

1 Calcolo delle probabilità

Esercizio 1 Supponiamo che dieci carte numerate da uno a dieci vengano introdotte in un cappello e che una carta venga estratta a caso. Vogliamo determinare:

- a) Qual è la probabilità che la carta estratta sia il dieci?
- b) Sapendo che la carta estratta è maggiore di quattro, qual è la probabilità che sia un dieci?

Esercizio 2 In una famiglia vi sono 2 figli. Qual è la probabilità che entrambi siano maschi dato che almeno uno di loro è maschio?

Esercizio 3 Si supponga che un'urna contenga 7 palline bianche e 5 nere. Supponiamo di estrarre due palline senza reimmissione.

Assumendo che ogni pallina possa essere estratta con egual probabilità, qual è la probabilità che entrambe le palline estratte siano bianche?

Esercizio 4 Supponiamo che tre amici ad una festa gettino il proprio cappello sulla stessa sedia. Questi cappelli vengono mescolati tra loro e, successivamente, i tre amici scelgono un cappello a caso.

Qual è la probabilità che nessuno di loro venga in possesso del proprio cappello?

Esercizio 5 Viene estratta una pallina da un'urna che contiene 4 palline numerate da 1 a 4.

Siano $E = \{1, 2\}$, $F = \{1, 3\}$, e $G = \{1, 4\}$, calcolare

- $P(E)$
- $P(E \cap G)$
- $P(F \cap G)$
- $P(E \cap F \cap G)$

Esercizio 6 Si considerino due urne. La prima contiene 2 palline bianche e 7 nere, mentre la seconda contiene 5 palline bianche e 6 nere. Lanciamo una moneta e se otteniamo testa estraiamo una pallina dalla prima urna, altrimenti estraiamo una pallina della seconda urna.

Qual è la probabilità che il lancio della moneta sia stato testa dato che la pallina estratta è bianca?

Esercizio 7 Uno studente risponde ad un quiz multivalore conoscendo la risposta o selezionando a caso. Si indichi con p la probabilità che lo studente conosca la risposta (ha studiato).

Supponiamo che uno studente che non conosca la risposta indovini con probabilità $1/m$, dove m è il numero di risposte possibili.

Qual è la probabilità che lo studente conoscesse la risposta dato che ha risposto correttamente?

Esercizio 8 Un test di laboratorio è efficace al 95% nel rilevare una certa malattia quando questa è presente. Il test fornisce comunque “falsi positivi” (dice che una persona sana è malata) per l'1% delle persone sane esaminate. Nel caso in cui lo 0.5% della popolazione sia affetto dalla suddetta malattia, qual è la probabilità che una persona sia effettivamente malata dato che il test dà risultato positivo?

Esercizio 9 Si sa che una data lettera ha eguale probabilità di essere contenuta in uno di tre folders. Si indichi con α_i la probabilità che la lettera venga trovata tramite una ricerca sommaria nel folder i . Pertanto con probabilità $(1 - \alpha_i)$ la lettera non viene trovata anche se si trova nel folder i . Supponiamo di cercare nel folder 1 senza trovare la lettera. Qual è la probabilità che la lettera sia nel folder 1?

Esercizio 10 Un'urna contiene 3 biglie bianche e 2 biglie nere.

- Calcolare la probabilità che estraendo in successione (senza reimbussolamento) 3 biglie almeno una sia nera.
- Ripetere il punto precedente supponendo il reimbussolamento.

Esercizio 11 Una squadra di calcio schiera ad ogni partita 1 portiere, 5 difensori e 5 attaccanti. La società “Aleas” sceglie in modo casuale ciascun gruppo di giocatori tra 2 portieri, 8 difensori e 12 attaccanti disponibili.

- Quante sono le formazioni possibili?
- Se Roberto e Ronaldo sono due attaccanti, quante sono le formazioni in cui giocano entrambi?
- Se Franco è un difensore, quante sono le formazioni in cui gioca con l'attaccante Roberto?

Esercizio 12 Date due urne indistinguibili esternamente. La prima contiene 3 palline rosse, 1 bianca e 2 verdi, mentre la seconda contiene 1 pallina rossa, 3 bianche e 2 verdi. Presa un'urna a caso, si estraggono due palline senza reimbussolamento da tale urna. Calcolare la probabilità che:

- Vengano estratte due palline rosse
- Venga estratta almeno una pallina rossa
- L'urna dalla quale stiamo estraendo sia la prima quando la prima pallina estratta è rossa
- La seconda pallina estratta dall'urna sia rossa quando la prima è rossa

Esercizio 13 Due urne indistinguibili esternamente. La prima contiene 2 palline rosse, 2 bianche e 4 verdi, mentre la seconda contiene 2 palline rosse, 4 bianche e 2 verdi. Selezioniamo un'urna a caso ed estraiamo da essa due palline senza reimbussolamento. Calcolare la probabilità che:

- Vengano estratte due palline verdi
- L'urna dalla quale stiamo estraendo sia la prima sapendo che la prima pallina estratta è verde
- Vengano estratte due palline di colore diverso

Esercizio 14 Si effettuano tre tiri verso un medesimo bersaglio. La probabilità di colpirlo al primo, al secondo e al terzo colpo sono rispettivamente uguali a $p_1 = 0.4$, $p_2 = 0.5$, $p_3 = 0.7$. Calcolare la probabilità:

- Di colpire il bersaglio una sola volta in tre tiri
- Di colpire il bersaglio al più una volta in tre tiri
- Di colpire il bersaglio almeno una volta in tre tiri
- Che il centro sia avvenuto al terzo colpo sapendo che il bersaglio viene colpito una sola volta

Esercizio 15 Un dado equilibrato viene lanciato 3 volte. Calcolare la probabilità che:

- Il punteggio del primo lancio non sia divisibile per 3
- Il punteggio somma dei primi due lanci non sia divisibile per 3
- Né il punteggio del primo lancio, né il punteggio somma dei primi due lanci siano divisibili per 3

Esercizio 16 Abbiamo dei componenti elettronici che provengono da tre diversi centri di produzione, C1, C2 e C3, la cui probabilità di funzionamento è, rispettivamente, $\frac{1}{10}$, $\frac{2}{10}$ e $\frac{3}{10}$. Sappiamo inoltre che il 30% dei componenti provengono da C1, altri 20% da C2 ed i restanti 50% provengono da C3. Calcolare:

- La probabilità che preso un componente a caso questo sia funzionante
- La probabilità che un componente preso a caso provenga da C1 avendo osservato che esso è funzionante
- La probabilità che almeno uno tra tre componenti presi a caso sia funzionante
- Il numero minimo di componenti da estrarre affinché la probabilità che almeno uno sia funzionante sia maggiore o uguale a 0.999

Esercizio 17 Si supponga di estrarre contemporaneamente 13 carte da un mazzo di 52 (13 carte di cuori, 13 di picche, 13 di fiori, 13 di quadri). Si calcoli la probabilità che le carte così estratte contengano:

- Il tre di cuori
- Una sola carta di quadri
- Tre carte di picche e cinque carte di quadri
- Solo tre figure

Esercizio 19 Da un mazzo regolare di 52 carte (13 cuori, 13 quadri, 13 fiori, e 13 picche) vengono estratte contemporaneamente 3 carte. Qual è la probabilità che le tre carte estratte siano tutte di picche?

Esercizio 20 Supponiamo che un'urna contenga 1 pallina rossa e 1 pallina bianca. Una pallina è estratta e se ne osserva il colore. Essa viene poi rimessa nell'urna insieme a 1 pallina dello stesso colore. Sia R_i l'evento All' i -esima estrazione viene estratta una pallina rossa, analogamente B_i sia l'evento All' i -esima estrazione viene estratta una pallina bianca. Si calcoli:

- $P(R_2)$ e $P(R_3)$
- Sapendo che la seconda estratta è una pallina rossa, è più probabile che la prima pallina estratta fosse rossa o bianca?
- Qual è la probabilità che la prima estratta sia rossa e la seconda bianca?

