*Si svolgano 3 esercizi a scelta sui 4 proposti. In nessun caso verranno assegnati punti per più di 3 esercizi.*

**Problema 5.1** (11 punti)**.** Uno studente delle superiori, ad ogni verifica scritta od orale prende un voto tra 0 e 10 che assumiamo avere distribuzione binomiale di parametri $n$ e $p$ con $n = 10$. Supponiamo che tutti i voti siano variabili aleatorie indipendenti e identicamente distribuite $V_1, V_2, \dots, V_m$, per un totale di $m$ voti.

**(7 punti)** Quanto vale $p$ se la media $E[V_i]$ di ogni voto è 7.2? Quanto valgono varianza e deviazione standard di ogni voto? Quanto vale approssimativamente la probabilità che alla fine dell'anno, con $m = 40$ voti, la media aritmetica degli stessi sia sufficiente (maggiore o uguale a 6)?

**(2 punti)** Se in una singola materia il numero di voti a fine anno è 8, quanto deve essere grande $p$ come minimo affinché lo studente abbia una probabilità dell'80% almeno di essere sufficiente in quella materia?

**(2 punti)** Se ci sono 4 materie con 8 voti ciascuna, e supponiamo per semplicità che $p$ sia lo stesso per tutte le materie, quanto deve essere grande $p$ come minino affiché lo studente abbia una probabilità dell'80% almeno di essere sufficiente in ciascuna materia?

**Problema 5.2** (12 punti)**.** Sia $X$ una variabile aleatoria continua con densità

$$f_X(t) = \begin{cases} \lambda e^{\lambda a} e^{-\lambda t} & t \geq a \\ 0 & t < a \end{cases}$$

dove $a \in \mathbb{R}$ e $\lambda > 0$ sono parametri.

**(7.5 punti)** Si determinino, in funzione di $a$ e $\lambda$, i valori di media, deviazione standard, moda e mediana di $X$ e si tracci il grafico di $f_X$.

**(2 punti)** Si mostri che $X - a$ ha distribuzione esponenziale.

**(2.5 punti)** Fissato $\lambda$ e al variare di $a$, quanto vale come minimo $E[X^2]$?

**Problema 5.3** (12 punti)**.** Una linea di produzione di cerniere sembra avere un problema di difettosità troppo elevata. Un recente controllo ha trovato il 18.75% di cerniere difettose. (Più precisamente, vi erano 3 cerniere difettose su un campione di 16.)

**(7 punti)** Usando la approssimazione gaussiana, si verifichi al 5% di significatività se questi dati sono compatibili con l'ipotesi che la difettosità reale sia del 7%. Sono soddisfatte le ipotesi di tale approssimazione?

**(2 punti)** Si determini la regione di accettazione del test del punto precedente non relativa alla statistica del test, come si fa di solito, ma relativa alla statistica *numero di pezzi difettosi*. Si calcoli inoltre il $p$-dei-dati.

**(3 punti)** Si ripetano i primi due punti non usando più l'approssimazione gaussiana, ma con calcoli esatti a partire dalla distribuzione binomiale.

**Problema 5.4** (12 punti)**.** Si vuole studiare quanto riproducibili sono le performance delle reti neurali, visto che tipicamente queste vengono addestrate una volta per tutte e i loro risultati misurati come se fossero deterministici.

Si addestra quindi per 6 volte una stessa rete neurale per classificazione di immagini (ogni volta da zero e in modo indipendente). Le performance delle 6 reti, testate su uno stesso set di immagini, sono le seguenti:

| | | | | | |
|---|---|---|---|---|---|
| 0.78 | 0.67 | 0.75 | 0.81 | 0.66 | 0.73 |

**(6 punti)** Si dia un intervallo di confidenza unilaterale sinistro[1] al 95% di confidenza per la deviazione standard delle performance.

**(3 punti)** Si cambiano alcuni parametri e si ripete il test, sempre con numerosità 6, ottenendo media campionaria 0.7667 e deviazione standard campionaria 0.0439. Si verifichi tramite calcolo del $p$-dei-dati se è plausibile che la performance media non sia cambiata.

**(2 punti)** Supponendo che si trovi che la media non è cambiata, è ragionevole fondere i due campioni in uno solo di numerosità maggiore, per ripetere il primo punto con una stima più accurata. È possibile farlo con i soli dati a disposizione?

[1] Quindi del tipo $\sigma \leq u$.