Serie Storiche

L'approccio moderno all'analisi delle serie storiche

Corso di Laurea in Statistica per le Tecnologie e le Scienze
Corso di Laurea in Statistica per l'Economia e l'Impresa
matricole dispari

Francesco Lisi

a.a. 2026/2027

Introduzione

- Processo generatore della serie storica osservata: $Y_t = f(t) + u_t$
- Nell'approccio classico l'attenzione è posta su $f(t)$ cioè sulla parte deterministica, u_t, la parte stocastica, assume invece un ruolo residuale
- Nell'approccio moderno, invece, si ipotizza che $f(t)$ sia assente o che sia stata rimossa e ci si concentra sulla parte stocastica, che ora assume un ruolo fondamentale
- Il modello di riferimento è ora $Y_t = u_t$, con u_t processo stocastico a componenti correlate
- La serie storica $\{y\}_{t=1}^{n}$ viene pensata come una realizzazione finita del processo stocastico che ha generato i dati.
- Dobbiamo definire cos'è è un processo stocastico.

Old faithful geyser

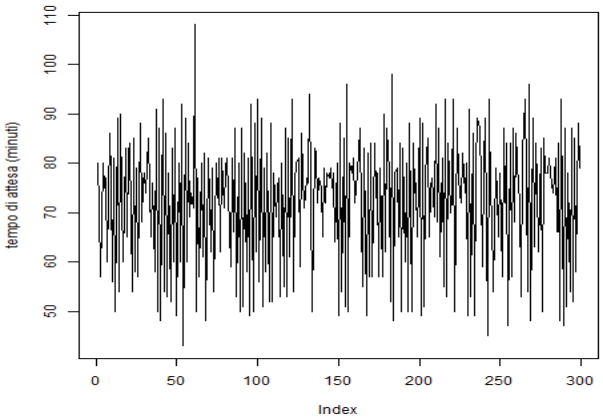


Figure: Old faithful: tempi di attesa (in minuti) tra due eruzioni consecutive.

Serie mensile delle precipitazioni piovose a Padova - stazione Orto Botanico

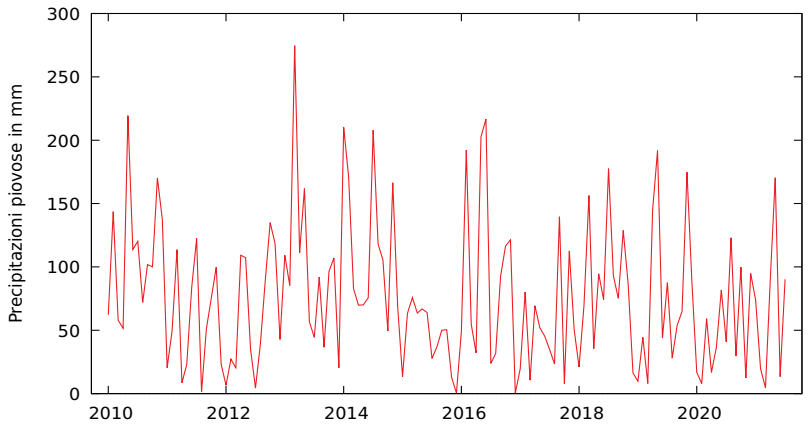


Figure: Dati mensili (gennaio 2010 - luglio 2020). Fonte: ARPAV

Processo stocastico

Siano:

- (Ω, F, P) spazio di probabilità
- Ω : spazio degli eventi elementari (insieme dei risultati possibili di un esperimento casuale)
- F : $\sigma-$algebra su Ω (insieme di insieme - gli eventi - per i quali si può calcolare una probabilità)
- P : misura di probabilità su Ω ($P : \Omega \to [0, 1]$)
- T : spazio parametrico

Un processo stocastico è una funzione finita a valori reali:

