

## LEZIONI DI R

### INTRODUZIONE

Un modello di politica economica contenente variabili endogene ( $y = y_1, \dots, y_2$ ), ovvero obiettivo, e variabili esogene ( $x = x_1, \dots, x_2$ ), ovvero date che raffigurano gli strumenti del policy maker. Il nostro obiettivo sarà quello di relazionare variabili endogene e variabili esogene che passeranno attraverso  $n$  funzioni matematiche in relazione alle nostre variabili esogene, assumendo che a destra dell'equazione vi siano gli strumenti e che questi siano legati all'obiettivo di politica economica da perseguire (sx dell'equazione).

### CRITICA DI LUCAS

La critica di Lucas è una critica alla macroeconometria, cioè a quella disciplina che si occupa di studiare tramite strumenti statistico-matematici i dati macroeconomici. La domanda di fondo è: il policy maker può riuscire ad indirizzare la macroeconomia? Può effettivamente riuscire a raggiungere i suoi obiettivi di politica economica?

Per rispondere a queste domanda, Lucas, scrisse un paper molto sofisticato. Di seguito viene presentata una versione semplificata della matematica dietro al paper:

Prendo un vettore  $Y$  che contiene gli  $m$  obiettivi del policy maker (cioè la volontà di raggiungere determinati valori di variabili come la spesa pubblica, il PIL, il tasso di inflazione, ecc...):

$$Y = Y_1, \dots, Y_m$$

Per raggiungere questi valori può utilizzare alcuni o tutti degli  $n$  strumenti di cui dispone all'interno del vettore  $X$ :

$$X = X_1, \dots, X_n$$

Dal punto di vista matematico questo significa trovare una qualche relazione matematica tra le  $X$  e le  $Y$ , che esplicheremo con  $f$ .

Ad esempio, la nostra prima variabile obiettivo potrebbe essere legare gli strumenti matemaici da una funzione chiamata  $f_1$ , la quale dipende dagli altri  $m$  obiettivi:

$$Y_1 = f_1(X_1, \dots, X_n)$$

La seconda variabile obiettivo potrebbe dipendere ancora dagli  $n$  strumenti secondo una funzione matematica del tipo  $f_2$ :

$$Y_2 = f_2(X_1, \dots, X_n)$$

E così via. Le funzioni matematiche potrebbero anche essere simili tra di loro; infatti, l'assunzione semplificatrice che facciamo è di assumere che queste siano una combinazione lineare dei parametri del modello:

$$Y_1 = f_1(X_1, \dots, X_n) \rightarrow Y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n$$

$$Y_2 = f_2(X_1, \dots, X_n) \rightarrow Y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n$$

...

$$Y_n = f_n(X_1, \dots, X_n) \rightarrow Y_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n$$

Questo sistema di equazioni è un sistema di equazioni in forma lineare. Si può esprimere in forma algebrica:

$Y = Ax$  con:  $Y$  ( $m \times 1$ ): **vettore colonna con  $m$  righe ed una colonna (raffigura l'obiettivo);**

$A$  ( $m \times n$ ): **matrice con  $m$  righe ed  $n$  colonne (raffigura i coefficienti che determineranno gli effetti delle  $X$  sulle  $Y$ );**

$x$  ( $n \times 1$ ): **vettore colonna con  $n$  righe ed una colonna (raffigura gli strumenti).**

Se  $m = n \rightarrow$  **Sistema determinato**

Se  $m < n \rightarrow$  **Sistema sottodeterminato (infinte soluzioni)**

Se  $m > n \rightarrow$  **Sistema non risolubile (più obiettivi che strumenti)**

Il prodotto tra  $A$  e  $x$  è possibile perché hanno  $n$  colonne di  $A$  e  $n$  righe di  $x$ .

$$\begin{bmatrix} Y1 \\ Y2 \\ Y3 \\ Y4 \\ Y5 \end{bmatrix} = \begin{bmatrix} a11 & a12 & a13 & a14 \\ a21 & a22 & a23 & a24 \\ a31 & a32 & a33 & a34 \\ a41 & a42 & a43 & a44 \\ a51 & a52 & a53 & a54 \end{bmatrix} \begin{bmatrix} X1 \\ X2 \\ X3 \\ X4 \end{bmatrix}$$

La matrice  $a$  racchiude i parametri comportamentali dell'economia, cioè fa capire la reazione del sistema economico quando vengono immesse in input le variabili  $X$  (degli strumenti esogeni). Perciò, se si fa variare gli strumenti quale sarà la reazione del sistema economico in termini di vettore  $Y$ ? Se si usano certi strumenti quali sono i risultati sugli obiettivi? Questa variazione è spiegata da questa matrice.

Lucas sostiene che il policy maker lavori con un sistema di politica economica dato; quindi, se vengono dati dei valori alla  $X$ , affinché possa stimare gli effetti sulla  $Y$ , bisogna assumere che i parametri della matrice  $A$  non cambino. L'assunzione di fondo del modello, perciò, è che la matrice  $a$  rimanga fissa.

Da qua nasce la critica di Lucas: nel momento in cui si mette in atto una politica economica, cioè si lavora sugli strumenti, gli agenti economici ne sono a conoscenza e, sapendolo, cambieranno i loro comportamenti che perciò faranno raggiungere un valore della  $Y$  che non era quello inizialmente previsto.

Ad esempio, il policy maker nella  $Y$  ha come obiettivo la riduzione del  $n$  di contagi; lo strumento che vuole usare è una politica restrittiva della socialità degli agenti, cioè il green pass. Inizialmente il governo assumerà una matrice  $a$ , cioè si pensa che con l'introduzione del green pass si ottenga la riduzione nel  $n$  dei contagi ma ciò non tiene conto del fatto che nel momento in cui avviene l'introduzione del green pass, gli agenti economici possono cambiare comportamenti, come il diventare meno prudenti perché hanno il green pass e questo non porterà quindi alla riduzione auspicata nel  $n$  dei contagi.

Perciò, la critica di Lucas afferma che nel momento in cui il policy maker mette in atto la politica, gli agenti economici cambiano comportamento e questo, quindi, impedisce di raggiungere il risultato desiderato. Questa critica si basa su un'assunzione molto forte: gli agenti sono perfettamente razionali, perché un agente perfettamente razionale è un agente che usa tutte le info che ha a disposizione per prendere le sue scelte; quindi, non appena viene a conoscenza di un cambiamento dello strumento, agirà di conseguenza e questo cambierà  $a$  e il risultato finale di  $Y$ . Questo assunto di perfetta razionalità è, però, molto criticato. Quindi il sunto è che per Lucas la politica economica ed i modelli macro-econometrici sono inutili, dal momento che il policy maker NON può prevedere ex ante i comportamenti degli agenti del mercato.

La *Regola aurea della politica economica* è che affinché un modello statico e deterministico di politica economica con obiettivi fissi sia controllabile è che il numero di strumenti sia maggiore o uguale al numero di obiettivi. Questo è il **Teorema di Tinbergen**. È necessario oltretutto che uno strumento non sia il duplicato dell'altro: ad esempio, una variazione del PIL ed una variazione dei contagi come obiettivi e come strumenti spesa pubblica e spesa pubblica poiché sarebbe solo uno strumento.

