UNIVERSITÀ DEGLI STUDI DI TRIESTE

Dipartimento di Ingegneria e Architettura

Corso di misure meccaniche, termiche e collaudi

Prof. Rodolfo Taccani
Prof. Lucia Parussini
Prof. Marco Bogar

a.a.2023-2024

Outline

- Serie storiche

Serie temporali

Che cos'è una serie temporale?

Una **serie temporale** (o **serie storica**) è una serie di punti di dati indicizzati in ordine temporale.
Nel caso in cui, siano rilevate k variabili in n istanti di tempo, i dati prendono quindi la forma

	variabili rilevate		
tempo	Y_1	$\cdots$	Y_k
$\boldsymbol{t_1}$	y_{11}	$\cdots$	y_{k1}
$\boldsymbol{t_2}$	y_{12}	$\cdots$	y_{k2}
$\vdots$	$\vdots$		$\vdots$
$\boldsymbol{t_n}$	y_{1n}	$\cdots$	y_{kn}

e costituiscono quello che è usualmente chiamata una serie temporale k-variata.

Esempio: misure di pressione in un altoforno, portata di un fiume, produzione di energia elettrica.

Serie temporali

I valori possono essere quantitativi o qualitativi.

La cadenza temporale di osservazione può fare riferimento a

- ✓ una successione di istanti temporali

(**serie di stato o posizionali**, ad es. la temperatura ambiente)

oppure

- ✓ intervalli temporali

(**serie di flusso**, ad es. la produzione giornaliera)

- ○ equispaziati

oppure

- ○ non equispaziati

Serie temporali

In base alla definizione si arguisce che per le serie storiche:

- il significato del dato si arricchisce del fatto di essere osservato prima e/o dopo di un altro
- l'ordinamento dei dati non può essere alterato

In un contesto di serie storiche, la naturale tendenza di molti fenomeni ad evolversi in modo più o meno regolare porta a pensare che il dato rilevato in un dato istante t sia più simile a quello rilevato all'istante $t-1$ piuttosto che in epoche distanti; si può dire, in un certo senso, che la serie storica che analizziamo ha *memoria di sè*. Questa caratteristica è generalmente indicata col nome di **persistenza**, e differenzia profondamente i campioni di serie storiche da quelli cross-section (è il caso in cui le osservazioni di cui disponiamo siano relative a individui diversi, ad esempio prendiamo un campione di n cerniere per computer portatili e ne misuriamo la coppia per verificare siano conformi a quanto dichiarato dal produttore), perchè nei primi l'ordine dei dati ha un'importanza fondamentale, mentre nei secondi esso del tutto irrilevante.

Serie temporali

Si assume che il processo possa essere scomposto in più componenti.

Le componenti *virtuali* della serie sono:

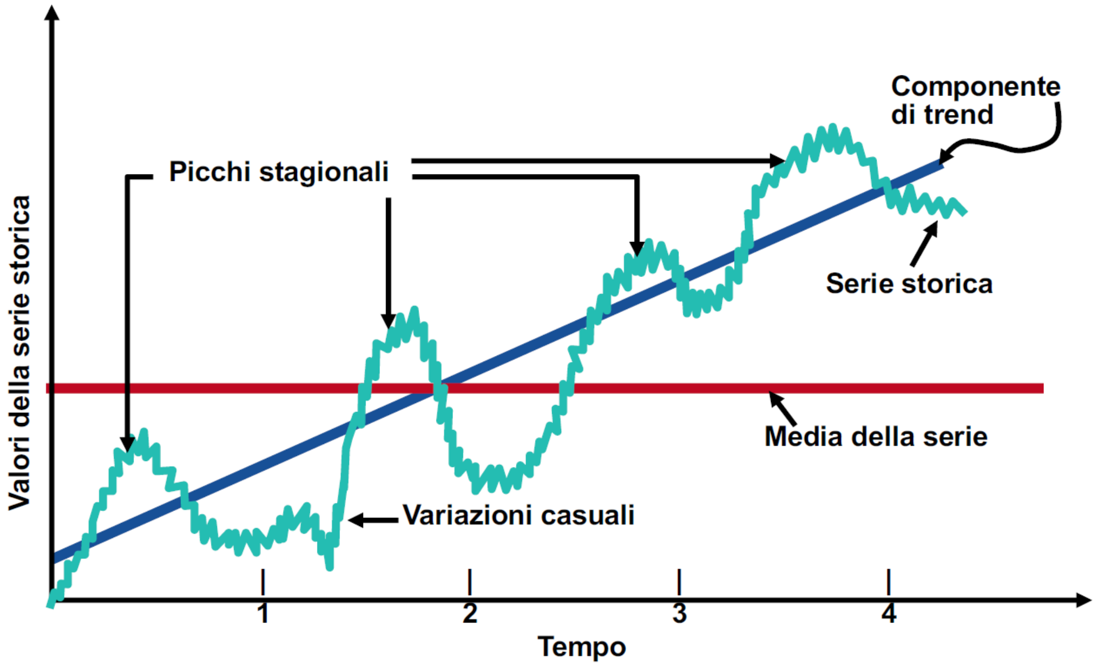
Trend (T): movimento tendenziale monotono di fondo, di lungo periodo, che mette in evidenza una evoluzione strutturale del fenomeno dovuta a cause che agiscono in modo sistematico sullo stesso, è una componente che varia *lentamente* nel tempo e che essenzialmente determina il *livello* della serie.

Stagionalità (S): una o più componenti *periodiche,* ovvero che si ritrovano uguali o quasi a distanza fissa nel tempo.

Accidentalità (e) o componente di disturbo: è data da movimenti irregolari, erratici o accidentali provocati da una serie di circostanze ciascuna di entità trascurabile, che determina nella serie delle oscillazioni tipicamente di breve periodo. Normalmente può essere assimilato ad un processo stocastico stazionario.

$$Y_t = f(T_t, S_t, e_t) \quad t = 1, \dots, n$$

Componenti di una serie temporale



Approcci di analisi

L'**analisi delle serie storiche** raggruppa una serie di metodi statistici atti a indagare una serie storica, determinare il processo alla base della stessa e a trarre previsioni.

Secondo l'approccio tradizionale, si assume che il processo abbia una parte deterministica, che consente di scomporlo in componenti tendenziali e/o stagionali, e che la differenza tra i dati teorici del modello deterministico ed i dati osservati sia attribuibile ad una componente casuale residuale.

Secondo l'approccio moderno, invece, si assume che il processo descritto sia stato generato da un processo stocastico descrivibile mediante un modello probabilistico di tipo parametrico.

APPROCCIO DETERMINISTICO

- ✓ scomporre la serie nelle sue componenti

APPROCCIO STOCASTICO

- ✓ trovare il processo generatore dei dati

Principali applicazioni

- **Previsione:** al tempo t_n vogliamo *prevedere* i valori che la serie temporale assumerà al tempo $t > t_n$

- **Controllo:** si supponga di avere a che fare, per semplicità, con due sole variabili ($k = 2$) e che:
  - le variazioni di y_{1t} influenzino y_{2t}
  - y_{1t} sia controllabile (ovvero possiamo fissarne i valori)
  - y_{2t} non sia controllabile, però desidereremmo che y_{2t} risulti uguale ad un valore prefissato, diciamo h per ogni t.

  L'analisi della serie storica consente di scegliere i valori per la prima variabile affinchè la seconda si discosti il meno possibile dal valore desiderato.

Modelli predittivi

Nel processo di creazione di un modello predittivo le fasi sono:

- Pulire i dati rimuovendo gli outlier e trattando i dati mancanti
- Identificare un approccio di modellazione predittiva parametrico o non parametrico da utilizzare
- Eseguire la pre-elaborazione dei dati in una forma adeguata all'algoritmo di modellazione prescelto
- Specificare un sottoinsieme dei dati da utilizzare per addestrare il modello
- Addestrare, o stimare, i parametri del modello a partire dall'insieme dei dati di training
- Eseguire test della performance del modello o della bontà del fitting per verificare l'adeguatezza del modello
- Validare l'accuratezza della modellazione predittiva su dati non impiegati per calibrare il modello
- Utilizzare il modello per una previsione se la sua performance è soddisfacente

Modelli predittivi

Diamo un introduzione all'analisi delle serie temporali: si vuole presentare solo alcune idee e tecniche di base.

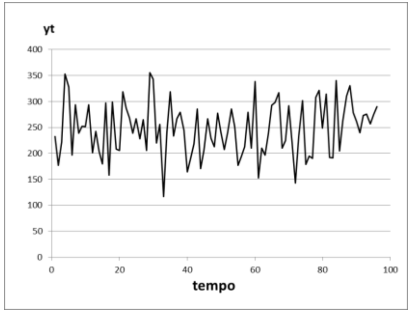
Considereremo solo:

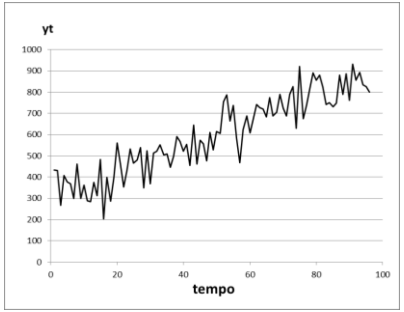
- dati equispaziati nel tempo ($t_i - t_{i-1} = \Delta$)
- situazioni in cui le variabili rilevate siano numeriche ed (almeno assimilabili a variabili) reali
- il caso di serie univariate
- solo relazioni dinamiche di tipo lineare.

Analisi grafiche preliminari

La prima cosa importante da fare quando vogliamo analizzare una serie storica per **verificare l'andamento** e l'eventuale presenza di **oscillazioni/componenti** è quella di predisporne un'opportuna **rappresentazione grafica**.

time plot o grafico di lungo periodo: vengono riportati i valori del fenomeno osservato Y (in ordinata) in corrispondenza di ciascun tempo t (in ascissa)

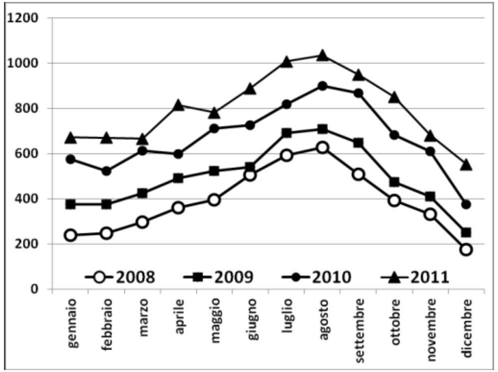




- **serie stazionaria in media**
- **serie evolutiva**

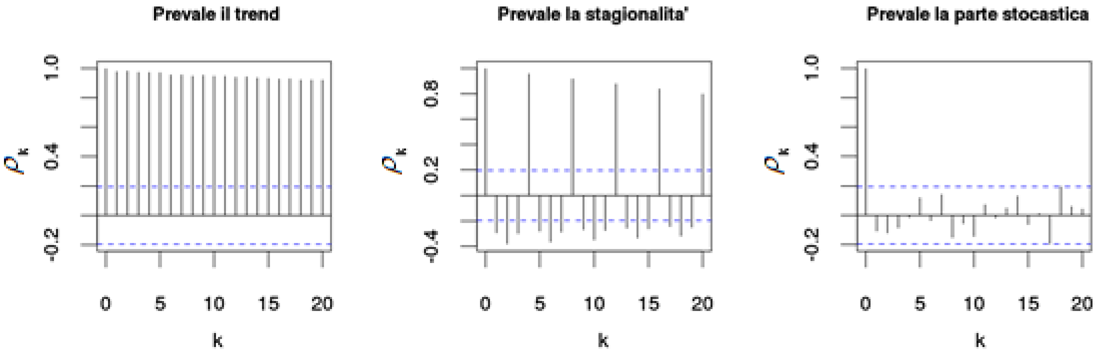
Analisi grafiche preliminari

seasonal plot o grafico a periodi sovrapposti per verificare la presenza di **oscillazioni stagionali**, che consiste nella rappresentazione dei valori della serie (in ordinata) con riferimento ad un solo periodo (in ascissa) scansionato nei sottoperiodi (ad esempio giornalieri o settimanali o mensili o trimestrali sul periodo di un anno).



