

N. matricola : 01766A

COGNOME - NOME: Iannetta Simone

<1> Utilizzando la formula disponibile al link: <https://homes.di.unimi.it/re/Corsi/TMP/tmp1/formulaPDensity.jpg> e i parametri  $\mu=0$  e  $\sigma=4$  calcolare le densità di probabilità dei valori contenuti in OGGETTO\_002 e salvarle in una variabile x, utilizzando un'unica istruzione R.

<2> In un campione di 12 persone esposte ad un determinato fattore ambientale è stato osservato che 7 di queste si sono ammalate. Data  $H_0$  "La proporzione di persone esposte che si ammalano è uguale alla proporzione di persone esposte che non si ammalano",  $H_A$  "La proporzione di persone esposte che si ammalano è maggiore rispetto alla proporzione di persone esposte che non si ammalano" e la distribuzione nulla della statistica test la cui distribuzione di probabilità è riportata di seguito ( $\Pr[0 \text{ malati}] = 0.000244$ ,  $\Pr[1 \text{ malato}] = 0.00293$ ,  $\Pr[2 \text{ malati}] = 0.016113$ ,  $\Pr[3 \text{ malati}] = 0.053711$ ,  $\Pr[4 \text{ malati}] = 0.12085$ ,  $\Pr[5 \text{ malati}] = 0.193359$ ,  $\Pr[6 \text{ malati}] = 0.225586$ ,  $\Pr[7 \text{ malati}] = 0.193359$ ,  $\Pr[8 \text{ malati}] = 0.12085$ ,  $\Pr[9 \text{ malati}] = 0.053711$ ,  $\Pr[10 \text{ malati}] = 0.016113$ ,  $\Pr[11 \text{ malati}] = 0.00293$ ,  $\Pr[12 \text{ malati}] = 0.000244$ ), calcolare il p-value ed indicare a quale dei seguenti valori corrisponde il p-value corretto e se sia possibile rifiutare l'ipotesi nulla dato un livello di significatività  $\alpha = 0.05$ . "A") 0.387207, rifiuto  $H_0$ ; "B") 0.387207, non rifiuto  $H_0$ ; "C") 0.003174, non rifiuto  $H_0$ ; "D") 0.003174, rifiuto  $H_0$ .

<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 varietà è caratterizzata da un certo grado di variabilità. Si suppone tuttavia che l'altezza media delle piante di tale varietà sia di 37.11 cm. Al fine di verificare tale ipotesi sono stati raccolti dati relativi all'altezza di un campione di 8 piante: l'altezza media delle piante appartenenti al campione è risultata pari a 37.78 cm con deviazione standard di 0.94 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 è sufficientemente forte da poter rifiutare l'ipotesi nulla ( $H_0$ : "l'altezza media è di 37.11 cm";  $H_A$ : "l'altezza media non è di 37.11 cm") assumendo un livello di significatività  $\alpha = 0.05$ ? "A") no; "B") sì.

<5> OGGETTO\_013\_b contiene le misurazioni di una variabile riferita ad un test ematologico in due gruppi di soggetti portatori e non portatori di una mutazione genetica X. Testare l'ipotesi che la media nelle due popolazioni sia uguale e salvare il valore assoluto della differenza della stima dei valori medi nelle due popolazioni in una variabile x utilizzando un'unica istruzione R. Suggerimento: indagate la struttura dell'oggetto restituito dal test statistico prima di costruire la soluzione da sottomettere.

<6> Il test t per un campione è stato applicato al fine di verificare se il valore medio di emoglobina in portatori di una mutazione genetica sia di 16 g/dl ( $H_0$ : "il valor medio di emoglobina nei portatori della mutazione è di 16 g/dl";  $H_A$ : "il valor medio di emoglobina nei portatori della mutazione non è di 16 g/dl"). Basandosi sul p-value ottenuto (p-value = 0.876), se assumessi un livello di significatività  $\alpha = 0.05$  incorrerei in errore nel prendere la decisione riguardo  $H_0$  sapendo che il valor medio di emoglobina nei portatori della mutazione è di 16 g/dl (realtà:  $H_0$  vera)? "A") Sì; "B") No.

<7> OGGETTO\_014\_c contiene dati relativi a 18 valori distribuiti su tre gruppi. Applicare un test ANOVA ad una via. Costruite una lista x contenente una data frame contenente le colonne 2 e 3 della tabella dei risultati restituiti dal test. Attribuite a questo elemento della lista il nome SommeEMedieSq. Il tutto utilizzando un'unica istruzione R.

<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 più 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).