensionale)}$$

$$P(Y \le y, X \le x) = F_{Y,X}(y,x)$$

PROBABILITÀ CONGIUNTA FUNZIONE DI RIPARTIZIONE (CDF)



Associata alla cdf, nel caso "continuo" esiste una funzione di densità congiunta

$$f_{Y,X}(y,x)$$

Il legame con la CDF è:

$$F_{Y,X}(y,x) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{Y,X}(u_1,u_2)\, du_1\, du_2$$

Dalla congiunta possiamo passare alle marginali:

marginale di Y:

$$P(Y \le y) = F_Y(y)$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{Y,X}(y,x)\, dx$$

funzione di densità marginale

marginale di X : $P(X \leq x) = F_X(x)$

$$f_X(x) = \int_{-\infty}^{+\infty} f_{Y,X}(y,x)\,dy$$

Dalle marginali possiamo calcolare diversi "operatori" relativi ad una (e una sola) delle due variabili

valore atteso : $E[Y] = \int_{-\infty}^{+\infty} y\, f_Y(y)\,dy$

momento secondo : $E[Y^2] = \int_{-\infty}^{+\infty} y^2 f_Y(y)\,dy$

varianza : $V[Y] := E[(Y - E(Y))^2] = E[Y^2] - E[Y]^2$

in generale : $E[g(Y)] = \int_{-\infty}^{+\infty} g(y)\, f_Y(y)\,dy$

Dalla congiunta alla condizionata :

$$f_{Y|X}(y|x) = \frac{f_{Y,X}(y,x)}{f_X(x)}$$

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

Si può anche scrivere come:

$$\underbrace{f_{Y,X}(y,x)}_{\text{congiunta}} = \underbrace{f_{Y|X}(y|x)}_{\text{condizionata}} \cdot \underbrace{f_X(x)}_{\text{marginale}}$$

Dalla condizionata possiamo ricavare i <u>momenti condizionati</u>:

$$E[Y|X] = \int_{-\infty}^{+\infty} y \, f_{Y|X}(y|x) \, dy$$

nota: $E[Y|X] = g(X)$ $\qquad E[Y|X=x]$

I valori attesi condizionati sono variabili <u>casuali</u>.

Alcune proprietà

1. "linearità" $E[aY+b\,|X] = aE[Y|X]+b$
2. "inclusione" $E[g(X)|X] = g(X)$ $\qquad E(X|X)=X$

③ "legge dei valori attesi iterati"

$$E_X\{\underbrace{E[Y|X]}_{E(g(x))}\} = E[Y]$$

<u>esempio</u>: la distribuzione gaussiana ("normale")

$$\omega = \begin{pmatrix} \underset{p\times 1}{Y} \\ \underset{k\times 1}{X} \end{pmatrix} \sim N\left(\begin{pmatrix} \underset{p\times 1}{\mu_Y} \\ \underset{k\times 1}{\mu_X} \end{pmatrix}, \begin{pmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{pmatrix} \right)$$

<u>Proprietà</u>: se $(Y, X) \sim N$ (congiuntamente) allora
$Y \sim N$, $X \sim N$ (marginali normali)

$$Y \sim N(\underbrace{\mu_Y}_{E[Y]}, \underbrace{\Sigma_{YY}}_{V[Y]}) \qquad X \sim N(\underbrace{\mu_X}_{E[X]}, \underbrace{\Sigma_{XX}}_{V[X]})$$

<u>Proprietà</u> Se $(Y, X) \sim N$ (congiuntamente) allora
$Y|X \sim N$ (condizionate normale)
$X|Y \sim N$

In particolare:

$$E[Y|X] = \mu_Y + \Sigma_{YX} \Sigma_{XX}^{-1} (X - \mu_X)$$

note: se $\Sigma_{YX} = 0$ (no covarianze), $E[Y|X] = E[Y]$

note:

$$E[Y|X] = \underbrace{\mu_Y - \Sigma_{YX} \Sigma_{XX}^{-1} \mu_X}_{} + \underbrace{\Sigma_{YX} \Sigma_{XX}^{-1}}_{} \cdot X$$

$$\underset{p\times 1}{E[Y|X]} = \underset{p\times 1}{\alpha_Y} + \underset{p\times K}{\beta_{Y|X}} \cdot \underset{K\times 1}{X}$$

$$V[Y|X] = \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} =: \Sigma_{YY.X}$$

