Register No:
Date:

ST. JOSEPH'S COLLEGE (AUTONOMOUS), BANGALORE-27
M.SC. BIG DATA ANALYTICS - II SEMESTER
SEMESTER EXAMINATION: APRIL 2018
BDA 2216: ADVANCED STATISTICAL METHODS

TIME: 2 ½ HRS **MAX MARKS 70**

This Question Paper Contains Two Printed Pages

Statistical tables will be provided on request

Answer any seven of the following **7 X 10 = 70**

1.
   a. Distinguish between Estimator and Estimate? What are the qualities of a good estimator? Explain
   b. If X_1, X_2, X_3 is a random sample from B (1,p) distribution. Let $T_1 = \frac{X_1+X_2}{2}$ and $T_2 = \frac{X_1+2X_2+X_3}{4}$ Show that T_1 and T_2 are unbiased estimators of p (6 + 4)
2.
   a. Suppose X_1, X_2… X_n are random sample from U (-θ, θ). Obtain MLE of θ.
   b. Let X_1, X_2… X_n be a random variable from gamma distribution with parameter (α, θ). Obtain an estimator for the parameter (α, θ) based on method of moments (4 + 6)
3.
   a. Define Parameter and Statistic with example
   b. Define likelihood function and MLE.
   c. Suppose X_1, X_2…X_{20} be a random sample from a distribution with density function $f(x, \beta) = \beta/x^{\beta+1}$ $x > 1$. Find the MLE of β (3 + 2 + 5)
4.
   a. The weight of grapes produced per vine has been measured on 10 vines selected at random in a vineyard. The results in kilograms are the following:
      2.4, 3.4, 3.6, 4.1, 4.3, 4.7, 5.4, 5.9, 6.5, 6.9
      The weight of grapes produced by each vine is modeled by N (μ, σ^2) distribution.
      (i) Give a 99% confidence interval for μ
      (ii) Give a 95% confidence interval for σ^2
   b. Define critical region, Acceptance region, simple hypothesis and composite hypothesis (6 + 4)
5.
   a. Define: Type I error, Type II error, Size and Power of a test. Also deduce relation between type II error and power
   b. For a random sample of size 55 from normal distribution with mean θ and variance 6 for testing
      H_0: θ ≥ 9 against H_1: θ < 9 test is given by
      Φ (Σx) = 1 if Σx < k
      0 if Σx ≥ k.
      Determine k such that size of the test is 5%. (5 + 5)

6.

a. Explain main steps involved in testing of hypothesis

b. Let X be a discrete random variable with P(X=1) =1-P(X=-1) = q. Given a random sample of size 2 from this distribution for testing H_0: p = 0.56 vs H_1: p = 0.86. Test is given by

$\Phi(x_1, x_2) = 1$ if $x_1 + x_2 = 2$

$\frac{1}{5}$ if $x_1 + x_2 = 0$

0 if $x_1 + x_2 = -2$

Compute size and power of test (5 + 5)

7.

a. Suppose we wanted to test whether or not the average SAT score at Brown University was significantly different from the average SAT score at Wassamatta University. Taking a random sample, we manage to get the following 7 scores from Brown students:
Brown University: 1340, 1500, 1430, 1440, 1380, 1470, 1290
and the following 9 scores from Wassamatta students:
Wassamatta University: 1540, 1480, 1390, 1450, 1440, 1350, 1520, 1400, 1600
Assuming that the variance in SAT scores for students at Brown and students at Wassamatta are the same, and assuming that both samples are taken from normal populations, test (at the α =0.01 significance level) whether or not there is a significant difference in the average SAT score between these two schools.

b. The following measurements of the diameters (in feet) of Indian mounds in southern Wisconsin were gathered by examining reports in Wisconsin Archeologist:
22, 24, 24, 30, 22, 20, 28, 30, 24, 34, 36, 15, 37
Does the data substantiate the conjecture that the population variation in diameter is larger than 7 feet?
Test at α =0.05 significance level. You may assume the sample was taken from a normal distribution (6 + 4)

8.

a. Explain Gauss Markov Model

b. What is ANOVA? Explain two way classification (5 + 5)

9.

a. Students were given different drug treatment before revising for their exams. Some were given drug A, some were given drug B, and some were given drug C. The exam scores (%) are shown below for three different groups:

Drug A	70, 77, 83, 90, 97
Drug B	37, 43, 50, 57, 63
Drug C	3, 10, 17, 23, 30

Test whether the different drugs will have different effects?

b. Explain multiple linear regression (6 + 4)

10.

a. Explain the difference between R square and adjusted R square

b. Write a note on Logistic Regression (6 + 4)