Analisi grafiche preliminari

correlogramma è uno strumento grafico per la valutazione della tendenza di una serie storica ad evolversi in modo più o meno regolare. Il correlogramma si costruisce a partire dall'**autocorrelazione** ρ_k di una serie storica in funzione del **ritardo** (lag) k con cui l'autocorrelazione è calcolata: nel grafico ogni barretta verticale riporta il valore dell'autocorrelazione ACF (sull'asse delle ordinate) in funzione del ritardo (sull'asse delle ascisse). I correlogrammi possono presentare gli andamenti più disparati, ma vengono normalmente confrontati con quelli illustrati negli grafici seguenti con le tre situazioni tipiche di andamento di ρ_k al variare di k.



In blu i limiti di confidenza superiore e inferiore. Gli intervalli di confidenza approssimati intorno al valore zero dell'ACF stimata su un campione. L'intervallo di confidenza approssimato al 95% nell'ipotesi che il processo sia decorrelato è definito dai due limiti $\pm 1.96\sqrt{1/N}$, in cui 1.96 è il 0.975-esimo quantile della distribuzione Gaussiana standard e N è il numero di osservazioni.

Analisi preliminare

Riparazioni della serie storica

Le rilevazioni dei dati nel tempo possono andare incontro a due problemi
- ❑ Valori anomali
- ❑ Valori mancanti

E' necessario intervenire per la riparazione della serie storica:
- ✓ Valori troppo diversi dall'ordinario possono avere effetti imprevedibili e duraturi.
- ✓ Valori mancanti possono oscurare la percezione di movimenti importanti.

L'aggiustamento della serie storica avviene con valori sostitutivi o di rimpiazzo di quelli sbagliati. Il presupposto con cui sono scelti è che vuoti ed anomalie non siano *sintomi* di comportamenti aberranti del fenomeno che si studia. In caso contrario si rischia di nascondere il vero problema.

Analisi preliminare

Valori mancanti

- Metodo della media: si sceglie una misura di centralità (media aritmetica, mediana, semisomma quartile) calcolata su tutta la serie storica. Questa tecnica ignora del tutto l'ordinamento temporale dei valori. Sottostima la variabilità; interruzione della periodicità, perché si aggiunge un valore artificialmente stabile.

- Mediana Hodges e Lehmann: mediana della semisomma di tutte le coppie di valori

$$\hat{y}_t = Me\left(\frac{y_i + y_j}{2} \quad i < j\right)$$

dove i valori nella variabile sono disposti in ordine. Stesse considerazioni fatte sopra.

- Interpolazione lineare tra due valori contigui o media ponderata di più elementi contigui: è una tecnica semplice e rispetta l'ordinamento temporale.

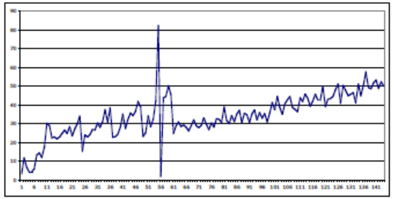
Analisi preliminare

Valori anomali

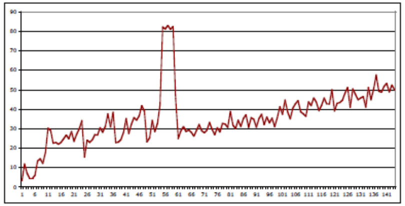
Alcune osservazioni possono risultare così lontane dalle altre da avere una influenza eccessiva sulla identificazione del modello di formazione dei dati.

Le anomalie nelle serie storiche spesso si presentano in coppie o in gruppi.

Un valore anomalo in eccesso è seguito da un valore anomalo in difetto (o viceversa), a causa dell’attrazione esercitata del trend.



In varie occasioni i valori anomali si presentano a pacchetti perché lo *shock* si estende in più periodi finché non interviene l’attrazione verso il trend.

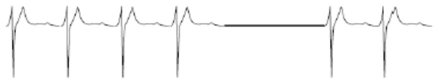


Analisi preliminare

Valori anomali

I valori anomali non sono per forza sbagliati o estranei ai dati. Possono invece essere molto utili: ciò che per qualcuno è solo un rumore, per altri può essere un segnale importante.

Ad esempio:



Tracciato di un elettrocardiogramma.

L'anomalia è costituita dal plateau al centro. Il valore osservato non è anomalo in sé, ma il suo perdurare per un periodo così lungo ingenera preoccupazione.

Valori anomali

Il valore anomalo nelle serie storiche è tale se si innalza o si abbassa in modo eccessivo rispetto ai valori adiacenti: precedenti e seguenti.

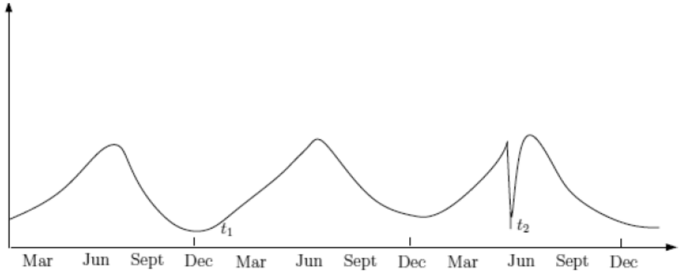
Consideriamo Y_t ed i valori vicini da Y_{t-r} a Y_{t+r} che coinvolge $2r$ valori con esclusione di Y_t. Calcoliamo la mediana dei valori vicini, m_t, e confrontiamola con il valore osservato al tempo t

$$|Y_t - m_t| \leq \tau \qquad\qquad |Y_t - m_t| > \tau$$

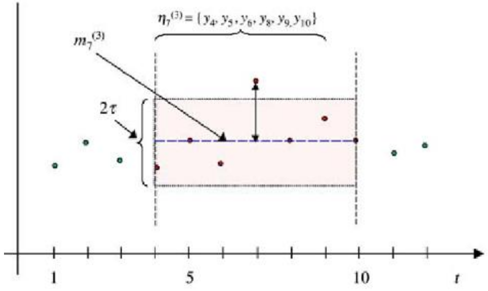
Se lo scarto assoluto è minore di una certa soglia allora Y_t non potrà essere considerato un valore anomalo. Se invece supera la soglia allora dovrà essere trattato come outlier.

Il trattamento consiste nel porre $Y_t = m_t$.

La difficoltà del test sta nella scelta del numero di vicini e della soglia.



L'evento t_2 produce lo stesso valore di t_1, ma t_1è del tutto normale nel suo contesto, non così il t_2.



Analisi preliminare

Valori anomali

Solitamente la soglia di anomalia non è fissata, ma si ordinano gli scarti dalle mediane e quelli maggiori sono studiati con altri test.

Calcoliamo la media dei valori da $(t-r)$ a $(t+r)$ con INCLUSIONE dell't-esimo valore: μ_1
Calcoliamo la media dei valori da $(t-r)$ a $(t+r)$ con ESCLUSIONE dell't-esimo valore: μ_2

La presenza di una anomalia in posizione t-esima dovrebbe comportare una marcata differenza tra le due medie dei $p = 2r + 1$ e $p - 1 = 2r$ dati.

Il controllo della significatività di tale condizione consente di individuare le anomalie.

Analisi preliminare

Valori anomali

Il confronto delle due medie avviene con un test t-Student:

$$t = \frac{\mu_1 - \mu_2}{\sqrt{\dfrac{s_1^2}{p} + \dfrac{s_2^2}{p-1}}}$$

con gradi di libertà

$$\nu = \frac{1}{\dfrac{c}{p-1} + \dfrac{(1-c)^2}{p-2}} \qquad c = \frac{\dfrac{s_1^2}{p}}{\dfrac{s_1^2}{p} + \dfrac{s_2^2}{p-1}}$$

Analisi preliminare

Valori anomali

I candidati all'anomalia sono i valori che danno luogo al più basso valor p-value del test della differenza tra medie.

L'ipotesi è che i campioni provengano da due gaussiane diverse sia per la media che per la varianza. Questa ipotesi non è mai vera. Solo qualche volta è vera in via approssimata.
I valori nei due campioni non dovrebbero essere in relazione tra di loro (in realtà sono interni ad una medesima serie storica).
I valori di un campione non dovrebbero essere singolarmente abbinabili con quelli dell'altro (in realtà i valori coincidono, tranne che per un elemento).

Le decisioni sull'anomalia andranno quindi prese con molta prudenza.

I modelli di composizione e scomposizione e i metodi per la stima delle componenti

L'approccio classico all'analisi delle serie temporali ipotizza, come si è detto, che la **serie** sia composta da **variazioni** (pattern) sistematiche o deterministiche (trend e stagionalità) e da **oscillazioni di disturbo o casuali**.

- Seasonality (S_t)
- Trend (T_t)
- Errore (e_t) (detto White Noise)

Con riferimento al modello $Y_t = f(T_t, S_t, e_t)$, si tratta quindi di stimare le singole componenti virtuali T_t, S_t.

I modelli di composizione e scomposizione e i metodi per la stima delle componenti

Al fine di stimare le componenti virtuali occorre:

- stabilire il modo con il quale le stesse interagiscono tra loro (si aggregano) per dar luogo alla serie effettiva (specificare la f).
- decidere il metodo con cui stimare le singole componenti.

