PUMDET-2018 **82260001**

Subject: Statistics **(Booklet Number)**

Duration: 90 minutes **Full Marks: 100**

Instructions

1. All questions are of objective type having four answer options for each. Only one option is correct. Correct answer will carry full marks 2. In case of incorrect answer or any combination of more than one answer, ½ marks will be deducted.
2. Questions must be answered on OMR sheet by darkening the appropriate bubble marked A, B, C, or D.
3. Use only Black/Blue ball point pen to mark the answer by complete filling up of the respective bubbles.
4. Do not make any stray mark on the OMR.
5. Write question booklet number and your roll number carefully in the specified locations of the OMR. Also fill appropriate bubbles.
6. Write your name (in block letter), name of the examination centre and put your full signature in appropriate boxes in the OMR.
7. The OMRs will be processed by electronic means. Hence it is liable to become invalid if there is any mistake in the question booklet number or roll number entered or if there is any mistake in filling corresponding bubbles. Also it may become invalid if there is any discrepancy in the name of the candidate, name of the examination centre or signature of the candidate vis-a-vis what is given in the candidate's admit card. The OMR may also become invalid due to folding or putting stray marks on it or any damage to it. The consequence of such invalidation due to incorrect marking or careless handling by the candidate will be sole responsibility of candidate.
8. Candidates are not allowed to carry any written or printed material, calculator, pen, docu-pen, log table, any communication device like mobile phones etc. inside the examination hall. Any candidate found with such items will be reported against & his/her candidature will be summarily cancelled.
9. Rough work must be done on the question paper itself. Additional blank pages are given in the question paper for rough work.
10. Hand over the OMR to the invigilator before leaving the Examination Hall.

ROUGH WORK ONLY

1.	Box plot usually gives ‘x’-number summary of data, where the value of ‘x’ is (A) 3 (B) 4 (C) 5 (D) 6
2.	A Stem-Leaf display closely resembles (A) an Ogive (B) a Pie Chart (C) a Histogram (D) a Divided bar chart
3.	For a set of 8 observations, if 3 and 23 are the minimum and maximum values respectively, then the minimum and maximum possible values of the variance are respectively (A) 25 and 100 (B) 50 and 150 (C) 20 and 80 (D) 50 and 80
4.	Suppose the standard deviations of two groups of observations having 100 and 400 observations are 10 and 20, respectively. If S is the standard deviation of the combined group then (A) $0 < S < 15.7$ (B) $15.7 < S < 18$ (C) $18 < S < 18.4$ (D) $S > 18.4$
5.	Given the regression lines of Y on X and X on Y, where X & Y represent the weight & height of a randomly chosen individual respectively, one can find (A) the variances of X and Y (B) the product of the variances of X and Y (C) the coefficient of variation of X (D) the ratio of the coefficient of variations of X and Y
6.	For two regression equations with two variables X and Y, given by $15x - 10y + 40 = 0$, and $10x - 6y - 20 = 0$ the correlation coefficient between X and Y is approximately (A) 0.95 (B) - 0.95 (C) 0.81 (D) - 0.81
7.	Let $p_{j\,i}$ and $q_{j\,i}$ respectively stand for price and quantity for the i^{th} commodity at the j^{th} period, where j = 0 and 1 correspond to base and current periods respectively. By choosing $w_i = p_{1i}.q_{1i}$ as weights, Paasche’s price index number formula may be obtained as (A) weighted arithmetic mean of price relatives (B) weighted harmonic mean of price relatives (C) weighted geometric mean of price relatives (D) median of price relatives