Siccome nella prassi la condizione  $m = n$  non è quasi mai verificata, allora quello che può fare il policy maker è di passare da un problema con obiettivi fissi a un problema con obiettivi flessibili. Ogni modello,

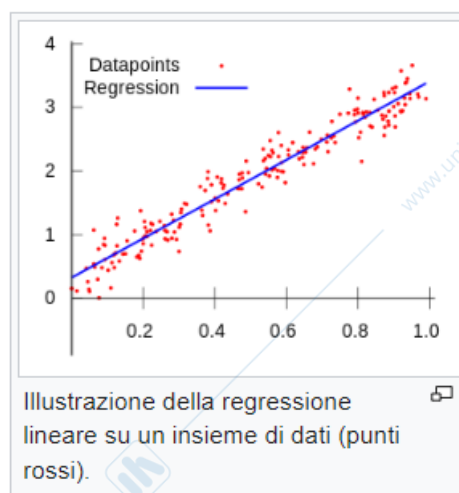
- Problema statico: trascura il fattore tempo nelle analisi;
- Problema deterministico: tutti gli outcomes son certi, quindi non ci sono probabilità nei modelli;
- Obiettivi fissi: si vogliono valori puntuali non probabilistici;
- Sistema controllabile: il sistema è risolvibile, perciò si hanno almeno tanti strumenti quanti gli obiettivi.

Perciò il policy maker costruisce una funzione di perdita che tiene conto non del valore puntuale ma dello scostamento del valore puntuale. Ad esempio, se si hanno due obiettivi ma uno strumento, allora il problema non è risolvibile con obiettivi fissi; quindi, si ambisce a risolverlo con obiettivi flessibili, perciò, invece di considerare  $Y_1^*$ , che è un obiettivo fisso, si considera lo scostamento che si può raggiungere da quell'obiettivo, cioè  $(Y_1 - Y_1^*)$ . Quindi, invece di ottenere quell'esatto valore, ci si accontenta di avvicinarsi. che sia statistico, microeconomico o macroeconomico, sarà pur sempre scomposto da due componenti: una sistematica, che potrà essere stimata, ed una erratica, o noise, che invece dovrà essere ottenuta per differenza.

Per farlo si costruisce una funzione con tutti questi scostamenti e la si minimizza poiché minimizzandola si minimizzano le distanze tra quanto si può effettivamente raggiungere e quanto si sarebbe potuto raggiungere.

## REGRESSIONE

Si osservano due variabili  $X$  e  $Y$  e si postula che sia una relazione di dipendenza della  $Y$  rispetto alla  $X$ , cioè dati certi valori della  $X$  la  $Y$  avrà certi valori.



Sembrerebbe ci sia una relazione positiva tra  $X$  e  $Y$  dato che ad un aumento della  $X$  corrisponde un aumento della  $Y$ . Bisogna però trovare una funzione che descriva bene questo modello ma di linee che passano tra i punti che ne sono infinite; quindi, bisogna cercare di capire quale sia la funzione matematica lineare che meglio descrive i dati. Per fare questo si utilizza un algoritmo che si chiama **OLS**, cioè Ordinary Least Square, cioè Metodo dei minimi quadrati.

Nel modello viene contemplato un margine di errore che è dato dalla distanza che si ha tra i punti sulla retta (cioè le predizioni) e i valori che effettivamente si osservano. Questa distanza viene chiamata Residuo, cioè  $e_i$ . Questo succede perché quando si osserva la variabile dipendente non si osserva il vero modello statistico

ma una versione rumorosa del modello. Questa versione è composta di due parti: una parte è la funzione matematica che lega la  $X$  alla  $Y$  e un'altra parte che è un errore e che è il rumore (random):

$$Y = f(X) + \varepsilon$$

In particolare, si è assunto anche che  $f(X)$  sia una retta, perciò:

$$Y = b(0) + b(1)X + \varepsilon$$

Dove:

- $b(0)$  è l'intercetta;
- $b(1)$  è il coefficiente angolare che misura l'inclinazione della retta;
- $\varepsilon$  è il termine di errore di cui non si è a conoscenza;
- $E(Y | X) = \beta_1 + \beta_2 X$ .

Si vuole stimare  $b(0) + b(1)X$ , cioè la componente sistematica deterministica del modello, la funzione che minimizza gli scarti:

$$\min \sum_{i=1}^n e_i^2$$

Si considera la sommatoria perché si vogliono prendere le distanze contemporaneamente e si considera il quadrato perché se si stesse sopra la retta  $e_i$ , sarebbe positivo e se si stesse sotto sarebbe negativo; quindi, se non si considerasse il quadrato si potrebbe avere il caso di positivo e negativo che si eliminano e quindi si potrebbe avere una misura della distanza pari a 0 quando non lo è. Non si usa il valore assoluto perché dal punto di vista matematico e di computazione è più complesso e perché, lavorando con i quadrati, si può dare un'interpretazione di queste grandezze in termini di variabilità.

L'equazione stimata con OLS sarà quindi:

$$\hat{y} = \hat{\beta}_1 + \hat{\beta}_2 x$$

Residui:

$$\hat{e}_i = y_i - \hat{y}_i$$

Con  $y_i$  osservazione e  $\hat{y}_i$  predizione (il cappello sulle variabili indica la stima di quella variabile).

Con le seguenti assunzioni:

- $E(Y | X) = \beta_1 + \beta_2 X$  ( $E$  componente di errore del modello);
- $Var(Y | X) = \sigma^2$ ;
- $Cov(Y_i, Y_j) = 0 \rightarrow Corr(Y_i, Y_j) = 0$ ;
- $X$  non è casuale e deve assumere almeno due valori distinti;
- (opzionale)  $y \sim N(\beta_1 + \beta_2 X, \sigma^2) \Leftrightarrow \varepsilon \sim N(0, \sigma^2)$

## REGRESSIONE MULTIPLA

Riguardo la **regressione multipla**, si può dire che sia la stessa cosa ma con più variabili esplicative:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon$$

$$\text{dove } E(y/x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

la cui equazione stimata con OLS:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_k x_k$$

Residui:

$$\hat{e}_i = y_i - \hat{y}_i$$

Possiamo verificare le assunzioni di base tramite analisi dei residui:

- Normalità;
- Omoschedasticità;
- Assenza di (auto)correlazione;
- Linearità;
- Presenza di outliers.

Con la retta lo scopo era di trovare la linea che meglio si adattava ai dati; ora lo scopo è trovare il piano che meglio si adatta ai dati. Il concetto rimane lo stesso, si vuole comunque sempre minimizzare la distanza tra l'osservazione e il modello (che può essere retta o piano).

## LE GRANDEZZE DELLA REGRESSIONE

Dopo aver stimato una regressione lineare, ci si può chiedere quanto bene la retta di regressione stimata descriva i dati. Il regressore cattura molto o poco della variazione nella variabile dipendente? Le osservazioni si concentrano intorno alla retta di regressione, o sono disperse?