Le due principali forme di f sono:

il **modello additivo**:

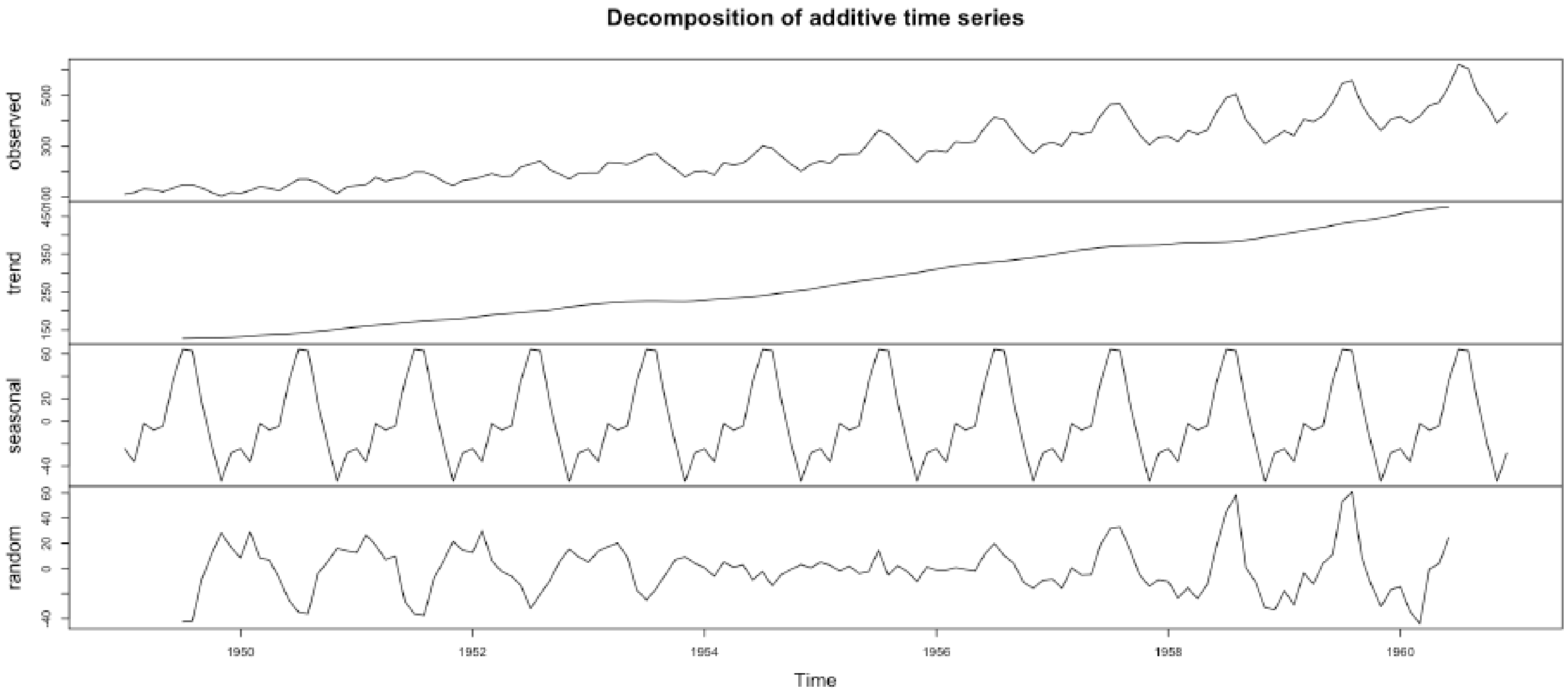
$$Y_t = T_t + S_t + e_t$$

il **modello moltiplicativo**:

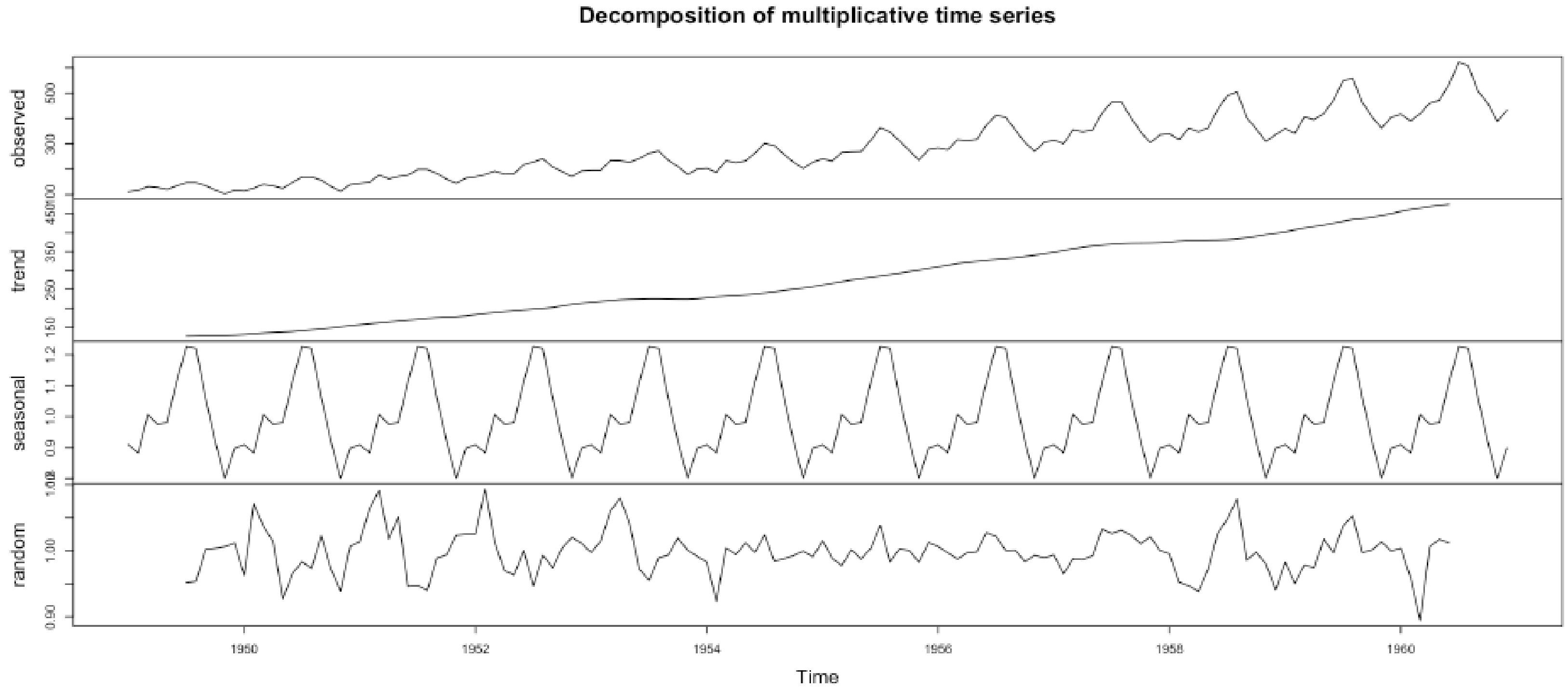
$$Y_t = T_t \, S_t \, e_t$$

NOTA: Una serie storica moltiplicativa può essere convertita in additiva calcolandone il logaritmo

I modelli di composizione e scomposizione e i metodi per la stima delle componenti

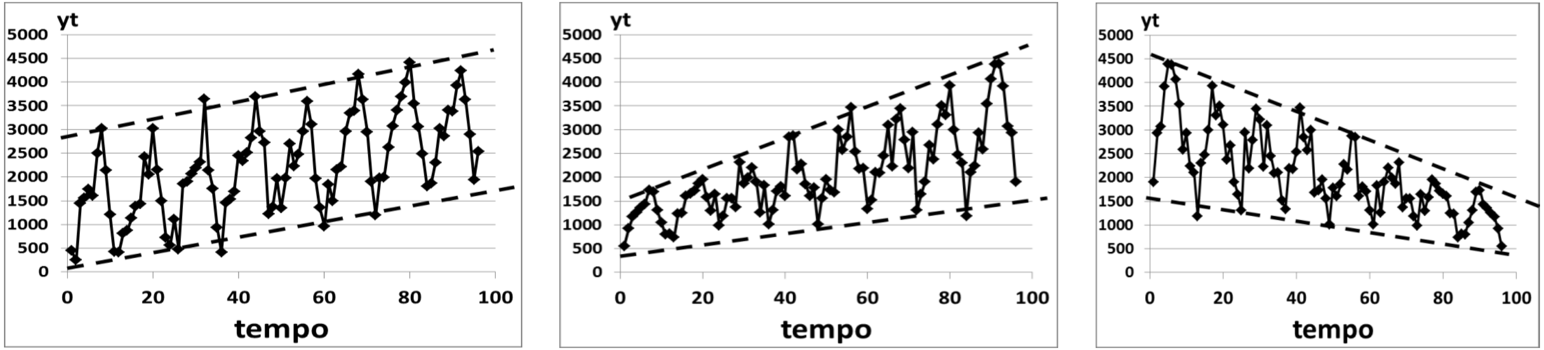


I modelli di composizione e scomposizione e i metodi per la stima delle componenti



I modelli di composizione e scomposizione e i metodi per la stima delle componenti

Un modello additivo è appropriato quando **l'ampiezza dell'oscillazione** stagionale **non varia** con il variare del livello della serie (si parla allora di serie additiva). Se invece la fluttuazione stagionale **aumenta (o diminuisce)** proporzionalmente all'aumento (diminuzione) del livello della serie, allora è più adeguato un modello moltiplicativo. A titolo esemplificativo si vedano i grafici che seguono



Nel **modello additivo**, le componenti T_t, S_t, e_t sono espresse nella **stessa unità di misura di** y_t.
Nel modello moltiplicativo, **solo** T_t (per convenzione) viene espresso **nell'unità di misura di** y_t mentre S_t e e_t sono espressi come numeri adimensionali.

I modelli di composizione e scomposizione e i metodi per la stima delle componenti

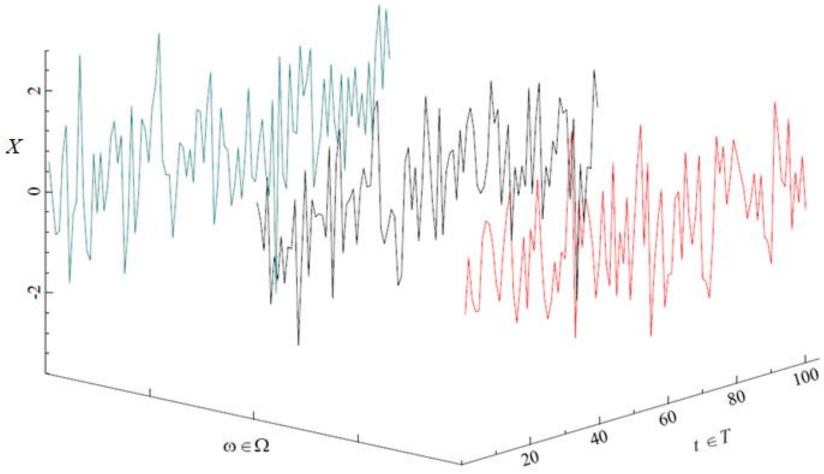
Le singole componenti possono essere stimate utilizzando **metodi empirici** (perequativi) oppure **metodi analitici** (ovvero di interpolazione).

Nel **primo caso** si utilizza il **metodo delle medie mobili** che stima i valori delle componenti, ma non consente di per sé di effettuare estrapolazioni (consiste nel sostituire a y_i la media aritmetica di due, quattro, sei, … dati di cui y_i è il dato centrale, nel caso più semplice si sostituisce y_i con la media aritmetica tra y_{i-1} e y_{i+1}).

Nel **secondo caso** si impiega una **funzione analitica** per la quale è possibile stimare i parametri e che consente di effettuare estrapolazioni al futuro.

Approccio stocastico

Realizations of a stochastic process



- al variare di t e di ω, $X(t, \omega)$ è un PROCESSO STOCASTICO
- al variare di ω fissato t, $X(t, \omega)$ è una VARIABILE CASUALE
- al variare di t fissato ω, $X(t, \omega)$ è una REALIZZAZIONE
- fissati t e ω, $X(t, \omega)$ è un VALORE REALE

Approccio stocastico

- Obiettivo determinare la legge che "governa" il processo
- Il passato spiega il presente e il futuro
- Modelli lineari che spiegano guardando il passato

L'obiettivo principale dell'analisi delle serie storiche e quello di sviluppare modelli matematici che riescano a descrivere in modo plausibile il campione osservato.
Punti adiacenti nel tempo sono correlati tra loro, pertanto il valore della serie osservato al tempo t dipenderà in qualche maniera dai valori passati $x_{t-1}, x_{t-2}, \ldots$
Tale considerazione e alla base della costruzione e analisi delle serie storiche.

