Test Booklet No. _____

This booklet consists of 100 questions and 12 printed pages.

RGUPET/____/__

Series NIL

RGUPET 2023
Ph.D. in STATISTICS

Full Marks: 100 **Time: 3 Hours**

Roll No. | | | | | | | | |

Day and Date of Examination :

Signature of Invigilator(s) :

Signature of Candidate :

General Instructions:

PLEASE READ ALL THE INSTRUCTIONS CAREFULLY BEFORE MAKING ANY ENTRY.

1. DO NOT OPEN THIS TEST BOOKLET UNTIL YOU ARE TOLD TO DO SO.
2. Candidate must write his/her Roll Number on the space provided.
3. This Test Booklet contains 100 Multiple Choice Questions (MCQs) from the concerned subject. Each question carries 1 mark.
4. Please check the Test Booklet to verify that the total pages and total number of questions contained in the test booklet are the same as those printed on the top of the first page. Also check whether the questions are in sequential order or not.
5. Candidates are not permitted to enter into the examination hall 15 minutes after the commencement of the entrance test or leave the examination hall before 30 minutes of end of examination.
6. Making any identification mark in the OMR Answer Sheet or writing Roll Number anywhere other than the specified places will lead to disqualification of the candidate.
7. Candidates shall maintain silence inside and outside the examination hall. If candidate(s) is/are found violating the instructions mentioned herein or announced in the examination hall, they will be summarily disqualified from the entrance test.
8. In case of any dispute, the decision of the Entrance Test Committee, RGU shall be final and binding.
9. The OMR Answer Sheet consists of two copies, the Original copy and the Student's copy.

1	$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$ is an example of					
	a) Distributed lag model	b) Auto regressive model	c) Regression model	d) All the above	a	Distributed lag model
2	In Almon distributed lag model; the second-degree polynomial in i is					
	a) $\beta_i = a_0 + a_1 i + a_2 i^2$	b) $\beta_i = a_0 + a_1 i + a_2 i^2 + a_3 i^3$	c) $\beta_i = a_2 i^2$	d) $\beta_i = a_0 + a_1 i + a_2 i + v$	a	$\beta_i = a_0 + a_1 i + a_2 i^2$
3	The variance-covariance matrix of residuals for the linear regression model is $\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. The given covariance matrix shows the problemof					
	a) Autocorrelation	b) Non autocorrelation	c) Hetroscedasticity	d) Homoscedasticity	c	Heteroscedasticity
4	In the case of perfect multicollinearity, the value of the determinant is equal to					
	a) 0	b) 1	c)>1	d)<1	a	0
5	Randomization is a process in which the treatments are allocated to the experimental units:					
	a) At the will of the investigator.	b) In a sequence	c) With equal probability.	d) With unequal probability.	c	With equal probability.
6	In a completely randomized design with 't' treatments and 'n' experimental units, error degrees of freedom is equal to:					
	a) n-t	b) n-t-1	c) n-t+1	d) t-n	a	n-t
7	The randomized block design is a					

	a) no classification	b) one-way classification	c) two-way classification	d) three-way classification	c	two-way classification
8	Error sum of squares in two-way classification (RBD) as compared to one-way classification (CRD) using the same material is					
	a) more	b) less	c) equal	d) not comparable	b	less
9	The additional effect gained due to the combined effect of two or more factors is known as:					
	a) main effect	b) interaction effect	c) partial effect	d) all	b	Interaction effect
10	The difference between the upper and lower limit of a class is called					
	a) class interval	b) mid value	c) lower value	d) upper value	a	class interval
11	Tabulation makes the data easily					
	a) analysis	b) understandable	c) quantified	d) qualified	b	understandable
12	In a column chart, the bars are:					
	a) horizontal	b) vertical	c) slanting	d) at an angle of 45 degree	b	vertical
13	The sum of the deviations about the mean is:					
	a) zero	b) minimum	c) maximum	d) one	a	zero