Torniamo ai modelli uniequazionali

$y \qquad x = (x_1, \dots, x_k)'$

$$\underbrace{f_{y,x}(y, x_1, \dots, x_k ; \theta)}_{\text{"congiunta"}} = \underbrace{f_{y|x}(y | x_1, \dots, x_k ; \lambda_1)}_{\text{condizionata}} \underbrace{f_x(x_1, \dots, x_k ; \lambda_2)}_{\text{marginale}}$$

esempio della gaussiana

parametri della congiunta : $\theta = \{ \mu_y, \mu_x, \Sigma_{yy}, \Sigma_{xx}, \Sigma_{yx} \}$

parametri della condizionata : $\lambda_1 = \{ \alpha_y, \beta_{y|x}, \Sigma_{yy.x} \}$

parametri della marginale : $\lambda_2 = \{ \mu_x, \Sigma_{xx} \}$

notiamo che λ_1 e λ_2 sono funzioni di θ.

nel caso gaussiano, si può dimostrare che λ_1 e λ_2 sono "liberi di variare" (VARIATION FREE):

pone dei vincoli su λ_1 non impone vincoli su λ_2

In generale, la scelta del modello da analizzare dipende dai parametri di interesse:

Se i parametri di interesse sono $\tau = \tau(\lambda_1)$ e se λ_1 e λ_2 sono liberi di variare, allora l'inferenza su τ può essere effettuata considerando il solo modello condizionato (senza perdita di informazione)
In questo caso si dice che $X = (X_1, \dots, X_k)'$ è debolmente esogeno per $\tau = \tau(\lambda_1)$

Supponiamo che ci interessino i parametri del modello condizionato (o una funzione di tali parametri)

L'inferenza viene effettuata usando

$$f_{Y|X}(y \mid x_1, \dots, x_k ; \lambda_1)$$

Sulla base di $f_{y|x}$, possiamo derivare il valore atteso condizionato:

$$E[y|x] = g(x_1, x_2, \ldots, x_k; \lambda_1)$$

La forma di $g(\cdot)$ dipende da $f_{y|x}$. Ad esempio, nel caso gaussiano:

$$E[y|x] = g(x_1, x_2, \ldots, x_k; \lambda_1) = \alpha_y + \beta_{y|x} \cdot x$$

non è sempre vero in generale — $g(\cdot)$ potrebbe essere una funzione più complicata

Un modello uniequazionale (per il valore atteso condizionato) è del tipo

$$\boxed{E[y|x] = g(x; \lambda_1)} = g(x_1, x_2, \ldots, x_k; \lambda_1)$$

Consideriamo il caso particolare del

$\boxed{\text{modello di regressione lineare}}$

Nelle slide precedenti abbiamo considerato la relazione tra y e X nella popolazione.

Consideriamo ora il caso in cui abbiamo n osservazioni:

$$\underset{1\times 1}{y_1}, y_2, \dots, y_n \quad ; \quad \underset{K\times 1}{X_1}, X_2, \dots, X_n$$

Il modello che abbiamo in mente è il seguente:

$$E[y_i \mid X_i] = E[y_i \mid X_{1i}, X_{2i}, \dots, X_{Ki}] = g(X_{1i}, \dots, X_{Ki} ; \beta)$$

che, nel caso lineare, diventa

$$E[y_i \mid X_i] = \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} = \sum_{j=1}^{K} \beta_j X_{ji}$$