La Somma dei quadrati di regressione è la somma delle deviazioni quadratiche dei valori predetti di  $y_i$  e  $\hat{y}_i$  dalla loro media. Quanto si avvicina realmente il modello rispetto al più semplice che ci possa essere (la media)? SSR misura la variabilità della mia stima rispetto alla media; quindi, dice quanta variabilità il modello riesce a spiegare.

$$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = SSR =$$

**= Somma dei quadrati di regressione o Somma dei quadrati dei residui**

$$\rightarrow df = k$$

(deve corrispondere ad un valore alto così da spiegare al meglio la variabilità della dipendente)

SSE la si può interpretare come la variabilità del modello che non si riesce a spiegare perché è dovuta al fatto che la stima non può essere perfetta. Potrebbe essere pari a 0 quando i dati formano già una retta di per sé;

quindi, quando non ci sono distanze da minimizzare. È la somma delle deviazioni quadratiche degli  $y_i$  dalla loro media:

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = SSE =$$

**= Somma dei quadrati degli errori o Somma dei quadrati spiegati**

**→  $df = n - k - 1$**

(deve essere il più piccolo possibile per trovare la retta migliore che rende minore la distanza tra la retta e la realtà)

La somma degli errori è:

$$SSR + SSE = SST = \text{Totale variabilità} = \text{Somma totale dei quadrati}$$

L' $R^2$  misura quanto bene la retta di regressione OLS si adatta ai dati. Varia tra 0 e 1 e misura la frazione della varianza di  $y_i$  che è spiegata da  $x_i$  ed è nota anche come coefficiente di determinazione, consente di calcolare la bontà di adattamento ai dati della mia regressione mettendo a rapporto la somma dei quadrati di regressione con la somma dei quadrati degli errori:

$$R^2 = \frac{SSR}{SST} = \frac{\text{Variabilità che spiego}}{\text{variabilità totale}} = 1 - \frac{SSE}{SST} = 1 - \frac{\text{Variabilità che non spiego}}{\text{variabilità totale}}$$

Perciò si ottiene la variabilità che il modello spiega. Ha un difetto però: dal punto di vista matematico tende ad aumentare quando si aggiungono al modello più variabili esplicative. Questo vuole dire che quando si confrontano due modelli con la stessa  $Y$ , se si usa  $R^2$  si va nella maggior parte dei casi a preferire il modello più grosso; però non sempre è vero che quello più grande sia quello migliore perché potrebbe esserci una variabile che non c'entra nulla ed  $R^2$  crescerebbe lo stesso, perciò,  $R^2$  non va bene quando si vogliono comparare modelli con  $n$  di variabili diverse. Nella regressione multipla, cresce ogni volta che si aggiunge un regressore, a meno che il coefficiente del regressore aggiunto sia esattamente pari a zero.

Poiché sono stimati due coefficienti ( $\beta_0, \beta_1$ ), si parla di correzione per i “gradi di libertà” essendo che due “gradi di libertà” dei dati sono perduti. I gradi di libertà sono valori che possono variare per ottenere comunque una stima puntuale. I gradi di libertà di una statistica sono il numero minimo di dati sufficienti a valutare la quantità d'informazione contenuta nella statistica. I gradi di libertà (gdl, in inglese degree of freedom = df) sono dei numeri, in genere interi positivi, che si utilizzano per poter verificare delle ipotesi sulla popolazione da cui è stato estratto un campione.

Se un campione contiene  $n$  osservazioni, possiamo dire che è costituito da  $n$  pezzi individuali di informazioni. I gradi di libertà corrispondono al numero di informazioni indipendenti che sono libere di variare nel calcolo di una determinata stima di un parametro: quante informazioni indipendenti possono variare affinché io ottenga sempre la stessa stima?

Poiché l' $R^2$  aumenta aggiungendo una nuova variabile, un aumento di questo non significa che aggiungere una variabile migliori realmente l'adattamento del modello. In questo senso, l' $R^2$  fornisce una stima in eccesso della bontà della regressione. Un modo per correggere questo effetto è quello di deflazionare o ridurre l' $R^2$ , e questo è proprio quello che fa l' $R^2$  corretto o  $R^2_{adj}$ . Ha la stessa interpretazione di  $R^2$  ma non aumenta necessariamente quando si aggiunge un nuovo regressore.

La sua formula è:

$$R_{adj}^2 = 1 - \frac{n-1}{n-k-1} (1 - R^2)$$

Ciò significa che l' $R^2$  corretto è pari a 1 meno il rapporto tra la varianza campionaria dei residui OLS. Ci sono 3 risultati utili da conoscere riguardo questo:

- $\frac{n-1}{n-k-1}$  è sempre maggiore di 1, perciò  $R_{adj}^2$  è sempre minore di  $R^2$ ;
- L'aggiunta di un regressore ha due effetti opposti sull' $R_{adj}^2$ . Da un lato l'SSR decresce, il che fa aumentare l' $R_{adj}^2$ . Dall'altro, il fattore  $\frac{n-1}{n-k-1}$  aumenta. L'aumento o diminuzione dell' $R_{adj}^2$  dipende da quale dei due effetti è più forte;
- L' $R_{adj}^2$  può essere negativo. Questo accade quando i regressori, presi nel complesso, riducono la somma dei quadrati dei residui di un ammontare così piccolo da non bilanciare il fattore  $\frac{n-1}{n-k-1}$ .

Ad esempio:  $R_{adj}^2 = 0,6 = 60\%$  significa che il modello di regressione riesce a spiegare il 60% della variabilità della dipendente, di Y. Generalmente il valore è  $0 \leq R_{adj}^2 \leq 1$ ; tuttavia vi sono casi matematici particolari in cui potrebbe avere valori negativi.

Più  $\frac{n-1}{n-k-1}$  aumenta, più aumenta la penalità. Man mano che si aggiungono variabili, il denominatore si riduce e la penalizzazione diventa più pesante. Funziona, appunto, come fattore di penalizzazione, cioè va a penalizzare i modelli sui cui si hanno più osservazioni o su cui si usano più variabili esplicative.

Quindi, affinché l' $R_{adj}^2$  si mantenga sui valori dell' $R^2$ , è necessario con più osservazioni e più variabili diminuire la percentuale di variabilità non spiegata poiché se è molto alta, l' $R_{adj}^2$  diventerà sempre più basso e così facendo si correggerebbe il problema dell' $R^2$  che cresceva al crescere di n e k.

Quando si devono paragonare modelli di regressione con n variabili diverse, bisogna usare l' $R_{adj}^2$ . Il confronto tra due modelli di regressione ha senso se la variabile dipendente (ad esempio Y) è la stessa, altrimenti non ha senso il confronto.

Ad esempio:

- $y = \beta_0 + \beta_1x + \beta_2z + \varepsilon \rightarrow R_{adj}^2 = 0,5 \rightarrow$  è il modello migliore perché ha valore più vicino ad 1 e spiega meglio la variabile y;
- $y = \beta_0 + \beta_1x + \varepsilon \rightarrow R_{adj}^2 = 0,2$ ;
- $Z = \beta_0 + \beta_1x + \varepsilon \rightarrow R_{adj}^2 = 0,7 \rightarrow$  non può essere confrontato con Y in quanto Z è un'altra misura.

## I TEST STATISTICI PER VALUTARE UN MODELLO DI REGRESSIONE

### Test F

Per valutare al meglio un modello di regressione occorre generalizzare i risultati ottenuti su un campione all'intera popolazione, questo procedimento viene detto Inferenza.

