Question Booklet Series: **A** *Question Booklet Serial No.* 311030

M. Phil./Ph. D. Entrance Test – 2015

Subject: Statistics

Paper – I

Important: *Please consult your Admit Card/Roll No. slip before filling your Roll Number on the Test Booklet and Answer Sheet.*

Roll No. ***In Figure*** ***In Words***

O.M.R. Answer Sheet Serial No.

Signature of Candidate: ______________ *Signature of Invigilator:* ______________

Time: 60 Minutes Number of Questions: 50 Maximum Marks: 50

DO NOT OPEN THE SEAL ON THE BOOKLET UNTIL ASKED TO DO SO.

INSTRUCTIONS:

1. Write your Roll No. on the Questions Booklet and also on the OMR Answer Sheet in the space provided and nowhere else.
2. Enter the Question Booklet Serial No. on the OMR Answer Sheet. Darken the corresponding bubbles with **Black Ball Point/Black Gel Pen.**
3. Do not make any identification mark on the Answer Sheet or Question Booklet.
4. Please check that this Question Booklet contains 60 Questions. In case of any discrepancy, inform the Assistant Superintendent within 10 minutes of the start of Test.
5. Each question has four alternative answer (A,B,C,D) of which only one is correct. For each question, darken only one bubble (A or B or C or D), whichever you think is the correct answer, on the Answer Sheet with **Black Ball Point/Black Gel Pen.** **There shall be no negative marking for wrong answers.**
6. If you do not want to answer a question, leave all the bubbles corresponding to that question blank in the Answer Booklet. No marks will be deducted in such cases.
7. Darken the bubbles in the OMR Answer Sheet according to the Serial No. of the question given in the Question Booklet.
8. If you want to change an already marked answer, erase the shade in the darkened bubble completely.
9. For rough work only the blank sheet at the end of the Question Booklet be used.
10. The Answer Sheet is designed for computer evaluation. Therefore, if you do not follow the instructions given on the Answer Sheet, it may make evaluation by the computer difficult. **Any resultant loss to the candidate on the above account, i.e. not following the instructions completely, shall be of the candidate only.**
11. After the test, hand over the Question Booklet and the Answer Sheet to the Assistant Superintendent on duty.
12. In no case the Answer Sheet, the Question Booklet, or its part or any material copied/noted from this Booklet is to be taken out of the examination hall. Any candidate found doing so would be expelled from the examination.
13. A candidate who creates disturbance of any kind or changes his/her seat or is found in possession of any paper possibly of any assistant or found giving or receiving assistant or found using any other unfair means during the examination will be expelled from the examination by the Centre Superintendent/Observer whose decision shall be final.
14. **Communication equipment such as mobile phones, pager, wireless set, scanner, camera or any electronic/digital gadget etc., is not permitted inside the examination hall. Use of calculators is not allowed.**
15. The candidates will not be allowed to leave the Examination Hall/Room before the expiry of the allotted time.

Let there be p levels of factor A and q levels of factor B in a two way ANOVA with m observations per cell. Answer questions **1 and 2** using this information.

1. The degrees of freedom associated with error are:

   A) (p-1)(q-1) B) mpq-1 C) (m-1)(p-1)(q-1) D) pq(m-1)

2. The degrees of freedom associated with interaction AB are:

   A) mpq-1 B) (p-1)(q-1) C) pq(m-1) D) (m-1)(p-1)(q-1)

3. A 2^n factorial experiment is carried out in r replications. The error degrees of freedom are:

   A) (r-1)(n-1) B) $(r-1)2^n$ C) $(r-1)(2^n-1)$ D) $r(2^n-1)$

4. Let S_E^2 be the mean square error associated with a Latin Square Design (LSD) of order m. The standard error associated with an elementary treatment contrast is:

   A) $(m-1)S_E$ B) $S_E/(m)^{1/2}$ C) $S_E(2/m)$ D) $[(2/m)S_E^2]^{1/2}$

Let $X_{(1)} \leq \ldots \leq X_{(n)}$ be the order statistics corresponding to a random sample of size n from a standard exponential distribution. Define $Y_i = (n-i+1)(X_{(i)} - X_{(i-1)})$, i= 1, …, n, $Y_i \geq 0$ with $X_{(0)} = 0$. The random variables $Y_1, \ldots, Y_n$ are independent with common standard exponential distribution. Use this information to answer questions 5, 6, 7, 8, 9.

