Entrance Examination – 2018
Ph.D. Statistics

Hall Ticket No.	

Time : 2 hours
Max. Marks : 80

PART A: 40 MARKS
PART B: 40 MARKS

Instructions

1. Write your Hall Ticket Number on the OMR Answer Sheet given to you. Also write the Hall Number in the space provided above.

2. Answers are to be marked on the OMR sheet.

3. Please read the instructions carefully before marking your answers on the OMR answer sheet.

4. Hand over the OMR answer sheet at the end of the examination to the Invigilator.

5. No additional sheets will be provided. Rough work can be done in the question paper itself/space provided at the end of the booklet.

6. Calculators are not allowed.

7. There are a total of 40 questions in **PART A** and **PART B** together.

8. Each correct answer carries 2 marks.

9. The appropriate answer(s) should be coloured with either a blue or black ball point or a sketch pen. DO NOT USE A PENCIL.

10. This book contains **8 pages** including this page and excluding pages for the rough work. Please check that your paper has all the pages.

11. Given below are the meanings of some symbols that may have appeared in the question paper:
$\mathbb{R}$-The set of all real numbers, $E(X)$-Expected value of the random variable X, $V(X)$-Variance of the random variable X, $Cov(X,Y)$-Covariance of the random variables X and Y, $\rho_{X,Y}$ denotes the correlation coefficient between X and Y, iid-independent and identically distributed, pdf-probability density function, $B(n,p)$ and $N(\mu,\sigma^2)$ denote respectively,the Binomial and the Normal distributions with the said parameters. $Rank(A)$ and $det(B)$ mean rank and determinant of the matrices A and B respectively.

S-54

Part-A

1. The batting average of a group of 8 cricketers was 20. Another player joins in with batting average 29. What is the new average of the 9 players?

   (A) 21. (B) 29. (C) 23. (D) 31.

2. What is the probability that the position in which the consonants appear remain unchanged when the letters of the word Math are re-arranged?

   (A) 1/4. (B) 1/6. (C) 1/3. (D) 1/24.

3. How many five digit numbers can be formed using digits 0,1,2,3,4,5, Which are divisible by 3, without any of the digits repeating?

   (A) 96. (B) 120. (C) 181. (D) 216.

4. Among the following categories providing information on the population which one is NOT provided by population census?

   (A) Demographic characteristics.

   (B) Social-cultural characteristics.

   (C) Economic characteristics.

   (D) Characteristics related to health of the population.

5. A is a 5×5 real non-singular matrix. B is a matrix whose first and fifth rows are the fifth and first rows of A respectively, the second and fourth rows are the fourth and second rows of A respectively, and the third row of B is the same as the third row of A. Then the rank of the matrix $C = A + B$

   (A) is 5. (B) can't be said definitely. (C) is 3. (D) < 3.

6. If x is a positive integer such that $(x-1)(x-3)(x-5)....(x-93) < 0$, how many values can x take?

   (A) 47. (B) 23. (C) 26. (D) 21.

7. If the mode is to the left of median and the mean is to the right of the median, then the skewness of the distribution is best described as:

   (A) Positively. (B) Negatively. (C) Not. (D) Highly.

8. Given IQ scores are approximately normally distributed with a mean of 100 and standard deviation of 15, the proportion of people with IQs above 130 is

   (A) 95%. (B) 68%. (C) 5%. (D) 2.5%.

9. Which score has a higher relative position, a score of 48 on a test for which the mean is 40 and standard deviation is 4, or a score of 288.6 on a test for which the mean is 260 and standard deviation is 26?

   (A) A score of 48. (B) Both scores have the same relative position.

   (C) A score of 288.6 . (D) Not enough information given.

10. Select the item which is odd in the collection below:

    SPSS Mathscinet R SAS

    (A) SPSS. (B) Mathscinet. (C) R. (D) SAS.

11. Which of the following is not an advantage of pooling sample variances?

    (A) Confidence intervals are a little narrower.

    (B) The number of degrees of freedom is a little higher.

    (C) Hypothesis tests have more power.

    (D) We often know that the population variances are equal.

12. What is the infinite sum $x + 2x^2 + 3x^3 + \ldots$ for $|x| < 1$?

    (A) $x/(1-x^2)$. (B) $x/(1+x)^2$. (C) $1/(1-x)^2$. (D) $x/(1-x)^2$.

13. Suppose we use a least squares linear regression model on a set of data points (x,y). We find the coefficient of determination is 0.81 and the regression line is given by y = -51x + 32. What is the value of the correlation coefficient?

    (A) 0.6561 (B) 0.9 (C) -0.6561 (D) -0.9

14. A residual is the difference between

    (A) the observed value of y and the predicted value of y.

    (B) the observed value of y and the predicted value of x.

    (C) the observed value of x and the predicted value of x.

    (D) the observed value of x and the predicted value of y.

15. All of the following increase the width of a confidence interval except:

    (A) Increased confidence level. (B) Increased variability.