Poiché $(x_{1i}, x_{2i}, \dots, x_{Ki})' =: \underset{K\times 1}{x_i}$ e $(\beta_1, \beta_2, \dots, \beta_K)' = \beta_{K\times 1}$ possiamo scrivere

$$\underset{1\times 1}{E[y_i|x_i]} = \underbrace{\underset{1\times K}{\beta'}\underset{K\times 1}{x_i}}_{1\times 1} = \underbrace{\underset{1\times K}{x_i'}\underset{K\times 1}{\beta}}_{1\times 1} \qquad i = 1, 2, \dots, n$$

note: intercetta – Il caso in cui vogliamo includere l'intercetta si ottiene ponendo

$$x_{1i} = 1 \quad (i = 1, 2, \dots n)$$

Il modello si può equivalentemente scrivere nella forma:

$$y_i = x_i'\beta + \varepsilon_i \qquad (i = 1, 2, \dots, n)$$

Infatti

$$y_i = \underbrace{y_i - E[y_i|x_i]}_{\varepsilon_i} + \underbrace{E[y_i|x_i]}_{x_i'\beta} = x_i'\beta + \varepsilon_i$$

dove abbiamo definito $\varepsilon_i = y_i - E[y_i | x_i] = y_i - x_i'\beta$

Nota: ε_i, così costruito, ha valore atteso condizionato uguale a 0:

$$E[\varepsilon_i | x_i] = E[y_i - x_i'\beta | x_i] = \underline{E[y_i|x_i]} - \underline{E[x_i'\beta | x_i]}$$

$$= x_i'\beta - x_i'\beta = 0$$

Quindi:

Il modello di regressione lineare può essere equivalentemente scritto come

(i) $E[y_i|x_i] = x_i'\beta$

(ii) $\begin{cases} y_i = x_i'\beta + \varepsilon_i \\ E[\varepsilon_i | x_i] = 0 \end{cases}$

in entrambi i casi per $i = 1, 2, \dots, n$.

nota: $E[\varepsilon_i | x_i] = 0$ implica che $E[\varepsilon_i] = 0$

(il viceversa non è necessariamente vero)

Perché? $E[\varepsilon_i] = E_{x_i}\Big[\underbrace{E[\varepsilon_i | x_i]}_{0}\Big] = E_{x_i}[0] = 0$

nota: $E[\varepsilon_i | x_i] = 0$ implica

$$\underset{1\times 1}{\varepsilon_i} \cdot \underset{1\times K}{x_i'} \to E[\varepsilon_i \cdot x_i'] = \underset{1\times K}{0}$$

[fare dimostrazione]

Sostituisco ad ε_i la sua espressione $\varepsilon_i = y_i - \beta' x_i$

$$E[(y_i - \beta' x_i) x_i'] = 0$$

$$E[y_i x_i' - \beta' x_i x_i'] = 0$$

$$E[y_i x_i'] - \beta' E[x_i x_i'] = 0$$

$$E[y_i x_i'] = \beta' E[x_i x_i']$$

$(AB)' = B'A'$

se $E[x_i x_i']$ è invertibile, posso scrivere

$$E[y_i x_i'] E[x_i x_i']^{-1} = \beta'$$

$\beta = E[x_i x_i']^{-1} E[x_i y_i]$ · coeff. di regressione nella popolazione

Non abbiamo detto nulla sulle varianze di y_i, o delle varianze condizionate, $V[y_i | x_i]$

In generale,

$$V[y_i | x_i] = \sigma_i^2 = f_i(x_i) \qquad (i = 1, 2, \dots, n)$$

Se queste varianze condizionate sono uguali tra loro,

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

diciamo che il modello assume omoschedasticità (condizionata). Viceversa, parliamo di eteroschedasticità.

Nota: nel modello scritto nella forma $y_i = x_i'\beta + \varepsilon_i$,
abbiamo

$$V[\varepsilon_i | x_i] = \sigma_i^2$$

ovvero:

$$V[\varepsilon_i | x_i] = V[y_i | x_i]$$

Infatti:

$$V[\varepsilon_i | x_i] = V[y_i - \cancel{x_i'\beta} | x_i] = V[y_i | x_i]$$

Riassumendo:

modello lineare con errori eteroschedastici

$$y_i = x_i'\beta + \varepsilon_i, \quad E[\varepsilon_i | x_i] = 0, \quad V[\varepsilon_i | x_i] = \sigma_i^2 \qquad (i = 1 \dots n)$$