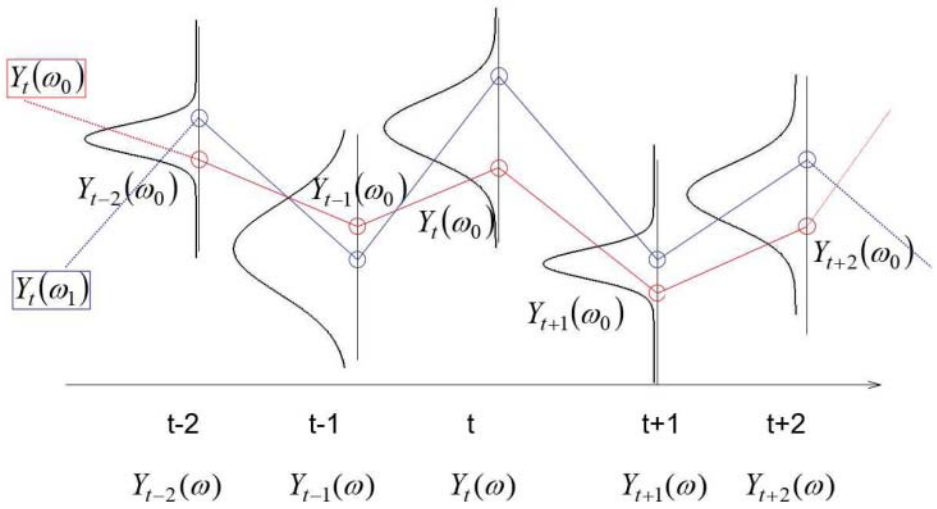
$$Y(t, \omega), \qquad t \in T, \qquad \omega \in \Omega$$

Processo stocastico

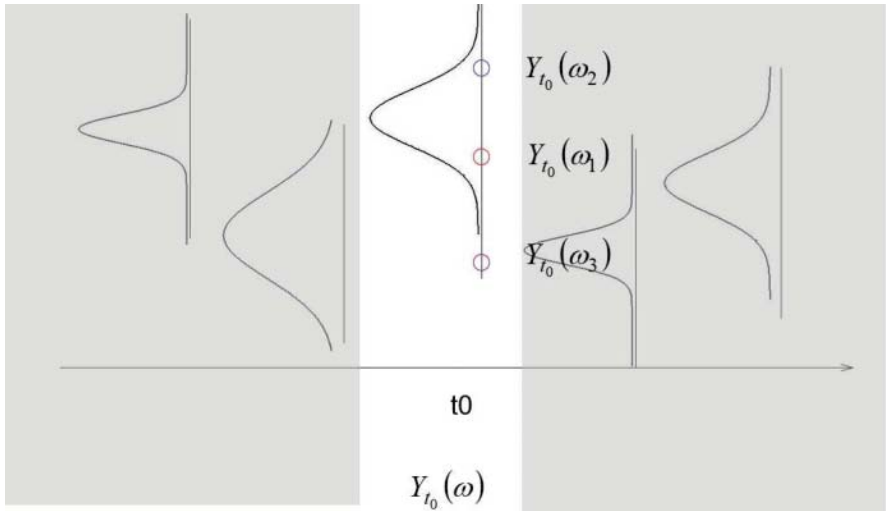
- $Y(t, \omega), \qquad t \in T, \qquad \omega \in \Omega$
- Per $t = t_0 \in T, \quad Y(t_0, \omega) \equiv Y_{t_0}(\omega)$ è una funzione misurabile di ω, cioè una variabile casuale.
- Per $\omega = \omega_0, \quad Y_t(\omega_0)$ è una realizzazione o traiettoria del processo stocastico
- Per $\omega = \omega_0, \quad t = t_0, \quad Y_{t_0}(\omega_0)$ è il valore di una variabile casuale

Più intuitivamente, si può anche dire che un processo stocastico è una successione di variabili casuali indicizzate da un parametro che di solito (ma non sempre) rappresenta il tempo.

Rappresentazione grafica di un p.s.