    (C) Increased sample size. (D) Decreased sample size.

16. In a Markov chain with state space {0, 1, 2} and one step transition matrix given by $P = \begin{bmatrix} 1/4 & 3/4 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/4 & 3/4 \end{bmatrix}$, the value of $p_{01}^{(2)}$ will be:

    (A) 3/4. (B)3/16. (C) 7/16. (D) 2/16.

17. Set S contains points whose abscissa and ordinate are both natural numbers. Point P, an element in set S has the property that the sum of the distances from point P to the point (8, 0) and the point (0, 12) is the lowest among all elements in set S. How many such points P exist in set S?

    (A) 1. (B) 5. (C) 8. (D) 3.

18. What will help insure that the effect of a treatment is not due to some characteristic of a single experimental unit?

    (A) Blinding. (B) Randomization. (C) Blocking. (D) Replication.

19. We wish to draw a sample of size 5 without replacement from a population 50 households. Suppose the households are numbered 01, 02, . . . , 50, and suppose that the relevant line of the random number table is:

    Digits 11362 35692 96237 90842 46843 62719 64049 17823.

    Then the households selected are

    (A) households 11 13 36 62 73.

    (B) households 11 36 23 08 42.

    (C) households 11 36 23 23 08.

    (D) households 11 36 23 56 92.

20. Let a, b and also $a - b$, be unit vectors. Then what is the angle between a and b?

    (A) 60 degrees. (B) 120 degrees.

    (C) 90 degrees. (D) None of the above.

Part-B

21. In a chi-square test of homogeneity of proportions, we test the claims that

(A) across a single sample the proportion of individuals with the same characteristic is the same as the population.

(B) different populations have the same proportions of individuals with the same characteristics.

(C) the proportion of individuals with a given characteristic does not change over time.

(D) the proportion of a population having a given characteristic is based on the homogeneity of the population.

22. $X_1, \ldots, X_n$ are i.i.d. random variables with absolutely continuous distribution function $F(x;\theta)$, then $-\sum_{i=1}^{n} logF(X_i;\theta)$ has

(A) Normal distribution. (B) Beta distribution.

(C) Gamma distribution. (D) Weibull distribution.

23. Let $X_1,\ X_2, \ldots, X_N$ be a random sample from $N_p(\mu,\ \Sigma)$ and $A = \sum_{i=1}^{N}(X_i - \bar{X})(X_i - \bar{X})'$. The MLE of Σ is:

(A) $\frac{1}{N-1}A$. (B) $\frac{1}{N}A$. (C) A. (D) $\frac{N}{N-1}A$.

24. $\lim_{n\to\infty} n\sum_{j=n}^{\infty}\frac{1}{j!} =$

(A) 2. (B) 1. (C) 0. (D) does not exist.

25. The one step transition probability matrix of a Markov chain with state space $\{0, 1, 2, \ldots\}$ is $p_{0,1} = 1$, $p_{0,j} = 0$, $j = 0, 2, 3, \ldots$; $p_{i,i+1} = p_{i,i-1}$=1/2, $\forall$ i $= 1, 2, \ldots$. Then this Markov chain is

(A) not irreducible and state 0 is recurrent, but the other states are not recurrent.

(B) irreducible and recurrent but not positive recurrent.

(C) irreducible and positive recurrent.

(D) irreducible and not aperiodic.

26. $(X_1, X_2, X_3)^T$ is a 3-variate normal random vector with mean vector $(0, 0, 0)^T$ and dispersion matrix $\begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. Then what can be said about the distribution of the random variable $\frac{1}{12}(X_1 + X_2 + X_3)^2$:

(A) it has normal with mean 0 and variance 12.

(B) it has central chi-square with 4 degrees of freedom.

(C) it has central chi-square with 2 degrees of freedom.

(D) none of the above is true.

27. A random sample from $U(\theta+1, \theta+2)$ distribution is 3.2, 2.8, 2.6, 2.9, 3.4. Based on the given sample

    (A) the maximum likelihood estimate of θ is 2.4.

    (B) any number less than 1.4 is a maximum likelihood estimate of θ.

    (C) the maximum likelihood estimate of θ is 3.

    (D) 1.45 as well as 1.5 are maximum likelihood estimates of θ.

28. A random sample from a Poisson random variable with parameter λ is given by 4, 2, 5, 5, 4. Based on this sample

    (A) e^4 is an unbiased estimate for e^{λ}.

    (B) 16 is an unbiased estimate for e^{λ}.

    (C) 20 is an unbiased estimate for e^{λ}.

    (D) 256 is an unbiased estimate for $e^{2\lambda}$.

29. X_1 and X_2 are i.i.d. $U(-1,1)$ random variables. Then the characteristics function $\phi_Y(t)$ of $Y = X_1 + X_2$ is given by

    (A) $sin(t)$. (B) $sin^2(t)$. (C) $1+it$. (D) $1-it$.