Esercizio 21 Un'urna contiene 6 palline di cui 3 bianche, 2 rosse ed 1 nera. Si estraggono senza reimmissione tre palline e si vince se una delle tre è nera.

- Si calcoli la probabilità di vincere
- Si calcoli la probabilità di vincere sapendo che la pallina nera non è uscita nelle prime due estrazioni
- Sapendo di aver vinto, qual è la probabilità che la pallina nera non sia uscita nelle prime due estrazioni?

Esercizio 22 Il modello «plus» di una chiave USB può presentare due tipi di difetto, difetto di tipo A con probabilità pari a 0.03 e difetto di tipo B con probabilità pari a 0.07, i due tipi di difetto sono indipendenti l'uno dall'altro. Qual è la probabilità che una generica chiave USB

- presenti entrambe i difetti?
- sia difettosa?
- presenti il difetto A , sapendo che è difettosa?
- presenti uno solo dei difetti, sapendo che è difettosa?

Esercizio 23 Una roulette semplificata è formata da 12 numeri che sono classificati rosso (R) e nero (N) in base allo schema seguente: Siano:

1	2	3	4	5	6	7	8	9	10	11	12
R	R	N	N	R	N	N	R	N	N	R	R

Tabella 1: Schema di estrazione

- $A = \{\text{esce un numero pari}\}$
- $B = \{\text{esce un numero rosso}\}$
- $C = \{\text{esce un numero} \leq 6\}$
- $D = \{\text{esce un numero} \leq 8\}$

Stabilire:

- Se gli eventi A , B e C sono a 2 a 2 indipendenti
- Se A , B e C costituiscono una famiglia di eventi indipendenti
- Stabilire se A , B e D costituiscono una famiglia di eventi indipendenti
- Ponendo $E = \{\text{esce un numero dispari} \leq 3\}$, E è indipendente da A e da D ?
- Supponiamo di sapere che esca un numero rosso, gli eventi A e C sono indipendenti?

Esercizio 24 Da un'urna che contiene 3 palline bianche e 2 palline nere vengono trasferite 2 palline, scelte a caso, in un'altra urna che contiene 4 palline bianche e 4 palline nere. Infine, si estrae una pallina dalla seconda urna.

- Trovare la probabilità di estrarre una pallina bianca dalla seconda urna
- Calcolare la probabilità di aver trasferito almeno una pallina bianca se estraiamo una pallina nera dalla seconda urna

Esercizio 25 Ci sono due urne, la prima urna U_1 contiene due dadi a sei facce (dadi onesti), la seconda urna U_2 contiene due dadi a sei facce (dadi truccati nel modo seguente: ognuno dei dadi ha tre facce che indicano il numero 6 e le rimanenti 3 il numero 5). Si lancia una moneta onesta e se l'esito del lancio è testa si prendono i dadi dalla prima urna U_1 mentre se l'esito è croce si prendono i dadi dalla seconda urna U_2 , poi, in ogni caso si lanciano i dadi.

- Calcolare la probabilità che la somma dei due dadi sia 11
- Sapendo di aver ottenuto un 11 lanciando i due dadi, calcolare la probabilità di aver ottenuto croce lanciando la moneta
- Gli eventi $D_1 = \{\text{ottengo 6 sul primo dado}\}$ e $D_2 = \{\text{ottengo 6 sul secondo dado}\}$ sono indipendenti?

Esercizio 26 Nel gioco del lotto ad ogni estrazione che avviene settimanalmente vengono pescate contemporaneamente 5 palline da un'urna contenente 90 palline numerate da 1 a 90 e per il resto indistinguibili. Giovanni ogni settimana gioca l'ambo $\{89, 90\}$, cioè Giovanni vince se fra le 5 palline estratte vi sono quelle numerate con 89 e 90.

- Calcolare la probabilità di vincere in una singola estrazione
- Se nelle prime 2 estrazioni Giovanni non ha vinto con quale probabilità vincerà almeno una volta nelle prime 10 estrazioni?
- Se nelle prime 10 giocate Giovanni ha vinto tre volte, con quale probabilità ha vinto nelle prime tre giocate?

Esercizio 27 Un'urna contiene 2 palline rosse e 4 palline nere. Due giocatori A e B giocano nel modo seguente: le palline vengono estratte ad una ad una e messe da parte. A vince se l'ultima pallina estratta è rossa, altrimenti vince B.

- Qual è la probabilità che A vinca?
- Qual è la probabilità che A vinca sapendo che la prima pallina estratta è rossa?
- Qual è la probabilità che A vinca e che la prima pallina estratta sia rossa?

Esercizio 28 Gli abitanti della località abc raggiungono ogni giorno la città xyz in treno o in macchina. Il 90% degli abitanti di abc arrivano nella città xyz in ritardo (sull'orario previsto). Il 40% di questi (cioè di quelli in ritardo) usano il treno, mentre, il 20% di quelli che NON arrivano in ritardo (sull'orario previsto) usano la macchina

- Qual è la probabilità che un abitante di abc usi il treno per raggiungere la città xyz (un dato giorno)?
- Se un abitante di abc usa il treno per raggiungere xyz , è più probabile che arrivi in ritardo o non in ritardo?

Esercizio 29 Un mazzo di carte napoletane costituito da 40 carte suddivise in 4 classi (detti “semi”) ciascuna contenente 10 carte numerate. Supponiamo di estrarre “a caso” 8 carte da un tale mazzo.

- Calcolare la probabilità di estrarre i 4 “sette”
- Calcolare la probabilità di estrarre al più 2 “sei”
- Calcolare la probabilità di estrarre “4 sette” e al più 2 “sei”

Esercizio 30 Consideriamo un mazzo di 40 carte suddivise in 4 classi (detti “semi”) ciascuna contenente 10 carte numerate da 1 a 10. Ogni mano servita è formata da cinque carte estratte dal mazzo.

- Calcolare la probabilità di ricevere una mano che contiene i numeri 6, 7, 8, 9, 10
- Calcolare la probabilità di ricevere una mano che contiene cinque numeri distinti

Esercizio 31 Ho due urne A e B : A contiene 6 palline di cui una numerata 1, due numerate 2 e tre numerate 3 e B contiene 9 palline di cui due numerate 2, tre numerate 3 e quattro numerate 4.

Lancio un dado regolare: se esce la faccia 6, estraggo una pallina dall'urna A . Invece, se non esce faccia 6, estraggo una pallina dall'urna B .

- Qual è la probabilità di estrarre una pallina numerata 2?
- Se la pallina estratta è numerata 2, qual è la probabilità che il dado lanciato abbia esibito la faccia 6?
- Sia X la variabile aleatoria discreta che rappresenta il numero della pallina estratta. Determinare la densità di X

Esercizio 32 Giulio, Carlo e Federico giocano al tiro con l'arco. Si sa che la probabilità che Giulio colpisca il bersaglio è pari a $3/4$, quella che Carlo colpisca il bersaglio è pari ad $1/2$ e quella di Federico è pari ad $1/4$. Si sa inoltre che i risultati dei lanci dei tre arcieri sono indipendenti.

- Calcolare la probabilità che nessuno di loro colpisca il bersaglio
- Calcolare la probabilità che esattamente due di loro colpiscano il bersaglio
- Calcolare la probabilità che Giulio colpisca il bersaglio sapendo che esattamente due di loro hanno colpito il bersaglio

Esercizio 33 Un supermercato accetta pagamenti con carte di credito di due soli tipi, A e B . Il 25% dei clienti possiede la carta A , il 50% la carta B e il 15% possiede entrambe le carte.