Modello di regressione e modello di serie storiche

Modelli di regressione $Y_t = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \varepsilon_t$

- assunzione di independenza delle osservazioni
- la variabile dipendente Y è funzione di variabili esplicative $X_1, \dots, X_k$
- si usa il coefficiente di correlazione r

Modelli di serie storiche $X_t = \alpha + \phi_1 X_{t-1} + \cdots + \phi_k X_{t-k} + \varepsilon_t$

- dipendenza tra osservazioni
- X spiegata dal suo passato
- uso dell'autocorrelazione

MODELLO REGRESSIONE
media
varianza
correlazione

MODELLO SERIE STORICHE
media
varianza
autocorrelazione

Modelli lineari univariati

Ci limitiamo a MODELLI **LINEARI** e osseviamo solo **UNA** variabile X nel tempo. Spieghiamo il valore attuale X_t osservando il passato

DEFINIZIONE DI VARIABILE RITARDATA

X_t	X_{t-1}	X_{t-2}	X_{t-3}
X_1			
X_2	X_1		
X_3	X_2	X_1	
X_4	X_3	X_2	X_1
X_5	X_4	X_3	X_2

L'operatore ritardo è indicato con L (**Lag**) ed è definito come segue:

$$L^1 X_t = X_{t-1}$$
$$L^2 X_t = X_{t-2}$$
$$L^3 X_t = X_{t-3}$$
...
$$L^p X_t = X_{t-p}$$

A volte viene indicato con B (= Backshift).

Trasformazioni

Molto spesso nella elaborazione dei dati statistici le variabili vengono trasformate.

Le trasformazioni utili a fini interpretativi dei fenomeni sono:

- variazione assoluta
- variazione percentuale

La variazione assoluta si ottiene immediatamente come differenza del valore X al tempo t rispetto allo stesso un tempo precedente:

$$\Delta X_t = X_t - X_{t-1}$$

Analizzando ΔX piuttosto che X il modello specificato spiegherà la variazione di X.

Introducendo l'operatore ritardo, avremo:

$$\Delta X_t = X_t - LX_t = (1 - L)X_t$$

Trasformazioni

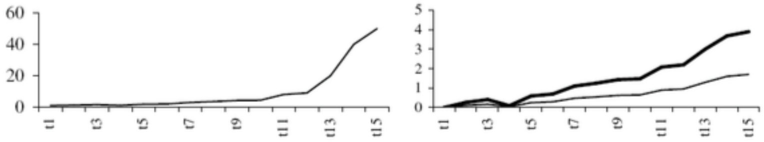
L'altra trasformazione importante è quella che porta al tasso di variazione della serie:

$$\frac{X_t - X_{t-1}}{X_{t-1}}$$

Si dimostra facilmente che questa è ben approssimata dalla seguente espressione:

$$\frac{X_t - X_{t-1}}{X_{t-1}} \approx \log X_t - \log X_{t-1} = \Delta \log X_t$$

Molto spesso quindi piuttosto che rilevare X_t, si studierà $\log X_t$ e poi, successivamente, la sua variazione. La trasformazione logaritmica viene ampiamente utilizzata in statistica.



Trasformazioni logaritmiche della serie storica a sinistra.

Alcuni concetti per l'analisi

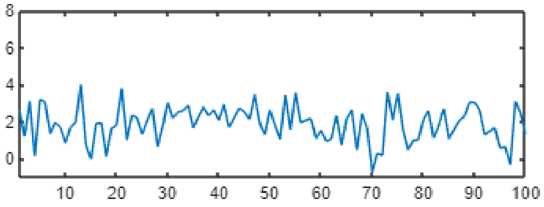
Stazionarietà

- definizione intuitiva: una serie è stazionaria se, qualora perturbata da shock, torna al suo equilibrio di lungo periodo una volta terminato l'effetto di questi shock
- definizione rigorosa: un processo stocastico si dice stazionario (in senso debole o in covarianza) se, per qualsiasi t e $t-k$, vale che:

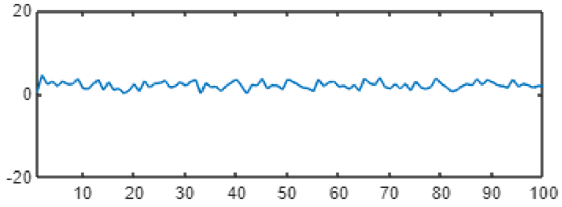
$E(X_t) = E(X_{t-k}) = \mu$ media finita
$Var(X_t) = Var(X_{t-k}) = \sigma^2$ varianza finita
$Cov(X_t, X_{t-k}) = Cov(X_{t-j}, X_{t-k-j}) = \gamma_k$

Alcuni concetti per l'analisi

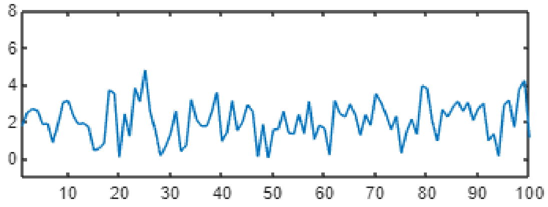
Processo stazionario nella media.

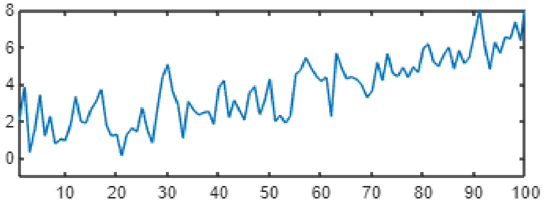


Processo stazionario nella varianza.

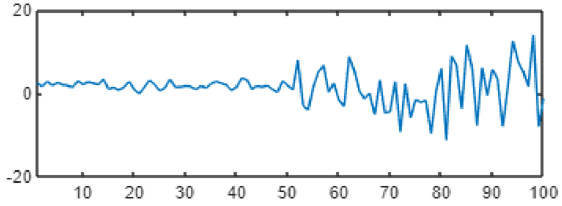


Processo stazionario nella covarianza.

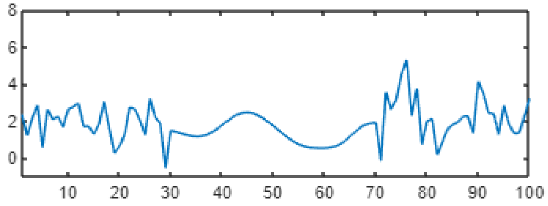




Processo non stazionario nella media.



Processo non stazionario nella varianza.



Processo non stazionario nella covarianza.

Alcuni concetti per l'analisi

Autocorrelazione (ACF)

La correlazione fra X_t e X_{t-k} è data da:

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

$-1 \leq \rho_k \leq 1$

$\rho_k = 0 \rightarrow$ nessuna relazione lineare

$\rho_k = 1 \rightarrow$ relazione lineare perfetta crescente

$\rho_k = -1 \rightarrow$ relazione lineare perfetta decrescente

Il grafico degli indici di autocorrelazione si dice *correlogramma.*

Alcuni concetti per l'analisi

Autocorrelazione parziale (PACF)

Misura la correlazione fra X_t e X_{t-k} al netto dell'effetto prodotto da altre variabili/fattori. Dato il processo X_t, definiamo ρ_k la correlazione tra X_t e X_{t-k}. Preso

$$I_t = \{X_{t-1}, X_{t-2}, \dots X_{t-k+1}\}$$

l'autocorrelazione parziale A_k è definita come

$$A_k = corr(X_t, X_{t-k} | I_t)$$

In particolare

$$A_k = \frac{|\mathbf{Q}_k|}{|\mathbf{P}_k|}$$

dove

$$\mathbf{Q}_k = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_2 \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_k \end{bmatrix} \qquad \boldsymbol{P}_k = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{bmatrix}$$

La matrice $\mathbf{P}_k$ è comunemente chiamata matrice di Toeplitz.

Modelli stazionari

Esistono diversi modelli per descrivere un processo stazionario:

- ❑ Moving average MA (*Media mobile)*
- ❑ Autoregressive Model AR (*Autoregressivo)*
- ❑ Mixed Autoregressive Moving Average model ARMA

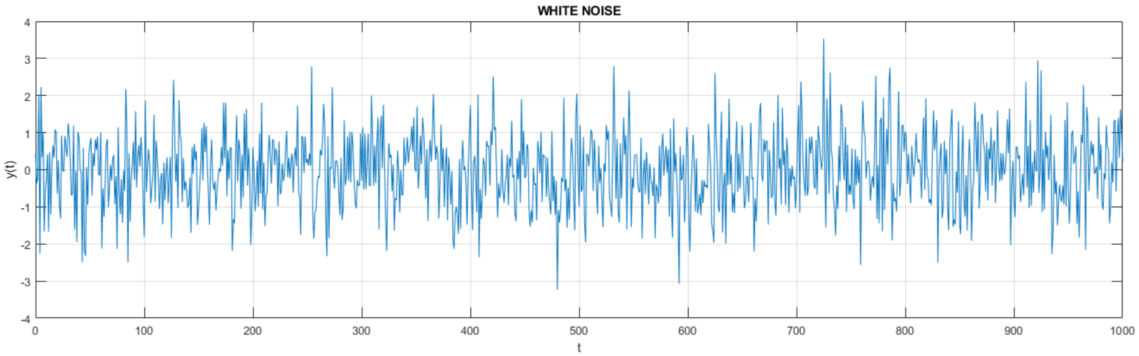
E poi esistono diversi modelli per descrivere processi non stazionari.

Processo White Noise

Un processo X_t si dice White Noise se:

$E(X_t) = 0$ (no trend)
$Var(X_t) = \sigma^2$ varianza finita
$Cov(X_t, X_{t-k}) = 0$ (no covarianza)

- ✓ Si tratta quindi di un processo casuale, senza relazione tra eventi in diversi istanti nel tempo.
- ✓ E' in ogni caso un processo stazionario.
- ✓ E' l'ipotesi tipica per la componente residua (o shock) ε_t.



Modelli AR(p)

AR(1) $X_t = \phi_1 X_{t-1} + \varepsilon_t$
AR(2) $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$
...
AR(p) $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t$

In un processo autoregressivo di ordine p, il valore osservato al tempo t è funzione lineare dei p valori precedenti sommati al disturbo corrente.

- ✓ ε_t è white noise
- ✓ a seconda dei valori di ϕ, la serie può "esplodere" o convergere verso un punto (equilibrio di lungo periodo)
- ✓ l'ampiezza di ϕ è misura della persistenza dell'effetto di shock passati.

Modelli MA(q)

MA(1) $X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$

MA(2) $X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$

...

MA(q) $X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$

In un processo a media mobile di ordine q, il valore osservato al tempo t è funzione lineare dei q shock precedenti sommati al disturbo corrente.

- ✓ ε_t è white noise
- ✓ a seconda dei valori di θ, la serie può "esplodere" o convergere verso un punto

Modelli ARMA(p,q)

Si tratta di una combinazione di processi AR(p) e MA(q):

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$$

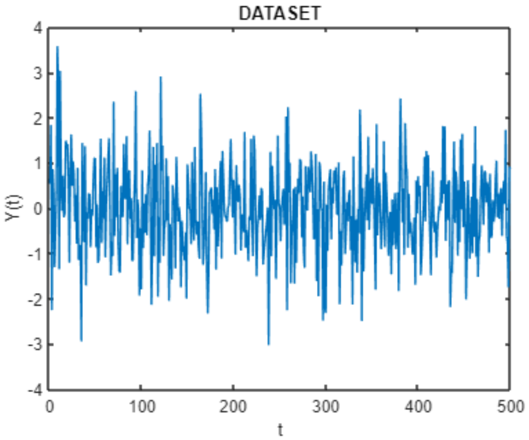
Nota che, anche in questo caso, siamo di fronte a serie stazionarie.