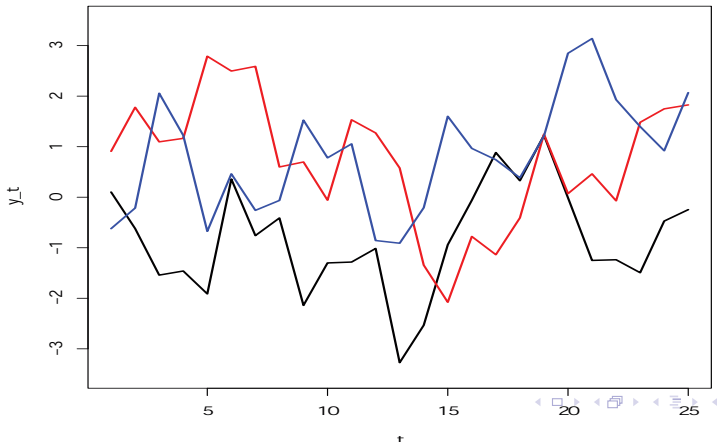


Rappresentazione grafica di un p.s.



Realizzazioni di un processo stocastico

- Consideriamo il PGD: $Y_t = 0.7Y_{t-1} + \varepsilon_t$ e tre diverse realizzazioni: $Y_1 = Y_t(\omega_1), \quad Y_2 = Y_t(\omega_2), \quad Y_3 = Y_t(\omega_3)$.
- Le tre serie, pur generate dallo stesso PGD, differiscono per l'evento elementare (ω_i) su cui opera il meccanismo casuale.



Caratterizzazione di un processo stocastico

- Un processo stocastico è caratterizzato dalle relazioni probabilistiche tra le v.c. che lo compongono

$$F(..., Y_{t-1}, Y_t, Y_{t+1}, ...)$$

- E' possibile marginalizzare la distribuzione per ogni sottoinsieme del processo, ottenendo le distribuzioni marginali. A rigore, per conoscere un processo stocastico sarebbe necessario conoscere tutte le distribuzioni di probabilità congiunta:

$$F(Y_{t_1}, Y_{t_2}, ..., Y_{t_n}) \qquad \forall (t_1, t_2, \ldots, t_n) \qquad \forall n$$

  In generale, ciò è molto difficile, se non impossibile.
- Se le distribuzioni marginali hanno i momenti si può scrivere, ad es.:

$$E(Y_t) = \mu_t, \qquad Var(Y_t) = \sigma_t^2, \qquad Cov(Y_t, Y_s) = \gamma_{t,s}$$

Caratterizzazione di un processo stocastico

- Allo stesso modo si possono definire le distribuzioni condizionate, ad es.

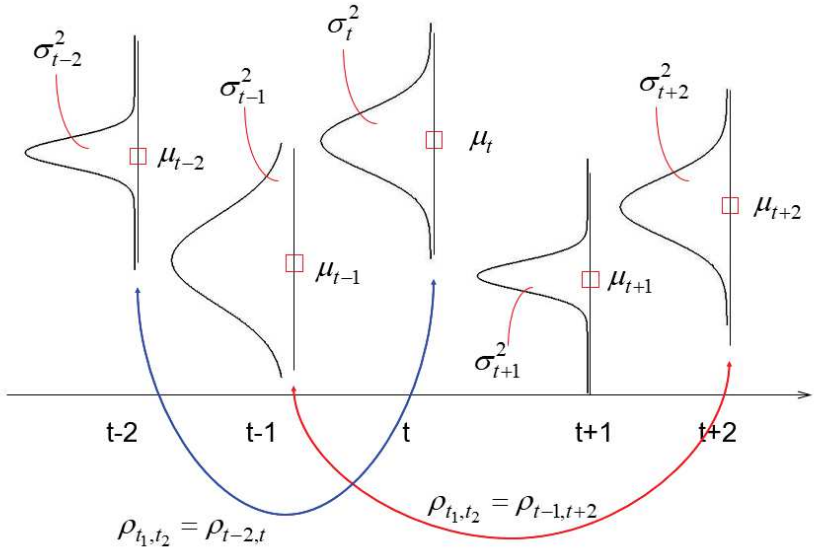
$$F(Y_t | Y_{t-1}, Y_{t-2}...)$$

- Anche per le distribuzioni condizionate si possono definire i corrispondenti momenti.