- Calcolare la probabilità che un cliente scelto a caso possieda almeno una carta di credito accettata dal supermercato
- Calcolare la probabilità che un cliente scelto a caso possieda esattamente una carta di credito accettata dal supermercato
- Se un cliente possiede almeno una carta di credito accettata dal supermercato, qual è la probabilità che sia una carta di credito di tipo A ?

Esercizio 34 Scegliamo tre carte a caso fra un mazzo di 52 carte da gioco (divise in quattro semi con ciascuno tre figure: fante, donna, re).

- Qual è la probabilità che due siano assi e una sia un dieci?
- Qual è la probabilità che almeno due siano figure?
- Qual è la probabilità che almeno due siano figure e nessuna sia un asso?

Esercizio 35 Uno scommettitore ha nel suo portafoglio una moneta equa ed una moneta con entrambe i lati testa. Egli seleziona in modo casuale una delle due monete, la lancia e ne osserva l'esito testa.

- Qual è la probabilità che si tratti della moneta equa?
- Supponiamo che lanci la stessa moneta una seconda volta e che l'esito sia ancora testa. Qual è la probabilità che si tratti della moneta equa?
- Supponiamo che lanci una terza volta la stessa moneta ottenendo croce. Qual è la probabilità che si tratti della moneta equa?

Esercizio 36 In un certo collegio, il 25% degli studenti è stato bocciato in matematica, il 15% è stato bocciato in chimica, e il 10% è stato bocciato sia in matematica che in chimica. Viene scelto a caso uno studente.

- Se egli è stato bocciato in chimica, qual è la probabilità che sia stato bocciato in matematica?
- Se egli è stato bocciato in matematica, qual è la probabilità che sia stato bocciato in chimica?
- Qual è la probabilità che sia stato bocciato in matematica o in chimica?

2 Distribuzioni notevoli

Esercizio 1 Quattro monete bilanciate vengono lanciate. Assumendo l'indipendenza dei risultati, qual è la probabilità di ottenere due testa e due croce?

Esercizio 2 E' noto che gli item prodotti da una macchina utensile saranno difettosi con probabilità 0.1
Qual è la probabilità che in un campione di 3 items al più uno sia difettoso?

Esercizio 3 Supponiamo che il colore degli occhi di una persona sia determinato in base ad una coppia di geni e che « d » rappresenti il gene dominante mentre « r » il gene recessivo. Pertanto una persona con la coppia di geni « dd » ha dominanza pura, una con « rr » ha recessione pura e una con « dr » o « rd » è ibrida.

Un bambino riceve un gene da ognuno dei genitori.

Se rispetto al colore degli occhi i due genitori «ibridi» hanno 4 figli, qual è la probabilità che esattamente 3 dei 4 figli abbiano almeno 1 gene dominante?

Esercizio 4 Supponiamo che il numero di errori tipografici presenti in una singola pagina di un libro sia distribuito secondo una Poisson con parametro $\lambda = 1$.

Si calcoli la probabilità che vi sia almeno un errore in una data pagina.

Esercizio 5 In media su un'autostrada si verificano 3 incidenti al giorno. Qual è la probabilità che non si verifichino incidenti oggi?

Esercizio 6 Un'azienda produce dischetti, la probabilità che un dischetto sia difettoso è pari a 0.01

L'azienda vende i dischetti in confezioni da 10 e rimborsa l'acquirente se più di 1 dischetto è difettoso.

- Quale proporzione delle confezioni sarà restituita?
- Se un acquirente acquista 3 scatole, qual è la probabilità che ne restituisca esattamente una?

Esercizio 7 Si supponga che la probabilità che un prodotto costruito da una macchina sia difettoso è pari a 0.1

Si trovi la probabilità che un campione di 10 prodotti contenga al più un prodotto difettoso.

Esercizio 8 Un bus arriva ad una data fermata ad intervalli di 15 minuti a partire dalle ore 7:00.

Poiché il bus passa ogni quarto d'ora, se un passeggero arriva alla fermata in un istante di tempo uniformemente distribuito nell'intervallo 7:00 - 7:30, si determinino:

- Probabilità che attenda meno di 5 minuti
- Probabilità che attenda almeno 12 minuti

Esercizio 9 Se X è una variabile distribuita secondo una Normale con $\mu = 3$ e $\sigma^2 = 16$, si determini:

- $P(X < 11)$
- $P(X > -1)$
- $P(2 < X < 7)$

Esercizio 10 Un'industria produce su commissione delle sbarre d'acciaio cilindriche, il cui diametro dovrebbe essere di 4 *cm*, ma che tuttavia sono accettabili se hanno diametro compreso fra 3.95 *cm* e 4.05 *cm*. Il cliente, nel controllare le sbarre fornitegli, constata che il 5% sono di diametro inferiore al minimo tollerato ed il 12% di diametro superiore al massimo tollerato

- Supponendo che le misure dei diametri seguano una legge normale, determinarne media e deviazione standard
- Mantenendo la media precedentemente calcolata, determinare quale dovrebbe essere il valore massimo della deviazione standard affinché la probabilità che le sbarre abbiano un diametro superiore al massimo tollerato sia minore 0.05

Esercizio 11 Un'azienda stipula un contratto per vendere barattoli di conserva da 500 g. La quantità di conserva X messa in ciascun barattolo è predeterminata meccanicamente ed è normalmente distribuita con media μ e deviazione standard 25 g.

- A quale valore minimo μ deve essere tarata la macchina, perché non più del 2% dei barattoli contenga meno di 500 g di conserva?
- Supponiamo che i barattoli siano di metallo e che il loro peso Y da vuoti segua una distribuzione $N(90, 64)$. Se un ispettore pesa i barattoli pieni e scarta quelli il cui peso è inferiore a 590 g, quale percentuale di barattoli non passerà l'ispezione?

NB: 64 è la Varianza

Esercizio 12 Un certo tipo di componenti elettronici viene prodotto da una ditta che utilizza due linee di produzione. La prima di queste produce il 40% dei pezzi ed i pezzi prodotti hanno un tempo di vita che segue una legge esponenziale di parametro $\lambda_1 = 1.5$. La seconda invece produce il 60% dei pezzi ed i pezzi prodotti hanno tempo di vita che segue una legge esponenziale di parametro $\lambda_2 = 2$. Un componente viene scelto a caso tra quelli prodotti dalle due linee e indichiamone con X il suo tempo di vita.

- Qual è la probabilità che sia ancora funzionante al tempo $t = 1$?
- Sapendo che è ancora funzionante al tempo $t = 1$, qual è la probabilità che esso sia ancora funzionante al tempo $t = 2$?
- Qual è la probabilità che su 7 componenti scelti a caso e in modo indipendente l'uno dall'altro, almeno 3 siano ancora funzionanti al tempo $t = 1$?
- Supponiamo di scegliere a caso ed in modo indipendente dei componenti. Qual è la probabilità che il primo pezzo non funzionante al tempo $t = 1$ sia il secondo esaminato?

Esercizio 13 Da un mazzo di 52 carte (13 di picche, 13 di cuori, 13 di fiori e 13 di quadri) ne vengono estratte cinque con reinserimento. Si è interessati alla variabile casuale X che descrive il numero di carte di cuori ottenute nelle estrazioni. Determinare:

- il valore atteso e la varianza della variabile X
- la probabilità di estrarre tre carte di cuori
- la probabilità di estrarre almeno tre carte di cuori
- la probabilità di estrarre al più tre carte di cuori

Esercizio 14 L'istante d'arrivo dell'autobus è uniformemente distribuito tra le 10:00 e le 10:30. Tu arrivi alla fermata dell'autobus alle 10:00.