14	The correct relation between variance and standard deviation (S.D.) of a variable X is:				d	S.D.= $[Var(X)]^{1/2}$
	a) S.D .= $Var($	b) S.D.=$[Var(X)]^2$	c) S.D.= $[Var(X)]^3$	d) S.D.=$[Var(X)]^{1/2}$		
15	In the case of positively skewed distribution, the extreme values lie in the				c	right tail
	a) left tail	b) middle	c) right tail	d) does not exist		
16	In the regression line of Y on X, the variable X is known as:				a	independent variable
	a) independent variable	b) dependent variable	c) response variable	d) all		
17	When β_{YX} is the slope for the regression line of Y on X and β_{XY} is the slope for the regression line of X on Y. What should be the value of β_{XY} if $\beta_{YX} > 1$				a	less than 1
	a) less than 1	b) greater than 1	c) equal to 1	d) equal to 0		
18	The range of correlation coefficient is:				d	-1 to 1
	a)$-\infty$ to 1	b) -1 to ∞	c) -∞ to ∞	d) -1 to 1		
19	The individual probabilities of occurrence of two events A and B are known, the probability of occurrence of both events together will be				b	decreased
	a) increased	b) decreased	c) one	d) zero		
20	One of the two events must happen; given that the chance of one is one-fourth of the other. The odd in favour of the other is:				b	1: 4
	a) 1: 3	b) 1: 4	c) 4: 1	d) 1: 5		
21	Let $\underline{X} \sim N_p\left(\underline{\mu}, \Sigma\right)$ and $\left(\underline{X}-\underline{\mu}\right)'\Sigma^{-1}\left(\underline{X}-\underline{\mu}\right)$ follow the Chi-square distribution with degrees of freedom					

	a)p-1	b) p	c) p+1	d) p+2	b	p
22	Suppose $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$ follow iid $N_p(\underline{0}, \Sigma)$ then the $p \times p$ matrix $W = \sum_{i=1}^{n} \underline{X}_i \underline{X}_i' = X'X$ has the distribution					
	a) $W \sim N_p(\underline{0}, \Sigma)$	b) $W \sim N_n(\underline{0}, \Sigma)$	c) $W \sim W_p(\underline{0}, \Sigma)$	d) $W \sim W_n(\underline{0}, \Sigma)$	c	$W \sim W_p(\underline{0}, \Sigma$
23	The _____ distribution is a multivariate generalization of the chi-square distribution.					
	a) Multivariate Normal	b) Hotelling's T^2	c) Wishart distribution	d) all	b	Hotelling's T^2
24	If the element c{1,2} of the covariance matrix C is 114, what is the value of c{2,1} and what is the meaning?					
	a)114, variance	b)1/114, variance	c)1/114, covariance	d)114, covariance	d	114, covariance
25	This process is performed after extraction to obtain a more interpretable factor solution.					
	a) factor Normalization	b) factor rotation	c) factor optimization	d) factor interpretation	b	factor rotation
26	It is a correlation coefficient, which tells us the extent to which a question is measuring that factor.					
	a) factor analysis	b) factor variable	c) factor rotation	d) factor loading	d	factor loading
27	Testing the overall significance of multiple regression could be done by:					
	a) F-test	b) t-test	c) Chi-square test	d) Chow test	a	F-test
28	The smaller the standard errors, the stronger is the evidence that the estimates are statistically					
	a) insignific ant	b) significant	c) in conclusion	d) wrong	b	significan t
29	Find the linear model from the following:					

	a)$Y = \beta_0 + e^{\beta_1 X} + \epsilon$	b)$Y = \frac{\beta_0}{e^{\beta_1 X}} + \epsilon$	c) $Y = \beta_0 + \log(\beta_1 X) + \epsilon$	d) $Y = \beta_0 + \beta_1 X + \epsilon$	d	$Y = \beta_0 + \beta_1 X + \epsilon$
30	For a Normal equation $X'X\hat{\beta} = X'Y$, the matrix $X'X$ has a full rank. The solution of $\hat{\beta}$ has					
	a) a unique	b) an infinite	c) 0	d) 2 times the rank of X'X	a	a unique
31	$100(1-\alpha)$% confidence interval of the slope β_1 in simple regression is given by					
	a)$\hat{\beta}_1 \pm t_{\frac{\alpha}{2},n-2} S.E.($	b)$\hat{\beta}_1 \pm t_{1-\frac{\alpha}{2},n-2} S.E.(\hat{\beta}_1)$	c)$\hat{\beta}_1 \pm t_{\frac{\alpha}{2},n-1} S.E.(\hat{\beta}_1)$	d)$\hat{\beta}_1 \pm t_{\frac{\alpha}{2},n} S.E.(\hat{\beta}_1)$	a	$\hat{\beta}_1 \pm t_{\frac{\alpha}{2},n-2} S.E$
32	**Logistic regression is used for:**					
	a) regression	b) classification	c) clustering	d) all	a	regression
33	The name of the link function for Poisson regression is:					
	a) logistic link	b) identity link	c) log link	d) reciprocal link	c	log link
34	The moving average (MA) process of order 1 is					
	a) $Y_t = \emptyset Z_{t-1} + Z_t$	b) $Y_t = Z_{t-1} + \emptyset Z_t$	c) $Y_t = \emptyset Y_{t-1} + Z_t$	d) $Y_t = Y_{t-1} + \emptyset Z_t$	a	$Y_t = \emptyset Z_{t-1} + Z_t$
35	A time series consists of					
	a) short term variations	b) long term variations	c) irregular variation	d) all	d	all
36	The increase in the number of patients in the hospital due to heat stroke is:					
	a) seasonal variation	b) secular trend	c) irregular variation	d) cyclical variation	a	seasonal variation