I momenti di un processo stocastico

- Un processo stocastico può anche essere caratterizzato dai suoi momenti.
- Una semplificazione è quella di considerare soltanto i primi due momenti:
  - funzione media: $\mu_t = E[Y_t]$
  - funzione varianza: $\sigma_t^2 = Var(Y_t) = E[(Y_t - \mu_t)^2]$
  - funzione di autocovarianza: $\gamma_{t,s} = E[(Y_t - \mu_t)(Y_s - \mu_s)]$
  - funzione di autocorrelazione (ACF): $\rho_{t,s} = \frac{\gamma_{t,s}}{\sigma_t \sigma_s}$

Rappresentazione grafica dei momenti di un p.s.



Alcune considerazioni sui momenti di un p.s.

- L'autocovarianza è espressa nel quadrato dell'unità di misura di Y.
- Consideriamo la covarianza tra v.c. dello stesso p.s. spaziate da uno sfasamento temporale (*lag*) $k = \mid t - s \mid$.
- Per $t = s$ si ha $k = 0$, che è la funzione varianza.
- La correlazione tra v.c. dello stesso processo stocastico viene detta autocorrelazione e rappresenta una misura di dipendenza lineare standardizzata.
- L'autocorrelazione è un numero puro, cioè non dipende dall'unità di misura.

Alcune considerazioni sui momenti di un p.s.

- Anche se si considerano solo i primi due momenti, caratterizzare un processo nella sua accezione più generale è troppo difficile.
- Per poter lavorare bisogna richiedere ulteriori condizioni di regolarità che semplifichino le cose.
- Condizioni di stazionarietà.

Processi stazionari in senso stretto

- Stazionarietà forte (o in senso stretto) - Un processo stocastico viene detto stazionario in senso stretto se

$$F(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}) = F(Y_{t_1+\tau}, Y_{t_2+\tau}, \ldots, Y_{t_n+\tau})$$

$$\forall (t_1, t_2, \ldots, t_n), \ \forall \tau \text{ e } \forall n$$

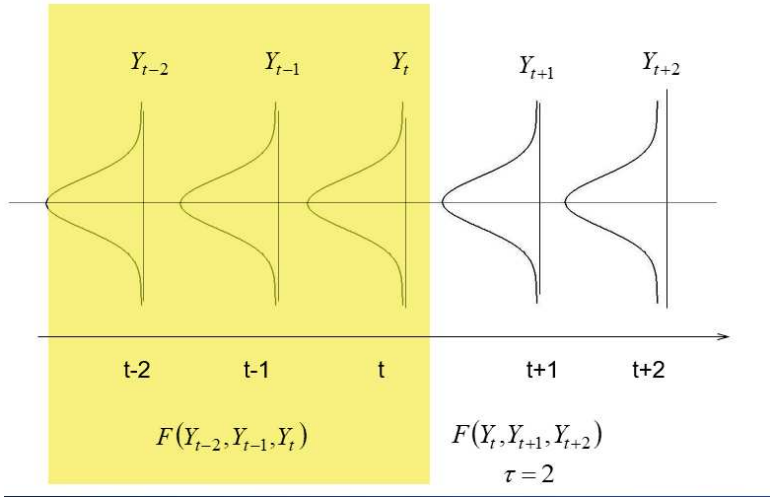
- In particolare, se $n = 1$, la stazionarietà forte implica che media e varianza, se esistono, non dipendono da t, ossia:

$$\mu_t = \mu \quad \forall t \qquad\qquad \sigma^2_t = \sigma^2 \quad \forall t$$