- Qual è la probabilità che tu debba aspettare più di 10 minuti?
- Quanto tempo aspetti in media? Con quale deviazione standard?
- Se l'autobus non è ancora passato alle 10:10, qual è la probabilità che tu debba aspettare almeno altri 10 minuti?
- Supponi di arrivare alla fermata dell'autobus alle 10:00 per 100 giorni. Calcola, in modo approssimato, la probabilità che il tempo totale speso ad attendere l'autobus sia superiore a 1 giorno

Esercizio 15 Il numero X di telefonate ricevute nell'intervallo di tempo $[0, t_0]$ è una variabile aleatoria distribuita secondo una distribuzione di Poisson con parametro $\lambda = t_0$.

Calcolare la probabilità $P(2 \leq X \leq 4)$ di ricevere da due a quattro telefonate (2 e 4 inclusi) entro l'istante $t_0 = 1$.

Esercizio 16 In un dato canale di comunicazione sappiamo che la probabilità di ricevere in modo errato un singolo messaggio è pari a 0.01. Sapendo che viene inviata una sequenza di 150 messaggi, e che i messaggi trasmessi sono stocasticamente indipendenti tra loro, ci si chiede quale sia la probabilità che due dei messaggi ricevuti siano errati.

Esercizio 17 In una certa provincia montuosa si può supporre che il numero X di frane al mese sia una variabile aleatoria con la legge di Poisson di parametro $\lambda = 2.3$.

- Calcolare la probabilità che ci siano almeno due frane in un dato mese.
- Quanto dovrebbe valere il parametro λ affinché la probabilità che in un mese non ci siano frane sia superiore a $\frac{1}{2}$?

Esercizio 18 Il “Crazy Boat” è un battello a due motori utilizzato per le crociere sul Tamigi. I due motori lavorano indipendentemente e il numero di piccoli guasti in una singola crociera è modellabile tramite una v.a. X di Poisson di parametro $\lambda = 1$ per il primo motore, e da una v.a. Y di Poisson di parametro $\lambda = 2$ per il secondo.

- Qual è la distribuzione della v.a. $X + Y$?
- Calcolare la probabilità che non avvenga alcun guasto in una data crociera
- Partono 10 modelli identici “Crazy Boat”. Calcolare la probabilità che almeno 2 battelli concludano la crociera senza guasti

3 Stima di parametri

Esercizio 1 Si desidera stimare il tempo medio di CPU utilizzato dagli utenti di un server in modo da affermare con il 95% di confidenza che il valore stimato non disti più di $\frac{1}{2}$ secondo dal valore vero.

Dall'esperienza passata possiamo ipotizzare che il tempo di CPU utilizzato sia distribuito secondo una normale con $\sigma^2 = 2.25$.

Quale deve essere la dimensione n del campione?

Esercizio 2 Un segnale μ viene trasmesso da una sorgente A ad una destinazione B . Il segnale ricevuto in B è distribuito normalmente con media μ e varianza σ^2 incognita. Se vengono trasmessi 9 segnali successivi con valore μ eguale e vengono ricevuti in B i seguenti valori:

5 - 8.5 - 12 - 15 - 7 - 9 - 7.5 - 6.5 - 10.5

Si calcoli un intervallo di confidenza con il 95% di confidenza per μ .

Esercizio 3 Si determini l'intervallo di confidenza al 95% del valore medio del battito cardiaco dei membri di un club di fitness su un campione casuale di 15 individui che ha portato ad ottenere:

94 - 63 - 58 - 72 - 49 - 92 - 70 - 73 - 69 - 104 - 48 - 66 - 80 - 64 - 77

Esercizio 4 Una fabbrica produce alzacristalli elettrici la cui durata di vita si distribuisce Normalmente con varianza $\sigma^2 = 0.25$. Si supponga di estrarre un campione casuale di 49 pezzi aventi durata media pari a 5.5 anni. Costruire un intervallo di confidenza al 95% per la durata media degli alzacristalli prodotti dalla fabbrica.

Esercizio 5 La durata di funzionamento degli utensili prodotti da una macchina segue la distribuzione normale con scarto quadratico medio $\sigma = 7$. Un campione casuale di 25 utensili ha presentato una durata di funzionamenti di 5 mesi. Costruire un intervallo di confidenza al 99% per la durata media della popolazione degli utensili prodotti dalla macchina.

Esercizio 6 Un segnale radio viene emesso con frequenza distribuita normalmente e con valore atteso μ e deviazione standard 30 kHz . Supponendo di osservare la seguente serie di frequenze in kHz:

610 - 601 - 578 - 615 - 640 - 630 - 618 - 602 - 613 - 610 - 625 - 585 - 622 - 608 - 597

- determinare una stima di μ e la probabilità che la frequenza stia nell'intervallo di estremi 590 kHz e 610 kHz.
- Determinare poi un intervallo di confidenza per μ al 95%.

Esercizio 7 Tra i pasticcini prodotti artigianalmente in una pasticceria se ne prelevano 100; risulta che il loro peso medio è pari a 35g. Si sa che lo scarto quadratico medio del peso di tutti i pasticcini prodotti dalla pasticceria è pari a 4g.

- Si trovi l'intervallo di confidenza per il peso medio di tutti i pasticcini prodotti a livello di confidenza del 98%.
- Di quanto deve aumentare la numerosità campionaria se si vuole che l'ampiezza dell'intervallo si dimezzi?
- Si determini quanti pasticcini occorre ancora estrarre se si vuole che lo stimatore del peso medio si discosti dal vero peso medio per meno di 1 grammo con probabilità del 96%.

Esercizio 8 Da un lotto di gelati se ne estraggono 100 e se ne calcola in 82g il valore del peso medio. Sapendo che la varianza è pari a 25

- si determini l'intervallo di confidenza per il peso medio μ dei gelati al livello di confidenza del 97%
- si determini la probabilità che la differenza in valore assoluto fra la media campionaria e il peso medio μ dei gelati sia inferiore a 3g.

Esercizio 9 Si è svolta un'indagine su 100 persone per saggiare l'opinione su una proposta politica. Avendo ottenuto 48 risposte favorevoli:

- si determini l'intervallo di confidenza per la proporzione di risposte favorevoli nella popolazione con un livello di confidenza del 97%
- si determini quanto deve essere l'ampiezza campionaria se si vuole che la varianza dello stimatore della suddetta proporzione non sia superiore a 0.001, tenendo conto dei risultati ottenuti dall'indagine sulle 100 persone.

Esercizio 10 In un vivaio ci sono 1000 alberi tra i quali una proporzione incognita p ha contratto una malattia.

- Stabilire quanti alberi occorre controllare affinché l'intervallo di confidenza a livello $1 - \alpha = 0.99$ per la proporzione incognita di alberi malati risulti ampio meno di 0.1.
- Si decide di selezionare 100 alberi con reposizione al fine di stimare la proporzione p di alberi che hanno contratto la malattia. Di questi 100 risulta che 40 hanno contratto la malattia. Si calcoli l'intervallo di confidenza per la proporzione p al livello di confidenza del 98%.
- Tenendo conto del risultato campionario di cui al punto 2), si determini la numerosità campionaria che assicura che la varianza dello stimatore della proporzione di alberi ammalati sia pari a 0.001.

Esercizio 11 Si vuole verificare se la quantità X di una sostanza inquinante emessa dalle marmitte prodotte da un'azienda è contenuta entro i limiti restabiliti. A tal fine, si estrae un campione di 10 marmitte dalla produzione settimanale dell'azienda e attraverso prove su strada si rilevano le seguenti quantità (in mg per Km) della sostanza nociva rilasciate:

895 902 894 903 901 893 897 908 906 891

Sapendo che la quantità emessa della sostanza in esame ha distribuzione normale di parametri μ e σ^2 incogniti, determinare la stima intervallare di σ^2 al livello di confidenza del 99%

Esercizio 12 Il responsabile del controllo qualità di un'azienda che produce manufatti cementizi è interessato a determinare il diametro dei tubi in calcestruzzo fabbricati in serie. Egli sa che, a causa degli errori di misurazione, la misura che può effettuare con i suoi strumenti è una realizzazione di una variabile casuale X distribuita secondo una normale di media μ e varianza σ^2 . Dopo aver effettuato 25 misurazioni ottiene che la somma dei valori osservati è pari a 4250 mm, e la varianza è 225.