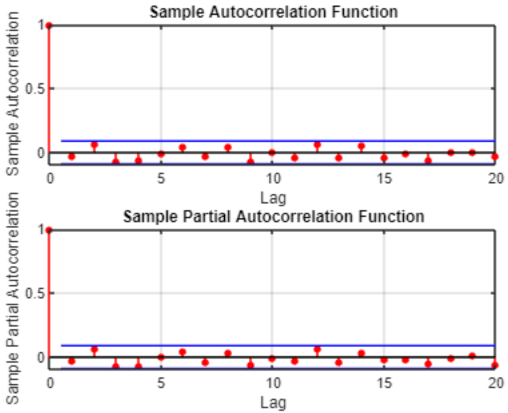
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

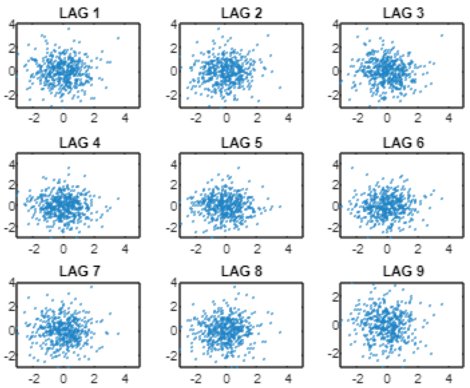
Processo	ACF	PACF
White Noise	Ovunque nulla	Ovunque nulla
AR(1), $\phi_1 > 0$	Decade esponenzialmente	Nulla per $k > 1$
AR(1), $\phi_1 < 0$	Decade oscillando	Nulla per $k > 1$
AR(p)	Decade a zero, eventualmente oscillando	Nulla per $k > p$
MA(1), $\theta_1 > 0$	Positiva per $k = 1$, nulla per $k > 1$	Decade oscillando
MA(1), $\theta_1 < 0$	Negativa per $k = 1$, nulla per $k > 1$	Decade a zero (valori negativi)
ARMA(1,1), $\phi_1 > 0, \theta_1 > 0$	Decade esponenzialmente	Decade oscillando
ARMA(1,1), $\phi_1 > 0, \theta_1 < 0$	Decade esponenzialmente	Decade esponenzialmente
ARMA(1,1), $\phi_1 < 0, \theta_1 > 0$	Decade oscillando	Decade oscillando
ARMA(1,1), $\phi_1 < 0, \theta_1 < 0$	Decade oscillando	Decade esponenzialmente
ARMA(p,q)	Decade (direttamente od oscillando) per $k > q$	Decade (direttamente od oscillando) per $k > p$

Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

White noise







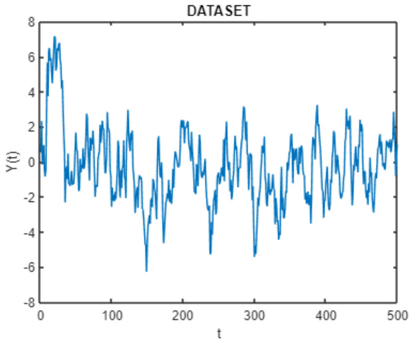
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

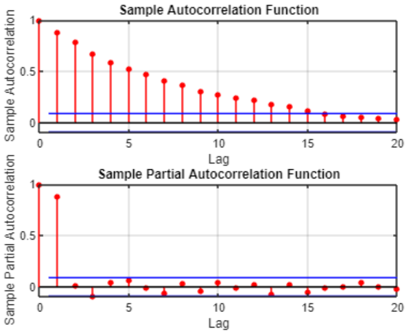
Processo	ACF	PACF
White Noise	Ovunque nulla	Ovunque nulla
AR(1), $\phi_1 > 0$	Decade esponenzialmente	Nulla per $k > 1$
AR(1), $\phi_1 < 0$	Decade oscillando	Nulla per $k > 1$
AR(p)	Decade a zero, eventualmente oscillando	Nulla per $k > p$
MA(1), $\theta_1 > 0$	Positiva per $k = 1$, nulla per $k > 1$	Decade oscillando
MA(1), $\theta_1 < 0$	Negativa per $k = 1$, nulla per $k > 1$	Decade a zero (valori negativi)
ARMA(1,1), $\phi_1 > 0, \theta_1 > 0$	Decade esponenzialmente	Decade oscillando
ARMA(1,1), $\phi_1 > 0, \theta_1 < 0$	Decade esponenzialmente	Decade esponenzialmente
ARMA(1,1), $\phi_1 < 0, \theta_1 > 0$	Decade oscillando	Decade oscillando
ARMA(1,1), $\phi_1 < 0, \theta_1 < 0$	Decade oscillando	Decade esponenzialmente
ARMA(p,q)	Decade (direttamente od oscillando) per $k > q$	Decade (direttamente od oscillando) per $k > p$

Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

AR(1) $X_t = \phi_1 X_{t-1} + \varepsilon_t$

Ci si potrebbe aspettare che ci sia autocorrelazione solo per $k = 1$.
Non è così perchè X_{t-1} sarà influenzata da X_{t-2} e quindi ci sarà una relazione indiretta anche tra X_t e X_{t-2}, etc. evidenziata da ACF. In PACF si evidenzia la correlazione tra X_t e X_{t-k}, tolti tutti i legami intermedi indotti, quindi è presente solo per $k = 1$.

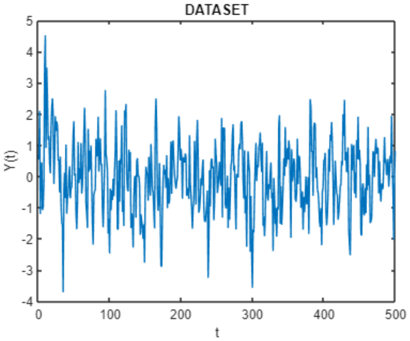


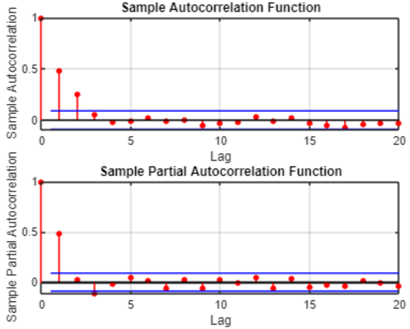


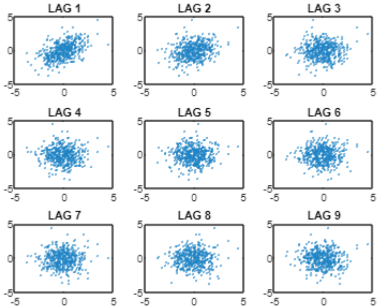
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

AR(1) $X_t = \phi_1 X_{t-1} + \varepsilon_t$

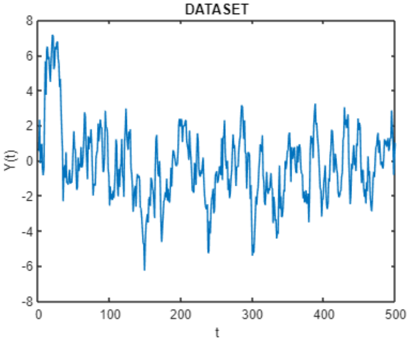
ϕ_1=0.5

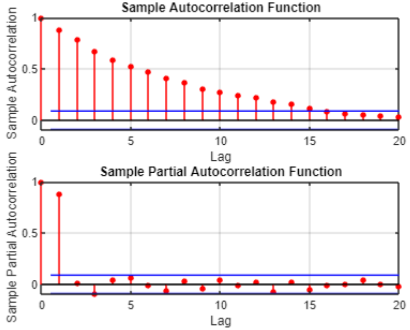


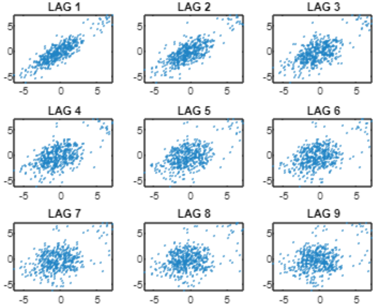




ϕ_1=0.9



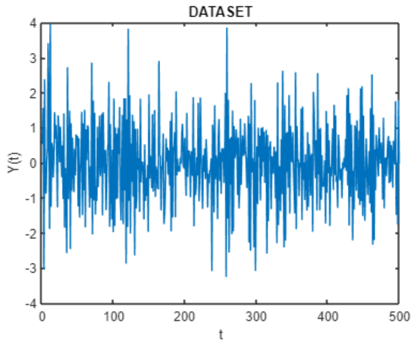


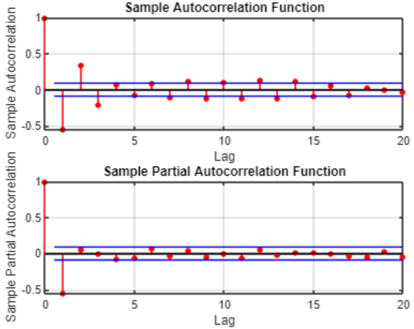


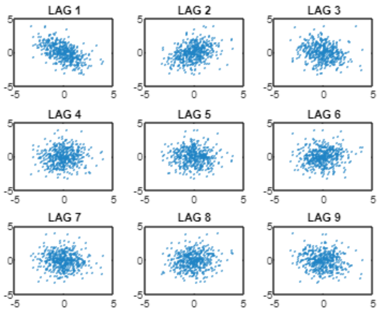
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

AR(1) $X_t = \phi_1 X_{t-1} + \varepsilon_t$

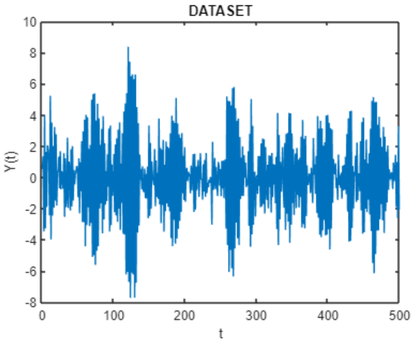
ϕ_1=-0.5

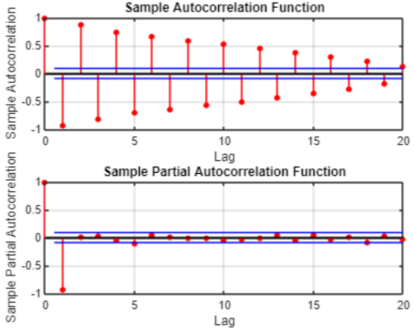


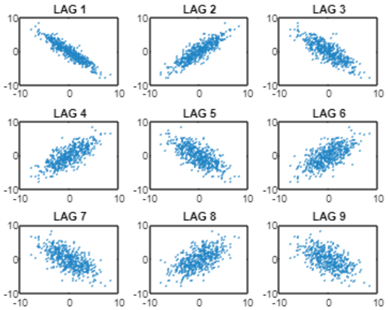




ϕ_1=-0.9



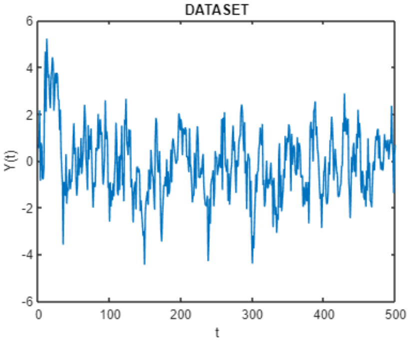


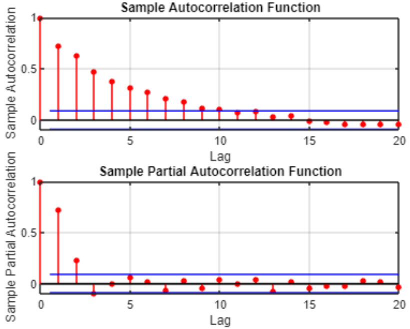


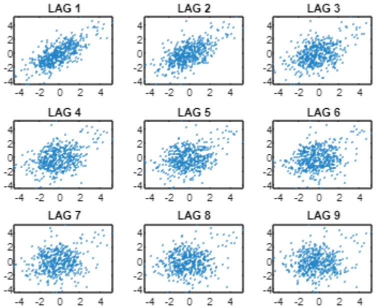
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

