



Register Number:

DATE:

ST. JOSEPH'S COLLEGE (AUTONOMOUS), BANGALORE – 27
BCA(DATA ANALYTICS) – IV SEMESTER
SEMESTER EXAMINATION – APRIL 2022
(EXAM CONDUCTED IN JULY 2022)
BCADA 4121: STATISTICAL FORECASTING AND INFERENCE

Time: 2.5 hrs.

Maximum marks: 70

This question paper contains **FOUR** printed pages and **THREE** parts

PART A

Answer ALL questions from the following

1 X 20 = 20

1. An estimator is considered to be the best if its distribution is
 - (a) Continuous
 - (b) discrete
 - (c) concentrated about true population parameter value
 - (d) normal
2. The estimator $\sum xi/n$ of population mean is
 - (a) Unbiased estimator
 - (b) Consistent estimator
 - (c) Both (a) and (b)
 - (d) Neither (a) nor (b)
3. What is the expectation of the sample mean?
 - (a) $\frac{\mu}{n-1}$
 - (b) $\frac{n\mu}{n-1}$
 - (c) μ
 - (d) $n\mu$
4. As the width of the confidence interval increases, the confidence level associated with the interval
 - (a) Tend to increase
 - (b) Tend to decrease
 - (c) Remains the same
 - (d) None of these
5. Critical region is a region of
 - (a) Rejection
 - (b) Acceptance
 - (c) Indecision
 - (d) None of these

6. The term $1-\beta$ is called
 - (a) Level of the test
 - (b) Power of the test
 - (c) Size of the test
 - (d) Level of significance
7. If a sufficient estimator exists it is a function of
 - (a) M.L.E
 - (b) Unbiased estimator
 - (c) Most efficient
 - (d) Unique
8. Interval estimation for a single population mean when σ is known is given by
 - (a) $\bar{x} \pm Z_{\alpha} \sigma(\bar{x})$
 - (b) $\bar{x} \pm z_{\frac{\alpha}{2}} \sigma(\bar{x})$
 - (c) $\bar{x} \pm Z_{\alpha} S(\bar{x})$
 - (d) $\bar{x} \pm z_{\frac{\alpha}{2}} S(\bar{x})$
9. When using the sample proportion \bar{p} to test the hypothesis, the S.E of \bar{p} is
 - (a) $\frac{\bar{p} \bar{q}}{n}$
 - (b) $\frac{pq}{n}$
 - (c) $\sqrt{\frac{\bar{p}\bar{q}}{n}}$
 - (d) $\sqrt{\frac{PQ}{n}}$
10. The test statistic to test $\mu_1=\mu_2$ for normal population is
 - (a) F-test
 - (b) Z-test
 - (c) t-test
 - (d) U-test
11. ABC is an Ed-Tech Company that needs to forecast the sales volume of its Data Science Courses. ABC hires you as a consultant and provides all the needed data. Which can be one of the ways to collect data?
 - (a) The number of people working in the sales team
 - (b) The number of working hours of the sales team
 - (c) The number of hours of DS courses to be consumed by learners
 - (d) Using the past data of the sales team
12. Amazon claimed that its total valuation in December 2020 was \$20 billion. What would be the alternate hypothesis for the given situation?
 - (a) H_1 : Total Valuation = \$20 billion
 - (b) H_1 : Total Valuation < \$20 billion
 - (c) H_1 : Total Valuation > \$20 billion
 - (d) H_1 : Total Valuation \neq \$20 billion
13. Forecasting time horizon includes
 - (a) Long range
 - (b) Medium range
 - (c) Short range
 - (d) All of these
14. A component of time series used for short term forecast is
 - (a) Trend
 - (b) Seasonal
 - (c) Cyclical
 - (d) Irregular

15. In which situations can Time Series Analysis be used?
- You are an agricultural Officer and want to know different crop yields in tons of the previous year district wise
 - You are a health Officer and want to know the infant mortality rate in your state since the last two years
 - You are the CM and want to know the unemployment rate for your state in each financial quarter
 - All the above
16. In case of unbiased estimators, the estimator having minimum variance is called an.
- Efficient estimator.
 - Sufficient
 - Both A and B
 - Consistent
17. Estimation is the branch of.
- Statistics
 - Statistical Method
 - Both A and B
 - Statistical Inference
18. The asymptotic distribution of t -statistic with n degrees of freedom is
- F
 - Normal
 - Z
 - T
19. When the value of population standard deviation σ is unknown , the values of t in the t-distribution are :
- More variable than for z
 - Less variable than for z
 - Equal to z
 - None of these
20. Method of least squares to fit in the trend is applicable only if the trend is
- Linear
 - Parabolic
 - Both (a) and (b)
 - Neither (a) nor (b)

PART B

Answer any SIX questions:

6 x 5 = 30

- Explain what do you understand by time series? Why is time series forecasting considered to be an effective tool for forecasting?
- Why is estimation important? What are the advantages of using interval estimates than point estimates?
- Let X_1, X_2, X_3 be a random sample of size $n=3$ taken from a normal population with mean μ and variance σ^2 . Also let

$$\widehat{\mu}_1 = (x_1 + 2x_2 + 3x_3)/6$$

$$\widehat{\mu}_2 = (x_1 + x_2 + x_3)/3$$

Be two estimates of μ .

Are $\widehat{\mu}_1$ and $\widehat{\mu}_2$ unbiased ? Explain

- Write a short note on Type I and Type II error .

25. The average monthly electricity consumption for a sample of 100 families is 1250 units. Assuming the SD of electric consumption of all families is 150 units, construct a 95 percent confidence interval estimate of the actual mean electric consumption.
26. A coin is tossed 400 times and it is seen that the head occurs 216 times. Test whether the coin is fair.
27. Explain the differences between qualitative and quantitative forecasting.
28. What is the utility of estimating the sample size? Write down the formula of sample size for estimating Population Proportion. A car manufacturing company received a shipment of petrol filters. These filters are to be sampled to estimate the proportion that it is unusable. From past experience, the proportion of unusable filter is estimated to be 10 per cent. How large a random sample should be taken to estimate the true proportion of unusable filters within 0.07 with 99 percent confidence.

SECTION C

Answer any TWO questions:

2 x 10 = 20

29. Define a sufficient statistic. If X_1, X_2, \dots, X_n are independent and follow Bernoulli distribution with parameter p , examine whether $T = \sum_{i=1}^n x_i$ is the sufficient estimator for p .

30. Define the term 'level of significance'. How is it related to probability of committing a Type I error? Explain the general steps needed to carry out a test of hypothesis.

A fertilizer mixing machine is set to give 12 kg of nitrate for every 100 kg of fertilizer. Ten bags of 100 kg each are examined. The percentage of nitrate so obtained is : 11 , 14 , 13 , 12, 13, 12, 13, 14, 11 and 12. Is there reason to believe that the machine is defective? [$t_{0.025,9}=2.262$]

31. Explain the advantages of exponential smoothing techniques over simple moving averages.

A firm uses simple exponential smoothing with $\alpha=0.1$ to forecast demand. The forecast for the week of February 1 was 500 units whereas actual demand turned out to be 450 units .

(a) Forecast the demand for the week of February 8

(b) Assume the actual demand during the week of February 8 turned out to be 505 units. Forecast the demand for the week of February 15. Continue forecasting through March 15, assuming that subsequent demands were actually 516, 488, 467, 554 and 510 units.