- Calcolare l'intervallo di confidenza con confidenza al 95% per la media μ .
- Determinare la numerosità campionaria necessaria affinché l'intervallo di confidenza per la media del diametro dei tubi al livello del 95% abbia un ampiezza $b = 10$ mm.

4 Test Parametrici

Esercizio 1 Per provare l'ipotesi che una moneta è buona viene adottata la seguente regola di decisione. Si accetta l'ipotesi se il numero di teste in un campione di 100 lanci è comprese tra 40 e 60 inclusi. In caso contrario si rifiuta l'ipotesi. Si determini la probabilità di rifiutare l'ipotesi quando essa è vera.

Esercizio 2 Progettare la regola di decisione per provare che l'ipotesi che una moneta è buona, se si prende in considerazione un campione di 64 lanci della moneta e se si considera un livello di significatività dello 0.05

Esercizio 3 Una fabbrica produce corde le cui resistenze alla rottura hanno una media di 300 N e uno scarto quadratico medio di 24 N . Si crede che per mezzo di un recente processo produttivo la resistenza media alla rottura possa essere aumentata. Si determini una regola di decisione per il rifiuto del vecchio processo produttivo al livello di significatività dello 0.01 se si voglia provare 64 corde.

Esercizio 4 Si consideri l'esercizio precedente.

Qual è la probabilità di accettare il vecchio processo produttivo con la regola di decisione stabilita quando in realtà il nuovo processo ha aumentato la resistenza media alla rottura a 310N?

Esercizio 5 Una macchina che controlla la qualità di nastri viene valutata efficiente se la deviazione standard della numerosità per cm^2 dei difetti sui nastri non è maggiore di 0.15 cm. Il numero di difetti è distribuito secondo una normale.

Se un campione di 20 nastri porta ad ottenere $\hat{S}_n^2 = 0.025 \text{ cm}^2$, si può concludere che la macchina sia efficace con il 90% di confidenza?

Esercizio 6 Sono possibili due scelte di un catalizzatore per effettuare un certo processo chimico. Per controllare se la varianza della resa è la medesima indipendentemente dal catalizzatore si estrae un campione di 10 elementi dove è stato utilizzato il primo catalizzatore, e un campione di 12 elementi dove è stato utilizzato il secondo catalizzatore. Si è ottenuto quanto segue: $\hat{S}_1^2 = 0.14$ e $\hat{S}_2^2 = 0.28$. E' possibile rifiutare l'ipotesi di uguale varianza con il 10% di significatività?

Esercizio 7 Una compagnia di assicurazioni vuole valutare l'entità media delle richieste di risarcimento danni per incidenti automobilistici. Un'indagine svolta su di un campione di 25 richieste ha dato i seguenti risultati (con X si indica la variabile "richiesta di risarcimento in migliaia di euro")

$$\sum_{i=1}^{25} X_i = 112.12$$

$$\sum_{i=1}^{25} X_i^2 = 629.89$$

- Stimare l'entità media della richiesta e la varianza delle richieste di risarcimento, giustificando la scelta degli stimatori usati
- Ipotizzando che X abbia distribuzione gaussiana, saggiare, ad un livello di significatività $\alpha = 0.05$ l'ipotesi $H_0: \mu = 3$ contro l'ipotesi alternativa $H_1: \mu > 3$

Esercizio 8 Il processo di riempimento delle confezioni di pasta di una azienda non è perfetto: il peso dichiarato è 500 grammi, ma è plausibile che vi siano confezioni con un peso superiore e altre con un peso inferiore. Un'associazione per la tutela del consumatore vuole verificare se il macchinario che riempie le confezioni è veramente tarato su 500 grammi oppure è tarato su un peso inferiore. Viene selezionato casualmente un campione di 40 confezioni, e ne viene pesato il contenuto; indicato con X il “peso in grammi di una confezione”. Si osserva

$$\sum_{i=1}^{40} X_i = 19840.02$$

$$\sum_{i=1}^{40} X_i^2 = 9841173.7$$

Ipotizzando la normalità distributiva del carattere X , si può affermare, ad un livello di significatività dell'1% che il macchinario è tarato su un peso inferiore a 500 grammi?

Esercizio 9 Si esegue un'indagine sulla spesa mensile in cancelleria di un campione di studenti e studentesse di età compresa tra i 15 ed i 18 anni e si osserva che:

$$n_M = 10 \text{ e } S_M^2 = 210$$

$$n_F = 12 \text{ e } S_F^2 = 225$$

Verificare ad un livello di significatività del 5% se la varianza della spesa mensile in cancelleria delle studentesse è significativamente più grande di quella degli studenti.

Esercizio 10 Vengono collezionati i seguenti valori di peso ed altezza per 10 individui

Peso	73	68	71	82	87	64	61	79	83	88
Altezza	172	165	173	184	190	162	164	181	185	190

Tabella 2: Tabella

Sotto ipotesi che peso e altezza siano variabili casuali distribuite secondo una distribuzione normale, verificare ad un livello di significatività del 5% se peso ed altezza risultano correlati.

Esercizio 11 Viene condotto uno studio finalizzato a verificare se l'ipnosi è in grado di ridurre il livello di dolore percepito. I risultati ottenuti su un campione casuale di soggetti sono mostrati di seguito:

	A	B	C	D	E	F	G	H
PRIMA	6,6	6,5	9,0	10,3	11,3	8,1	6,3	11,6
DOPO	6,8	2,4	7,4	8,5	8,1	6,1	3,4	2,0

Tabella 3: Schema

Per ogni soggetto vengono riportati i livelli di dolore percepiti prima e dopo ipnosi.

Si determini con livello di significatività del 5% se il livello di dolore medio è inferiore dopo aver ipnotizzato il soggetto.

Esercizio 12 Due meli, A e B, producono frutti con pesi variabili. Si sa che il peso delle mele prodotte è una variabile aleatoria distribuita normalmente. Vengono raccolti e pesati 10 frutti dal melo A e 15 frutti dal melo B, ottenendo le seguenti stime puntuali

$$\overline{x_A} = 75 \text{ e } s_A^2 = 2.8$$

$$\overline{x_B} = 70 \text{ e } s_B^2 = 1.5$$

Si determini con livello di significatività del 5% se i pesi medi dei frutti prodotti dai due meli sono differenti.

5 Test non parametrici

Esercizio 1 Un campione di dimensione 10, collezionato sul valore assunto dalla variabile aleatoria X ha fornito i seguenti valori:

21 - 23 - 20 - 18 - 28 - 30 - 20 - 15 - 13 - 12

Controllare con Kolmogorov-Smirnov, con significatività 0.1, l'ipotesi che X sia distribuita secondo una distribuzione uniforme $[10,40]$.

Esercizio 2 Un campione di numerosità 5 estratto da una popolazione X ha fornito i seguenti valori:

6.2 - 8.8 - 9.4 - 13.6 - 10

Controllare con Kolmogorov-Smirnov e livello di significatività 0.1 l'ipotesi che X sia distribuita:

- Secondo una uniforme $[0, 20]$
- Secondo un'esponenziale di parametro $\lambda = 0.1$.

