

*Si svolgano 3 esercizi a scelta sui 4 proposti.*

*Il punteggio finale sarà la somma dei punti dei 3 esercizi riusciti meglio.*

**Problema 3.1** (12 punti). Le bustine di zucchero di una scatola hanno un contenuto medio di 5.5 g e una deviazione standard di 0.87 g.

**(7 punti)** Sia  $S$  il peso totale di zucchero contenuto in 40 bustine prese a caso. Quanto valgono media e deviazione standard di  $S$ ? Quanto vale la probabilità che  $S$  sia inferiore a 200 g? E che sia superiore a 250 g?

**(3 punti)** Un'altra scatola contiene bustine di zucchero di 3 tipi: un 20% di bustine da 4 g, un 50% di bustine da 5 g e un 30% di bustine da 6 g. Si pesca a caso una bustina, e sia  $Y$  il peso del suo contenuto. Quanto valgono media e deviazione standard di  $Y$ ? Quante bustine occorre prendere per accumulare un peso di almeno 200 g con una probabilità di almeno il 75%?

**(2 punti)** In realtà i tre tipi di bustine della seconda scatola hanno *medie* di 4, 5 e 6 g, ma deviazioni standard  $\sigma$  (ogni tipo uguale). Quanto valgono media e deviazione standard di  $Y$  in questo caso? (Si richiede di rispondere in funzione di  $\sigma$ .)

**Problema 3.2** (12 punti). Sia  $X$  una variabile aleatoria continua con funzione di densità

$$f(t) = \begin{cases} c(1 - t^2) & -1 < t < 1 \\ 0 & \text{altrimenti} \end{cases}$$

**(7 punti)** Si determinino il valore della costante  $c$ , la media, la varianza, la moda e la mediana di  $X$ . Si trovi la corrispondente funzione di ripartizione  $F$  e si traccino i grafici approssimativi di  $f$  e di  $F$ .

**(2 punti)** Determinare la legge di  $\sqrt{|X|}$ .

**(3 punti)** Sia  $Y$  indipendente da  $X$  e con la stessa legge. Determinare il valore atteso di  $\max(X, Y)$  e  $\min(X, Y)$ .

**Problema 3.3** (12 punti). In un esperimento di biologia sono state analizzate 850 cellule, misurando in ciascuna il livello di espressione del gene *Sfrp1* (il livello di espressione è per ogni cellula un numero intero non negativo), trovando media campionaria 4.247 e deviazione standard campionaria 6.079.

**(6 punti)** Si stimino la media e la deviazione standard della popolazione con intervalli bilaterali al 90% di confidenza.

**(3 punti)** Sia  $p$  la probabilità che una cellula abbia espressione 0. Sia  $q$  un numero reale in  $(0, 1)$ . Se le cellule con espressione 0 nel campione sono 273, si verifichi al 1% di significatività se sia plausibile che  $p = q$ : per quali valori di  $q$  il test accetta l'ipotesi nulla?

**(3 punti)** Si stimino puntualmente media e deviazione standard dell'espressione, restringendosi alle cellule con espressione non nulla. (*Suggerimento*: i dati sono sufficienti, e si ricordi che l'espressione è un intero non negativo.)

**Problema 3.4** (13 punti). Supponiamo che il tempo che passa dall'acquisto di un cellulare nuovo alla rottura che ne causa la sostituzione, sia una variabile aleatoria **esponenziale** di media  $\theta$ . Una azienda basa le sue strategie di mercato sull'ipotesi che  $\theta$  sia pari a 13 mesi, e vorrebbe ora verificare tale ipotesi. Viene intervistato un campione di utenti ottenendo 40 dati con media campionaria di 7.60 mesi e deviazione standard campionaria di 8.71 mesi.

**(7 punti)** Si verifichi tramite calcolo del  $p$ -value se l'ipotesi dell'azienda sia plausibile.

**(3 punti)** Si determini la potenza del test per livello di significatività  $\bar{\alpha} = 5\%$  e media vera  $\theta = 9$  mesi.

**(3 punti)** Il campione raccolto ha il difetto che non tiene conto che molti utenti hanno cambiato cellulare prima di rompere lo schermo e tutti questi (60 utenti) sono stati esclusi dal campione (che inizialmente contava 100 utenti). Se si suppone che pure il tempo da acquisto a sostituzione (senza rotture, per obsolescenza) sia una variabile aleatoria esponenziale, indipendente dal tempo di eventuale rottura, come si può correggere la stima di  $\theta$ ? (Si richiede solo la stima puntuale.)