8.	A population growing exponentially reaches 1000 in 1970 and 3000 in 2010. Then its doubling time (in years) is (A) $40 \log_e 2$ (B) $40 \log_e 3$ (C) $40\frac{\log_e 3}{\log_e 2}$ (D) $40\frac{\log_e 2}{\log_e 3}$
9.	For detrending a time series with cubic trend, differencing should be carried out (A) once (B) twice (C) thrice (D) four times
10.	In a control chart, even though all the points are inside the control limits, indications of presence of assignable causes of variation in the process are sometimes evidenced from unusual patterns or arrangements of points corresponding to rational subgroups. Which of the following patterns does not indicate such evidence? (A) Alternate short series on either sides of the central line (B) A series of points all falling close to one of the control limits (C) A long series predominantly on one side of the central line (D) A series of points exhibiting a trend or a cyclical pattern
11.	If the letters of the word ARRANGE are permuted at random, the probability that all the vowels will be together is (A) $\frac{1}{7}$ (B) $\frac{3}{35}$ (C) $\frac{2}{35}$ (D) $\frac{1}{35}$
12.	Let A and B be two events with probabilities $P(A)=\frac{1}{3}$ and $P(B)=\frac{3}{4}$. Then which of the following is a possible value of $P(A\cap B)$? (A) $\frac{5}{24}$ (B) $\frac{11}{24}$ (C) $\frac{1}{24}$ (D) $\frac{1}{13}$
13.	The rules of a chess championship are as follows: in the title match, if the defending champion loses to the challenger, then there is an immediate rematch; but otherwise he retains the title. The challenger wins the title only if he beats the defending champion in the rematch too. The probability of the challenger beating the defending champion in any match is ⅓ independently of all other matches. If the defending champion retains his title, the probability that he did so in the first match itself is (A) $\frac{1}{2}$ (B) $\frac{2}{3}$ (C) $\frac{3}{4}$ (D) $\frac{8}{9}$
14.	If X has a binomial distribution with parameters n = 10, p = 0.4, then the distribution of 2X is (A) binomial (n = 20, p = 0.4) (B) binomial (n = 10, p = 0.8) (C) neither binomial nor truncated binomial (D) truncated binomial

15.	An airline sells 250 tickets for a flight on a plane with 248 seats, based on past experience that on average, 1 % of passengers do not show up on time to board the flight and miss it. Assuming all passengers travel independently of others, the probability that the airline will have to refuse a seat to at least one passenger coming in time is best approximated by (A) $\frac{1}{251}$ (B) $\frac{1}{250}$ (C) $e^{-2.5}$ (D) $3.5e^{-2.5}$
16.	Let X be a random variable and G(x) = P (X < x) for all real x. Then which of the following may not be true? (A) $\lim_{x\to\infty} G(x)=1$ (B) $\lim_{x\to-\infty} G(x)=0$ (C) G is monotonically non-decreasing (D) G is right continuous
17.	Let X be a random variable with probability mass function $f(x)=\begin{cases} qp^{x-1}, x=1,2,....., \\ 0, \quad \text{otherwise} \end{cases}, \quad 0<p<1, p+q=1$. Then which of the following is true? (A) $\text{Mean}(X)=\frac{1}{p}, \text{Variance}(X)=\frac{q}{p^2}$ (B) $\text{Mean}(X)=\frac{q}{p}, \text{Variance}(X)=\frac{q}{p^2}$ (C) $\text{Mean}(X)=\frac{1}{q}, \text{Variance}(X)=\frac{p}{q^2}$ (D) $\text{Mean}(X)=\frac{p}{q}, \text{Variance}(X)=\frac{p}{q^2}$
18.	If X, Y are random variables, $X\|Y=y$ follows binomial$\left(y,\frac{1}{2}\right)$ and Y follows Poisson(λ) where $\lambda>2$, then the distribution of X is (A) Geometric $\left(\frac{1}{2\lambda}\right)$ (B) Geometric $\left(\frac{2}{\lambda}\right)$ (C) Poisson $\left(\frac{2}{\lambda}\right)$ (D) Poisson $\left(\frac{\lambda}{2}\right)$
19.	Suppose X and Y are independent random variables each following $F_{6,1}$ distribution. Then the distribution of X+Y is (A) $F_{12,1}$ (B) $F_{6,2}$ (C) $F_{12,2}$ (D) not an F distribution
20.	Let X be an exponential random variable with mean $\frac{1}{2}$. What is the median of X? (A) $\frac{1}{2}$ (B) $\log_e\sqrt{2}$ (C) $\log_e 2$ (D) $\log_e 4$