Esercizio 3 Un campione di dimensione 30 estratto da una popolazione X ha fornito i seguenti valori:

2.03 - 9.82 - 2.50 - 6.14 - 4.44 - 4.62 - 10.10 - 7.34 - 8.31 - 8.32 - 16.46 - 3.61
- 16.55 - 15.49 - 17.87 - 6.93 - 10.96 - 4.53 - 17.72 - 7.52 - 4.69 - 6.70 - 16.24
- 6.19 - 13.14 - 16.73 - 10.64 - 11.53 - 2.78 - 15.36

Controllare con il metodo del Chi-Quadro, con significatività 0.01, l'ipotesi che X sia distribuito secondo un'uniforme con parametri $[0,20]$. Dividere l'intervallo in 4.

Esercizio 4 Un campione di dimensione 80 estratto da una popolazione X ha fornito i conteggi indicati nella tabella seguente, conteggio di casi che assumono valore in 4 intervalli da 1 a 4 aventi egual ampiezza:

I_i	1	2	3	4
n_i	15	22	18	25

Tabella 4: Tabella

Controllare con il metodo del Chi-Quadro, con significatività 0.05, l'ipotesi che X sia uniformemente distribuita su $[1,4]$.

Esercizio 5 Si vuole controllare l'effetto di una terapia su due gruppi di animali da laboratorio ammalati. I dati di cui si dispone sono i seguenti:

	Animali non guariti	Animali guariti
Con terapia	22	48
Senza terapia	24	26

Tabella 5: Tabella

Controllare l'ipotesi che la terapia sia efficace, con un livello di significatività $\alpha = 0.05$.

Esercizio 6 Si desidera controllare l'ipotesi di indipendenza tra due caratteri discreti X ed Y di una popolazione. Il campione, di numerosità 150, ha fornito i conteggi riportati in tabella:

$X \backslash Y$	0	1	2
-1	25	15	10
0	20	30	25
1	5	15	5

Tabella 6: Tabella

Controllare l'ipotesi di indipendenza di X ed Y con un livello di significatività 0.05

Esercizio 7 Sono stati ottenuti dall'anagrafe i dati relativi ai primi tre figli di 1054 famiglie con almeno tre figli. Ne risulta che 157 famiglie hanno tre figli maschi, 371 due maschi e una femmina, 362 un maschio e due femmine, 164 tre femmine.

Questi dati sono compatibili con l'ipotesi che ogni figlio nasca maschio (o femmina) con probabilità $1/2$ indipendentemente dagli altri? (significatività $\alpha = 0.1$)

Esercizio 8 Sospettiamo che un dado utilizzato ad un tavolo da gioco di un famoso Casinò non sia equo. Per verificare il sospetto ci appostiamo al tavolo in questione e ne osserviamo il seguente esito di 60 lanci:

ESITO	1	2	3	4	5	6
OCCORRENZA	9	14	7	9	10	11

Tabella 7: Tabella

Cosa possiamo concludere circa il nostro sospetto che il dado non sia equo? (significatività $\alpha=0.05$)

Esercizio 9 Una azienda afferma che un integratore di sua ideazione è in grado di influire sul livello di colesterolo nel sangue, per tale ragione viene prodotto uno studio dell'azienda medesima dove a 10 individui è stato misurato il livello del colesterolo prima di assumere l'integratore in questione e dopo 2 settimane di assunzione del medesimo. Si è ottenuto quanto segue:

	1	2	3	4	5	6	7	8	9	10
PRIMA	225	310	287	249	345	288	247	268	213	332
DOPO	210	301	291	212	307	290	216	245	195	301

Tabella 8: Tabella

Cosa possiamo concludere circa l'affermazione dell'azienda? (significatività $\alpha = 0.05$)

Esercizio 10 Una azienda afferma che un integratore di sua ideazione è in grado di influire sul livello di colesterolo nel sangue, per tale ragione viene prodotto uno studio dell'azienda medesima dove è stato misurato il livello del colesterolo a 15 individui che non assumono l'integratore e a 10 individui che invece assumono l'integratore. Si è ottenuto quanto segue:

	1	2	3	4	5	6	7	8	9	10
Int	225	310	287	249	345	288	247	277	267	194
No int	210	301	291	212	307	290	216	312	206	185

	11	12	13	14	15
Int					
No int	199	204	245	195	301

Tabella 9: Tabella

Cosa possiamo concludere circa l'affermazione dell'azienda? (significatività $\alpha = 0.1$)

Esercizio 11 La tabella sottostante riporta il rango dello Human Development Index (HDI) e dell'Income per 10 Paesi.

	RANGO	
Paese	Income	HDI
Norvegia	3	1
Islanda	2	2
Svezia	10	3
Australia	8	4
Olanda	5	5
Belgium	7	6
USA	1	7
Canada	6	8
Giappone	9	9
Svizzera	4	10

Tabella 10: Tabella

Cosa possiamo concludere circa la correlazione tra Income e HDI? (significatività $\alpha = 0.1$)

Esercizio 12 Pazienti leucemici vengono sottoposti a due trattamenti A e B . Il conteggio delle piastrine (103) viene riportato nella tabella.

Trattamento	PAZIENTI					
A	20,50	22,53	25,70	13,23	29,67	24,46
B	10,56	28,13	19,94	11,03	8,09	12,95

Trattamento	PAZIENTI					
A	26,07	19,35	17,81	16,00	13,50	32,90
B	21,14	32,50	10,90			

Tabella 11: Tabella

Possiamo affermare che tali conteggi provengano dalla stessa distribuzione?
(significatività $\alpha = 0.05$)

6 Formulario

6.1 Test parametrico su binomiale

Statistica $Z_n = \frac{X - \mu}{\sigma} \sim N(0, 1)$

Regione critica

$$C'_Z = (-\infty, -z_{1-\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, +\infty)$$

$$C''_Z = (-\infty, -z_{1-\alpha})$$

$$C'''_Z = (z_{1-\alpha}, +\infty)$$

6.2 Test parametrico sulla media μ

6.2.1 Normalmente Distribuita e σ^2 nota

Statistica $Z_n = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$

Regione critica

$$C'_{Z_n} = (-\infty, -z_{1-\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, +\infty)$$

$$C''_{Z_n} = (-\infty, -z_{1-\alpha})$$

$$C'''_{Z_n} = (z_{1-\alpha}, +\infty)$$

Statistica $\bar{X}_n \sim N(\mu, \frac{\sigma}{\sqrt{n}})$

Regione critica

$$C'_{\bar{X}_n} = (-\infty, -z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} + \mu_0) \cup (z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} + \mu_0, +\infty)$$

$$C''_{\bar{X}_n} = (-\infty, -z_{1-\alpha} \frac{\sigma}{\sqrt{n}} + \mu_0)$$

$$C'''_{\bar{X}_n} = (-z_{1-\alpha} \frac{\sigma}{\sqrt{n}} + \mu_0, +\infty)$$

6.2.2 Normalmente Distribuita e σ^2 non nota

Statistica $T_n = \frac{\overline{X_n} - \mu}{\frac{\hat{S}_n}{\sqrt{n}}} \sim t^{n-1}$

Regione critica

$$C'_{T_n} = (-\infty, -t_{1-\frac{\alpha}{2}}^{n-1}) \cup (t_{1-\frac{\alpha}{2}}^{n-1}, +\infty)$$

$$C''_{T_n} = (-\infty, -t_{1-\alpha}^{n-1})$$

$$C'''_{T_n} = (t_{1-\alpha}^{n-1}, +\infty)$$

Statistica $\overline{X_n} \sim N(\mu, \frac{\hat{S}_n}{\sqrt{n}})$

Regione critica

$$C'_{\overline{X_n}} = (-\infty, -t_{1-\frac{\alpha}{2}}^{n-1} \frac{\hat{S}_n}{\sqrt{n}} + \mu_0) \cup (t_{1-\frac{\alpha}{2}}^{n-1} \frac{\hat{S}_n}{\sqrt{n}} + \mu_0, +\infty)$$

$$C''_{\overline{X_n}} = (-\infty, -t_{1-\alpha}^{n-1} \frac{\hat{S}_n}{\sqrt{n}} + \mu_0)$$

$$C'''_{\overline{X_n}} = (t_{1-\alpha}^{n-1} \frac{\hat{S}_n}{\sqrt{n}} + \mu_0, +\infty)$$