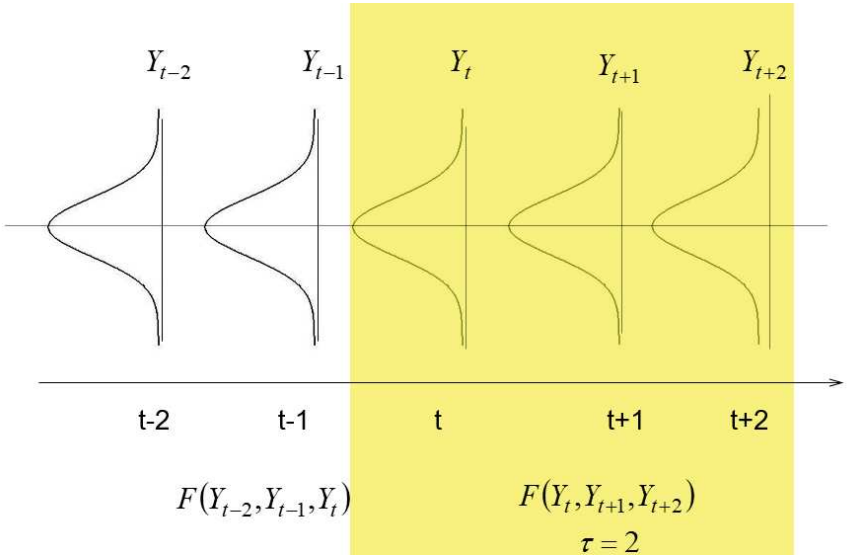
- La funzione di autocovarianza dipende solo dalla distanza temporale (per istanti temporali equispaziati), cioè:

$$\gamma_{t,s} = \gamma_{|t-s|} \qquad \Rightarrow \qquad \gamma_{t,t-k} = \gamma_k, \quad k = 0, \pm 1, \pm 2, \ldots$$

Stazionarietà forte



Stazionarietà forte



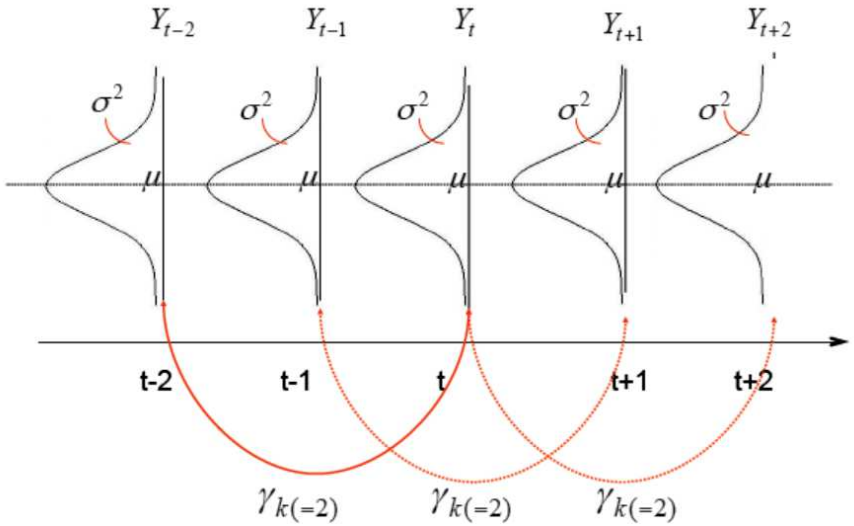
Stazionarietà debole (o del secondo ordine)

- La condizione di stazionarietà in senso stretto impone vincoli sull'intera distribuzione del p.s.
- In pratica, è utile definire la stazionarietà in modo meno restrittivo di quello descritto.
- Un p.s. Y_t, è detto debolmente stazionario (o stazionario del secondo ordine) se:

  i. $E[Y_t] = \mu < \infty \qquad \forall t$
  ii. $Var(Y_t) = \sigma^2 < \infty \qquad \forall t$
  iii. $Cov(Y_t, Y_{t-k}) = \gamma_k \qquad \forall t, \forall k$

  Nessuna assunzione viene fatta sui momenti di ordine superiore.