21.	Let X & Y be independent and identically distributed exponential random variable with mean 1. Then the distribution of $\frac{X}{X+Y}$ is (A) Beta (1, 1) (B) Beta $\left(1,\frac{1}{2}\right)$ (C) Beta $\left(\frac{1}{2},1\right)$ (D) Beta $\left(\frac{1}{2},\frac{1}{2}\right)$
22.	If the distribution of the maximum of a random sample of size 5 from a distribution is Uniform (0,1), then the parent distribution is (A) Uniform $\left(0,\frac{1}{5}\right)$ (B) Uniform (0,5) (C) Beta (5, 1) (D) Beta $\left(\frac{1}{5},1\right)$
23.	If X has a Uniform distribution between 1 and 2, then the distribution of 2X+10 is (A) Uniform (0, 10) (B) Uniform (10, 12) (C) not uniform (D) Uniform(12, 14)
24.	If A, B are non empty subsets of $\mathbb{R}$ (Real line) with $A\subseteq[1,2]$ and $B\subseteq[-2,-1]$, then $\sup(\{a.b : a\in A, b\in B\})$ equals (A) $-\inf(A).\sup(B)$ (B) $\inf(A).\sup(B)$ (C) $-\inf(A).\inf(B)$ (D) $\sup(A).\inf(B)$
25.	The infinite series $\sum_{n=1}^{\infty}\frac{n^{\alpha}+1}{n^{3\alpha}+2}$ converges if and only if (A) $\alpha<0$ (B) $\alpha>0$ (C) $\alpha>\frac{1}{2}$ (D) $\alpha<\frac{1}{2}$
26.	The equation $x^7+x^5-2017=0$ has (A) exactly one real root (B) exactly five real roots (C) exactly seven real roots (D) no real root
27.	The dimension of the null space of a 4 x 5 matrix of rank 3 is (A) 1 (B) 2 (C) 3 (D) 4
28.	If A and B are matrices of the same order and ranks 4 and 2 respectively, then the set of possible ranks of A – 2B is (A) {0, 2, 4} (B) {2, 4, 6} (C) {2, 3, 4, 5, 6} (D) {0, 1, 2, 3, 4, 5, 6}

29.	The determinant of the matrix $\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 3 & 0 \\ 0 & 4 & 1 & 0 \\ 4 & 0 & 0 & 1 \end{pmatrix}$ is (A) 0 (B) - 1 (C) - 11 (D) 121
30.	If $\rho_{xy} = 0, -1 < \rho_{yz} < 0$ and $-1 < \rho_{xz} < 0$ then the partial correlation coefficient $\rho_{xy.z}$ (A) is positive (B) is zero (C) is negative (D) can be positive, negative or zero
31.	The multiple correlation coefficient between X_1 and (X_2, X_3, X_4) and that between X_1 and (X_2, X_3) are respectively denoted as $r_{1.234}$ and $r_{1.23}$. Which of the following is possible for a particular data set? (A) $r_{1.23}$ = 0.86 and $r_{1.234}$ = 0.72 (B) $r_{1.23}$ = - 0.86 and $r_{1.234}$ = 0.72 (C) $r_{1.23}$ = 0.86 and $r_{1.234}$ = - 0.72 (D) $r_{1.23}$ = 0.72 and $r_{1.234}$ = 0.86
32.	Suppose X, Y and Z are jointly distributed non-degenerate random variables with finite second moments, $Z_1 = \alpha_1 + \beta_1 Z$ and $Z_2 = \alpha_2 + \beta_2 Z$ are the linear regressions of X on Z and Y on Z respectively, and $R_1 = X - Z_1$, $R_2 = Y - Z_2$. Then the partial correlation coefficient $\rho_{XY.Z}$ is the ordinary correlation coefficient between (A) Z_1 and Z_2 (B) Z and a linear combination of X and Y (C) R_1 and R_2 (D) Z and a linear combination of R_1 and R_2
33.	Suppose the variance –covariance matrix of a random vector $\underset{\sim}{X} = (X_1, X_2, X_3)'$ is $\sum = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 8 & 2 \\ 0 & 2 & 8 \end{pmatrix}$ Then covariance of $X_1 - X_3$ and $X_1 + 2X_2 + 3X_3$ is (A) 4 (B) - 16 (C) - 24 (D) - 32
34.	Suppose the sample first quartile $\hat{Q}_1$ based on a random sample of size n is used to estimate the population first quartile Q_1 of a distribution with a strictly positive continuous density f. Then for large n, the distribution of $\hat{Q}_1$ is approximately normal with variance (A) $\frac{1}{16nf^2(Q_1)}$ (B) $\frac{1}{4nf^2(Q_1)}$ (C) $\frac{3}{16nf^2(Q_1)}$ (D) $\frac{3}{4nf^2(Q_1)}$