6.2.3 Non normalmente distribuita

$n \geq 30$

6.3 Test parametrico sulla varianza σ^2

Statistica $Q_n = \frac{(n-1)\hat{S}_n^2}{\sigma^2} \sim \chi_{n-1}^2$

Regione critica

$$C'_{Q_n} = (0, q_{\frac{\alpha}{2}}^{n-1}) \cup (q_{1-\frac{\alpha}{2}}^{n-1}, +\infty)$$

$$C''_{Q_n} = (0, q_{\alpha}^{n-1})$$

$$C'''_{Q_n} = (q_{1-\alpha}^{n-1}, +\infty)$$

Statistica \hat{S}_n^2

Regione critica

$$C'_{\hat{S}_n^2} = (0, q_{\frac{\alpha}{2}}^{n-1} \frac{\sigma_o^2}{n-1}) \cup (q_{1-\frac{\alpha}{2}}^{n-1} \frac{\sigma_o^2}{n-1}, +\infty)$$

$$C''_{\hat{S}_n^2} = (0, q_{\alpha}^{n-1} \frac{\sigma_o^2}{n-1})$$

$$C'''_{\hat{S}_n^2} = (q_{1-\alpha}^{n-1} \frac{\sigma_o^2}{n-1}, +\infty)$$

Statistica S_n^2

Regione critica

$$C'_{S_n^2} = (0, q_{\frac{\alpha}{2}}^{n-1} \frac{\sigma_o^2}{n}) \cup (q_{1-\frac{\alpha}{2}}^{n-1} \frac{\sigma_o^2}{n}, +\infty)$$

$$C''_{S_n^2} = (0, q_{\alpha}^{n-1} \frac{\sigma_o^2}{n})$$

$$C'''_{S_n^2} = (q_{1-\alpha}^{n-1} \frac{\sigma_o^2}{n}, +\infty)$$

6.4 Test parametrico sulla differenza delle medie di due popolazioni $\mu_X - \mu_Y$

Statistica

$$T_n = \frac{(\bar{X}_n - \bar{Y}_m) - \mu}{\sqrt{\frac{(n+m)(nS_{X,n}^2 + mS_{Y,m}^2)}{nm(n+m-2)}}} \sim t^{n+m-2}$$

NB: $S_{X,n}^2$ e $S_{Y,m}^2$ sono varianze campionarie **NON** corrette

NB: Le due popolazioni abbiano **identica varianza** anche se sconosciuta

$$Z_{n,m} = \frac{\bar{X}_n - \bar{Y}_m - \mu}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim N(0, 1)$$

Regione critica

$$C'_{T_{n,m}=Z_{n,m}} = (-\infty, -t_{1-\frac{\alpha}{2}}^{n+m-2}) \cup (t_{1-\frac{\alpha}{2}}^{n+m-2}, +\infty)$$

$$C''_{T_{n,m}=Z_{n,m}} = (-\infty, -t_{1-\alpha}^{n+m-2})$$

$$C'''_{T_{n,m}=Z_{n,m}} = (t_{1-\alpha}^{n+m-2}, +\infty)$$

Intervallo di confidenza

$$\begin{aligned} & (\bar{X}_n - \bar{Y}_m) - t_{1-\frac{\alpha}{2}}^{n+m-2} \sqrt{\frac{(n+m)(nS_{X,n}^2 + mS_{Y,m}^2)}{nm(n+m-2)}} \\ & \leq \mu \leq \\ & (\bar{X}_n - \bar{Y}_m) + t_{1-\frac{\alpha}{2}}^{n+m-2} \sqrt{\frac{(n+m)(nS_{X,n}^2 + mS_{Y,m}^2)}{nm(n+m-2)}} \end{aligned}$$

6.5 Test parametrico sull'uguaglianza delle varianze di due popolazioni

Statistica $V_{n,m} = \frac{S_{X,n}^2}{S_{Y,m}^2} \sim F_{n-1,m-1}$

Regione critica

$$C'_{V_{n,m}} = \left[0, F_{\frac{\alpha}{2}}^{n-1,m-1}\right) \cup \left(F_{1-\frac{\alpha}{2}}^{n-1,m-1}, +\infty\right)$$

$$C''_{V_{n,m}} = \left[0, F_{\alpha}^{n-1,m-1}\right)$$

$$C'''_{V_{n,m}} = \left(F_{1-\alpha}^{n-1,m-1}, +\infty\right)$$

6.6 Test parametrico di incorrelazione

Statistica

$$R_n = \frac{\sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{n S_{X,n} S_{Y,n}}$$

NB: $S_{X,n}$ e $S_{Y,n}$ sono le **radici** delle varianze campionarie $S_{X,n}^2$ e $S_{Y,n}^2$

$$\hat{T}_n = R_n \cdot \sqrt{\frac{n-2}{1-R_n^2}} \sim t^{n-2}$$

Regione critica

$$C_{\hat{T}_n} = (-\infty, -t_{1-\frac{\alpha}{2}}^{n-2}) \cup (t_{1-\frac{\alpha}{2}}^{n-2}, +\infty)$$

6.7 Test Bontà Adattamento di Kolmogorov-Smirnov

1. Ordinare dati - Riga 1
2. $F(t)$ - Riga 2
3. $\hat{F}_{X,n}(t)$ - Riga 3, $\hat{F}_{X,n}(< t)$ - Riga 4
4. $D_n = \sup|F(t) - \hat{F}_{X,n}(t)|$
5. $C_{D_n} = (d_{1-\alpha}, 1]$

6.8 Test Bontà Adattamento Chi-Quadro

1. Dividere in intervalli $I_k = [t_k, t_{k+1})$
2. Per ogni intervallo determinare quante realizzazioni ci cadono n_k
3. Per ogni intervallo determinare $p_k = F(t_{k+1}) - F(t_k)$
4. $W = \sum_{i=1}^k \frac{(n_k - np_k)^2}{np_k}$
5. $C_W = (\chi_{1-\alpha, k-1}^2, +\infty]$

6.9 Test confr. distr. due pop. Segni

1. $S_n = S^+ - S^-$
2. $C_{S_n} = (-\infty, -z_{1-\frac{\alpha}{2}} \frac{n-S^=}{2}) \cup (z_{1-\frac{\alpha}{2}} \frac{n-S^=}{2}, +\infty)$

6.10 Test confr. distr. due pop. Ranghi di Wilcoxon

1. Ordinare dati e metterli tutti nella stessa riga
2. Segnare POS, RANGO, PROVENIENZA (X, Y)
3. $U_x = nm + \frac{n(n+1)}{2} - R_x$, R_x Somma ranghi X
4. $U_y = nm + \frac{m(m+1)}{2} - R_y$, R_y Somma ranghi Y
5. $U = \min\{U_x, U_y\}$
6. $\hat{Z}_{n,m} = \frac{U - \frac{nm}{2}}{\sqrt{\frac{n \cdot m \cdot (n+m+1)}{12}}}$
7. $C = (-\infty, -z_{1-\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, +\infty)$

