

Si svolgano 3 esercizi a scelta sui 4 proposti.

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

Problema 5.1 (12 punti). In una famiglia di 5 persone, ciascuno dei membri consuma ogni giorno un numero casuale di mascherine con le seguenti probabilità

mascherine	0	1	2	3
probabilità	20%	65%	10%	5%

(7 punti) Si calcolino media e varianza del numero di mascherine usate al giorno. Si calcoli in modo approssimato, con il teorema del limite centrale, la probabilità che il numero di mascherine usate da tutta la famiglia in 28 giorni sia superiore a 180.

(2 punti) Si determini il minimo numero di mascherine che sia sufficiente al fabbisogno di tutta la famiglia per 28 giorni, con probabilità del 90% almeno.

(3 punti) Sia X la somma di due vv.aa. indipendenti diverse, di medie μ_1, μ_2 e varianze σ_1^2, σ_2^2 . Sia invece Y la somma di due vv.aa. indipendenti e identicamente distribuite, ciascuna delle quali generata con probabilità $\frac{1}{2}$ con media μ_1 e varianza σ_1^2 e con probabilità $\frac{1}{2}$ con media μ_2 e varianza σ_2^2 . Si verifichi se X e Y hanno media e/o varianza uguali¹.

Problema 5.2 (11 punti). Sia X una variabile aleatoria continua con funzione di densità

$$f_X(t) = \frac{t}{2} - 5, \quad t \in [10, 12]$$

(7 punti) Si determinino la media, la deviazione standard, la moda e la mediana di X . Si traccino i grafici della densità e della funzione di ripartizione² corrispondente.

(2 punti) Sia $Y = aX + b$ una trasformazione lineare di X con media 0 e varianza 1. Quanto valgono a e b ? Si tracci il grafico della pdf di Y .

(2 punti) Si descriva come si può generare³ una v.a. con la distribuzione di X , e si verifichi che effettivamente la sua pdf coincida con f_X .

Problema 5.3 (12 punti). In un sondaggio riguardo ad un referendum, su 2000 persone intervistate, 987 si sono dette a favore del sì, 567 a favore del no e i restanti incerti.

(7 punti) Usando solo le risposte sì e no, si stimi al 95% di confidenza, con un intervallo bilaterale, la percentuale di favorevoli al sì nella popolazione. Si ripeta con un intervallo unilaterale al 95% di confidenza, che indichi un limite inferiore alla percentuale in favore del sì che si può stimare.

(2 punti) Si verifichi tramite calcolo del p -value se vi sia evidenza significativa che vinceranno i sì, ipotizzando quattro diversi scenari: a) escludendo gli incerti, b) contandoli metà favorevoli e metà contrari, c) contandoli tutti favorevoli, d) contandoli tutti contrari.

(3 punti) Consideriamo il test bilaterale che i sì e i no siano pari. Si fissi una numerosità di 400 intervistati (in assenza di indecisi) e il 5% di significatività. Si determini un'espressione per la curva OC, calcolando in particolare la potenza del test in caso di ripartizione 60%-40% e 55%-45%.

Problema 5.4 (12 punti). Una serie di 12 esperimenti di agenti artificiali ha dato i risultati seguenti (in percentuale di successi),

73.33%	86.67%	76.67%	70.00%	63.33%	86.67%
63.33%	63.33%	76.67%	83.33%	93.33%	80.00%

(7 punti) Si verifichi tramite calcolo del p -value se vi sia evidenza statistica che la percentuale di successi media sia superiore al 70%.

(2 punti) Si determinino intervalli di confidenza bilaterali al 95%, per media e deviazione standard della percentuale di successi.

(3 punti) Le percentuali del campione sembrano provenire da rapporti tra numeri interi (successi/tentativi). Si lavori in questa ipotesi migliorando la risposta al primo quesito. Quali rischi comporta questo approccio?

¹Scopo di questo punto è capire se sarebbe utile in questo problema distinguere i vari membri della famiglia e i vari giorni della settimana.

²La CdF si può ricavare anche solo graficamente e rappresentare in modo approssimativo.

³Al calcolatore, ovvero a partire da una v.a. uniforme su $[0, 1]$.