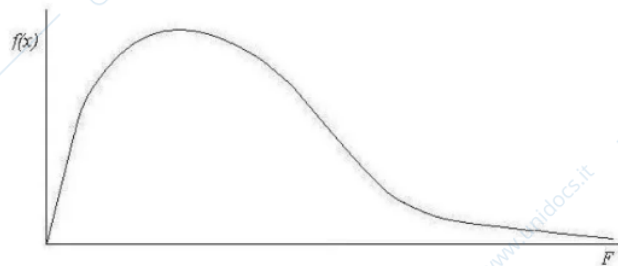
Ad esempio, tenendo in considerazione un sondaggio fatto in università, il professore detiene un campione del dipartimento di management ma deve capire se sia possibile generalizzare i risultati trovati in quel dipartimento a tutta l'università. Per fare questo esiste un processo, chiamato "di inferenza statistica" che permette di attuare questo procedimento sulla base di distribuzioni statistiche e regole decisionali.

Il test statistico funziona così:

- $H_0: \theta = 0 \rightarrow$  ipotesi nulla: è l'ipotesi di partenza, è lo stato del mondo che assumo essere vero fino a prova contraria;
- $H_1: \theta \neq 0 \rightarrow$  corrisponde all'ipotesi opposta ad  $H_0$ .
- $t$ : statistica test calcolata sui dati o campioni per capire quale ipotesi sia più probabile, è una variabile casuale;
- $t \sim D \rightarrow$  la variabile si distribuisce secondo una variabile  $D$ .

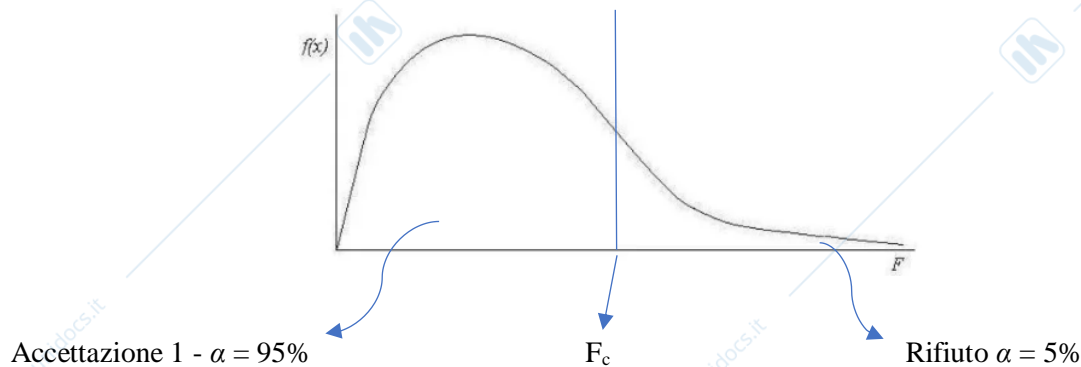
Viste le regole di base, occorre testare questi dati attraverso risultati del campione. Lo statistico, per decidere se rifiutare o meno  $H_0$ , va a costruire una formula detta "**F**" (**di Fisher**) e ripete questo procedimento per diversi campioni capendo, poi, come questa  $F$  si distribuisce. Questo test ha a che fare con la regressione multipla, cioè con più variabili esplicative.

La distribuzione  $F$  è, così, rappresentata:



A questo punto, lo statistico, sceglie un certo valore di errore che è disposto a tollerare. Essendo che si sta lavorando con dei campioni e non si ha visione della popolazione, non si sa realmente se la risposta che verrà data sarà corretta o no, ci sarà per cui un margine di errore. Questo margine di errore è  $\alpha$  ed è detto **errore del primo tipo** ed è anche detto **livello di significatività**. Generalmente è il 5%. La scelta del margine di errore dipende dal tipo di problema con cui si sta lavorando. Ad esempio, se un medico testa un vaccino molto probabilmente è disposto ad accettare un margine di errore più basso possibile. Perché, allora, non imporre  $\alpha = 0$ ? Perché settare questo errore a 0 fa alzare un altro errore detto **errore del secondo tipo**.

Sulla base di questo errore lo statistico sa che osserverà un dato valore di  $F$  che si chiama  $F_c$ . Una volta impostato  $\alpha$  la distribuzione di  $F$  viene divisa in due regioni: una regione chiamata di **accettazione** che vale  $1 - \alpha$  e che in questo caso sarà al 95% e un'altra detta regione di rifiuto e che vale 5%. Quando si va a calcolare  $F$  sui dati, si sa che essa può assumere qualunque valore sulla linea e, inoltre, si sa che il valore che definisce la regione di rifiuto è  $F_c$ , che è un valore teorico.



Ora con i dati in possesso si può calcolare  $F$ , che è il valore empirico calcolato sui dati. Se  $F$  cade nella regione di accettazione allora si ha abbastanza evidenza per dire che ci si trova nella regione di accettazione e quindi non si può rifiutare  $H_0$ . Se, invece, si trova nella regione di rifiuto, si ha abbastanza evidenza per rifiutare  $H_0$ .

Perciò:

- Se  $F > F_c \rightarrow$  si rifiuta  $H_0$ , in questo caso si dice che il modello di regressione è statisticamente significativo e valido e lo si può usare per fare predizione perché almeno un coefficiente è diverso da 0;
- Se  $F < F_c \rightarrow$  si accetta  $H_0$ ;
- Se  $F = F_c \rightarrow$  è controverso perché si è sulla linea di rifiuto.

Ad esempio:

- $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0 \rightarrow H_0$ : modello di regressione non valido, infatti in caso di regressione multipla  $Y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + e$  se  $\beta_1 = \dots = \beta_k = 0$  allora tutte le variabili esplicative sparirebbero dal modello e perciò, a cosa servirebbe il modello?
- $H_1: B_j$  diverso 0 per almeno un  $j = 1, \dots, k \rightarrow H_1$ : il modello è valido, cioè si può usare per fare predizioni perché almeno una variabile esplicativa aiuta a spiegare la  $Y$ .

È necessaria una regola per capire quando e quando non rifiutare  $H_0$ . Perciò, in questo caso, la formula è:

$$F = \frac{SSR / k}{SSE / [n - (k + 1)]} \sim F_{k, n-k-1} = F_c$$

Questa è la statistica test empirica, cioè quella che si calcola sulla base dei dati. Gli statistici hanno calcolato questa statistica molte volte ed hanno scoperto che si distribuisce secondo la distribuzione  $F$  di Fisher che è definita da due parametri  $k$  (1° grado di libertà) e  $n-k-1$  (2° grado di libertà).

## P value

Per semplificare il procedimento ed evitare di calcolare la formula tutte le volte, calcolare  $F_c$  ed andare a vedere se  $F > F_c$ , gli statistici hanno introdotto una regola più semplice: se  $p\text{-value} \leq \alpha$  si rifiuta  $H_0$ .

Il **p-value** è una probabilità e corrisponde all'area alla destra di  $F$ , perciò non può essere negativo.