6.11 Test Chi-Quadro indipendenza

1. Tabella frequenze relative
2. Frequenze marginali $f_a^X, f_b^X, f_a^Y, f_b^Y$
3. Tabella incroci $f_a^X \cdot f_a^Y, f_b^X \cdot f_b^Y \dots$
4. $W = n \sum_{k=1}^{M_X} \sum_{j=1}^{M_Y} \frac{(f_{k,j} - f_k^X \cdot f_j^Y)^2}{f_k^X \cdot f_j^Y}$
5. $n \cdot \frac{(\text{Primacella1T} - \text{Primacella2T})^2}{\text{Primacella2T}} + \frac{(\text{Secondacella1T} - \text{Secondacella2T})^2}{\text{Secondacella2T}} \dots$
6. $C = (\chi_{1-\alpha, (M_x-1)(M_y-1)}^2, +\infty)$

6.12 Test di Correlazione dei ranghi R_s di Spearman

1. Ad ogni coppia di dati assegnare la coppia di ranghi
2. $d_i^2 = (r_i^X - r_i^Y)^2$
3. $R_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}$
4. $\tilde{T}_n = R_s \cdot \sqrt{\frac{n-2}{1-R_s^2}}$
5. $C = (-\infty, -t_{1-\frac{\alpha}{2}}^{n-2}) \cup (t_{1-\frac{\alpha}{2}}^{n-2}, +\infty)$

6.13 Regressione Lineare

$$\begin{aligned}A_0 &= \bar{Y} - A_1 \bar{X} \\A_1 &= \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \\S_{RES}^2 &= \frac{1}{n-2} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \\c_{x,y} &= \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n} \\r_{x,y} &= \frac{c_{x,y}}{s_x s_y} \\R^2 &= r_{x,y}^2\end{aligned}$$

6.13.1 Test

Statistica

$$T_0 = \frac{A_0 - \alpha_0}{\sqrt{S_{RES}^2 \cdot \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right]}} \sim t_{n-2}$$

Regione critica per T_0

$$C'_{T_0} = (-\infty, -t_{1-\frac{\alpha}{2}}^{n-2}) \cup (t_{1-\frac{\alpha}{2}}^{n-2}, +\infty)$$

$$C''_{T_0} = (-\infty, -t_{1-\alpha}^{n-2})$$

$$C'''_{T_0} = (t_{1-\alpha}^{n-2}, +\infty)$$

Regione di confidenza per A_0

$$\left[\alpha_0 - t_{1-\frac{\alpha}{2}} \sqrt{S_{RES}^2 \left[\frac{1}{n} + \frac{\bar{X}^2}{S_x^2} \right]}, \alpha_0 + t_{1-\frac{\alpha}{2}} \sqrt{S_{RES}^2 \left[\frac{1}{n} + \frac{\bar{X}^2}{S_x^2} \right]} \right]$$

Statistica

$$T_1 = \frac{A_1 - \alpha_1}{\sqrt{S_{RES}^2 \cdot \left[\frac{1}{\sum_{i=1}^n (X_i - \bar{X})^2} \right]}} \sim t_{n-2}$$

Regione critica per T_1

$$C'_{T_1} = (-\infty, -t_{1-\frac{\alpha}{2}}^{n-2}) \cup (t_{1-\frac{\alpha}{2}}^{n-2}, +\infty)$$

$$C''_{T_1} = (-\infty, -t_{1-\alpha}^{n-2})$$

$$C'''_{T_1} = (t_{1-\alpha}^{n-2}, +\infty)$$

Regione di confidenza per A_1

$$\left[\alpha_1 - t_{1-\frac{\alpha}{2}} \sqrt{S_{RES}^2 \frac{1}{S_x^2}}, \alpha_1 + t_{1-\frac{\alpha}{2}} \sqrt{S_{RES}^2 \frac{1}{S_x^2}} \right]$$

6.13.2 Attendibilità del modello lineare

Statistica $\tilde{F} = (n-2) \frac{D_{SP}}{D_{RES}} = \frac{D_{SP}}{S_{RES}^2} \sim F_{(1, n-2)}$

7 Domande aperte

Il candidato introduca le rappresentazioni per caratteri bidimensionali con riferimento alla statistica descrittiva

Sia $S = \{(s_j, u_k) | j = 1, \dots, n; k = 1, \dots, m\}$ l'insieme delle coppie dei caratteri analizzati

Il candidato descriva la legge dei grandi numeri ed il teorema centrale limite fornendone il campo di applicazione (eventualmente con esempio numerico)

Sia $\{X_i \mid i \in N\}$ una successione di V.A. indipendenti ed identicamente distribuite (I.I.D.)

Se esistono sia valore atteso che varianza finiti, allora la successione delle corrispondenti medie aritmetiche $\{\bar{X}_n \mid n \in N\}$ tende al crescere di n ad una variabile che assume il valore $E[X_i] = \mu$.

Considerata poi una variabile $S_n = X_1 + X_2 + \dots + X_n$ vale che S_n converge in distribuzione a $X \sim N(n \cdot \mu, \sqrt{n} \cdot \sigma)$

Il candidato introduca e descriva le rappresentazioni numeriche di dati statistici con riferimento alla statistica descrittiva, frequenza assoluta, frequenza cumulata assoluta ecc.

L'insieme dei dati viene denotato con $E = \{x_1, \dots, x_n\}$, mentre quello dei caratteri assumibili (le "modalità" del carattere) viene denotato con $S = \{s_1, \dots, s_m\}$.

- **Frequenza assoluta** f_j della modalità s_j : Numero di elementi di E aventi valore s_j
- **Frequenza cumulata assoluta** F_j della modalità s_j : Somma di f_j e delle frequenze assolute che la precedono
- **Frequenza relativa** p_j della modalità s_j : Rapporto tra f_j e n
- **Frequenza cumulata relativa** P_j della modalità s_j : Somma di p_j e delle frequenze relative che la precedono

Il candidato illustri gli indici di tendenza centrale e variabilità

- Moda: Valore per cui è massima la frequenza assoluta
- Media: Momento primo rispetto l'origine $\bar{x} = \frac{X_1 + \dots + X_n}{n}$
- Mediana: Valore che, dopo aver ordinato i dati, divide in parti uguali la serie
- Varianza: Valore che indica la dispersione dei dati rispetto la media
- Deviazione standard: Radice quadrata della varianza

Il candidato definisca il concetto di indipendenza stocastica sia nel caso di eventi che nel caso di variabili aleatorie

Due eventi si dicono stocasticamente indipendenti se $P(A \cap B) = P(A)P(B)$ oppure $P(A|B) = P(A)$ e $P(B|A) = P(B)$

Due variabili aleatorie si dicono stocasticamente indipendenti se $F_{X,Y}(t, s) = F_X(t)F_Y(s)$

Si introduca il concetto di test di ipotesi, se ne elenchino ed illustrino le principali componenti, si ponga particolare attenzione al concetto di errore di I e II tipo discutendone la relazione esistente

La “verifica di ipotesi” si occupa di validare, sulla base di un campione, un’ipotesi relativa ad una caratteristica della popolazione, la quale può essere rifiutata o no. Le sue caratteristiche principali sono:

- H_0 : Ipotesi nulla
- H_1 : Ipotesi alternativa
- X : Popolazione statistica
- $T = T(X_1, \dots, X_n)$: Statistica campionaria di cui è nota la distribuzione quando H_0 è falsa
- \bar{C} : Regione di accettazione, insieme di valori assumibili da T che portano a non rifiutare T
- C : Regione critica, insieme di valori assumibili da T che portano a rifiutare T

- α : Livello di significatività ed è tale che quando H_0 è vera allora T assume valori in C con probabilità α

Vi sono due tipi di errore dovuti alle osservazioni sul campione in uso:

- **Errori di prima specie:** Se si rifiuta H_0 quando, in realtà, l'ipotesi è vera. La probabilità di compiere tali errori coincide col livello di significatività α .
- **Errori di seconda specie:** Se non si rifiuta H_0 quando, in realtà, l'ipotesi è falsa. La probabilità di compiere tali errori si denota con β .