

Corso Metodi Matematici e Statistici per Giuristi (002NP)

A.A. 2021/22 - Prova scritta 2022-02-10

La durata della prova è di 120 minuti. Fornire risposte dettagliate.

Problema 1 (10 punti)

Un sondaggio riguardante la distanza del proprio domicilio dal luogo di lavoro ha ottenuto le seguenti risposte (in Km) tra i 15 dipendenti (10 maschi e 5 femmine) di una start-up. Le risposte sono raccolte nella seguente tabella:

Genere	Distanza domicilio/lavoro
M	12, 16, 18, 29, 18, 16, 15, 20, 11, 1
F	3, 9, 16, 30, 12

1. Determinarne moda, mediana, media ed errore standard dei dati complessivamente raccolti (maschi e femmine insieme).
2. Si interroga a caso uno (o una) dei 15 dipendenti e si viene a sapere che abita a meno di 10 Km dal luogo di lavoro: è più probabile che sia maschio o femmina?

Problema 2 (10 punti)

Vi rivolgete alla vostra banca e chiedete un prestito di 15000 euro. La banca vi accorda il prestito e vi chiede di rimborsarlo con 4 rate semestrali. Il piano di ammortamento prevede quote capitale costanti, mentre i tassi di interesse annui nominali convertibili 2 volte all'anno applicati sono 10% per il primo anno (prime 2 rate) e 12% per il secondo anno (ultime 2 rate). Redigere il piano di ammortamento del prestito.

Ora dovete investire la cifra che avete reperito grazie al prestito. Avete le seguenti opzioni:

- a) acquistare un titolo finanziario che in cambio della cifra totale che avete ottenuto vi pagherà due rate: la prima di 6000 euro tra 9 mesi e la seconda di 12000 euro tra 18 mesi,
- b) investire la cifra totale che avete ottenuto per iscrivervi ad un master. Stimate in 22000 euro i profitti che riuscirete ad ottenere tra 30 mesi grazie alle vostre maggiori competenze.

Scegliete l'opzione più conveniente sulla base del criterio del tasso interno di rendimento (TIR).

Corso Metodi Matematici e Statistici per Giuristi (002NP)

A.A. 2021/22 - Prima prova in itinere 2021-11-17

La durata della prova è di 90 minuti. Fornire risposte dettagliate.

Problema 1 (10 punti)

Un'azienda farmaceutica dichiara che un test di screening per una malattia è positivo con probabilità pari a 95% (*sensibilità*), in caso di presenza della malattia, ed è negativo con una probabilità pari a 98%, in caso di assenza della malattia (*specificità*). Supponiamo che il 98% della popolazione *non* sia affetta da tale malattia.

1. Qual è la probabilità che una persona presa a caso nella popolazione sia negativa al test?
2. Qual è la probabilità che una persona risultata negativa al test abbia in realtà la malattia (sia cioè un falso negativo)?

Una soluzione:

1. Abbiamo $P(\text{Negativa}|\text{Sana}) = 98\%$ e

$$P(\text{Negativa}|\text{Malata}) = 1 - P(\text{Positiva}|\text{Malata}) = 1 - 95\% = 5\%.$$

Di conseguenza,

$$\begin{aligned} P(\text{Negativa}) &= P(\text{Negativa}|\text{Sana})P(\text{Sana}) + P(\text{Negativa}|\text{Malata})P(\text{Malata}) \\ &= 98\% \cdot 98\% + 5\% \cdot 2\% = 96.14\%. \end{aligned}$$

2. Dobbiamo calcolare, usando la formula di Bayes,

$$\begin{aligned} P(\text{Malata}|\text{Negativa}) &= P(\text{Malata}) \frac{P(\text{Negativa}|\text{Malata})}{P(\text{Negativa})} \\ &= 2\% \cdot \frac{5\%}{96.14\%} \approx 0.1\%. \end{aligned}$$

Problema 2 (7 punti)

Un istituto di vigilanza deve valutare indipendentemente le probabilità dichiarate dall'azienda farmaceutica riguardo al test del problema precedente. Per fare questo ha analizzato una popolazione di 100 individui sani e 50 individui malati (lo stato di salute è determinato tramite analisi più approfondite), effettuando il test su di essi. I risultati sono riportati nella seguente tabella di contingenza.

	Sano	Malato
Test Positivo	10	45
Test Negativo	90	5

1. Scrivere la tabella di contingenza con le frequenze relative alle colonne.
2. Determinare se vi è associazione tra le due variabili.
3. Calcolare l'indice chi-quadrato di Pearson.