## Intuizione dietro la formula test F

$$F = \frac{SSR/k}{SSE/[n - (k + 1)]} \approx \frac{SSR}{SSE}$$

A numeratore si ha la variabilità spiegata e a denominatore si ha la variabilità che non si riesce a spiegare; perciò, se si aumenta il numeratore la frazione aumenta. Quindi, se cresce la percentuale di variabilità si riesce a spiegare la  $F$  che diventa più grande e ci si muove perciò più a destra verso l'area di rifiuto (dato che sono state raccolte più evidenze a favore del fatto che il modello di regressione è valido e funziona). Se aumenta il denominatore la frazione diminuisce, infatti aumenta la variabilità che non si riesce a spiegare perciò la variabilità si sposta verso sinistra più lontano dall'area di rifiuto (perché il modello è meno utile).

## Limiti del test F

- 1) L'ipotesi alternativa è che almeno un beta sia diverso da 0. Quale informazione si perde? Non si sa quanti e quali, cioè non dice quali siano i coefficienti significativi (è un test omnibus, cioè che testa più ipotesi contemporaneamente). Per sapere quali sono i coefficienti significativi si può usare il test  $t$ ;
- 2) Questo test dice se il modello è valido, ma non dice se è il migliore. Per trovare il modello migliore si può usare  $R_{adj}^2$ .

## Test t

L'obiettivo del test t è andare a vedere coefficiente per coefficiente quali sono quelli diversi da 0. La formula è:

$$t = \frac{\hat{\beta}_j - 0}{se_{\hat{\beta}_j}} \sim t_{n-k-1}$$

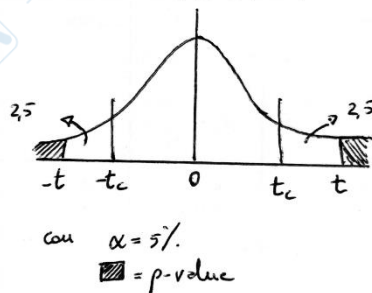
Le ipotesi alla base sono:

- $H_0: \beta_j = 0$  con  $j=0, 1, \dots, k$  compresa l'intercetta;
- $H_0: \beta_j \neq 0$ ;

Si basa su:

- Stima del coefficiente di regressione: ad esempio, si immagini di aver stimato la retta o il piano dei minimi quadrati e si trova che  $\hat{\beta}_2 = 5$ ;
- Standard error della stima: cioè il grado di incertezza sulla stima fatta; più è alto lo standard error e più si è incerti sulla stima;
- $n-k-1$ : grado di libertà.

La distribuzione seguita si chiama **t di Student**: è simmetrica rispetto a 0: la differenza rispetto al test F è che in questo, siccome la distribuzione ha due code, la statistica critica si può trovare sia a destra che a sinistra e perciò l'area di rifiuto si divide nelle due code.



In questo caso:

- Si rifiuta  $H_0$  se  $p\text{-value} \leq \alpha$ ;
- Si rifiuta  $H_0$  se  $|t| \geq |t_c|$ .

Ad esempio:

- $t = 3,5$       $|t_c| = 2$       $\rightarrow$  si rifiuta  $H_0$ ;
- $t = -9$       $|t_c| = 2$       $\rightarrow$  si rifiuta  $H_0$  perché  $|-9| = 9 > 2$ ;

Quando  $\alpha$  è al 5% ed il campione ha una numerosità  $(n) > 30$ , il valore critico di t è circa uguale a 2.

Nella formula si trova la stima al numeratore e l'incertezza al denominatore. Se aumenta il numeratore, t cresce in valore assoluto e sarà più probabile che si trovi in regione di rifiuto. Se, invece, aumenta il denominatore, con più probabilità ci si trova nella regione di accettazione perché si è meno sicuri di cosa si sia calcolato e perciò t si sposta verso lo 0.

**Reset (Ramsey's test)**

Ipotesi:

- $H_0$ : modello originario adeguato;
- $H_1$ : modello originario non adeguato.

Si rifiuta  $H_0$  se  $p\text{-value} \leq \alpha$ **Jarque-Bera test**

Ipotesi:

- $H_0$ : campione estratto da una popolazione di dati distribuiti come una variabile casuale normale;
- $H_1$ : non-normalità.

Rifiuta  $H_0$  se  $p\text{-value} \leq \alpha$ .**INTRODUZIONE ALLA MACROECONOMETRIA**

**Econometria:** L'uso di metodi quantitativi (principalmente matematico-statistici) per analizzare problemi economici, o più in generale problemi nelle scienze sociali.

**Macroeconometria:** Un'applicazione dell'econometria per lo studio dei dati e dei modelli macroeconomici.

**SERIE STORICHE:**

**Definizione:** un insieme di osservazioni su un certo fenomeno ordinate nel tempo.

Una serie storica è una variabile osservata in diversi istanti temporali. Gli istanti temporali possono essere giorni, mesi, anni, ore, secondi, tic data (è intervallo più piccolo che c'è nel trading).

Non è necessario che le serie storiche siano equidistanziate, cioè non è necessario che io osservi  $Y$  ogni mese, possono essere campionate a frequenze irregolari quindi un'osservazione un mese e un'altra due mesi dopo. Noi assumiamo che le serie storiche siano equidistanziate, cioè che la  $Y$  è osservata a intervalli di tempo regolari (istanti di tempo equispaziati). Useremo anche il termine 'processo stocastico' o semplicemente 'processo'.

- **Notazione generica:**  $\{y_t; t = 1, 2, \dots, n\}$

Esempio di  $Y$  misurata annualmente:  $Y_{1999}, Y_{2000}, \dots, Y_{2022}$ .

La prima cosa da fare quando ho una serie storica è fare il grafico detto "time series plot", cioè il grafico della serie storica, in cui nelle  $X$  ho il tempo e sulle  $Y$  osservo la variabile. Una serie storica, quindi, sarà una serie di osservazioni nei diversi istanti temporali, ad esempio:

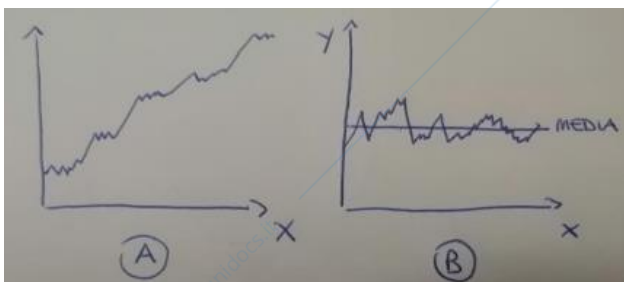


Grafico A): non stazionario

Grafico B): stazionario

A fini econometrici è fondamentale, prima di applicare qualsiasi tipo di modello, verificare che la nostra serie storica presenti la proprietà di **stazionarietà**.

Una serie storica  $y_t$  è **stazionaria** se:

- $E(y_t) = \mu \rightarrow$  Media costante
- $Var(y_t) = \sigma^2 \rightarrow$  Varianza costante
- $Cov(y_t, y_{t+s}) = Cov(y_t, y_{t-s}) = \gamma_s \rightarrow$  la covarianza tra due osservazioni della serie storica dipende solamente dall'intervallo che separa le due osservazioni e **non** dalle date in cui sono osservate le due variabili.

Dove  $\mu \in \mathbf{R}$ ,  $\sigma^2 \in (0, \infty)$ , e la terza indica che la covarianza dipende da  $s$  e non da  $t$ , cioè dall'ampiezza (distanza) che separa  $t$  e  $t \pm s$ , e non dalle date in cui osservo la variabile.

