Esame di INFORMATICA E STATISTICA per le Biotecnologie

N. matricola : 06034A
COGNOME - NOME: Benetti Giulia

<1> Utilizzando la funzione integrate calcolare l'interale di probabilita' sotto la curva della distribuzione normale tra i valori -1 e 0, dal risulato estrarre il valore calcolato e salvarlo in una variabile x, utilizzando un'unica istruzione R. Suggerimento: leggere il manuale della funzione integrate().

<2> Data una variabile casuale discreta X che puo' assumere valori pari a 0 (probabilita' = 0.28), 1 (probabilita' = 0.11), 2 (probabilita' = 0.31), 3 (probabilita' = 0.24), 4 (probabilita' = 0.06), indicare a quale delle seguenti combinazioni di valori corrisponda il valore atteso e la varianza della variabile (i valori arrotondati alla seconda cifra decimale sono riportati nell'ordine: valore atteso, varianza). "A") 1.69, 0.87; "B") 2.03, 1.61; "C") 2.03, 1.53; "D") 1.69, 1.61.

<3> Ipotizzando che il valore medio di una popolazione sia pari a mu=8 (ipotesi nulla) effettuare un test t, calcolare il p value (a 2 code) a partire dal set di dati (campione singolo) contenuto in OGGETTO_007_b, e salvare il p value in una variabile x, utilizzando un'unica istruzione R.

<4> L'altezza delle piante di una determinata varieta' e' caratterizzata da un certo grado di variabilita'. Si suppone tuttavia che l'altezza media delle piante di tale varieta' sia di 46.86 cm. Al fine di verificare tale ipotesi sono stati raccolti dati relativi all'altezza di un campione di 9 piante: l'altezza media delle piante appartenenti al campione e' risultata pari a 47.1 cm con deviazione standard di 1.26 cm. Applicando il test t per un campione e facendo riferimento alla tavola statistica della distribuzione t (tavola_statistica_distribuzione_t.jpg), l'evidenza derivante dai dati e' sufficientemente forte da poter rifiutare l'ipotesi nulla (H0: "l'altezza media e' di 46.86 cm"; HA: "l'altezza media non e' di 46.86 cm") assumendo un livello di significativita' alfa = 0.05? "A") no; "B") si'.

<5> OGGETTO_013_c contiene la quantita' di energia assunta da un campione casuale di 11 donne adulte. Calcolare l'intervallo di confidenza al 95% della quantita' di energia assunta e salvarlo in un vettore x contenente, in quest'ordine, l'estremo inferiore e l'estremo superiore. Il tutto utlizzando un'unica istruzione R.

<6> Il test esatto di Fisher e' stato applicato al fine di verificare se le variabili X ed Y siano indipendenti (H0: "le variabili sono indipendenti"; HA: "le variabili non sono indipendenti"). Basandosi sul p-value ottenuto (p-value = 0.0001), se assumessi un livello di significativita' alfa = 0.001 incorrerei in errore nel prendere la decisione riguardo H0 sapendo che le due variabili non sono indipendenti (realta': H0 falsa)? "A") Si'; "B") No.

<7> OGGETTO_014_b contiene i conteggi degli individui con una determinata allergia in due gruppi. Effettuare un test di Fisher per comparare la proporzione di individui allergigi nei due gruppi usando un livello di confidenza do 90% e salvare in una variabile x l'estremo superiore dell'intervallo di confidenza calcolato. Suggerimento: indagate la struttura dell'oggetto restituito dalla funzione che realizza il test statistico.

<8> Quale tra i valori di odds ratio stimati su dati raccolti nel contesto di quattro studi sperimentali indipendenti (studio 1: OR = 0.99, studio 2: OR = 3.19, studio 3: OR = 1.5, studio 4: OR = 9.18) indicherebbe evidenza piu' forte in merito all'efficacia di una tecnica chirurgica innovativa (valori variabile tecnica chirurgica: innovativa, standard) sulla guarigione da una determinata patologia (valori variabile guarigione: guarito, non guarito), considerando come successo l'evento "guarito" (gruppo di trattamento mediante tecnica chirurgica innovativa rispetto al gruppo di trattamento mediante tecnica chirurgica standard)? "A") OR = 0.99 (studio 1); "B") OR = 3.19 (studio 2); "C") OR = 1.5 (studio 3); "D") OR = 9.18 (studio 4).