5. $X_{(r)}$ can be written as

   A) $X_{(r)} = \sum_{i=1}^{r}(X_{(i)} - X_{(i-1)})$ B) $X_{(r)} = \sum_{i=1}^{r} Y_i/(n-i-1)$

   C) $X_{(r)} = Y_{(r)}/(n-r-1)$ D) Both A and B

6. The value of $E(X_{(r)})$ is:

   A. $1/(n-r-1)$ B) $\sum_{i=1}^{r} 1/(n-i+1)$ C) $\sum_{i=1}^{r}(n-i+1)$ D) (n-r-1)

7. The value of $Var(X_{(r)})$ is:

   A. $\sum_{i=1}^{r} 1/(n-i-1)$ B) $(n-i-1)^2$ C) $\sum_{i=1}^{r} 1/(n-i-1)^2$ D) 1

8. The value of $E(X_{(n)} - X_{(1)})$ is:

   A. $\sum_{i=2}^{r} 1/(n-i-1)$ B) $\sum_{i=1}^{r} 1/(n-i-1)$ C) 0 D) 1/(n-r-1) - 1/(n-2)

9. Let $X_{(1)} \leq \ldots \leq X_{(n)}$ be the order statistics corresponding to a random sample of size n from a continuous distribution with cdf F(x) and pdf f(x). The pdf of $X_{(n)}$ is :

   A) $n[F(x)]^{n-1} f(x)$ B) $n[F(x)]^{n-1}$ C) $n[F(x)] f(x)^{n-1}$ D) $F(x)^n$

10. Let AX = 0 be the system of m homogeneous linear equations, where A is mxn matrix, X is nx1 vector and 0 is mx1 zero vector. Let n> m. Then a nontrivial solution:

    A) Does not exist B) Always exists

    C) May or may not exist D) Exists if n<m.

11. The matrix

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \text{ is:}$$

A) Positive semi-definite
B) Negative semi-definite
C) Positive Definite
D) Negative Definite.

12. For the matrix A with A^t its transpose, the transformation Y=AX is orthogonal if:

A) $A=A^t$
B) $AA^t = O$ (zero matrix)
C) $AA^t = I$ (identity matrix)
D) $AXA^t = Y$

13. The inverse of a positive definite matrix is

A) Positive definite
B) Negative definite
C) Positive semidefinite
D) Negative semidefinite

14. Let $Y_1, \ldots, Y_n$ be independent random variables such that $E(Y_i) = \mu_i$, $Var(Y_i) = \sigma^2$, i = 1, …, n. Let tr(A) be the trace of matrix A, $Y^t = (Y_1, \ldots, Y_n)$ and $\mu^t = (\mu_1, \ldots, \mu_n)$. Then $E(Y^tAY)$ is:

A) $\sigma^2 A$
B) $\sigma^2 tr(A)$
C) $\mu^t A\mu$
D) $\sigma^2 tr(A) + \mu^t A\mu$

Let $Y_1, \ldots, Y_n$ be independent normal random variables such that $E(Y_i) = \mu_i$ and $Var(Y_i) = \sigma^2$, i= 1, …, n. Define the Quadratic form $Y^t M Y$, where $Y^t = (Y_1, \ldots, Y_n)$, $\mu^t = (\mu_1, \ldots, \mu_n)$ and $M = M^2$. With this information answer questions **15, 16 and 17.**

15. The distribution of this quadratic form is:

A) Central Chi-square
B) Non Central Chi square
C) Central F
D) Non central t

16. The degrees of freedom of the distribution of this quadratic form are:

A) n-1
B) n-2
C) tr(M)
D) tr(M) – (n-2)

17. The non-centrality parameter of the distribution of $(Y^tMY)/\sigma^2$ is:

A) $\mu^t M\mu$
B) $(\mu^t M\mu)/\sigma^2$
C) 0
D) $tr(M) + (\mu^t M\mu)/\sigma^2$