35.	If $\overline{X}$ and $\overline{Y}$ are sample means of two independent samples of respective sizes m and n from two Poisson populations with respective parameters λ_1 and λ_2, then for large m and n, an asymptotic test for the null hypothesis $H_0 : \lambda_1 = \lambda_2$ against $H_1 : \lambda_1 \neq \lambda_2$ can be carried out using the statistic (A) $\overline{X} - \overline{Y}$ (B) $m\overline{X} - n\overline{Y}$ (C) $\sqrt{\overline{X}} - \sqrt{\overline{Y}}$ (D) $\sqrt{m\overline{X}} - \sqrt{n\overline{Y}}$
36.	If T_n is consistent for estimating the parameter ϴ, then which of the following is necessarily true? (A) T_n is unbiased for ϴ (B) T_n is the only estimator consistent for ϴ (C) T_n is asymptotically unbiased for ϴ (D) $4T_n$ is consistent for estimating 4ϴ
37.	To estimate σ^2 based on a random sample $X_1, X_2, \ldots. X_n$, $n \geq 2$ from Normal (τ, σ^2) with both parameters unknown, an estimator T_n of the form $T_n = c.\sum_{j=1}^{n}(X_j - \overline{X})^2$ is used where c is constant and $\overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$. Among the following choices, the one with the smallest mean squared error is given by choosing (A) $c = \frac{1}{n+2}$ (B) $c = \frac{1}{n+1}$ (C) $c = \frac{1}{n-1}$ (D) $c = \frac{1}{n}$
38.	Suppose X_1 and X_2 are independent normally distributed random variables with $E(X_1) = E(X_2) = \mu$ and $Var(X_1) = 1, Var(X_2) = 4$. Then $\frac{X_1 + X_2}{2}$ is (A) the minimum variance unbiased estimator of μ (B) the best linear unbiased estimator of μ (C) not an MLE (D) not an unbiased estimator of μ
39.	Suppose X has a binomial distribution with n = 10 but p unknown. Suppose one needs to test the null hypothesis p = 0.5 against the alternatives that it is greater than 0.5. If the observed value of X = 10, then the p-value of the UMP test is closest to (A) 0.0001 (B) 0.05 (C) 0.00098 (D) 0.009