Stazionarietà debole



Rapporto tra stazionarietà forte e debole

- Se due v.c. hanno la stessa distribuzione, hanno momenti (se esistono) uguali, quindi:

  stazionarietà forte $\Rightarrow$ stazionarietà debole
- In generale non vale l'implicazione in senso inverso
- Per processi gaussiani, cioè p.s. tali che:

$$(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_k}) \sim \mathcal{N}_k$$

  stazionarietà in senso debole e in senso forte coincidono
- D'ora in poi si useranno i termini *stazionarietà* e *stazionario* come sinonimi di *stazionarietà debole* e di *stazionario in senso debole.*

La funzione di autocorrelazione (globale)

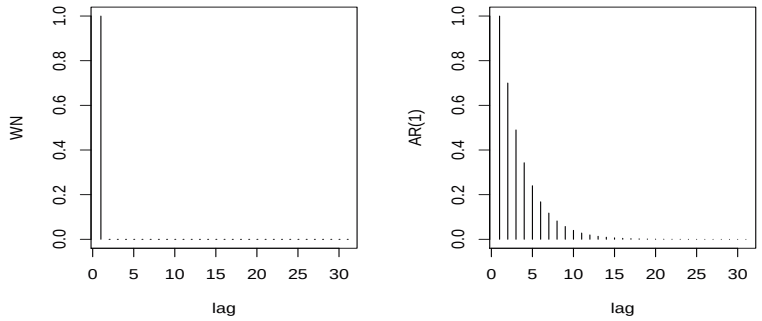
- Supponiamo che un p.s. stazionario Y_t abbia media μ, varianza σ^2 e funzione di autocovarianza γ_k con $k = 0, \pm 1, \pm 2, \ldots$
- La funzione di autocorrelazione è definita come:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_k}{\sigma^2}, \quad k = 0, \pm 1, \pm 2, \ldots$$

- Proprietà della funzione di autocorrelazione:
  1. $\rho_0 = 1$
  2. $\rho_k = \rho_{-k}$, cioè l'ACF è una funzione pari
  3. $\mid \rho_k \mid \leq 1$
- Date le proprietà 1 e 2 (simmetria) è sufficiente conoscere la funzione di autocorrelazione solo per $k = 1, 2, \ldots$.

Il correlogramma

- Il correlogramma è il grafico della funzione di autocorrelazione $\rho_k, \quad k = 0, 1, 2, \ldots$



ACF teoriche per un processo WN e AR(1)

La funzione di autocorrelazione parziale

- La funzione di autocorrelazione parziale è definita come

$$P_k = Corr(Y_t, Y_{t-k} \mid Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1}),$$

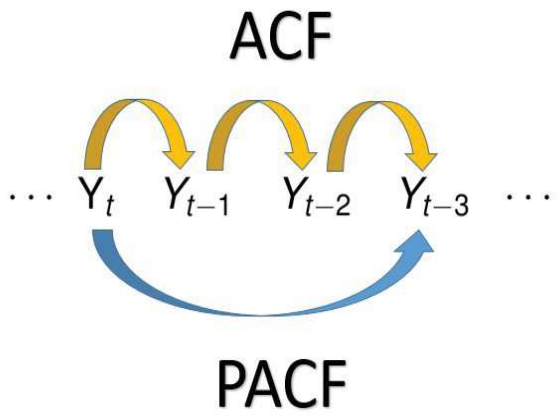
  e misura la correlazione fra Y_t e Y_{t-k} al netto delle variabili intermedie $Y_t, \underbrace{Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1}}, Y_{t-k}$
- **Esempio:** si consideri il processo $Y_t = \phi Y_{t-1} + \varepsilon_t$
  - Y_t dipende da Y_{t-1}
  - Ma, dato che

$$Y_t = \phi Y_{t-1} + \varepsilon_t = \phi(\phi Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \phi^2 Y_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t$$

$$Y_t = \phi^2(\phi Y_{t-3} + \varepsilon_{t-2}) + \phi\varepsilon_{t-1} + \varepsilon_t = \phi^3 Y_{t-3} + \phi^2\varepsilon_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t$$

  Y_t dipende anche da Y_{t-2}, Y_{t-3},...