18. Let $T_n(X_1, \ldots, X_n)$ be an unbiased estimator of parameter θ (scalar) based on a random sample $X_1, \ldots, X_n$ of size n. Let I(θ) be the fisher information of the sample and $f(x_1, \ldots x_n|\theta)$ be the joint density of $X_1, \ldots, X_n$. If variance of $T_n(X_1, \ldots, X_n)$ attains lower bound then partial derivative of joint density with respect to θ is equal to:

A) I (θ)
B) $(T_n(X_1, \ldots, X_n) - \theta)$
C) $I(\theta)(T_n(X_1, \ldots, X_n) - \theta)$
D) $[(T_n(X_1, \ldots, X_n) - \theta)]/I(\theta)$

Let $X_1, \ldots X_5$ be a random sample from a continuous distribution with median θ. Define $Y_i = 1$ if $X_i > \theta$ and zero otherwise. Let $Y = Y_1 + \ldots + Y_5$. Answer questions **19, 20 and 21** based on this information:

19. P(Y = 1) is equal to

A) ½ B) 5/32 C) 1/32 D) ¾

20. P(Y≥2) is equal to

A) 13/16 B) 5/16 C) 3/16 D) 1/16

21. The mean and variance of Y respectively are:

A) 2.5 and ¼ B) 2.5 and 1/2 C) ¼ and 2.5 D) 10/4, 5/4

22. Let A be the limit sup of sequence of events $\{A_n\}$. If $\sum_{n=1}^{\infty} P(A_n) < 1$, then the value of P(A) is :

A) 1 B) 0 C) <1 D) >1/2

23. For the sequence of random variables $\{X_n\}$, define $nS_n = \sum_{i=1}^{n} X_i$. Define the sequence $\{Z_n\}$, where $Z_n = S_n^2/(1+S_n^2)$. The sequence $\{X_n\}$ satisfies the weak law of large numbers if and only if:

A) $Z_n \to 0$ as $n \to \infty$

B) $Z_n \to 1$ as $n \to \infty$

C) $E(Z_n) \to 0$ as $n \to \infty$

D) $E(Z_n) \to 1$ as $n \to \infty$

24. For any positive ε, the sequence of estimator $T_n = T_n(X_1, \ldots, X_n)$ is consistent for the parameter θ if :

A) $E(T_n) = \theta$

B) $P(T_n = \theta) = 1$

C) $P[|T_n - \theta| > \varepsilon] \to 1$ as $n \to \infty$

D) $P[|T_n - \theta| > \varepsilon] \to 0$ as $n \to \infty$

25. Let F(x) and f(x) be the cdf and pdf, respectively, of a continuous random variable X with support R. If this support is truncated to the left at d then the pdf associated with this truncated support is:

A) f(x)/ F(d) B) f(x) / {1-F(d)} C) f(x) F(d) D) f(x) {1-F(d)}

26. Let F(x) be the cdf of a continuous random variable. Define a new random variable as Y =F(X), where the support of Y is $R_Y = [0, 1]$. The cdf of Y , say G(y) with $y \in R_Y$, is equal to:

A) $F^{-1}(y)$ B) F(y) C) $FG^{-1}(x)$ D) y

27. Let $U_1, \ldots, U_n$ be a random sample from uniform distribution over the interval (0, 1). Define the random variable $Y_i = U_i / \sum_{i=1}^{n} U_i$, i = 1, ..., n. The random variables $Y_1, \ldots, Y_n$ are:

A) Dependent

B) Independent

C) Undefined

D) Linearly dependent

28. The randomized test is used when the distribution of test statistics is:

A) Continuous B) Degenerate C) Rectangular D) Discrete.