Una soluzione:

1. Dobbiamo dividere la prima colonna per 100 (la popolazione sana) e la seconda per 50. Troviamo

	Sano	Malato
Test Positivo	0.1	0.9
Test Negativo	0.9	0.1

2. Vi è associazione tra le variabili, come si vede dalla tabella con le frequenze relative, perché le colonne sono diverse.

3. Per calcolare l'indice chi-quadrato, scriviamo prima in una tabella le frequenze $n_{ij}^* = n_{.j}n_{i.}/n$ (dove $n = 150$)

	Sano	Malato
Test Positivo	36.6	18.3
Test Negativo	63.3	31.6

Scriviamo in una nuova tabella i termini $(n_{ij} - n_{ij}^*)^2/n_{ij}^*$ (che poi basterà sommare per ottenere l'indice χ^2)

	Sano	Malato
Test Positivo	19.3	39
Test Negativo	11.3	22.4

Sommando si trova $\chi^2 = 92$. Ricordando che il massimo possibile nel caso 2×2 è $n = 150$, si trova una associazione molto forte (in questo caso maggiore è l'associazione, migliore è l'affidabilità del test).

Problema 3 (7 punti)

Siete impiegati in una banca la quale vi affida la redazione di un piano di ammortamento progressivo indicizzato per un cliente che necessita un prestito di 20000 € da rimborsare con 4 rate quadrimestrali al tasso inizialmente concordato del 12% annuo nominale convertibile 3 volte all'anno. Redigere il piano di ammortamento considerando che dopo il pagamento della terza rata il tasso effettivo aumenta di un punto percentuale.

Il compito successivo che vi viene affidato riguarda la scelta tra investimenti per mezzo del criterio del rendimento economico attualizzato (REA). In particolare, considerate i progetti

- a) prestare oggi 100000 € ad un cliente per ricevere 116000 € tra due anni,
- b) investire 100000 € oggi in un titolo finanziario che vi restituisce 60000 € tra un anno e 50000 € tra due anni.

Calcolate per quali valori del tasso di interesse usato per la valutazione il progetto a) risulta più conveniente del progetto b).

Problema 4 (10 punti)

Decidete di abbandonare l'impiego in banca e fondare la vostra società di consulenza. Avete bisogno di un server per poter svolgere il vostro lavoro. Vi rivolgete all'impresa fornitrice chiedendo un preventivo per il server. L'impresa fornitrice vi comunica che il costo del bene è pari a 10000 €. Non disponete in cassa della somma necessaria, quindi valutate tre diverse alternative di finanziamento riportate di seguito.

- a) Il fornitore propone di consegnarvi subito il bene in cambio del pagamento tra 90 giorni di 10364 €.
- b) La banca presso cui avete il conto corrente propone di prestarvi la somma necessaria ad acquistare subito il bene, questo prestito sarà ammortizzato in 10 rate mensili con metodo a quote capitale costanti (metodo italiano). La prima rata, da versare tra un mese, sarà di 1100 €.
- c) La finanziaria a cui il fornitore si affida vi propone di comprare il bene attraverso un leasing consistente in due canoni semestrali posticipati di 4448 € ciascuno ed un valore di riscatto pari a 2000 €. La finanziaria richiede inoltre il pagamento di commissioni pari a 100 € all'atto del versamento di ogni canone.

Valutate le alternative calcolando il tasso interno di rendimento (TIR) e scegliete quella più conveniente.

Corso Metodi Matematici e Statistici per Giuristi (002NP)

A.A. 2022/23 - Prima prova in itinere 2022-11-21

La durata della prova è di **90 minuti**. I punteggi sono indicativi. Fornire risposte dettagliate.

Problema 1 (6 punti)

Un'indagine statistica sul sistema giuridico americano ha selezionato un campione tra la popolazione degli imputati, ed è risultato che il 30% degli individui considerati sono stati poi incarcerati. Tra quelli incarcerati, il 40% aveva deciso di dichiararsi colpevole in sede di processo. Invece, tra quelli non incarcerati, il 60% aveva deciso di dichiararsi colpevole.

Avendo scelto completamente a caso un individuo tra quelli del campione,

1. calcolare la probabilità che sia dichiarato *colpevole* in sede di processo;
2. avendo inoltre osservato che il soggetto si era dichiarato *colpevole* in sede di processo, calcolare la probabilità che sia stato incarcerato.

Problema 2 (9 punti)

La seguente tabella di contingenza mostra la frequenza assoluta delle nascite avvenute in Italia nel 2021, per classe d'età e area geografica di provenienza della madre – per semplicità distinguiamo solo se la madre è di provenienza *Italia* o *Resto del Mondo* (fonte: *salute.gov.it*, Analisi dell'evento nascita - Anno 2021).

