

SECONDA PROVA IN ITINERE

Prova a risolvere i seguenti problemi, giustificando il ragionamento seguito.

La motivazione del processo è molto più importante della risposta numerica.

Puoi usare una calcolatrice o un regolo per i conti, così come R sul calcolatore dell'aula. PUOI CONTROLLARE I TUOI APPUNTI, LE NOTE DEL CORSO O UN LIBRO DI TESTO TRA QUELLI CONSIGLIATI. Lavora per tuo conto, senza aiuto esterno, ma discuti pure i problemi e le tue soluzioni **finita la prova**. I problemi **non** sono in ordine di difficoltà. Indica nome e cognome (e numero di matricola) sui fogli, così come il numero del problema o della domanda. **Lascia un po' di spazio per i commenti**. Le parti in R possono essere copiate sul foglio, oppure salvate come file (`.R`, `.txt`, `.Rhistory`) indicando con un commento (introdotto dal carattere `#`) a quale esercizio e domanda il codice si riferisce. Se parte di un problema è svolta in R, indicalo sul foglio in corrispondenza del punto dell'esercizio corrispondente. Non dimenticare di caricare il file nella risorsa *esamionline* al termine dell'esame.

Buon lavoro!

Problema 1. Un'azienda di e-commerce spedisce ogni giorno centinaia di pacchi. Il peso (in kg) di ciascun pacco può variare molto a seconda del contenuto, ma l'azienda sa per esperienza che il peso di un singolo pacco ha media 6 kg e deviazione standard 5 kg.

In un certo giorno, vengono spediti n pacchi.

1. Sappiamo che il camion che trasporta i pacchi ha una capacità massima di 760 kg. Qual è la probabilità che in un giorno serva più di un camion per trasportare tutti i pacchi?
 - a. Scrivere la risposta in funzione di n .
 - b. Determinarne il valore per $n = 100$.
2. Qual è il massimo numero di pacchi per cui la probabilità che partano almeno 2 camion sia al più uguale a 0.1?
3. Dopo alcuni studi, si scopre che il numero di ordini ricevuti in un giorno dall'azienda segue una distribuzione di Poisson di media 130. Ricalcolare la probabilità che in un giorno serva più di un camion per trasportare tutti i pacchi, tenendo conto di questa nuova informazione.

Problema 2. A tutta la studenti del corso di Probabilità è stato chiesto di compilare un questionario sugli esercizi proposti da due esercitatori, che indicheremo come E1 ed E2. Ogni studente ha attribuito un punteggio, chiamato OdS (Odio degli Studenti), che rappresenta in modo sintetico la difficoltà percepita degli esercizi svolti. Si assume che l'indice OdS sia una variabile quantitativa continua a valori in $[0, 20]$.

Indichiamo μ_1 e μ_2 le difficoltà medie associate agli esercizi proposti da E1 e da E2, rispettivamente. I dati sono raccolti nel file `dat1_ods.csv`.

1. Esplora il dataset tramite grafici adeguati. Cosa si può osservare sulla distribuzione degli indici OdS per i due gruppi?

2. Osserva che nel dataset ci sono valori mancanti (NA). Come possono essere gestiti in modo coerente con il problema? Giustifica la tua scelta.
3. Definisci formalmente un test di ipotesi per verificare se la difficoltà media degli esercizi di E1 ed E2 sono uguali.
4. Applica il test di ipotesi scelto. Riporta il valore della statistica test, il p-value e un intervallo di confidenza al 97.5% per la differenza tra le medie.
5. Sulla base del risultato del test e dell'intervallo di confidenza, cosa puoi concludere? C'è evidenza che una dei due esecutori proponga esercizi mediamente più difficili?

Problema 3. Sia $\{X_i\}_{i=1}^{n^2}$ una successione di variabili aleatorie iid uniformi su $[0, 1]$. Vogliamo studiare l'istogramma delle frequenze relative del campione casuale (X_1, \dots, X_{n^2}) con n rettangoli, tutti della stessa larghezza $\frac{1}{n}$. Indichiamo con H_k l'altezza del rettangolo k -esimo, ovvero della barra che copre l'intervallo $I_k = \left(\frac{k-1}{n}, \frac{k}{n}\right]$, ove $k \in \{1, \dots, n\}$.

Nota: per "istogramma delle frequenze relative" si intende quello la cui area totale è pari ad 1 (Vedere Fig. 4 per esempio di riferimento).

Istogramma delle frequenze relative

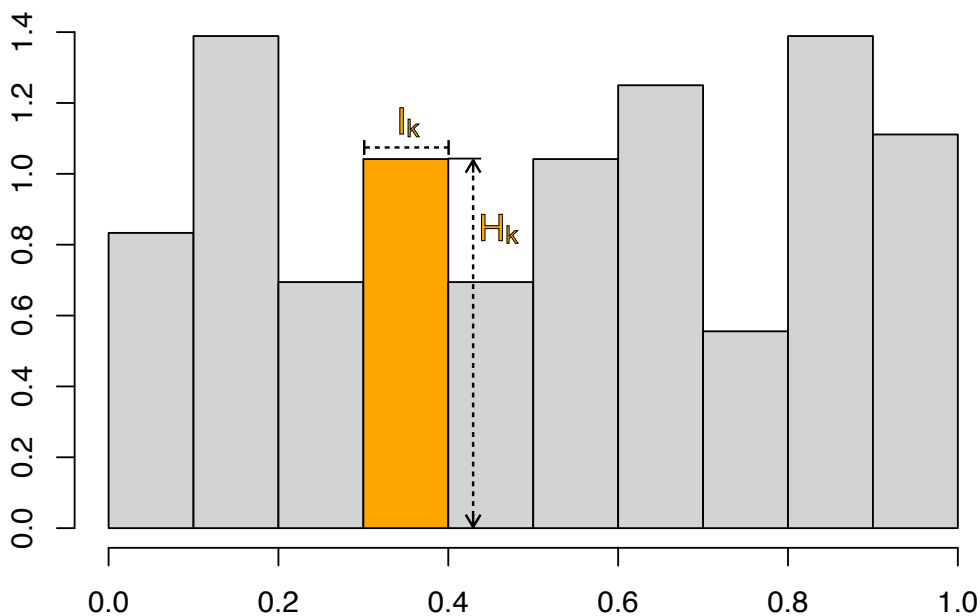


Figura 4. Istogramma descritto.

1. Scrivere H_k in funzione del campione (X_1, \dots, X_{n^2}) .
2. Determinare la distribuzione di H_k .
3. Dimostrare che H_k converge in legge alla variabile aleatoria degenerata 1. Commenta questo risultato.
4. Calcolare $\text{Corr}(H_\ell, H_k)$ per $k \neq \ell$. Commentare il risultato trovato e il suo comportamento per $n \rightarrow +\infty$.