Esame Scritto del Primo Appello

Tempo a disposizione: 2 ore

Le soluzioni devono includere il procedimento dettagliato che porta alle risposte. Risposte corrette non adeguatamente motivate saranno penalizzate. La soluzione di ogni esercizio deve essere scritta in modo chiaro e ordinato, risposte illeggibili non saranno valutate. Le soluzioni degli esercizi devono essere riportate sul foglio protocollo nell'ordine proposto, la soluzione di ogni esercizio deve iniziare in una nuova pagina. Non è permesso l'uso di appunti, manuali o materiale didattico di alcun tipo. Non è permesso l'uso di dispositivi elettronici ad esclusiva eccezione di una calcolatrice non programmabile.

1. Tre programmatori collaborano a un progetto. Anita scrive un programma che funziona correttamente. Anche Bartolo scrive un programma, ma non lo prova, ed esso può contenere un errore con probabilità $p \in [0,1]$. Corinna riceve il codice dai colleghi, sceglie uno dei due programmi a caso e ne verifica il funzionamento. Se il programma testato da Corinna funziona correttamente, qual è la probabilità che anche il programma *non* testato da Corinna funzioni correttamente?
   Introduciamo gli eventi B "il programma di Bartolo funziona", $\mathbb{P}(B) = 1-p$, e A "Corinna estrae il programma di Anita", $\mathbb{P}(A) = 1/2$. Gli eventi A e B sono assunti indipendenti. L'evento "il programma estratto funziona" è dato da $A \cup (A^c \cap B)$, mentre l'evento "anche il programma non testato funziona correttamente" semplicemente coincide con B.
   Calcoliamo dunque,
   $$\mathbb{P}(A \cup (A^c \cap B)) = \mathbb{P}(A) + \mathbb{P}(A^c \cap B) = \frac{1}{2} + \frac{1-p}{2},$$
   $$\mathbb{P}(B \mid A \cup (A^c \cap B)) = \frac{\mathbb{P}(B \cap (A \cup (A^c \cap B)))}{\mathbb{P}(A \cup (A^c \cap B))} = \frac{\mathbb{P}(B)}{\mathbb{P}(A \cup (A^c \cap B))} = \frac{1-p}{\frac{1}{2} + \frac{1-p}{2}} = \frac{2-2p}{2-p}.$$

2. Una unità di *parallel computing* ha n processori, ad ognuno dei quali è associata una unità di memoria (*distributed memory model*). Dobbiamo assegnare k compiti all'unità, ma controllare quali processori sono già occupati impiega tempo: si sceglie di assegnare ogni compito a un processore scegliendo quest'ultimo casualmente, in modo uniforme tra tutti gli n presenti (*random allocation*). Qual è il valore atteso del numero di processori a cui non viene assegnato alcun compito?
   Suggerimento: per ogni processore si consideri una variabile X_j che assume valore 0 se al processore viene assegnato almeno un compito, e 1 altrimenti.
   Indichiamo con $X_1, \dots, X_n : \Omega \to \{0,1\}$ variabili aleatorie (di Bernoulli) tali che X_j assume valore 0 se al processore j viene assegnato almeno un compito, e valore 1 altrimenti. Siamo dunque interessati alla variabile aleatoria $Y = X_1 + \dots + X_n$, per la quale vale
   $$\mathbb{E}[Y] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n].$$

Siamo dunque ricondotti a calcolare i valori attesi delle singole X_j, ma per ogni $j = 1, \dots, n$ vale

$$\mathbb{E}[X_j] = 1 \cdot \mathbb{P}(X_j = 1) + 0 \cdot \mathbb{P}(X_j = 0) = \mathbb{P}(X_j = 1) = \left(\frac{n-1}{n}\right)^k,$$

in cui l'ultimo passaggio segue dal fatto che un processore non ha compiti se ogni assegnazione (indipendente) di un compito ha scelto uno degli altri processori, con probabilità $\frac{n-1}{n}$. La risposta è dunque

$$\mathbb{E}[Y] = n\left(\frac{n-1}{n}\right)^k.$$

3. Si consideri la seguente funzione,

$$f(x) = \frac{e^{-x/\sigma}}{\sigma(1+e^{-x/\sigma})^2},$$

dipendente dal parametro $\sigma > 0$.