Età madre	Madre: Italia	Madre: Resto del Mondo
< 20	2 500	736
20 - 29	72 218	30 572
30 - 39	207 256	39 389
40 - 49	34 217	5 501
50 - 65	358	27

1. Completare la tabella con le frequenze marginali delle due variabili (età della madre, provenienza della madre).
2. Calcolare le frequenze condizionate per colonna (approssimare a due cifre decimali dopo la virgola). Dire se vi è associazione tra le due variabili considerate.
3. Calcolare moda e media della variabile *età della madre*, considerata su tutte le nascite (ossia usando la corrispondente frequenza marginale).

Problema 3 (6 punti)

Volendo investire 10.000 € per 30 mesi la banca mi offre queste possibilità:

- a) investirli in regime di capitalizzazione composta al tasso annuo del 3,2 %,
- b) investirli in regime di capitalizzazione semplice al tasso semestrale dello 1,7 %,
- c) ricevere al termine dei 30 mesi la somma di 10.880 €.

Dire quale di queste tre proposte è la più conveniente (argomentare la risposta).

Problema 4 (9 punti)

Per l'acquisto di uno scooter del costo di 4.500 € da pagare in due anni, la concessionaria offre queste tre possibilità:

- a) versare subito 1.500 € e poi due rate annuali posticipate di 1.700 €,
- b) un versamento immediato di 2.000 € ed un versamento di 3.000 € al termine dei due anni,
- c) ammortamento in rate mensili posticipate (cioè con le regole di un mutuo) calcolate al tasso di interesse annuo nominale dello 8,5 %.

Indagare, secondo il criterio del TIR, quale di questi pagamenti è più conveniente.

Osservazione: riflettere se, nella possibilità di pagamento c), è necessario calcolare la rata mensile per individuare il TIR.

Corso Metodi Matematici e Statistici per Giuristi (002NP)

A.A. 2022/23 - Prima prova in itinere 2022-11-21

La durata della prova è di **90 minuti**. I punteggi sono indicativi. Fornire risposte dettagliate.

Problema 1 (6 punti)

Un'indagine statistica sul sistema giuridico americano ha selezionato un campione tra la popolazione degli imputati, ed è risultato che il 30% degli individui considerati sono stati poi incarcerati. Tra quelli incarcerati, il 40% aveva deciso di dichiararsi colpevole in sede di processo. Invece, tra quelli non incarcerati, il 60% aveva deciso di dichiararsi colpevole.

Avendo scelto completamente a caso un individuo tra quelli del campione,

1. calcolare la probabilità che sia dichiarato *colpevole* in sede di processo;
2. avendo inoltre osservato che il soggetto si era dichiarato *colpevole* in sede di processo, calcolare la probabilità che sia stato incarcerato.

Problema 2 (9 punti)

La seguente tabella di contingenza mostra la frequenza assoluta delle nascite avvenute in Italia nel 2021, per classe d'età e area geografica di provenienza della madre – per semplicità distinguiamo solo se la madre è di provenienza *Italia* o *Resto del Mondo* (fonte: *salute.gov.it*, Analisi dell'evento nascita - Anno 2021).

Età madre	Madre: Italia	Madre: Resto del Mondo
< 20	2 500	736
20 - 29	72 218	30 572
30 - 39	207 256	39 389
40 - 49	34 217	5 501
50 - 65	358	27

1. Completare la tabella con le frequenze marginali delle due variabili (età della madre, provenienza della madre).
2. Calcolare le frequenze condizionate per colonna (approssimare a due cifre decimali dopo la virgola). Dire se vi è associazione tra le due variabili considerate.
3. Calcolare moda e media della variabile *età della madre*, considerata su tutte le nascite (ossia usando la corrispondente frequenza marginale).

Problema 3 (6 punti)

Volendo investire 10.000 € per 30 mesi la banca mi offre queste possibilità:

- a) investirli in regime di capitalizzazione composta al tasso annuo del 3,2 %,
- b) investirli in regime di capitalizzazione semplice al tasso semestrale dello 1,7 %,
- c) ricevere al termine dei 30 mesi la somma di 10.880 €.

Dire quale di queste tre proposte è la più conveniente (argomentare la risposta).

Problema 4 (9 punti)

Per l'acquisto di uno scooter del costo di 4.500 € da pagare in due anni, la concessionaria offre queste tre possibilità:

- a) versare subito 1.500 € e poi due rate annuali posticipate di 1.700 €,
- b) un versamento immediato di 2.000 € ed un versamento di 3.000 € al termine dei due anni,
- c) ammortamento in rate mensili posticipate (cioè con le regole di un mutuo) calcolate al tasso di interesse annuo nominale dello 8,5 %.

Indagare, secondo il criterio del TIR, quale di questi pagamenti è più conveniente.

Osservazione: riflettere se, nella possibilità di pagamento c), è necessario calcolare la rata mensile per individuare il TIR.