30. 9 heads showed up in 15 tosses of a coin. An unbiased estimate of $1+q+q^2+\ldots+q^{15}$, where q is 1- the probability of head, is

    (A) 8/5. (B) $\frac{5}{3}\left(1-\frac{2^{16}}{5^{16}}\right)$. (C) 7/5. (D) 5/3.

31. Two surveys were conducted before and after the recent increase in tax rate to find the proportion of voters who state they would vote for the current government. The results were as follows:

	Before	After	Total
No. surveyed	400	600	1000
No. in favor of current govt	150	150	300

An approximate 95% confidence interval for the change in support is:

(A) $(.375-.250) \pm 1.96\sqrt{\frac{(.375)(.625)}{400}+\frac{(.250)(.750)}{600}}$.

(B) $(.375-.250) \pm 1.96\sqrt{\frac{(.375)(.625)}{1000}+\frac{(.250)(.750)}{1000}}$.

(C) $(.375-.250) \pm 1.96\sqrt{\frac{(.300)(.700)}{400}+\frac{(.300)(.700)}{600}}$.

(D) $(.375-.250) \pm 1.96\sqrt{\frac{(.375)(.625)}{500}+\frac{(.250)(.750)}{500}}$.

32. A political poll was conducted to investigate opinions on capital punishment. The survey found that 25% of people contacted were not in favor of capital punishment. These results were accurate to within 3 percentage points, 19 times out of 20. Which of the following is not correct?

(A) The 95% confidence interval is approximately from 22% to 28%.

(B) We are 95% confidence that the true proportion of people not in favor is within 3 percentage points of 25%.

(C) In approximately 95% of polls on this issue, the confidence interval will include 25%.

(D) If another poll of similar size were taken, the percentage of people favouring capital punishment would likely range from 72% to 78%.

33. Which of the following is CORRECT?

(A) We do not need to randomize if our sample size is sufficiently large.

(B) A large sample size always ensures that our sample is representative of the population.

(C) In a properly chosen sample, an estimate will be less variable with a large sample size and hence more precise.

(D) In random samples, the randomization ensures that we get precise and accurate estimates.

34. Which type of bias occurs because we do not obtain complete information about a population?

(A) Nonresponse bias (B) Sampling bias

(C) Response bias (D) None of the above

35. The Central Limit Theorem states that

(A) if n is large, and if the population is normal, then the sampling distribution of the sample mean can be approximated closely by a normal curve.

(B) if n is large, and if the population is normal, then the variance of the sample mean must be small.

(C) if n is large, then the sampling distribution of the sample mean can be approximated closely by a normal curve.

(D) if n is large then the distribution of the sample can be approximated closely by a normal curve.

36. In order for the Poisson to give good approximate values for binomial probabilities we must have the condition(s) that:

(A) the population size is large relative to the sample size.

(B) the sample size is large.

(C) the probability, p, is small and the sample size is large.

(D) the probability, p, is close to .5 and the sample size is large.

37. If $\begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix} A \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, then the matrix A is

(A) $\begin{bmatrix} 3 & -4 \\ 3/4 & -1 \end{bmatrix}$. (B) $\begin{bmatrix} -13/4 & 3/2 \\ 5/4 & -1/2 \end{bmatrix}$.

(C) $\begin{bmatrix} -17/4 & 3/4 \\ -7/4 & -1/4 \end{bmatrix}$. (D) $\begin{bmatrix} 5/4 & 11/4 \\ 3 & -9/4 \end{bmatrix}$.

38. Let $N(t)$ follows a Poisson process and $s < t$. Then the probability $P[N(s) = k|N(t) = n]$ will be

(A) $e^{-s/t}\frac{(s/t)^k}{k!}$.

(B) $\binom{n}{k}\left(\frac{t}{s}\right)^n\left(1-\frac{t}{s}\right)^{n-k}$.

(C) $\binom{n}{k}\left(\frac{s}{t}\right)^k\left(1-\frac{s}{t}\right)^{n-k}$.

(D) $\left(\frac{s}{t}\right)^k\left(1-\frac{s}{t}\right)^{n-k}$.

39. Let X_1, X_2 and X_3 be independent random variables with $X_k, k = 1, 2, 3$ having probability density function $f_k(x) = k\theta e^{-k\theta x}, 0 < x < \infty, \theta > 0$. Then a sufficient statistic for θ is

(A) $X_1 + X_2 + X_3$. (B) $X_1 + 2X_2 + 3X_3$.

(C) $X_1X_2X_3$. (D) $3X_1 + 2X_2 + X_3$.

40. The 95% asymptotic confidence interval for θ of the Poisson distribution is given by

(A) $\bar{x} \pm 2.58\sqrt{\frac{\bar{x}}{n}}$. (B) $\bar{x} \pm 1.96\sqrt{\frac{\bar{x}}{n}}$.

(C) $\bar{x} \pm 1.96\sqrt{\frac{n}{\bar{x}}}$. (D) None of these.