29. Let the distribution of random vector X be multivariate normal with mean vector μ and non-singular dispersion matrix $\sum$. The distribution of the linear combination l'X is:

A) Chi square
B) Univariate Normal
C) Degenerate
D) Multivariate normal

30. Let -3, 9, 8, -10, 6, -8, -5, -3, 4 be the realization of a random sample from uniform distribution over the interval [-θ, θ]. The maximum likelihood estimate, based on the observations, is:

A) 10 B) -10 C) 9.5 D) 9

31. Let $X_1, \ldots, X_n$ be a random sample from a continuous distribution with cdf F(x) . For fixed x_0, let p =F(x_0). The degree of the kernel associated with the U-statistic which is an estimator of p is:

A) 1 B) 2 C) n D) n+1

32. U-statistic as an estimator is :

A) Unbiased
B) Consistent
C) Biased
D) Consistent and unbiased

33. The Wilk's λ criteria in a multi-sample problem, is based on:

A) Likelihood function
B) Sufficient Statistics
C) Efficient estimators
D) SPRT

34. Two stage sampling is a compromise between:

A) Stratified and Systematic
B) Cluster and pps
C) Stratified and cluster
D) Stratified and simple random

35. Ratio estimator is preferred when the line of regression of variable of interest (Y) on the auxiliary variable (X) passes through:

A) Origin
B) (Mean of X, Mean of Y)
B) A point on positive X-axis
D) A point on positive Y-axis

36. The distribution of the test statistic used to test the significance of a multiple linear regression model is:

A) Student's t
B) Chi square
C) F
D) Standard Normal

37. A sample of 4 items is selected from a lot containing 4 bad and 6 good items using simple random sampling with replacement. The probability that the sample contains 3 good and 1 bad items is:

A) .0864 B) .3456 C) 8/21 D) .5184

38. A bottle contains 3 under weight, 4 normal weight and 3 over weight tablets. A person selects a sample of 3 tablets from the bottle using simple random sampling without replacement. The probability of having each type of tablet in the sample is:

A) 24/329 B) 144/329 C) 2/7 D) 3/10

39. Probability is .4 that an industrial worker is a smoker. The probability is .3 that an industrial worker is smoker and suffers from respiratory problem. The probability that an industrial worker suffers from respiratory problem if he smokes is:

A) .12 B) ¾ C) 1/12 D) ¼

40. The value of k for which the function $f(x,y) = kx(x-y)$, $0<x<2$, $-x<y<x$ and zero otherwise is a joint density function is :

A) ½ B) 8 C) ¾ D) 1/8

41. The joint pdf $f(x,y)$ of a two dimensional random variable (X,Y) is uniform over the region R { ((x,y): $y=x$, $y=x^2$, $0<x<1$, $0<y<1$}. Then for (x,y) in R, the value of f(x,y) is:

A) 2 B) 6 C) 1/6 D) ½

42. The joint pdf of (X,Y) is $f(x,y) = 2$ $x>0$, $y>0$, $x+y<1$ and zero otherwise. $P[X+Y>2/3]$ is:

A) 1/2 B) ¾ C) 4/9 D) 5/9

43. The value of constant c for which the function $f(x,y) = cxy$, $0<x<1$, $0<y<1$, $x+y<1$ and zero otherwise is:

A) 12 B) 24 C) 1/12 D) 4

44. The number of orthogonal Latin Square Designs (LSD) associated with an LSD of order v is:

A) V^2 B) v(v-1) C) v-1 D) v(v-1)/2

45. Let $f(x) = x$ and $g(x) = 1-e^{-x}$, $x>0$. Then $\int_0^\infty f(x)dg(x)$ is equal to:

A) 1 B) 0 C) ½ D) 2

46. For i = 1, 2, 3 let the quadratic forms $X'A_iX$ follow chi square distribution such that $X'A_1X = X'A_2X + X'A_3X$. Then

A) $tr(A_1) = tr(A_2A_3)$ B) $tr(A_1) = Min(tr(A_2), tr(A_3))$

C) $tr(A_1) = Max(tr(A_2), tr(A_3))$ D) $tr(A_1) = tr(A_2) + tr(A_3)$

47. In a two phase simplex method, the second phase is used to find the solution which is:

A) Basic feasible B) Optimum

C) Optimum basic feasible D) Feasible

48. Let there exist a bounded optimum solution to a linear programming problem. Then a feasible solution exits to the:

A) Primal
B) Dual
C) Dual of dual
D) Both primal and dual

49. If maximum likelihood estimator exists MLE exists then it is a function of:

A) Unbiased estimator
B) Sufficient Statistic
C) Outliers
D) Sample moments

50. The Mahanalobis distance estimates the distance between:

A) Population Mean Vectors
B) Whishart matrices
C) Chi squares
D) Hotelling's T^2

x-x-x