oppure:

$$E[y_i | x_i] = x_i'\beta, \quad V[y_i | x_i] = \sigma_i^2 \qquad (i = 1, \dots, n)$$

modello lineare con errori omoschedastici

$$y_i = x_i'\beta + \varepsilon_i, \quad E[\varepsilon_i|x_i]=0, \quad V[\varepsilon_i|x_i]=\sigma^2 \quad (i=1\dots n)$$

oppure:

$$E[y_i|x_i]=x_i'\beta, \quad V[y_i|x_i]=\sigma^2 \quad (i=1,\dots,n)$$

Dobbiamo caratterizzare l'eventuale dipendenza tra unità statistiche

Ipotesi di indipendenza: $\{y_i, x_i\}_{i=1,2,\dots,n}$ sono indipendenti stocasticamente

Un'ipotesi meno forte (ma sempre forte!) è la seguente

$$X = \begin{bmatrix} - & x_1' & - \\ - & x_2' & - \\ & \vdots & \\ - & x_n' & - \end{bmatrix}$$

$E[y_i|x_i]=x_i'\beta$ viene sostituita da

$$E[y_i|\underset{K\times 1}{x_1}, x_2,\dots,x_i,\dots,x_n]=E[y_i|\underset{n\times K}{X}]=x_i'\beta$$

Il condizionamento è rispetto a <u>tutta</u> la matrice delle X

Analogamente, la varianza condizionata è del tipo

$$V[y_i | X_1, X_2, \dots, X_n] = V[y_i | X] = \begin{cases} \sigma_i^2 & \text{eteroschedasticità} \\ \sigma^2 & \text{omoschedastico} \end{cases}$$

Usando la notazione $y_i = x_i'\beta + \varepsilon_i$, il modello diventa

$$y_i = x_i'\beta + \varepsilon_i, \qquad E[\varepsilon_i | X_1 \dots X_n] = E[\varepsilon_i | X] = 0$$

$$V[\varepsilon_i | X_1 \dots X_n] = V[\varepsilon_i | X] = \begin{cases} \sigma_i^2 & \text{eterosch.} \\ \sigma^2 & \text{omosch.} \end{cases}$$

Non abbiamo detto nulla sulla covarianza (condizionata) tra le y_i delle diverse unità statistiche

$$Cov(y_i, y_j | X) = \sigma_{ij} \qquad (i,j = 1, 2, \dots, n)$$

nota: $\mathrm{Cov}(y_i, y_j \mid X) = \mathrm{Cov}(\varepsilon_i + \cancel{x_i'\beta}, \varepsilon_j + \cancel{x_j'\beta} \mid X)$

$= \mathrm{Cov}(\varepsilon_i, \varepsilon_j \mid X) = \sigma_{ij}$

Se $\sigma_{ij} = 0$ per ogni $i \neq j = 1, 2, \ldots, n$ si dice che il modello ha errori incorrelati

Se $\sigma_{ij} \neq 0$ per qualche $i \neq j$, si dice che il modello ha errori correlati, e si parla di modello lineare generalizzato

Modelli lineari uniequazionali

Modello lineare generalizzato

$$E[y_i \mid x_1, x_2, \dots, x_n] = E[y_i \mid X] = x_i' \beta \qquad i = 1, 2, \dots, n$$

$$V[y_i \mid x_1, x_2, \dots, x_n] = V[y_i \mid X] = \sigma_i^2 \qquad \left(0 < \underline{\underline{\sigma_i^2 < \infty}} \right)$$

$$Cov[y_i, y_j \mid x_1, x_2, \dots, x_n] = Cov[y_i, y_j \mid X] = \sigma_{ij}$$

equivalentemente (formulazione con la componente stocastica)

$$y_i = x_i' \beta + \varepsilon_i \qquad , \; i = 1, 2, \dots, n$$

$$E[\varepsilon_i \mid x_1 \dots x_n] = E[\varepsilon_i \mid X] = 0$$

$$V[\varepsilon_i \mid x_1 \dots x_n] = V[\varepsilon_i \mid X] = \sigma_i^2 \qquad (0 < \sigma_i^2 < \infty)$$

$$Cov[\varepsilon_i, \varepsilon_j \mid x_1 \dots x_n] = Cov[\varepsilon_i, \varepsilon_j \mid X] = \sigma_{ij}$$