La funzione di autocorrelazione parziale



La funzione di autocorrelazione parziale

- P_k è una funzione delle prime k correlazioni semplici:

$$P_k = \frac{\begin{vmatrix} 1 & \rho_1 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_1 & 1 \end{vmatrix}}$$

Calcolo del determinante

- Utilizzando lo sviluppo di Laplace per righe, per una matrice quadrata A $n \times n$, e in riferimento alla riga $i-$esima si ha

$$|A| = \sum_{j=1}^{n} a_{i,j}(-1)^{i+j} \cdot |A_{i,j}|$$

dove:

$a_{i,j}$ = elemento di posto (i, j) della matrice A;
$A_{i,j}$ = matrice ottenuta cancellando la riga $i-$esima e la colonna $j-$esima della matrice A.

Calcolo della PACF

- k=1 $\quad P_1 = \rho_1$
- k=2

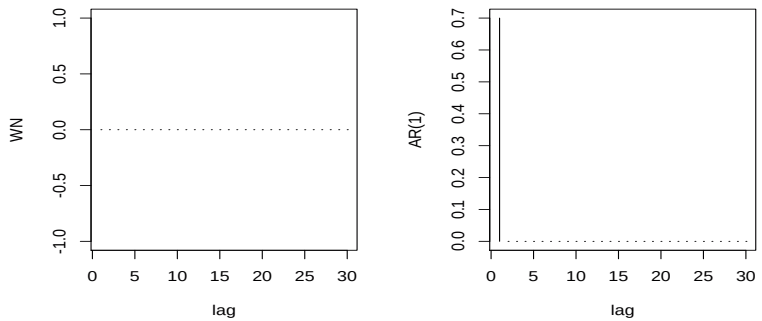
$$P_2 = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

- k=3

$$P_3 = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\rho_3 - \rho_2\rho_1 - \rho_1^2\rho_3 + \rho_1\rho_2^2 + \rho_1\rho_2 + \rho_2^2}{1 - 2\rho_2^2 + 2\rho_1^2\rho_2 - \rho_2^2}$$

Il correlogramma parziale

- Il correlogramma parziale è il grafico della funzione di autocorrelazione parziale $P_k, \quad k = 1, 2, \ldots$



ACF parziali teoriche per un processo WN e AR(1)

Esercizi

1. Verificare se i seguenti processi sono (debolmente) stazionari:
   i. $Y_t = \varepsilon_t - 0.5\varepsilon_{t-1} - 0.2\varepsilon_{t-2} \qquad \varepsilon_t \sim WN(0, 4)$
   ii. $Y_t = (-1)^t(1 - \theta B)\varepsilon_t \qquad \varepsilon_t \sim WN(0, \sigma^2) \qquad |\theta| < 1$

2. Verificare per quali valori di α il seguente processo è (debolmente) stazionario

$$Y_t = t^\alpha(1 - \theta B)\varepsilon_t \qquad \varepsilon_t \sim WN(0, \sigma^2) \qquad |\theta| < 1 \qquad \alpha \in \mathcal{R}$$

Problema

- Per ogni processo stocastico stazionario, le funzioni γ_k e ρ_k sono determinate in modo univoco.
- Il contrario è vero? Cioè: data una funzione di autocovarianza è unico il processo stazionario che la possiede? Purtroppo no!
- Si può dimostrare che esistono più p.s. stazionari con la stessa funzione di autocovarianza. Uno solo, però, invertibile.

