

7. The NCHS reported that the mean total cholesterol level in 2002 for all adults was 203. Total cholesterol levels in participants who attended the seventh examination of the Offspring in the Framingham Heart Study are summarized as follows : $n = 3,310$, $\bar{x} = 200.3$, and $s = 36.8$.

Is there statistical evidence of a difference in mean cholesterol levels in the Framingham Offspring? Frame suitable hypothesis symbolically and numerically. Apply a suitable test to validate the claim. List reasons for the type of test used to validate the hypothesis.

($t_{\text{crit}} = 1.94$, $Z_{\text{crit}} = 1.96$ at $\alpha = 0.05$) (15)

Roll No.

Total Pages : 4

758105

January 2023

M.Sc. (Life Science) Ist SEMESTER

Biostatistics (MLS-105)

Time : 3 Hours]

[Max. Marks : 75

Instructions :

1. *It is compulsory to answer all the questions (1.5 marks each) of Part-A in short.*
2. *Answer any four questions from Part-B in detail.*
3. *Different sub-parts of a question are to be attempted adjacent to each other.*

PART-A

1. (a) Define Type-II error. (1.5)
(b) Define null hypothesis. (1.5)
(c) Value of probability can never go beyond (1.5)
(d) distribution is a limiting case of binomial distribution. (1.5)
(e) Name *three* statistical software. (1.5)
(f) Variance of binomial distribution is (1.5)
(g) The test statistic value should be (less/more) than the critical value to accept the null hypothesis. (1.5)

- (h) Differentiate between correlation and regression. (1.5)
- (i) and are the measures of dispersion of data. (1.5)
- (j) The test score is 190. The test has a mean of 130 and a standard deviation of 30. Calculate Z-score assuming it to be normally distributed. (1.5)

PART-B

2. (a) The sensitivity of the COVID-19 testing rapid antigen self-test kit for detecting COVID-19 positive patients is 0.80. Its specificity is 0.96. The prevalence of prostatic cancer in the white male population is 25 per 10000. What is the probability that a person who tests positive has COVID-19 infection? (10)
- (b) What do you mean by variable in statistics? Discuss its different types. (5)
3. (a) In how many ways can 12 different amino acids be arranged into a polypeptide chain of five amino acids? Under what special conditions median is considered a more significant parameter of location as compared to the mean. Discuss with a suitable example. (10)
- (b) In a kit manufacturing company, the proportion of obtaining a defective kit is 0.01. A random sample of 100 items is selected. What is the probability that there are two defective kits in the selected sample? ($e=2.718$) (5)

4.

Plot of Land	Per acre production data			
	Variety of Rice			
	X1	X2	X3	X4
A	6	5	5	7
B	7	5	4	5
C	3	3	3	2
D	8	7	4	3

(15)

Table 1: The table represents the per acre production of different varieties of rice (X1, X2, X3, and X4) when grown on different plots of land (A, B, C, and D).

Formulate H_0 and H_a both in words and symbolically. Apply two-way ANOVA considering plot of land and variety of rice as two variables to investigate and validate the claim. ($F_{crit} = 3.86$ at $\alpha = 0.05$)

5. (a) Calculate mean and mode for both X2 and X4 variety of rice in Table 1. (5)
- (b) Is there any significant difference between variety X1 and X3 of rice (Refer Table 1)? Apply student t-test to investigate. ($t_{crit} = 1.94$ at $\alpha = 0.05$) (10)
6. (a) 5, 10, 1, 3, 6, 9, 2, 6, 7, 4, 13, 14, 17, 19, 25, 21, 16. (10)

Calculate the mean, standard deviation, median, first quartile and third quartile of the given data set.

- (b) A dice is thrown 6 times. Consider "getting an odd number" as a success. What is the probability of getting at least five successes. (5)