(a) Si mostri che, per ogni $\sigma > 0$, f è una densità di probabilità, e si calcoli la funzione di ripartizione di una variabile aleatoria con tale densità.
La funzione è chiaramente positiva, per mostrare che è una densità di probabilità basta mostrare che integra ad 1:

$$\int_{-\infty}^{\infty} \frac{e^{-x/\sigma}dx}{\sigma(1+e^{-(x-\mu)/\sigma})^2} = \int_0^{\infty} \frac{y}{\sigma(1+y)^2} \cdot \frac{\sigma}{y} dy = \int_0^{\infty} \frac{dy}{(1+y)^2} = \left[-\frac{1}{1+y}\right]_0^{\infty} = 1,$$

in cui abbiamo usato la sostituzione $y = e^{-x/\sigma}$. Con lo stesso calcolo si ottiene:

$$F(t) = \int_{-\infty}^{t} f(t)dt = \int_{\exp(-(t-\mu)/\sigma)}^{\infty} \frac{dy}{(1+y)^2} = \left[-\frac{1}{1+y}\right]_{\exp(-t/\sigma)}^{\infty} = \frac{1}{1+e^{-x/\sigma}}.$$

(b) Si calcoli il valore atteso di una variabile aleatoria di densità f.
Suggerimento: si ricordi che l'integrale su tutto $\mathbb{R}$ di una funzione dispari (integrabile) è nullo.
Usiamo la sostituzione $y = \frac{x}{\sigma}$:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{\infty} \frac{\sigma y e^{-y} dy}{(1+e^{-y})^2} = 0,$$

perché la funzione integranda è dispari.

(c) Sia X una variabile aleatoria con densità $f_X = f$. Si determini se la variabile aleatoria $Y = e^X$ ha densità, e in caso positivo la si calcoli.
La variabile X può assumere qualsiasi valore reale, e l'esponenziale $x \mapsto e^x$ è una

bigezione differenziabile tra $\mathbb{R}$ e $(0, \infty)$, per cui possiamo applicare la formula di cambio di variabile. Abbiamo dunque, posto $h(x) = e^x$,

$$f_Y(y) = \begin{cases} f_X(h^{-1}(y)) \left| \frac{dh^{-1}}{dy}(y) \right| & y > 0 \\ 0 & y \leq 0 \end{cases},$$

in cui, usando $h^{-1}(y) = \log y$ e sviluppando i calcoli,

$$f_X(h^{-1}(y)) \left| \frac{dh^{-1}}{dy}(y) \right| = f_X(\log y) \left| \frac{1}{y} \right| = \frac{y^{-1/\sigma}}{y\sigma(1 + y^{-1/\sigma})^2}, \quad y \geq 0.$$

4. Si vuole determinare la durata (in ore) di un certo modello di batteria per PC sotto stress. Vengono testati 16 esemplari di batteria, ottenendo (dai dati del campione) una media di 5.2. Supponiamo che la distribuzione della durata delle batterie segua una distribuzione Gaussiana, con deviazione standard pari a 0.5.

   (a) Determinare un intervallo di fiducia di livello 95% per la durata media della batteria.
   Stiamo cercando un intervallo di fiducia per la media di una popolazione gaussiana, deviazione standard nota $\sigma = 0.5$. L'invervallo di fiducia cercato è ($n = 16$, $\alpha = 0.05$)

   $$\left[\bar{X} \pm \frac{\sigma}{\sqrt{n}} q_{1-\alpha/2} \right] = \left[\bar{X} \pm 0.245 \right].$$

   Inserendo il valore $\bar{x} = 5.2$, troviamo $[4.955, 5.425]$.

   (b) Qual è la taglia minima per avere una precisione della stima della media (cioè metà dell'ampiezza dell'intervallo per la media) inferiore a 0.1, con fiducia al 95%?
   La precisione della stima è pari a

   $$d = \frac{\sigma}{\sqrt{n}} q_{0.975} = \frac{0.98}{\sqrt{n}}.$$

   Imponendo $d \leq 0.1$, otteniamo $n \geq 96.04$, quindi $n \geq 97$.

   Supponiamo ora di non avere informazioni di alcun tipo sulla distribuzione della durata della batteria. Vengono testati questa volta 144 esemplari, ottendendo (dai dati del campione) una media di 5.3 e una deviazione standard di 0.4.

   (c) Determinare (eventualmente in modo approssimato) un intervallo di fiducia di livello 95% per la durata media della batteria.
   Stiamo cercando un intervallo di fiducia per la media di una popolazione, varianza non nota, con grande taglia del campione. L'invervallo di fiducia approssimato cercato è ($n = 144$, $\alpha = 0.05$)

   $$\left[\bar{X} \pm \frac{S}{\sqrt{n}} q_{1-\alpha/2} \right] = \left[\bar{X} \pm 0.163 \cdot S \right].$$

   Inserendo $\bar{x} = 5.3$, $s = 0.4$, otteniamo $[5.2348, 5.3652]$.

Valori dei quantili che possono essere utili:

$$q_{0.95} = 1.64, \quad q_{0.975} = 1.96, \quad q_{0.995} = 2.58$$
$$t_{15,0.95} = 1.75, \quad t_{15,0.975} = 2.13, \quad t_{16,0.95} = 1.75, \quad t_{16,0.975} = 2.12$$