Caso particolare del modello lineare generalizzato è il
Modello lineare con errori eteroschedastici

$$E[y_i | X_1 \dots X_n] = E[y_i | X] = X_i' \beta \qquad i = 1, 2, \dots, n$$

$$V[y_i | X_1 \dots X_n] = V[y_i | X] = \sigma_i^2 \qquad (0 < \sigma_i^2 < \infty)$$

$$Cov[y_i, y_j | X_1 \dots X_n] = Cov[y_i, y_j | X] = 0 \qquad i \neq j = 1, 2, \dots, n$$

(data X, le y_i sono incorrelate tra loro)

equivalentemente (formulazione con la componente stocastica)

$$y_i = x_i' \beta + \varepsilon_i \qquad i = 1, 2, \dots, n$$

$$E[\varepsilon_i | X_1, \dots, X_n] = E[\varepsilon_i | X] = 0$$

$$V[\varepsilon_i | X_1 \dots X_n] = V[\varepsilon_i | X] = \sigma_i^2 \qquad 0 < \sigma_i^2 < \infty$$

$$Cov[\varepsilon_i, \varepsilon_j | X_1, \dots, X_n] = Cov[\varepsilon_i, \varepsilon_j | X] = 0 \qquad i \neq j = 1, 2, \dots n$$

Caso particolare del modello lineare generalizzato e del modello lineare con errori eteroschedastici è il **modello lineare classico**

$$E[y_i | x_1 \dots x_n] = E[y_i | X] = x_i' \beta \qquad i = 1, 2, \dots, n$$

$$V[y_i | x_1 \dots x_n] = V[y_i | X] = \sigma^2 \qquad 0 < \sigma^2 < \infty \quad \text{(omoschedasticità)}$$

$$Cov[y_i, y_j | x_1 \dots x_n] = Cov[y_i, y_j | X] = 0 \qquad i \neq j = 1, 2, \dots, n$$

equivalentemente (formulazione con la componente stocastica)

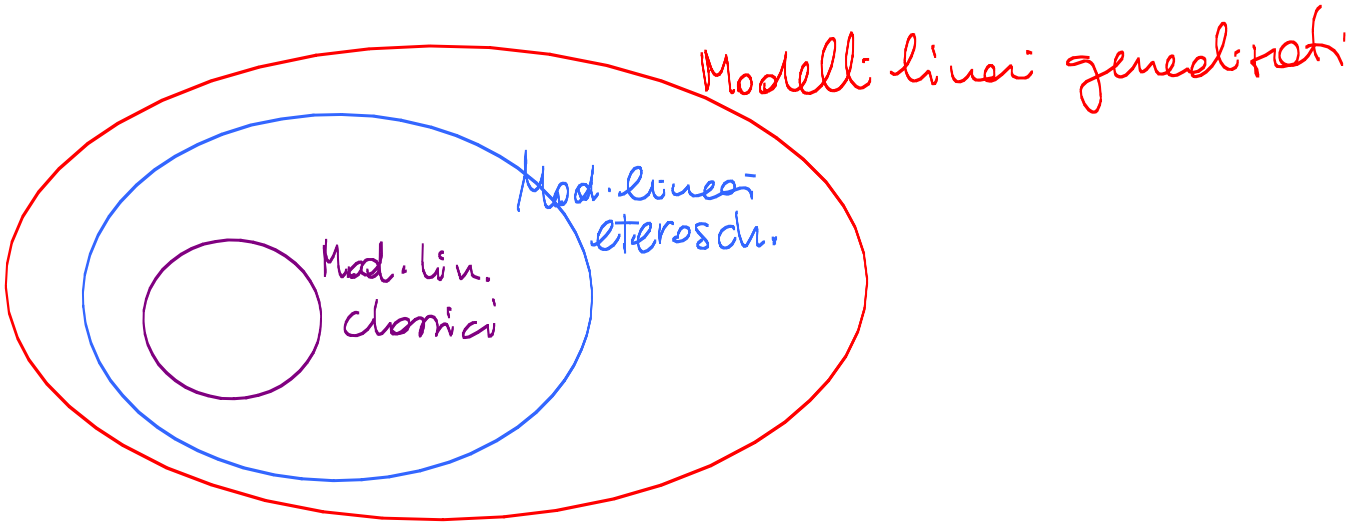
$$y_i = x_i' \beta + \varepsilon_i \qquad i = 1, 2, \dots, n$$