37	For AR(1) model $Y_t = \emptyset Y_{t-1} + Z_t$, the random variable Z_t follows:					
	a) iid $N(0, \sigma^2)$	b) $WN(0, \sigma^2)$	c) $N(0, \sigma^2)$	d) independent $N(0, \sigma^2)$	b	$WN(0, \sigma^2)$
38	If the slope of the trend line $\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 t$ is positive, it shows					
	a) rising trend	b) declining trend	c) stagnation	d) any of them	a	rising trend
39	Previous probabilities in Bayes Theorem that are changed with the newly available information are called					
	a)independent probabilities	b)dependent probabilities	c)interior probabilities	d)posterior probabilities	d	posterior probabilities
40	The formula for Bayes theorem is					
	a)$P(A\|B) = \frac{P(B\|A)P(A)}{P(B)}$	b)$P(A\|B) = \frac{P(A)}{P(B)}$	c)$P(A\|B) = \frac{P(B\|A)}{P(B)}$	d)$P(A\|B) = \frac{1}{P(B)}$	a	$P(A\|B) = \frac{P(B\|A)P}{P(B)}$
41	The method in which the previously calculated probabilities are revised with values of new probability is called					
	a) Revision theorem	b) Bayes theorem	c) Dependent theorem	d) Updation theorem	b	Bayes theorem
42	Let $X_1, X_2, \ldots, X_n$ be the iid $B(\theta)$. Then conjugate prior distribution of θ is					
	a) Student's t	b) Normal	c) Beta	d) Poisson	c	Beta
43	Which one is not the distribution of exponential family from the given distribution					
	a) Poisson	b) Normal	c) Binomial	d) Student-t	d	Student-t

44	A family of parametric distribution in which mean is equal to		

	variance is					
	a) Binomial distribution	b) Gamma distribution	c) Normal distribution	d) Poisson distribution	d)	poisson distribution
45	The distribution possessing the memoryless property is					
	a)gamma distribution	b)geometric distribution	c) hypergeometric distribution	d) all the above	b)	geometric distribution
46	The distribution in which the probability at success draw varies is					
	a) hypergeometric distribution	b) geometric distribution	c) binomial distribution	d) discrete uniform distribution	a)	hypergeometric distribution
47	The distribution for which the mode does not exist is					
	a) normal distribution	b) t-distribution	c)continuous rectangular distribution	d) F-distribution	c)	continuous rectangular distribution
48	If $X \sim N(\mu, \sigma^2)$, the maximum probability at the point of inflexion of normal distribution is					
	a) $\pm\mu$	b) $\mu\pm\sigma$	c) $\sigma\pm\mu$	d) $\pm\sigma$	b)	$\mu\pm\sigma$
49	An approximate relation between Q.D. and S.D. of normal distribution is					
	a) 5 Q.D.=4 S.D.	b) 4 Q.D.=5 S.D.	c) 2 Q.D.= 3 S.D.	d) 3 Q.D. = 2 S.D.	d)	3 Q.D. = 2 S.D.
50	The area under the standard normal curve beyond the lines $z = \pm 1.96$ is					
	a) 95 per cent	b) 90 per cent	c) 5 per cent	d) 10 per	c)	5 per cent