Corso Metodi Matematici e Statistici per Giuristi (002NP)

A.A. 2022/23 - Appello 2023-04-27

La durata della prova è di 120 minuti. Fornire risposte dettagliate.

Problema 1

Dare le definizioni di media e di mediana per una variabile statistica quantitativa. Fornire un esempio esplicito in cui la media è diversa dalla mediana.

Una soluzione:

La media di una variabile statistica è una misura di tendenza centrale che viene calcolata sommando tutti i valori in una popolazione e poi dividendo per il numero di individui. Matematicamente, possiamo esprimere la media campionaria come:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

dove \bar{x} è la media campionaria, n è la dimensione della popolazione e x_i rappresenta l' i -esima osservazione.

La mediana, d'altra parte, è il valore centrale in un insieme di dati quando i dati sono ordinati in ordine crescente o decrescente. Se ci sono un numero pari di osservazioni, la mediana è la media dei due valori centrali. Ad esempio, se abbiamo il seguente insieme di dati:

$$\{2, 5, 7, 10, 12\}$$

La mediana sarebbe 7, poiché è il valore centrale nell'insieme ordinato.

Ora, consideriamo un esempio in cui la media e la mediana differiscono. Supponiamo di avere il seguente insieme di dati:

$$\{1, 2, 3, 4, 100\}$$

La media sarebbe:

$$\bar{x} = \frac{1}{5}(1 + 2 + 3 + 4 + 100) = \frac{110}{5} = 22$$

Tuttavia, la mediana sarebbe 3, poiché è il valore centrale quando i dati sono ordinati in ordine crescente. In questo caso, la media è significativamente più alta della mediana, il che indica che la presenza del valore anomalo (100) ha distorto la media verso l'alto. Questo esempio illustra l'importanza di considerare sia la media che la mediana nell'analisi dei dati.

Problema 2

Supponiamo di condurre un sondaggio per determinare se esiste una relazione tra il genere

e il gusto preferito di gelato. Si intervistano 200 persone, chiedendo loro di identificare il loro genere e il loro gusto preferito di gelato. I risultati sono riassunti nella seguente tabella di contingenza:

	Cioccolato	Vaniglia	Fragola
Maschio	30	35	25
Femmina	25	40	45

1. Completare la tabella con le frequenze assolute marginali delle due variabili.
2. Calcolare le frequenze condizionate per colonna (approssimare a due cifre decimali)
3. Esiste qualche evidenza che suggerisca l'esistenza di una relazione tra il genere e il gusto preferito di gelato?

Una soluzione:

1.

	Cioccolato	Vaniglia	Fragola	Totale riga
Maschio	30	35	25	90
Femmina	25	40	45	110
Totale colonna	55	75	70	200

2.

	Cioccolato	Vaniglia	Fragola	
Maschio	$30/55=0.55$	$35/75=0.45$	$25/70=0.36$	
Femmina	$25/55=0.45$	$40/75=0.55$	$45/70=0.64$	
Totale colonna	$55/55=1$	$75/75=1$	$70/70=1$	

3. Le distribuzioni condizionate per colonna risultano diverse (al variare delle colonne). Questo suggerisce che vi sia associazione. Per quantificare tale associazione si potrebbe calcolare l'indice χ -quadrato di Pearson (risulta $\chi^2 \approx 0.69$).

Corso Metodi Matematici e Statistici per Giuristi (002NP)

A.A. 2022/23 - Prova 2023-05-23

La durata della prova è di 120 minuti. Fornire risposte dettagliate.

Problema 1

Descrivere la formula di Bayes e dare un esempio di una possibile applicazione (*eventualmente anche senza svolgere calcoli*).

Problema 2

Supponiamo di condurre un sondaggio per determinare se esiste una relazione tra l'età e il livello di istruzione di una popolazione. Si intervista un certo numero di persone, chiedendo loro di identificare la loro età (in anni) e il loro livello di istruzione (elementare, medio, superiore, universitario). I risultati sono riassunti nella seguente tabella di contingenza:

	Elementare	Medio	Superiore	Universitario
18-30	20	25	40	35
31-50	15	30	50	45
51 - 70	10	20	30	25

1. Descrivere l'età media della popolazione intervistata (*ricordare come si approssima la media di una variabile distribuita in classi*).
2. Completare la tabella con le frequenze assolute marginali delle due variabili.
3. Calcolare le frequenze condizionate per riga (approssimando a due cifre decimali dopo la virgola).
4. Dire se vi è associazione tra le due variabili statistiche rilevate. Indicare un possibile metodo per quantificare tale associazione (*non è richiesto di portare a termine tutti i calcoli, eventualmente descrivere solo i passaggi da svolgere*).