$$E[\varepsilon_i | x_1 \dots x_n] = E[\varepsilon_i | X] = 0$$

$$V[\varepsilon_i | x_1 \dots x_n] = V[\varepsilon_i | X] = \sigma^2$$

$$Cov[\varepsilon_i, \varepsilon_j | x_1, \dots, x_n] = Cov[\varepsilon_i, \varepsilon_j | X] = 0 \qquad i \neq j = 1, 2, \dots, n$$

nota: $Cov[\varepsilon_i, \varepsilon_j | X] = E[\varepsilon_i \varepsilon_j | X] - E\cancel{[\varepsilon_i | X]} E\cancel{[\varepsilon_j | X]}$



Notazione compatta del modello lineare

$$\underset{1\times 1}{y_i} = \underset{1\times K}{x_i'}\,\underset{K\times 1}{\beta} + \underset{1\times 1}{\varepsilon_i} \qquad i = 1, 2, \dots, n$$

$$E[\varepsilon_i \mid X] = 0$$

definiamo:

$$\underset{n\times 1}{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad \underset{n\times K}{X} = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{K1} \\ x_{12} & x_{22} & \dots & x_{K2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & \dots & x_{Kn} \end{bmatrix} \qquad \underset{n\times 1}{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

possiamo scrivere (forma compatta):

$$\boxed{\underset{n\times 1}{y} = \underset{n\times k}{X}\,\underset{k\times 1}{\beta} + \underset{n\times 1}{\varepsilon}}$$

infatti:

$$\begin{array}{l} i=1 \\ i=2 \\ \vdots \\ i=n \end{array} \qquad \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \begin{array}{c} = \\ = \\ \vdots \\ = \end{array} \begin{bmatrix} x_1' \\ x_2' \\ \\ x_n' \end{bmatrix} \begin{array}{c} \beta \\ \beta \\ \\ \beta \end{array} \begin{array}{c} + \\ + \\ \\ + \end{array} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \\ \varepsilon_n \end{bmatrix}$$

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

nota: $X\beta = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} \beta = \begin{bmatrix} x_1'\beta \\ x_2'\beta \\ \vdots \\ x_n'\beta \end{bmatrix}$

L'ipotesi $E[\varepsilon_i | X] = 0$, $i = 1, 2, \dots, n$ diventa, nella forma compatta:

$$\boxed{\underset{n\times 1}{E[\varepsilon | X]} = \underset{n\times 1}{0}}$$

Infatti:

$$E[\varepsilon | X] = E\left[\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \middle| X\right] = \begin{bmatrix} E[\varepsilon_1 | X] \\ E[\varepsilon_2 | X] \\ \vdots \\ E[\varepsilon_n | X] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \underset{n\times 1}{0}$$

nota importante: la forma compatta del modello scritto sente ε_f è:

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

dimostrazione: fare per esercizio.