40.	A pivot for setting up a confidence interval for λ based on a single observation from a Gamma (α,λ) population with density $f(x)=\frac{\lambda^{\alpha}}{\Gamma\alpha}x^{\alpha-1}e^{-\lambda x}$, $x>0$, where α is known, is given by (A) $\frac{X}{\lambda}$ (B) λX (C) λX^{α} (D) X^{α}/λ
41.	Suppose you are testing the null hypothesis that a person is innocent against the alternative hypothesis that the person is guilty. Also suppose the chance that you will conclude that the person is guilty if she is innocent is 0.1 and the chance that you will conclude the person is innocent if she is guilty is 0.6. Then which of the following is true? (A) P [Type-I error] = 0.1, P [Type-II error] = 0.6 (B) P [Type-I error] = 0.1, P [Type-II error] = 0.4 (C) P [Type-I error] = 0.9, P [Type-II error] = 0.6 (D) P [Type-I error] = 0.9, P [Type-II error] = 0.4
42.	Sign test is used for testing the hypothesis about the (A) mean of the distribution (B) median of the distribution (C) comparison of the means of two distributions (D) comparison of the medians of two distributions
43.	Let X and Y be independent and identically distributed random variables according to the exponential distribution with mean $\frac{1}{\theta}$, $\theta>0$. For testing $H_0:\theta=1$ against $H_1:\theta<1$, which of the following is true? (A) UMP test does not exist (B) UMP test rejects H_0 when $\frac{1}{X}+\frac{1}{Y}$ is small (C) UMP test rejects H_0 when X + Y is small (D) UMP test rejects H_0 when X + Y is large

44.	For testing equality of means in one-way analysis of variance against all alternatives, the distribution of the test statistic is F under the null hypothesis if (A) the observations at each level represent a random sample from a normal distribution. (B) the variance of the normal distributions across levels should be equal. (C) the observations across different levels are generated independently. (D) the statements (A), (B) and (C) are all true.
45.	The purpose of one-way analysis of covariance (ANCOVA) with one concomitant variable is to examine (A) whether population variances of the response variable are the same across levels of a factor, adjusting for differences on the concomitant variable . (B) whether population means of the response variable are the same across levels of a factor, adjusting for differences on the concomitant variable. (C) whether population variances of the response variable are the same across levels of a factor, without adjusting for differences on the concomitant variable . (D) whether population means of the response variable are the same across levels of a factor, without adjusting for differences on the concomitant variable .
46.	RBD is an improvement over CRD with respect to the principle of (A) local control (B) randomization (C) replication (D) factorization
47.	Suppose a 2^3 experiment with 3 factors (P, K, D) each at 2 levels are conducted in 4 replicates each containing 2 blocks, where the principal blocks respectively contain the treatment combinations ((1), pkd, kd, p), (pd, (1), k pkd), (pkd, pk, d, (1)) and ((1), pk, kd, pd). Then in this design (A) KD, PD, PK and PKD are all partially confounded (B) KD, PD, PKD and D are all partially confounded (C) PK, PKD, D and KD are all partially confounded (D) KD, PKD, P and PK are all partially confounded
48.	Suppose in a split plot design, we have a factor A at p levels, which are arranged in an RBD using r blocks and a second factor B at q levels, which are applied to the plots of a block after subdividing each whole plot into q subplots. Then the whole plot and subplot error degrees of freedom are respectively given by (A) (r - 1) (p - 1) and p (q – 1) (r – 1) (B) (r - 1) (q - 1) and q (p – 1) (r – 1) (C) (p - 1) (q - 1) and p (q – 1) (r – 1) (D) (r - 1) (p - 1) and q (p – 1) (r – 1)

49.	Sampling error in sample survey arises due to (A) non-response (B) under-coverage (C) faulty questionnaire design (D) selecting a subset of individuals from the population
50.	A simple random sample of size 5 is drawn with replacement (SRSWR) from a population of 100 units. The expected number of distinct units in the sample is (A) $100\left[1-\left(\frac{99}{100}\right)^5\right]$ (B) $100\left[1-\left(\frac{98}{100}\right)^5\right]$ (C) $100\left[1-\left(\frac{95}{100}\right)^5\right]$ (D) $100\left[1-\left(\frac{94}{100}\right)^5\right]$