				cent		
51	If the sample size n=2, the student's t-distribution reduces to					
	a) normal distribution	b) F-distribution	c) Cauchy distribution	d) none of the above	c)	Cauchy distribution
52	The relation between the mean and variance of χ^2 with n d.f. is					
	a) mean=2 variance	b) 2 mean=variance	c) mean = variance	d) none of the above	b)	2 mean=variance
53	The distribution for which the moment generating function does not exist but moments exists is					
	a) Pareto distribution	b) t-distribution	c) F-distribution	d) all the above	d)	all the above
54	In a multivariate study, the correlation between any two variable eliminating the effect of all other variables is called					
	a) simple correlation	b)multiple correlation	c) partial correlation	d)partial regression	c)	partial correlation
55	Parameters are those constants which occur in					
	a) Samples	b)probability density function	c)a formula	d) none of the above	b)	probability density function
56	Estimation of parameters in all scientific investigations is of					
	a)prime importance	b)secondary importance	c)no use	d)deceptive nature	a)	prime importance
57	Factorisation theorem for sufficiency is known as					
	a) Rao-	b) Crammer-	c) Chapman-	d)Fisher-	d)	Fisher-Neyman

	Blackwell theorem	Rao theorem	Robin theorem	Neyman theorem		theorem
58	Crammer-Rao inequality is based on					
	a) stringent conditions	b) mild conditions	c)no conditions	d)none of the above	a)	stringent conditions
59	The inequality for the lower bound of the variance of an estimator which is not based on stringent conditions was given by					
	a) Aitken and Silverstone	b)Neyman-Pearson	c)Chapman-Robins	d)none of the above	c)	Chapman-Robins
60	Minimum Chi-square estimators are					
	a) consistent	b) asymptotically normal	c) efficient	d) all the above	d)	all the above
61	The maximum likelihood estimators are necessarily					
	a) unbiased	b) sufficient	c)most efficient	d)unique	b)	sufficient
62	A wrong decision about H_0 leads to					
	a) one kind of error	b) two kinds of error	c)three kinds of error	d) four kinds of error	b)	two kinds of error
63	Neyman-Pearson lemma provides					
	a) an unbiased test	b) a most powerful test	c)an admissible test	d) minimax test	b)	a most powerful test
64	The ratio of the likelihood function under H_0 and under the entire parametric space is called					

	a) probability ratio	b)sequential probability ratio	c)likelihood ratio	d)none of the above	c)	likelihood ratio
65	The degrees of freedom for statistic-t for paired t-test based on n pairs of observation is					
	a) 2(n-1)	b)n-1	c)2n-1	d) n-2	b)	n-1
66	Equality of several normal population means can be tested by					
	a) Bartlett's test	b) F-test	c) $\chi^2 - test$	d) t-test	b)	F-test
67	In sequential probability test (SPRT) the sample size is					
	a) treatments are equally effective	b)treatments differ significantly	c)no conclusion	d) none of the above	b)	treatments differ significantly
68	Least square estimators of the parameters of linear model are					
	a) unbiased	b) BLUE	c) UMVU	d) all the above	d)	all the above
69	A uniformly most powerful test among the class of unbiased test is termed as					
	a) minimax test	b) minimax unbiased test	c) uniformly most powerful test	d) all the above	c)	uniformly most powerful test
70	SPRT was initiated by					
	a) R. A. Fisher	b)A. Wald	c)G. W. Snedecor	d)Thomas Bayes	b)	A. Wald
71	A contingency table having a zero count is called					

	a)a complete contingency table	b) an incomplete contingency table	c)abnormal contingency table	d)none of the above	b)	an incomplete contingency table
72	Probability of selection varies at each subsequent draw in					
	a) sampling without replacement	b)sampling with replacement	c)both a) and b)	d)neither a) nor b)	a)	sampling without replacement
73	A population consisting of results of the conceptually repeated trails is known as					
	a)hypothetical population	b)finite population	c)infinite population	d)real population	a)	hypothetical population
74	Under equal allocation in stratified sampling the sample from each stratum is					
	a) proportional to stratum size	b) of same size from each stratum	c) in proportion to the per unit cost of survey of the stratum	d)all the above	b)	of same size from each stratum
75	Systematic sampling means					
	a)selection of n contiguous units	b) selection of n units situated at equal distance	c)selection of n large units	d) selection of n middle units in a sequence	b)	selection of n units situated at equal distance
76	Double sampling has its utility in					
	a)stratified	b)ratio method	c)regression	d)all the	d)	all the above