AR(2) $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$

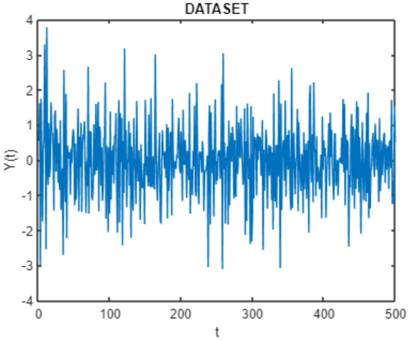
ϕ_1=0.6, ϕ_2=0.2

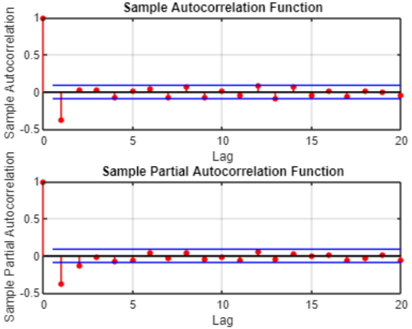


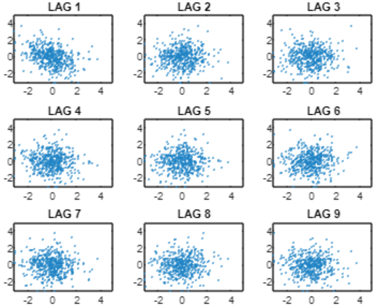




ϕ_1=-0.4 , ϕ_2=-0.2



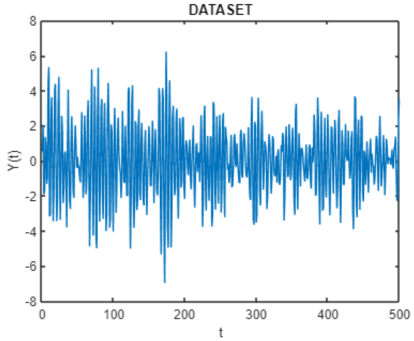


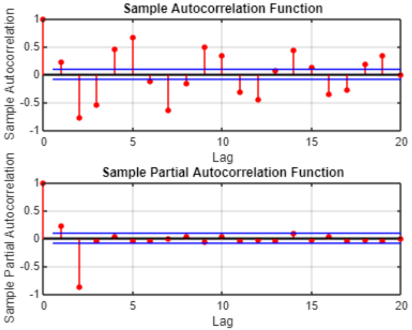


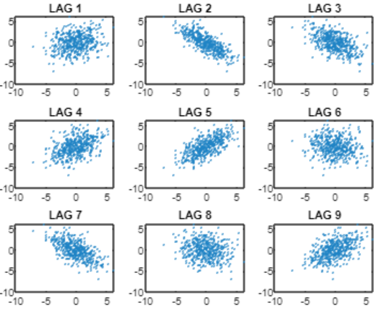
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

AR(2) $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$

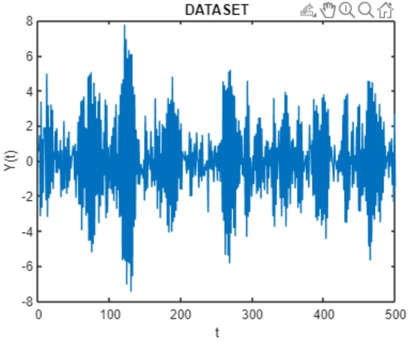
ϕ_1=0.4, ϕ_2=-0.9

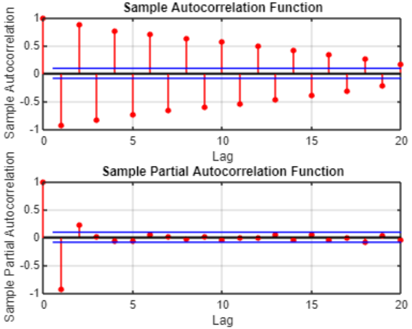


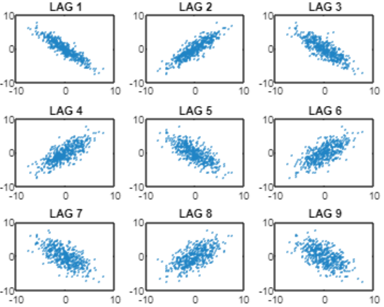




ϕ_1=-0.7 , ϕ_2=0.2







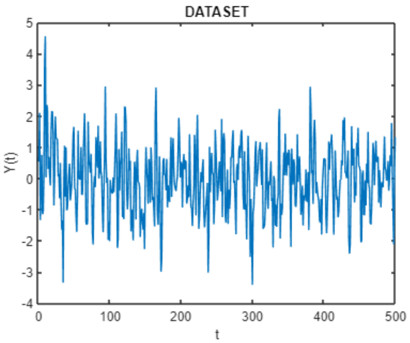
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

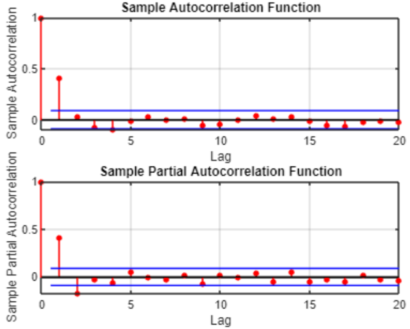
Processo	ACF	PACF
White Noise	Ovunque nulla	Ovunque nulla
AR(1), $\phi_1 > 0$	Decade esponenzialmente	Nulla per $k > 1$
AR(1), $\phi_1 < 0$	Decade oscillando	Nulla per $k > 1$
AR(p)	Decade a zero, eventualmente oscillando	Nulla per $k > p$
MA(1), $\theta_1 > 0$	Positiva per $k = 1$, nulla per $k > 1$	Decade oscillando
MA(1), $\theta_1 < 0$	Negativa per $k = 1$, nulla per $k > 1$	Decade a zero (valori negativi)
ARMA(1,1), $\phi_1 > 0, \theta_1 > 0$	Decade esponenzialmente	Decade oscillando
ARMA(1,1), $\phi_1 > 0, \theta_1 < 0$	Decade esponenzialmente	Decade esponenzialmente
ARMA(1,1), $\phi_1 < 0, \theta_1 > 0$	Decade oscillando	Decade oscillando
ARMA(1,1), $\phi_1 < 0, \theta_1 < 0$	Decade oscillando	Decade esponenzialmente
ARMA(p,q)	Decade (direttamente od oscillando) per $k > q$	Decade (direttamente od oscillando) per $k > p$

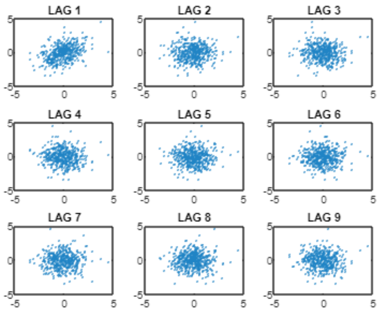
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

MA(1) $X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$

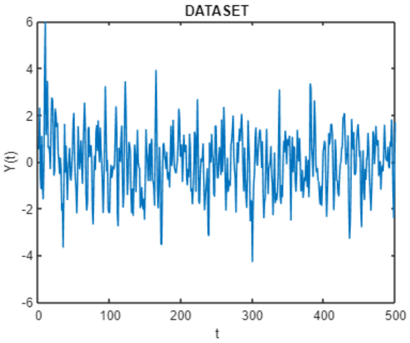
θ_1=0.5

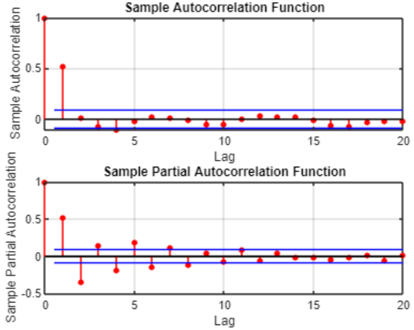


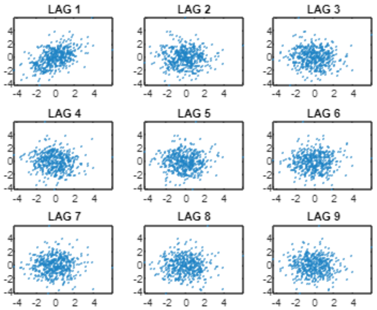




θ_1=0.9



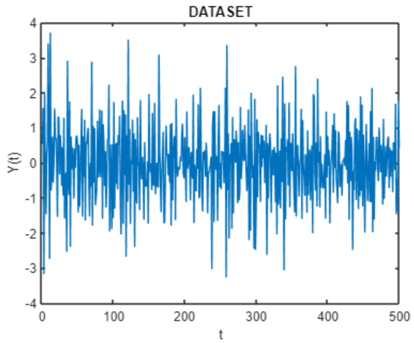


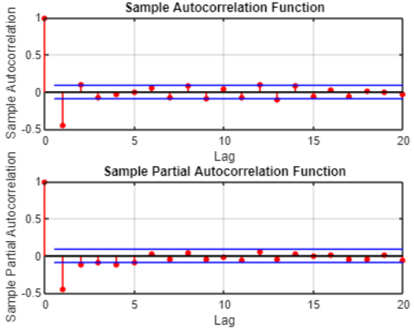


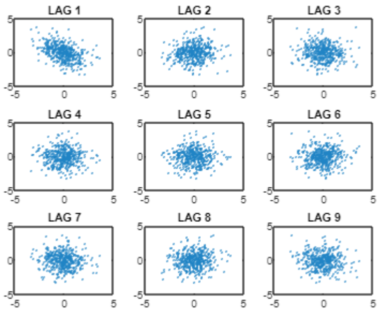
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