**Serie stazionaria:** vuol dire che la serie storica **ritorna sempre alla sua media**. Cioè io ho un valore medio della  $Y$ , se osservo che nel tempo la mia  $Y$  torna a quel valore medio, allora sarà stazionaria  $\rightarrow$  fluttua in maniera stazionaria attorno alla sua media. La serie nel grafico **A** non soddisfa questa proprietà perché si allontana sempre di più dalla media  $\rightarrow$  è una serie non stazionaria (le variabili macroeconomiche sono quasi tutte non stazionarie: variazione del pil, serie storica dei prezzi, debito pubblico). La serie del grafico **B** è una serie **stazionaria** (tassi di interesse, rendimenti asset, variazione PIL).

È importante che venga soddisfatta questa proprietà perché in questo modo è possibile stimare i momenti della serie. Se io sono interessato a modellizzare la media di  $Y$ , per poterla modellizzare, questa media deve essere definita, se la serie fluttua verso l'infinito non ho una media fissa e quindi non è possibile stimarla.

**Autocovarianza:**

**Funzione di autocovarianza:** La funzione di **autocovarianza** al ritardo  $k$ ,  $\gamma_k$ , di un processo stazionario  $\{X_t\}$  di **media** costante  $\mu$  è definita come la **covarianza** tra le variabili casuali  $X_t$  e  $X_{t-k}$ , al variare di  $k = 0, 1, 2, \dots$ , ossia

$$\bullet \quad \gamma_k = Cov(X_t, X_{t-k}) = E[(X_t - \mu)(X_{t-k} - \mu)]$$

dove:

- $X_t$  = valore oggi
- $X_{t-k}$  = valore ieri;

Ovviamente si ha che:

$$\bullet \quad \gamma_0 = Cov(X_t, X_t) = Var(X_t) = E[(X_t - \mu)^2]$$

**Autocorrelazione:**

**Funzione di autocorrelazione globale:** La funzione di **autocorrelazione** globale al ritardo  $k$ ,  $\rho_k$  di un processo stazionario  $\{X_t\}$  di media costante  $\mu$  è definita come il coefficiente di correlazione lineare tra le variabili casuali  $X_t$  e  $X_{t-k}$  al variare di  $k = 0, 1, 2, \dots$ , ossia:

$$\rho_k = Corr(X_t, X_{t-k}) = \frac{\gamma_k}{\gamma_0}$$

Notiamo che:

$$\bullet \quad \rho_0 = Corr(X_t, X_t) = \frac{\gamma_0}{\gamma_0} = 1$$

dove:

- $\gamma_0 = \text{varianza di } X_t$

Data una variabile (serie storica), si considerano una serie di ritardi della serie stessa (nr. di periodi passati) e si va a vedere per ogni ritardo osservato se c'è una correlazione tra l'osservazione odierna e l'osservazione in quel periodo passato.

Bisogna analizzare il valore del **pvalue** che, se  $<$  del 5%, dovremo rifiutare l'ipotesi nulla ed avremo evidenza del fatto che per quel lag, presi insieme i coefficienti fino a quel momento, questi sono statisticamente significativi (coefficienti diversi da zero<sup>39</sup>) e quindi vi è correlazione; viceversa, se  $>$  al 5%, si accetta l'ipotesi nulla ed i coefficienti diventano statisticamente NON significativi (ovvero coefficiente uguale a zero<sup>40</sup>).

### Funzione di autocorrelazione parziale:

Problema di  **$\rho_k$** : se faccio una misura di dati del 2006 e del 2020,  **$\rho_k$**  non tiene conto di cosa è successo all'interno del periodo. Il **coefficiente di autocorrelazione parziale** tiene conto di tutto il periodo  $\rightarrow$  togliamo da  **$Z_t$**  le informazioni parziali fino al tempo **t**. autocorrelazione parziale

La **funzione di autocorrelazione parziale** al ritardo **k**,  $\phi_{kk}$ , di un processo stazionario  $\{X_t\}$  di media costante  $\mu$ , è definita come la **correlazione lineare** tra le variabili casuali  $X_t$  e  $X_{t-k}$ , al netto di tutte le correlazioni lineari intermedie, ossia quelle tra  $X_t$  e  $X_{t-1}$ , tra  $X_t$  e  $X_{t-2}$ , e così via fino alla correlazione lineare tra  $X_t$  e  $X_{t-k+1}$ .

Ponendo  $\{Z_t\} = \{X_t - \mu\}$ , matematicamente:

- $\phi_{kk} = \text{Corr}(Z_t - E(Z_t | Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}))$
- $Z_{t-k} - E(Z_{t-k} | Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1})$

### IL PROCESSO AR (1):

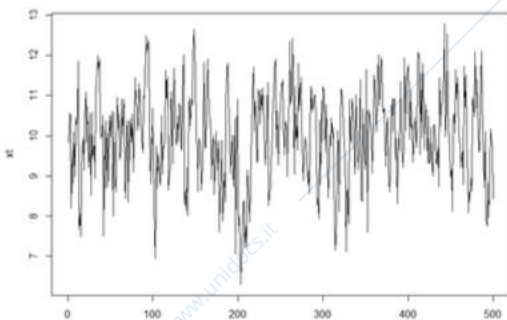
Il processo AR (1) è un esempio di **processo stazionario**: processo autoregressivo di ordine 1.

Si chiama processo autoregressivo perché stiamo cercando di spiegare la Y usando l'osservazione precedente. Quindi io sto postulando che quello che osservo oggi dipende solo da ieri e da una componente di errore che è imprevedibile e generata casualmente e su cui, quindi, non posso dire nulla. Affinchè sia un processo autoregressivo di ordine 1 è necessario che il coefficiente **rho** in valore assoluto sia **minore** di 1 e che l'errore sia imprevedibile, quindi che sia estratto in maniera casuale e indipendente da una distribuzione normale con media zero e varianza  $\sigma^2$ . Se rho non è  $<1$  allora la serie **non è stazionaria**.

$$AR(1) \rightarrow Y_t = \rho Y_{t-1} + V_t$$

Dove:

- $|\rho| < 1$  e  $v_t \sim iid \sim N(0, \sigma^2_v)$  (iid=identicamente e indipendentemente distribuiti)



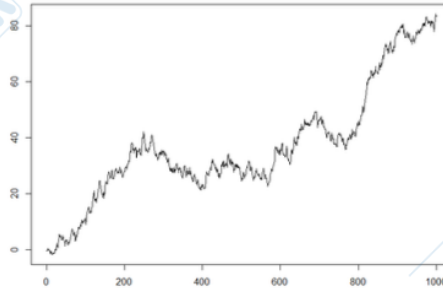
Nel tempo la serie torna al suo valore, è limitata in un range ben definito.

## IL PROCESSO RANDOM WALK:

Il processo random walk è un esempio **processo non stazionario** (random walk)

$$\text{Random walk} \rightarrow y_t = y_{t-1} + v_t$$

Dove:  $V_t \sim iid \sim N(0, \sigma^2_v)$



Identico ad AR(1) ma  $\rho$  è assunto = 1 ) è una serie storica non stazionaria.

