Identificazione dei Modelli e Analisi dei Dati MN

Prof. G. De Nicolao

30 Giugno 2010

Cognome .... **Nome**..............................

Matricola.............................. **Firma**..............................

- Compilare a penna questo foglio all'inizio della prova.
- Durante lo svolgimento della prova, non è consentito l'uso di materiale diverso dai comuni strumenti di calcolo, scrittura e disegno.
- Le risposte devono essere scritte in modo chiaramente leggibile nello spazio immediatamente seguente ogni domanda (se necessario, a seguito di cancellature, passare sul retro).
- Le uniche risposte valide sono quelle riportate nel presente fascicolo, che va consegnato, senza fogli addizionali, al termine della prova.

1.	
2.	
3.	
4.	

1. Dire se le seguenti affermazioni sono vere o false: (Punteggio: risposta esatta =1, errore=-1, non risponde =0).

V F

(a) Se A e B sono eventi indipendenti, allora $P(A+B) = P(A) + P(B)$.

☐ ☐

(b) Per una moneta onesta, la probabilità di ottenere 2 teste su 4 lanci è 3/8

☐ ☐

(c) Sia $Z \sim N(0,1)$. Allora, $E[Z^2] = 1$.

☐ ☐

(d) Sia $Y = aX + b$, $a \neq 0$. Allora, $f_Y(y) = |a| f_X((y-b)/a)$

☐ ☐

(e) La media e la mediana della V.C esponenziale coincidono.

☐ ☐

2. Siano X e Y le coordinate di un punto scelto in modo equiprobabile nel cerchio di raggio unitario centrato nell'origine e si definisca $Z = X^2 + Y^2$

   (a) Ricavare, riportando i passaggi, l'espressione di $F_Z(z)$.

   (b) Ricavare, riportando i passaggi, l'espressione di $f_Z(z)$ e disegnarne il grafico.

3. Il seguente codice MATLAB calcola le medie, la matrice di covarianza e il coefficiente di correlazione per le variabili casuali W de H (peso e altezza), rispettivamente. I vettori colonna h e w contengono altezze e pesi di soggetti estratti a caso dalla popolazione. Trovare gli eventuali errori presenti nel codice e correggerli.

```
medie=mean([h;w]);

EW=medie(1);

EH=medie(2);

cov_WH=cov([h;w]);

corr_WH=cov_WH(1,2)/(cov_WH(2,2)*cov_WH(1,1));
```

4. Si considerino i seguenti dati

$$\begin{array}{lll} y(1) = 0 & y(2) = 0 & y(3) = 2 \\ u(1) = -2 & u(2) = 0 & u(3) = 2 \end{array}$$

Si ipotizza che i dati siano generati dal seguente modello

$$y(t) = \theta u(t) + v(t), t = 1, 2, 3$$

dove $v(t)$ sono errori di misura i.i.d. $v(t) \sim N(0, \sigma^2)$, σ^2 incognita.

(a) Calcolare la stima di Gauss-Markov di θ.

(b) Stimare σ^2.

(c) Calcolare l'intervallo di confidenza per θ.

(d) Si consideri il modello alternativo (solo rumore).

$$y(t) = v(t), t = 1, 2, 3$$

ed i dati di validazione

$$\begin{array}{lll} y^V(1) = 1 & y^V(2) = -1 & y^V(3) = 0 \\ u^V(1) = -2 & u^V(2) = 0 & u^V(3) = 2 \end{array}$$

Dire quale modello è preferibile in base ai dati di validazione.

5. Dire se le seguenti affermazioni sono vere o false: (Punteggio: risposta esatta =1, errore=-1, non risponde =0)

V F

(a) Il prodotto di due V.C. congiuntamente gaussiane è ancora gaussiano.

□ □

(b) Date X ed Y congiuntamente gaussiane, risulta che $f_{X|Y}(x|Y = y)$ è una ddp gaussiana.

□ □

(c) Se due V.C incorrelate sono anche indipendenti, allora sono congiuntamente gaussiane.

□ □

(d) La varianza della media campionaria M_1 può essere maggiore della media della varianza campionaria S^2.

□ □

(e) La consistenza dei momenti campionari di V.C. i.i.d. è una conseguenza della Legge dei Grandi Numeri.

□ □

(f) Per V.C. i.i.d. gaussiane, la media campionaria è lo stimatore che minimizza l'errore quadratico medio.

□ □

(g) Uno stimatore si dice asintoticamente non polarizzato se $\lim_{n\to\infty} E[\hat{\theta}] = \theta^o$.

□ □

(h) Sotto l'ipotesi $I1$, se σ^2 non è nota, allora θ^M non è gaussiana.

□ □

(i) La somma dei quadrati dei residui in crossvalidazione può essere inferiore a quella ottenuta sui dati di identificazione.

□ □

(j) Se due colonne di Φ sono tra loro proporzionali, la condizione di identificabilità non è soddisfatta.

□ □