MA(1) $X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$

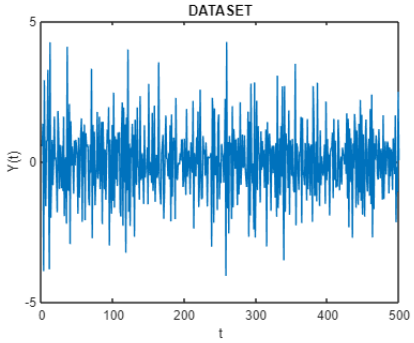
θ_1=-0.5

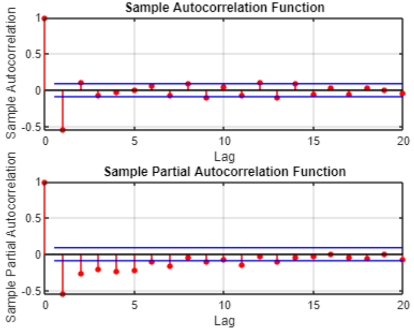


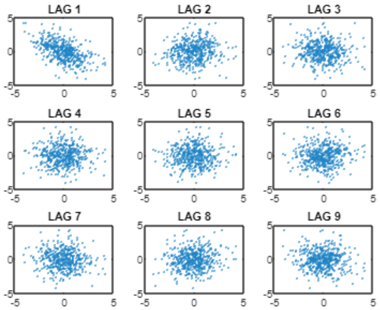




θ_1=-0.9



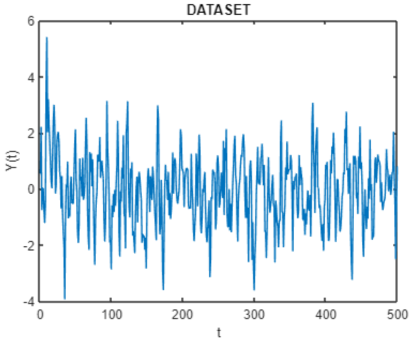


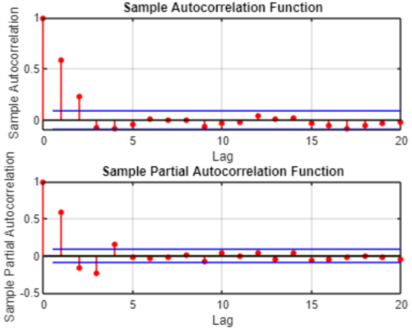


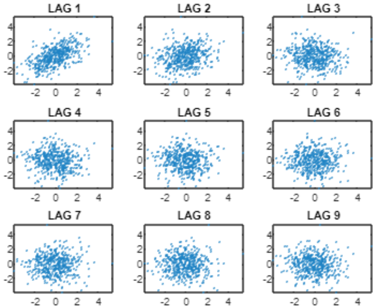
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

MA(2) $X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}$

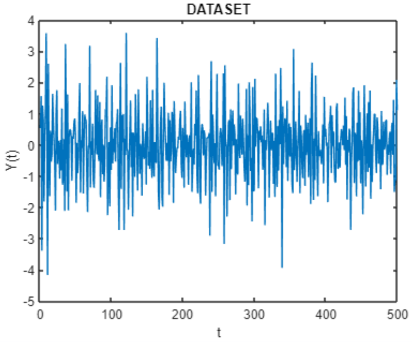
θ_1=0.7, θ_2=0.4

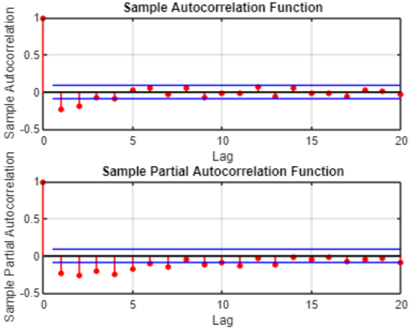


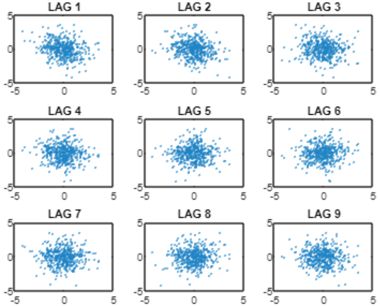




θ_1=-0.5 , θ_2=-0.4







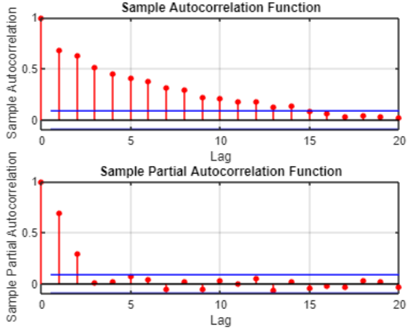
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

Processo	ACF	PACF
White Noise	Ovunque nulla	Ovunque nulla
AR(1), $\phi_1 > 0$	Decade esponenzialmente	Nulla per $k > 1$
AR(1), $\phi_1 < 0$	Decade oscillando	Nulla per $k > 1$
AR(p)	Decade a zero, eventualmente oscillando	Nulla per $k > p$
MA(1), $\theta_1 > 0$	Positiva per $k = 1$, nulla per $k > 1$	Decade oscillando
MA(1), $\theta_1 < 0$	Negativa per $k = 1$, nulla per $k > 1$	Decade a zero (valori negativi)
ARMA(1,1), $\phi_1 > 0, \theta_1 > 0$	Decade esponenzialmente	Decade oscillando
ARMA(1,1), $\phi_1 > 0, \theta_1 < 0$	Decade esponenzialmente	Decade esponenzialmente
ARMA(1,1), $\phi_1 < 0, \theta_1 > 0$	Decade oscillando	Decade oscillando
ARMA(1,1), $\phi_1 < 0, \theta_1 < 0$	Decade oscillando	Decade esponenzialmente
ARMA(p,q)	Decade (direttamente od oscillando) per $k > q$	Decade (direttamente od oscillando) per $k > p$

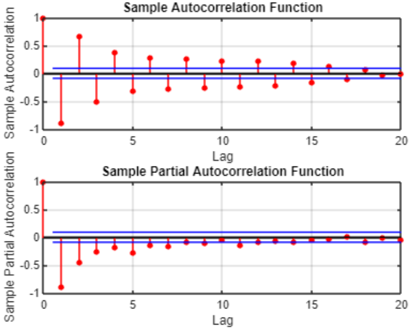
Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

ARMA(1,1) $X_t = \phi_1 X_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$

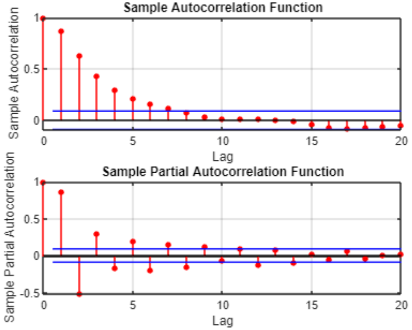
$\phi_1 > 0, \theta_1 < 0$



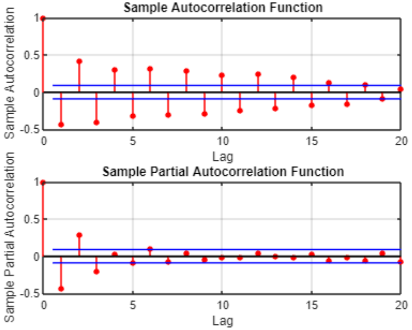
$\phi_1 < 0, \theta_1 < 0$



$\phi_1 > 0, \theta_1 > 0$



$\phi_1 < 0, \theta_1 > 0$



Proprietà delle autocorrelazioni (ACF) e autocorrelazioni parziali (PACF)

Processo	ACF	PACF
White Noise	Ovunque nulla	Ovunque nulla
AR(1), $\phi_1 > 0$	Decade esponenzialmente	Nulla per $k > 1$
AR(1), $\phi_1 < 0$	Decade oscillando	Nulla per $k > 1$
AR(p)	Decade a zero, eventualmente oscillando	Nulla per $k > p$
MA(1), $\theta_1 > 0$	Positiva per $k = 1$, nulla per $k > 1$	Decade oscillando
MA(1), $\theta_1 < 0$	Negativa per $k = 1$, nulla per $k > 1$	Decade a zero (valori negativi)
ARMA(1,1), $\phi_1 > 0, \theta_1 > 0$	Decade esponenzialmente	Decade oscillando
ARMA(1,1), $\phi_1 > 0, \theta_1 < 0$	Decade esponenzialmente	Decade esponenzialmente
ARMA(1,1), $\phi_1 < 0, \theta_1 > 0$	Decade oscillando	Decade oscillando
ARMA(1,1), $\phi_1 < 0, \theta_1 < 0$	Decade oscillando	Decade esponenzialmente
ARMA(p,q)	Decade (direttamente od oscillando) per $k > q$	Decade (direttamente od oscillando) per $k > p$

Modelli ARIMA(p,d,q)

In genere le serie storiche non sono stazionarie. È possibile rendere una serie storica stazionaria effettuando una trasformazione opportuna.

E' importante individuare se la non stazionarietà è causata da un elemento di natura stocastica (radice unitaria) o deterministica (trend esogeno al processo generatore della serie).

In entrambi i casi, uno shock ha effetto permanente.

Ciò implica che la serie presenta uno o più trend che cambiano continuamente la media delle serie.

Modelli ARIMA(p,d,q)

Esistono dei test che permettono di individuare la presenza di una radice unitaria nel processo generatore della serie.

Il più utilizzato è il **test di Dickey e Fuller** che ha le seguenti ipotesi:
H0 : il processo ha una radice unitaria
H1 : il processo non ha una radice unitaria
Il test ha numerose varianti, che permettono di individuare anche le componenti deterministiche e gli effetti di shock temporanei.

Se il processo generatore della serie include un trend di natura deterministica, si costruisce un modello di regressione con la variabile tempo come regressore, tramite la procedura OLS (minimi quadrati ordinari).

Se il processo generatore della serie include una radice unitaria, è possibile includerla nella stima del modello, il modello ARIMA, dove si specifica l'ordine della differenza.

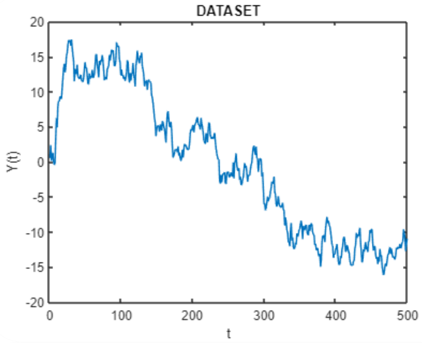
Modelli ARIMA(p,d,q)

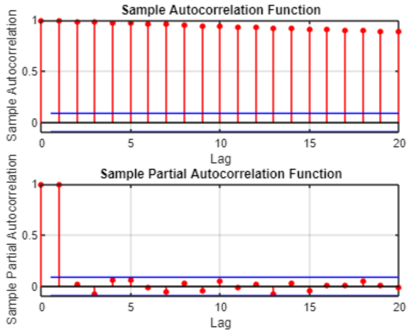
Un esempio di serie non stazionaria:

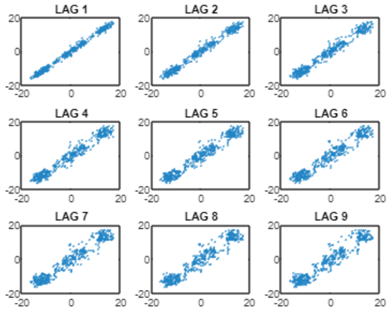
$$X_t = X_{t-1} + \varepsilon_t$$

Una serie di questo tipo è determinata da un processo ARIMA(p,d,q), dove d indica l'ordine delle differenze (d=1 equivale a $X_t - X_{t-1}$)