La random walk è l'equazione base dei modelli, se la scrivo:  $y_t = y_{t-1} + v_t$  → non posso prevedere il prezzo delle azioni. È l'assunzione base di tutti i modelli sui prezzi delle azioni. Cioè il prezzo dell'azione al tempo  $t$  è dato da prezzo azione a tempo  $t-1$  + termine d'errore. Questa è l'ipotesi che riteniamo vera se crediamo che i mercati siano efficienti → l'assunzione che i mercati siano efficienti è l'assunzione secondo cui tutte le informazioni disponibili sul mercato finanziario sono incorporate nei prezzi e quindi non posso usare nessuna info per predire i movimenti futuri dei prezzi e fare profitti superiori agli altri operatori di mercato perché tutti abbiamo accesso alla stessa informazione. (Questa è la ragione per cui esiste la norma sul divieto di insider trading).

### Differenziazione:

Un modo per avere una serie storica stazionaria consiste nel calcolare le differenze tra osservazioni consecutive.

Esempio: **differenziamo il random walk:**

$$Y_t - Y_{t-1} = Y_{t-1} - Y_{t-1} + v_t$$

se porto  $y_{t-1}$  a sinistra trasformo la serie storica in  $\Delta y = v_t$  che è una serie stazionaria

$$\Delta y_t = v_t$$

Dove  $v_t$  è stazionario → quindi  $y_t$  si dirà integrata di ordine 1, o  $\mathcal{L}(1)$ , mentre  $v_t$  si dirà integrata di ordine 0, o  $\mathcal{L}(0)$ .

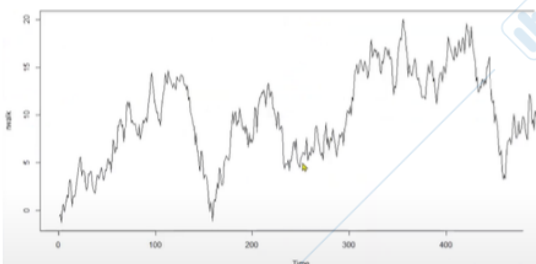


Grafico prima di applicare la differenziazione (ovvero grafico senza il ritorno alla media).

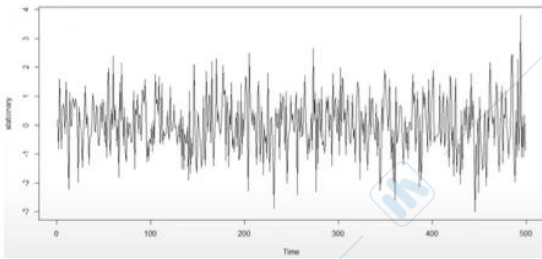


Grafico dopo aver applicato la differenziazione (ovvero grafico con ritorno alla media).

## DICKEY-FULLER TEST

In statistica non si fa nulla a occhio, non posso dire se torna alla media o meno solo guardando il grafico. Mi serve un test per vedere se la serie è stazionaria o meno, uno è il test di **dickey-fuller**.

Noi sappiamo che un processo AR(1) è stazionario se  $|\rho| < 1$  ma se  $\rho = 1$  allora è non stazionaria perché è random walk. Quindi il test di dickey-fuller mi dirà quale delle due ipotesi è più plausibile.

Prendo l'equazione del processo generale:

- $Y = \rho y_{t-1} + v_t$

Dove:

- se  $|\rho| < 1 \rightarrow y_t$  stazionaria
- se  $\rho = 1 \rightarrow y_t$  NON stazionaria

differenzio:

- $y_t - y_{t-1} = \rho y_{t-1} - y_{t-1} + v_t$
- $\Delta y_t = (\rho - 1)y_{t-1} + v_t$
- $\Delta y_t = \gamma y_{t-1} + v_t$

Quest'equazione è simile a un modello di regressione semplice:

**v.dipendente = coefficiente angolare \* variabilità esplicativa + termine errore causale**

Ora posso fare il test: **ipotesi**:

- $H_0: \rho = 1 \leftrightarrow H_0: \gamma = 0$  **non** stazionarietà
- $H_0: \rho < 1 \leftrightarrow H_0: \gamma < 0$  stazionarietà

La regola decisionale è sempre basata sul p-value

Testo le due ipotesi con il solito test del **p-value**. Esempio:

- **p-value = 0.02 e  $\alpha = 5\%$   $\rightarrow$  serie è stazionaria  $\rightarrow$  perchè **p-value  $<$   $\alpha$  rifiuta  $H_0$ .****

In questo test voglio rifiutare  $H_0$  così la serie è stazionaria.

Se il p-value è **superiore** al **5%**, allora si può concludere che **non** si rifiuta l'ipotesi nulla che in un test di Dickey-Fuller corrisponde alla **non stazionarietà** del processo, quindi il processo sarà **non** stazionario. Viceversa, se assumiamo alfa pari al 5% e se il p-value è **inferiore al 5%**, allora si può concludere che si rifiuta l'ipotesi nulla che in un test di Dickey-Fuller corrisponde alla **non stazionarietà** del processo, quindi il processo sarà **stazionario**!!

## TEST DI LJUNG-BOX

Il test di Ljung-Box è un tipo di test statistico per verificare se un qualunque gruppo di autocorrelazioni di una serie storica (congiuntamente) sono diverse da zero. Invece di testare la casualità per ogni singolo ritardo, il test analizza la casualità complessiva sulla base di un certo numero di ritardi.

In altre parole, non siamo interessati alla significatività isolata di  $\rho_k$ , ma alla significatività all'interno di un gruppo di autocorrelazioni  $\rho_1, \rho_2, \dots, \rho_k$ .

## APPLICAZIONE: RENDIMENTI FINANZIARI

Sia  $P_t$  la serie storica del 'prezzo' di variabile finanziaria (azione, indice, opzione, ecc.). Si definisce la serie dei rendimenti (returns) finanziari come segue:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \rightarrow \text{tasso di crescita prezzo} = \text{rendimento azione}$$

mentre definiamo i log-rendimenti (log-returns) come segue:

$$\log R_t = \log\left(\frac{P_t}{P_{t-1}}\right) \rightarrow \text{rendimento logaritmico}$$

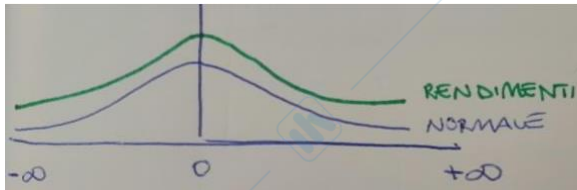
## FATTI STILIZZATI DELLE SERIE STORICHE FINANZIARIE:

Sono evidenze empiriche che sono costanti per diversi paesi, diversi indici e diversi periodi di tempo. Noi vedremo 6 tipi di fatti stilizzati.

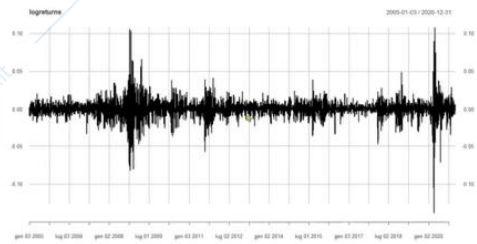
- 1)  **$P_t$  non stazionaria** (serie dei prezzi azionari)  $\rightarrow$  random walk;
- 2)  **$R_t$  e  $\log R_t$  stazionarie**: il rendimento comunque calcolato è stazionario. I rendimenti sono processi stazionari, quindi, è meglio lavorare sui rendimenti e **non** sui prezzi, poiché le proprietà statistiche dei rendimenti sono migliori di quelle dei prezzi;
- 3)  **$R_t$  e  $\log R_t$  esibiscono distribuzioni con code grasse**  $\rightarrow$  i rendimenti hanno le code grasse.