	sampling	of estimation	method of estimation	above		
77	In which of the following situation(s) cluster sampling is appropriate					
	a)when the units are situated for apart	b) when sampling frame is not available	c)when all the elementary units are easily identifiable	d) all the above	d)	all the above
78	In what situation two stage sampling is better than single sampling					
	a)when the elements in the same stage are positively correlated	b)when the elements in the same stage are negatively correlated	c)when the elements in the same stage are uncorrelated	None of the above	b)	when the elements in the same stage are negatively correlated
79	Non-response in survey mean					
	a) non-availability of respondents	b) non-return of questionnaire by the respondents	c) refusal to give information by the respondents	d)all the above	d)	all the above
80	Two-stage sampling design is more efficient than single stage sampling if the correlation between units in the first stage is					
	a)negative	b)positive	c)zero	d)none of the above	b)	positive
81	Circular systematic sampling was first used by					

	a)W .G. Cochran	b) M.H. Hansen	c) D.B. Lahiri	d) P.C. Mahalonobis	c)	D.B. Lahiri
82	Which of the following sampling designs will be categorised as non-probability sampling					
	a) haphazard sampling	b) convenience sampling	c) judgement sampling	d)all the above	d)	all the above
83	There are more chance of non-sampling errors than sampling errors in case of					
	a) studies of large sample	b) complete enumeration	c) insufficient investigation	d) all the above	d)	all the above
84	Which one of the following is an example of random process in communication?					
	a) Channel noise	b) Interference	c) Both a) and b)	d) None of the above	c)	Both a) and b)
85	The random walk is an example of					
	a) Non-deterministic process	b) Deterministic process	c) Both a and b	d) None of the above	a)	Non-deterministic process
86	Stochastic process are					
	a) Random in nature	b) Are function of time	c) Random in nature and are a function of time	d) None of the above	c)	Random in nature and are a function of time
87	In post-independence India, the registration of Births and Deaths Act was passed in					

	a) 1948	b)1959	c)1969	d)1979	c)	1969
88	The fertility of a women in India is maximum in the age group					
	a)15-20	b)20-24	c)25-29	d)15-29	c)	25-29
89	The age-specific death rate for the babies of age less than one year is specifically called					
	a)neonatal death rate	b)infant mortality rate	c)maternal mortality rate	d)foetal death rate	b)	infant mortality rate
90	A life table based on the experience of actual cohort is called					
	a) generation life table	b) fluent life table	c) both a) and b)	d) neither a) nor b)	C)	both a) and b)
91	Chance variation is respect of quality control of a product is					
	a) tolerable	b) not effecting the quality of a product	c) uncontrollable	d)all the above	d)	all the above
92	The cause leading to vast variation in the specification of a product are usually due to					
	a) random process	b) assignable causes	c) non-traceable causes	d) all the above	b)	assignable causes
93	R-charts are preferable over σ -charts because					
	a) R and S.D. fluctuate together in case of small samples	b) R can be easily calculated	c) R-charts are economical	d) all the above	d)	all the above
94	The graph of the proportion of defectives in the lot against					

	average sample number is					
	a) OC curve	b) A.S.N. curve	c) power curve	d) all the above	b)	A.S.N. curve
95	OC curve reveals the ability of the sampling plan to distinguish between					
	a) good and bad lots	b) good and bad sampling plans	c) good and bad product	d) all the above	a)	good and bad lots
96	A sampling plan is good for use provided					
	a) it ensures correct decision about the acceptance or rejection of a lot	b) it requires an adequate number of inspection	c) it is not complicated	d) all the above	d)	all the above
97	In sequential probability ratio test, the lies dividing the total space into regions are					
	a) perpendicular to each other	b) passing through the origin	c) parallel to each other	d) all the above	c)	parallel to each other
98	In a sequential sampling plan, the sample size is					
	a) a discrete random variable	b) a continuous random variable	c) both a) and b)	d)neither a) nor b)	a)	a discrete random variable
99	Who is regarded as the father of biostatistics					
	a) Fisher	b) Karl	c) Francis	d) Francis	c)	Francis Galton

		Pearson	Galton	Bacon		
100	The branch of biostatistics that deals with methods of collection, organization and presentation of data is called as					
	a) inferential biostatistics	b) descriptive biostatistics	c) both a) and b)	d) comparative biostatistics	b)	descriptive biostatistics

SPACE FOR ROUGH WORK