Invertibilità

- Un processo stocastico Y_t è invertibile se esiste una funzione lineare h tale che

$$Y_t = h(Y_{t-1}, Y_{t-2}, \ldots) + \varepsilon_t$$

  con $\varepsilon_t \sim WN(0, \sigma^2)$
- Per processi stazionari e invertibili si ha univocamente:

$$\{Y_t\} \iff \gamma_k, \rho_k$$

Stima dei momenti di un processo stocastico stazionario

- Si desidera fare inferenza sui parametri del processo stocastico partendo da

$$y = (y_1, y_2, \ldots, y_n)$$

- Se il processo stocastico è stazionario, e sotto condizioni di *ergodicità*, è possibile stimare in modo consistente i suoi momenti

$$\mu, \quad \gamma_k, \quad \rho_k, \quad P_k$$

Consistenza

- Uno stimatore $\hat{\theta}_n$ si dice consistente se converge in probabilità, ovvero se:

$$\lim_{n\to\infty} Pr(|\ \hat{\theta}_n - \theta\ | > \varepsilon) = 0$$

- Una condizione sufficiente per la consistenza è che

1. $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$ e
2. $\lim_{n\to\infty} Var(\hat{\theta}_n) = 0$

Ergodicità - 1

- Proprietà in base alla quale è possibile stimare tutte le caratteristiche di un processo a partire da una sola realizzazione dello stesso.
- L'ergodicità permette di sostituire medie spaziali con medie temporali
- Condizione necessaria affinché un p.s. sia ergodico è che il p.s. sia stazionario almeno in senso debole.



- Quindi, solo i processi stazionari possono essere processi ergodici ma, essendo una condizione solo necessaria, non è detto che un processo stazionario sia ergodico.

Ergodicità - 2

- Esistono processi che sono ergodici relativamente ad una sola caratteristica (ad es. la media)
- Esistono quindi condizioni specifiche di ergodicità rispetto alle varie caratteristiche

Stima della media di un p.s. stazionario

- Uno stimatore della media corretto e consistente è dato da

$$\hat{\mu} = \bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$$

Per esso si ha che:

- $E[\hat{\mu}] = \mu$
- $Var(\hat{\mu}) = \frac{\sigma^2}{n}\left[1 + 2\sum_{r=1}^{n-1}\left(1 - \frac{r}{n}\right)\rho_r\right]$
- Se $\sum_{r=1}^{\infty}\rho_r < \infty \Longrightarrow Var(\hat{\mu}) \to 0$
- $lim_{r\to\infty}\rho_r = 0$ è la condizione di ergodicità rispetto alla media

Stima dell'autocovarianza (di un p.s. stazionario)

- Uno stimatore, distorto, ma asintoticamente corretto e consistente della covarianza è dato da:

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n-k}(y_t - \hat{\mu})(y_{t+k} - \hat{\mu})$$

per il quale si ha che:

$$E[\hat{\gamma}_k] = \gamma_k - \frac{k}{n}\gamma_k - \frac{n-k}{n}Var(\hat{\mu}).$$

- Stime della varianza e della funzione di autocorrelazione sono, rispettivamente da:

$$\hat{\sigma}^2 = \hat{\gamma}_0 = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{\mu})^2 \quad \text{e} \quad \hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$$

Approssimazione di Bartlett

- Approssimazione di Bartlett per la varianza di $\hat{\rho}_k$ in processi stocastici gaussiani

$$Var(\hat{\rho}_k) \cong \frac{1}{n}\sum_{i=-\infty}^{\infty}(\rho_i^2 + \rho_{i-k}\rho_{i+k} + 2\rho_i^2\rho_k^2 - 4\rho_i\rho_k\rho_{i-k})$$

- Se $\rho_k \cong 0$ per $k > q \quad \Rightarrow \quad Var(\hat{\rho}_k) \cong \frac{1}{n}\sum_{i=-q}^{q}\rho_i^2$
- Lo stimatore di P_k, $\hat{P}_k$, si ottiene sostituendo $\hat{\rho}_k$ a ρ_k nell'espressione della correlazione parziale.
- Vale, inoltre, l'approssimazione di Quenouille:

$$Var(\hat{P}_k) \cong \frac{1}{n}$$