IDENTIFICAZIONE DEI MODELLI E ANALISI DEI DATI (Prof. M. Mazzoleni)
Appello del 14 Febbraio 2018

Cognome	Nome	Matricola
....	....................................………	..................…….……

Verificare che il fascicolo sia costituito da 6 pagine.

Scrivere le risposte ai singoli esercizi negli spazi che seguono ogni domanda.

Non consegnare fogli addizionali.

Non si possono consultare libri, appunti, dispense etc.

La **chiarezza e l'ordine** nelle risposte saranno oggetto di valutazione

==

1. Si consideri il processo stocastico generato dalla seguente equazione:

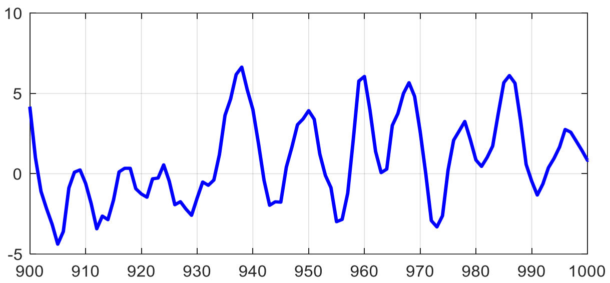
$$y(t) = \frac{3}{2}y(t-1) - \frac{3}{4}y(t-2) + e(t),\ \ e(t) \sim WN(0,1)$$

a) Dire perché il processo stocastico $y(t)$ è stazionario.

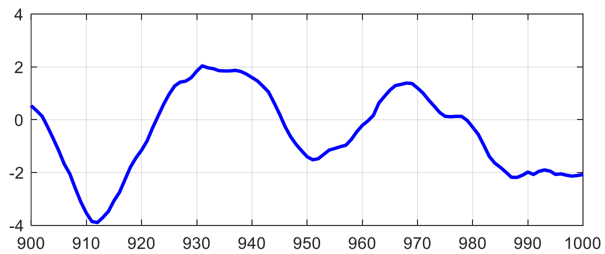
b) Calcolare media m_y e varianza $\gamma_y(0)$ di $y(t)$.

c) Dire, motivando la risposta, quale tra le seguenti realizzazioni temporali potrebbe corrispondere al processo $y(t)$.

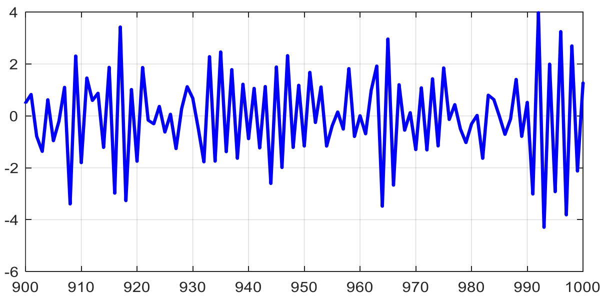
(Suggerimento: si calcoli lo spettro per i seguenti valori di omega $\omega = 0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{2}, \pi$)



A



B

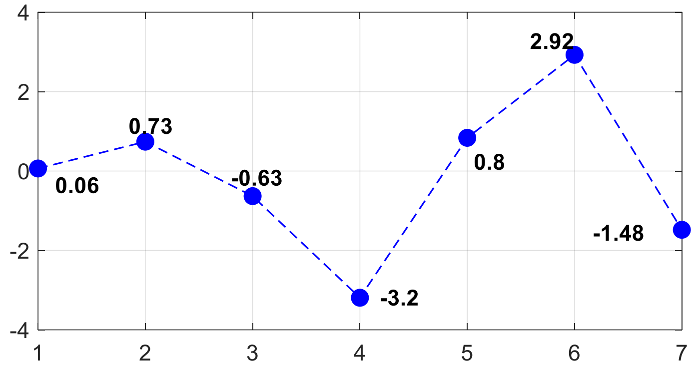


C

2. Si consideri il processo ARMA a media nulla:

$$y(t) = -0.3y(t-2) + e(t) - 2e(t-2), \qquad e(t) \sim WN(0,1)$$

Nella figura sottostante, sono rappresentati i primi 7 campioni di una realizzazione del processo $y(t)$.



a) Determinare il predittore ottimo ad un passo dai dati.

b) Si calcoli la predizione ottima ad un passo dell'ottavo valore della realizzazione, ovvero $\hat{y}(8|7)$.

c) Si valuti la varianza dell'errore di predizione ad un passo.

3. Si supponga di avere a disposizione N dati di un processo stocastico stazionario $y(t)$. Come famiglia di modelli per la descrizione della dinamica di $y(t)$, si consideri la famiglia AR(1):

$$M(\theta): y(t) = ay(t-1) + e(t)$$

Denotando con $\hat{a}(N)$ la stima a minimi quadrati del parametro a, calcolare il valore a cui tende $\hat{a}(N)$ per N tendente all'infinito, nei due casi seguenti:

a) I dati sono generati dal sistema $S: y(t) = 0.3y(t-1) + \eta(t)$ $\qquad \eta(t) \sim WN(0,1)$

b) I dati sono generati dal sistema $S: y(t) = 0.3y(t-1) + \eta(t) + 0.5\eta(t-1)$ $\qquad \eta(t) \sim WN(0,1)$

4. Descrivere il metodo della massima verosimiglianza per la stima di modelli ARMAX.

5. Si consideri un semplificato gioco del baseball. Solo un numero ristretto delle squadre partecipano ai playoff a fine stagione. Per partecipare, bisogna vincere più di 95 partite durante la stagione, facendo più punti dell'avversario. Vi sono due modi per fare punti durante una partita: o ottenere dei "runs", oppure non concederli all'avversario. Forti battitori fan si che sia più facile ottenere dei runs, mentre forti lanciatori fanno sia che l'avversario non li ottenga (si veda il seguente schema):

Si supponga di avere a disposizione diversi anni di risultati e statistiche della squadra X. In particolare, si hanno le seguenti variabili:

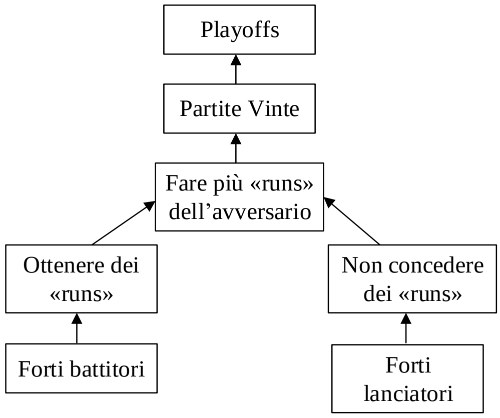
- Partite vinte: W
- Runs effettuati: E
- Runs concessi: C

Le seguenti variabili influenzano il numero di runs effettuati:

- Percentuale di volte che un giocatore va alla "base": P
- Strada percorsa dai giocatori in media: S

Le seguenti variabili influenzano il numero di runs concessi:

- Percentuale di volte che un avversario va alla "base": AP
- Strada percorsa dagli avversari in media: AS



Si ipotizzi un modello di machine learning per prevedere se la squadra X andrà ai playoff l'anno prossimo, usando i dati ottenuti alla fine della stagione di quest'anno.

6. Si dimostri che, se il sistema da identificare appartiene alla classe dei modelli e il minimo globale della cifra di merito è unico, tale minimo globale corrisponde al sistema che genera i dati.