ARIMA(0,1,0)

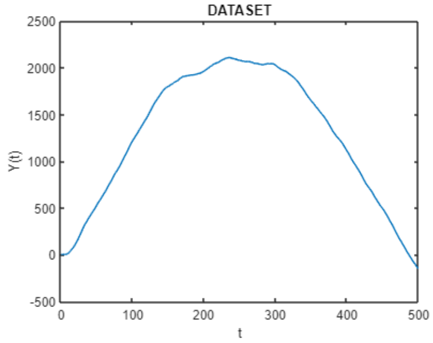


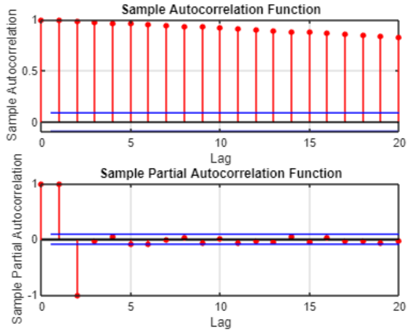


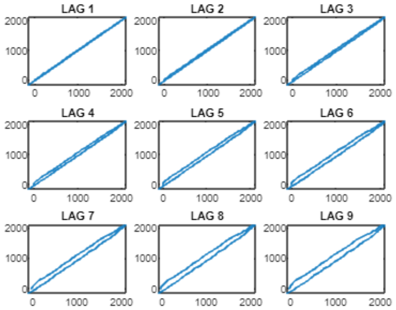


Modelli ARIMA(p,d,q)

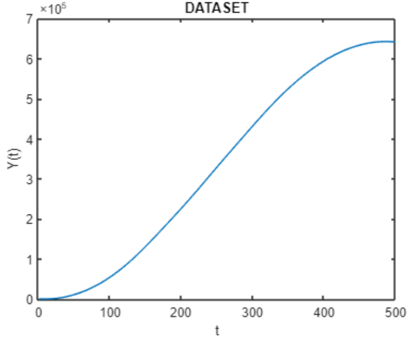
ARIMA(0,2,0)

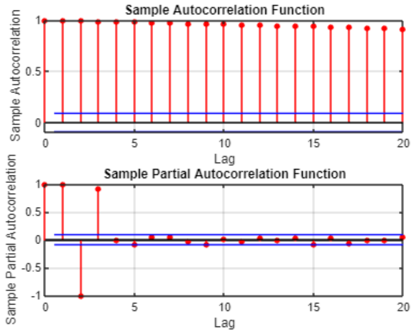


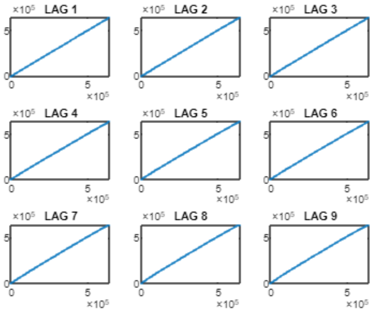




ARIMA(0,3,0)

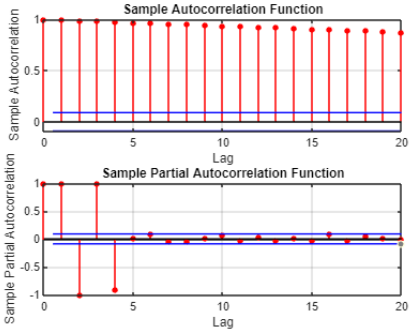




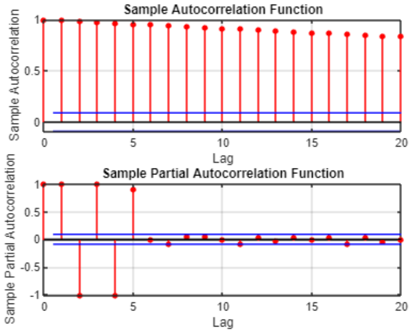


Modelli ARIMA(p,d,q)

ARIMA(0,4,0)



ARIMA(0,5,0)



Altri modelli non stazionari

Modelli SARIMA

Tengono conto della stagionalità

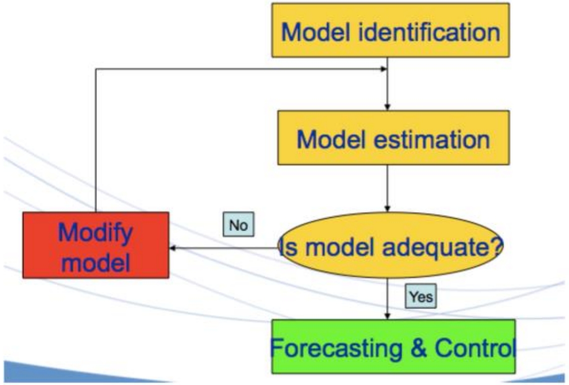
Modelli GARCH

Tengono conto della eteroschedaticità della serie

Selezione del processo/modello

Metodo Box-Jenkins, basato su tre fasi:
1) Identificazione del/dei possibili processi (analisi grafica, confronto ACF-PACF)
2) Stima (di più modelli) e check dei risultati (test su stazionarietà, goodness of fit, analisi dei residui)
3) Scelta del modello preferito

E' una procedura iterativa.



Identificazione

Necessario capire dai dati, quale è il modello più appropriato a descrive la serie osservata

Capire l'ordine del modello
Tale fase si basa sullo studio delle statistiche campionarie:
ACF (ρ_k), PACF (λ_k)

Ogni modello presenta un andamento teorico di ACF e PACF ben preciso. Sulla base della corrispondenza tra ACF e PACF campionaria e teorica e possibile abbinare l'appropriato modello ai dati osservati.

"The art of a time series analyst's model identification is very much like the method of an FBI agent's criminal search. Most criminals disguise themselves to avoid being recognized. This is also true of ACF and PACF" (Wei, 1990)

Stima modello

Si utilizza la stima MLE che richiede un'ipotesi sulla distribuzione dei termini di disturbo e procedure numeriche di massimizzazione (Algoritmo di Newton).

Disgnostic checking e criteri di selezione

Dopo aver stimato il modello, dobbiamo verificare la bontà del modello stesso.

- L'assunzione è che il residuo stimato sia un white noise
- No autocorrelazione
- Distribuzione Normale
- Se abbiamo più modelli possibili applicare criterio di selezione

Disgnostic checking e criteri di selezione

Una volta fatta la stima, si controlla se i residui sono white noise, quasi sempre col test di Ljung-Box o con statistiche equivalenti.
Un'altra classe di statistiche che si usano in questo contesto sono i cosiddetti criteri di informazione, come ad esempio quello di **Akaike** (spesso abbreviato in AIC) o quello di **Schwartz** (spesso abbreviato in BIC *Bayes Information Criterion*); l'uso di queste statistiche è motivato con concetti presi dalla teoria dell'informazione.
Fra due modelli, quello *migliore* dovrebbe avere un indice AIC o BIC più basso, in quanto tutti questi criteri possono essere scritti nella forma

$$C = -2L(\theta) + c(k, T)$$

dove $L(\theta)$ è la log-verosimiglianza calcolata in corrispondenza del valore stimato per il vettore dei parametri del modello, k è il numero di parametri stimati e T è l'ampiezza campionaria; la funzione $c(k, T)$ è crescente in k, per cui a parità di verosimiglianza viene scelto il modello più parsimonioso. Ad esempio, per il criterio di Schwartz, $c(k, T) = k \log(T)$.

Valutazione della bontà del modello e della sua capacità previsiva

Nel campo della analisi delle serie storiche a fini predittivi si può parlare di **due tipi di valutazione** riguardanti un modello scelto per rappresentare la serie di interesse.

In primo luogo, si possono stimare ($\hat{y}_t$), sulla base del modello scelto, i valori teorici della serie e confrontare i dati stimati con i valori osservati verificando come il modello riesce a riprodurre i dati storici. Si parla in questo caso di **goodness of fit** e in termini formali si ha che **l'errore di stima** è:

$$r_t = y_t - \hat{y}_t$$

In secondo luogo, interessa verificare come il modello stimato riesce a riprodurre i dati futuri, e si effettua perciò il confronto tra i valori futuri e le previsioni del fenomeno. In questo caso si misura la **goodness of forecast**, che in termini formali indica un **errore di previsione** (ep) al tempo $t + h$ e cioè:

$ep_{t+h} = y_{t+h} - F_{t+h}$

Ovviamente **l'errore di previsione**, come sopra definito, si può calcolare solo quando **i dati per i tempi futuri** saranno **disponibili**. Si pone perciò il problema di come si possa **valutare la goodness of forecast in anticipo**, fin dal tempo in cui si fa la previsione, per capire la capacità previsiva del modello.

Valutazione della bontà del modello e della sua capacità previsiva

In genere si avrà:

- una serie di dati disponibili: $y_1, y_2, \dots, y_N$
- una serie di dati precedenti al tempo n sui quali si adatta il modello da utilizzare per le previsioni: $y_1, y_2, \dots, y_m$ con $m < N$
- una serie di stime per i periodi da 1 a m, che si possono confrontare con i valori osservati $y_1, y_2, \dots, y_m$ per valutare il goodness of fit
- una serie di valori previsti: $F_{m+1}, F_{m+2}, \dots, F_N$ che possono essere confrontati con i valori osservati per lo stesso periodo $y_{m+1}, y_{m+2}, \dots, y_N$ fornendo la possibilità di valutare il goodness of forecast.
- La valutazione delle capacità previsive del modello è effettuata per il periodo passato e quindi nell'utilizzare il modello a fini previsivi occorre accettare l'ipotesi, o sperare, che gli errori di previsione abbiano la stessa intensità anche per il futuro.

Valutazione della bontà del modello e della sua capacità previsiva

Ovviamente interessa avere una valutazione sintetica degli **errori di stima (r_t)** e di quelli **di previsione (ep_t)** e a questo fine si utilizzano frequentemente le seguenti misure:

Errore medio (mean error): media aritmetica degli errori

$$ME = \frac{1}{m}\sum_{t=1}^{m} r_t \qquad ME = \frac{1}{N-m}\sum_{t=m+1}^{N} ep_t$$

Errore quadratico medio (*mean square error*): media aritmetica dei quadrati degli errori

$$MSE = \frac{1}{m}\sum_{t=1}^{m} r_t^2 \qquad MSE = \frac{1}{N-m}\sum_{t=m+1}^{N} ep_t^2$$

Errore medio assoluto (mean absolute error): media aritmetica degli errori presi in valore assoluto

$$MAE = \frac{1}{m}\sum_{t=1}^{m} |r_t| \qquad MAE = \frac{1}{N-m}\sum_{t=m+1}^{N} |ep_t|$$

Valutazione della bontà del modello e della sua capacità previsiva

Per evitare che le suddette misure dipendano dall'unità di misura della serie, gli errori possono essere trasformati in **errori relativi** (solitamente espressi in percentuale rispetto ai valori osservati), sui quali si calcolano le medie sopra indicate. In particolare, dal MAE si ottiene il MAPE, mean absolute percentage error):

$$MAPE = \frac{100}{m}\sum_{t=1}^{m}\frac{|r_t|}{y_t} \qquad MAPE = \frac{100}{N-m}\sum_{t=m+1}^{N}\frac{|ep_t|}{y_t}$$

UNIVERSITÀ
DEGLI STUDI
DI TRIESTE

Dipartimento di
Ingegneria
e Architettura