Il rendimento di un'azione è una variabile che prende valore da  $-\infty$  a  $+\infty$ . Generalmente la media dei rendimenti è centrata sullo 0, ma i rendimenti non si distribuiscono come una normale (cioè simmetrica rispetto allo 0 e con code piatte). Code piatte vuol dire che la probabilità di beccare un evento estremo è molto bassa. La distribuzione dei rendimenti finanziari invece ha code più alte di quelle della normale. Questo vuole dire che per i rendimenti finanziari c'è maggiore probabilità di osservare eventi estremi. quindi non solo posso fare grosse perdite ma posso fare anche grossi profitti.

La definizione di code grasse è sempre in riferimento alla distribuzione normale. quindi noi useremo la curtosi per verificare queste ipotesi, cioè per vedere se la distribuzione dei rendimenti finanziari ha code più grasse della normale.



- 4) **Volatility clustering:** fatto che la variabilità della serie storica dei rendimenti tende a seguire delle fluttuazioni molto simili l'una all'altra in periodi di tempo limitati.  
Ex. Se osservo serie storica di un rendimento azionario io vedo che fluttua attorno allo 0 e periodi in cui fluttua molto moderatamente sono seguiti da momenti dove fluttua eccessivamente e così via. Cioè la volatilità si agglomera, ho periodi dove c'è bassa volatilità e periodi dove c'è alta variabilità. Quindi, se ho osservato bassa volatilità è possibile che anche prima abbia osservato bassa volatilità e idem per l'alta, perché la volatilità tende ad agglomerarsi. Quando il mercato è al ribasso la variabilità è più alta, quando il mercato è al rialzo la variabilità è più bassa;



- 5) **Price overreaction:** i prezzi delle azioni tendono a reagire in maniera eccessiva alle notizie. Soprattutto quando si verificano cattive notizie il prezzo dell'azione può crollare molto di più di quello che sarebbe opportuno aspettarsi. Questo perché quando c'è cattiva notizia, gli uomini tendono tutti a vendere le azioni e questo fa crollare i prezzi.



Ex. reazione al covid è stata molto marcata, invece, con le buone notizie c'è sempre stata una over reaction ma più contenuta;

- 6) **Presenza di molti outliers:** presenza di molti valori estremi (questo è in relazione al fatto stilizzato 3).

## MODELLI ARMA (p, q)

Arma model è un modello di base per cercare di modellizzare la dinamica di Y in serie storica. Modello arma si chiama così perché è composto da una parte **AR**, quindi una parte a modello auto-regressive e una parte **MA** (media mobile - moving average), cioè a modello a media mobile. Quindi il modello arma è la combinazione di due modelli diversi.

Forma funzionale generale modello ARMA: equazione generale (con drift):

$$ARMA(p, q)Y = c + \sum_{i=1}^p \phi_i y_{t-1} + \sum_{i=1}^q \theta_i \epsilon_{t-1} + \epsilon_t$$

dove  $\epsilon_t \sim iid \sim N(0, \sigma^2_v)$

questo modello ha una componente AR di ordine temporale p e una componente MA di ordine q.

Y = variabile dipendente

Esempio di ARMA (3,2) con drift:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \epsilon_t$$

C = DRIFT che è un intercetta

La scelta del modello migliore può essere fatta usando i criteri di informazione quali AIC o BIC.

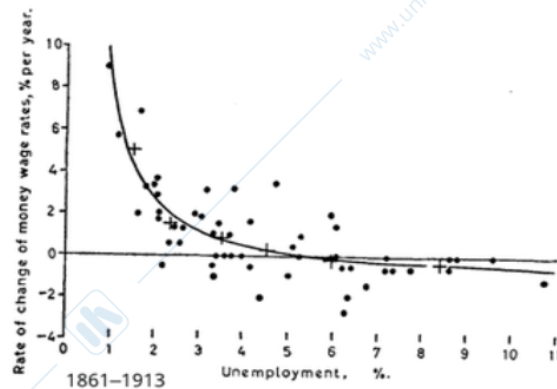
### Introduzione alla regressione semplice: la Curva di Phillips

la curva di Phillips è la relazione inversa tra il tasso di inflazione e il tasso di disoccupazione.

questa però non è l'originale curva di Philips. L'originale curva di Philips considera la relazione tra tasso di disoccupazione e tasso di variazione dei salari nominali.

se assumo che l'inflazione viene spinta dai costi di produzione del produttore, tra cui anche i costi salariali, allora la crescita del tasso nominale del salario ha un ruolo nello spiegare l'inflazione. quindi posso passare da una curva all'altra.

Il paper originale di phillips è del 1958. Va a raccogliere dati per il regno unito dal 1861 al 1957 e dopo fa un grafico per periodi. Il primo periodo considerato è 1861-1913 → scatter plot.



Phillips nota che c'è relazione inversa tra tasso di disoccupazione e tasso di variazione dei salari nominali. Quindi quando il tasso di disoccupazione è alto la dinamica di crescita dei salari nominali è bassa e viceversa. Phillips nota che la relazione tra queste due variabili non è lineare ma è curvilinea.

Phillips va a crearsi un modello di regressione semplice per spiegare il tasso di crescita dei salari in funzione del tasso di disoccupazione.

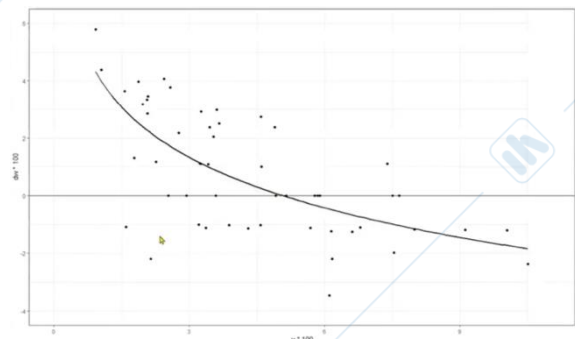
lui per modellizzare la non-linearità usa delle funzioni logaritmiche.

I dati della curva sono: data("PhillipsCurve"):

**BASSA DISOCCUPAZIONE** → forza contrattuale dei lavoratori più alta e potranno quindi chiedere salari più elevati

**ALTA DISOCCUPAZIONE** → forza contrattuale dei lavoratori più bassa

**NAIRU** → tasso di disoccupazione che NON accelera l'inflazione (Not-Accelerating Inflation Rate of Unemployment) rappresenta in questo caso l'intersezione tra l'asse orizzontale e la curva



Tasso di crescita dei salari nominali sarà uguale a intercetta ( $\beta_0$ ) + coefficiente angolare ( $\beta_1$ ) che sarà relazionato al tasso di disoccupazione + errore NON stimabile.

Se curva di Phillips è corretta  $\beta_1$  dovrebbe essere minore di zero per trovare relazione inversa tra le due variabili.