Relativamente alla componente stocastica, dobbiamo rappresentare in forma compatta le varianze e le covarianze (condizionate):

$$E[\varepsilon_i^2 \mid X] = \sigma_i^2 \qquad i=1,2,\dots,n$$

$$E[\varepsilon_i \varepsilon_j \mid X] = \sigma_{ij} \qquad i,j=1,2,\dots,n \quad (i \neq j)$$

Partiamo dalla rappresentazione in forma compatta di ε_i, $i=1,2,\dots,n$

$$\underset{n\times 1}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$$

La matrice di varianze e covarianze di ε è:

$$V[\underset{n\times 1}{\varepsilon} \mid X] = E[\underbrace{\underset{n\times 1}{\varepsilon}\,\underset{1\times n}{\varepsilon'}}_{n\times n} \mid X]$$

In particolare:

$$E[\varepsilon\varepsilon'|X] = E\left[\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_n\end{pmatrix}\begin{pmatrix}\varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_n\end{pmatrix} \Bigg| X\right] = E\left[\begin{bmatrix}\varepsilon_1^2 & \varepsilon_1\varepsilon_2 & \cdots & \varepsilon_1\varepsilon_n\\ \varepsilon_2\varepsilon_1 & \varepsilon_2^2 & \cdots & \varepsilon_2\varepsilon_n\\ \vdots & \vdots & \ddots & \vdots\\ \varepsilon_n\varepsilon_1 & \varepsilon_n\varepsilon_2 & \cdots & \varepsilon_n^2\end{bmatrix} \Bigg| X\right]$$

$$= \begin{bmatrix} E[\varepsilon_1^2|X] & E[\varepsilon_1\varepsilon_2|X] & \cdots & E[\varepsilon_1\varepsilon_n|X]\\ E[\varepsilon_2\varepsilon_1|X] & E[\varepsilon_2^2|X] & \cdots & E[\varepsilon_2\varepsilon_n|X]\\ \vdots & & & \\ E[\varepsilon_n\varepsilon_1|X] & E[\varepsilon_n\varepsilon_2|X] & \cdots & E[\varepsilon_n^2|X]\end{bmatrix}$$

varianze

covarianze

$$= \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n}\\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2\end{bmatrix}$$

in quanto $E[\varepsilon_i^2|X] = \sigma_i^2$

$E[\varepsilon_i\varepsilon_j|X] = \sigma_{ij} \quad i \neq j$

$$=: \underset{n\times n}{\Sigma}$$

nota: Σ è una matrice simmetrica, in quanto $\sigma_{ij} = \sigma_{ji}$ (la covarianze è simmetrica)

Riassumendo:

$$V[\varepsilon|X] = E[\varepsilon\varepsilon'|X] = \underset{n\times n}{\Sigma}$$

nota: per effetto della simmetria Σ, pur essendo $n \times n$, ha $n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}$ elementi liberi

Quindi, la forma compatta del mod. lineare generalizzato è

$$y = X\beta + \varepsilon$$
$$E[\varepsilon | X] = 0$$
$$V[\varepsilon | X] = \Sigma$$

$\max_{i=1 \dots n} \sigma_i^2 < \infty$

Σ definita positiva (non stocastica)

nella versione senza componente stocastica:

$$E[y | X] = X\beta$$
$$V[y | X] = \Sigma,$$

Σ definita positiva (non stocastica) $\max_{i=1 \dots n} \sigma_i^2 < \infty$

Il modello lineare con errori eteroschedastici diventa

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

$$E[\varepsilon | X] = 0$$

$$V[\varepsilon | X] = \Sigma \quad \underline{\text{diagonale}} \qquad 0 < \sigma_i^2 < \infty$$

oppure

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

$$V[y | X] = \Sigma \quad \underline{\text{diagonale}}$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

$$\boxed{\text{nota sulla notazione:} \quad \Sigma = \operatorname{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2\}}$$

Infine, il modello lineare classico si rappresenta come:

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

$$E[\varepsilon | X] = 0$$

$$V[\varepsilon | X] = \sigma^2 I_n \quad , \quad 0 < \sigma^2 < \infty \quad , \quad \underset{n\times n}{I_n} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

infatti, in questo caso:

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \sigma^2 I_n$$

Analogamente:

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

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

<u>nota</u>: una matrice nella forma $a \cdot I_n$ si dice "sferica"
l'ipotesi $V[\varepsilon | X] = V[y | X] = \sigma^2 I_n$ viene detta "di sfericità" "degli errori"