PUNJAB PUBLIC SERVICE COMMISSION

COMBINED COMPETITIVE EXAMINATION FOR RECRUITMENT TO THE POSTS OF PROVINCIAL MANAGEMENT SERVICE, ETC -2021

CASE NO. 3C2022

SUBJECT: STATISTICS (PAPER-I)

TIME ALLOWED: THREE HOURS **MAXIMUM MARKS: 100**

NOTE:

i. All the parts (if any) of each Question must be attempted at one place instead of at different places.
ii. Write Q. No. in the Answer Book in accordance with Q. No. in the Q. Paper.
iii. No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed.
iv. Extra attempt of any question or any part of the question will not be considered.

ATTEMPT FIVE QUESTIONS IN ALL. CALCULATOR IS ALLOWED (NOT PROGRAMMABLE)

Q.No.1 (a) Write down empirical relation between mean, median and mode.

(b) Find the median, quartiles and the 8^{th} decile for the distribution of examination marks given below:

Marks	30-39	40-49	50-59	60-69	70-79	80-89	90-99
No of student	8	87	190	304	211	85	20

(4+16=20 Marks)

Q.No.2 (a). Describe the properties of standard deviation.

(b) By multiplying each of the number 3,6,2,1,7,5 by 2 and then adding 5 we get 11,17,9,7,19,15 what is the relation between the standard deviation and the mean for the two sets.

(c) The breaking strength of 20 test pieces of a certain allow is given as under

95,103,97,130,96,73,78,95,89,68,82,79,69,67,83,108,94,87,93,117

Show that Mean deviation=(4/5)Standard deviation.

(4+8+8=20 Marks)

Q.NO.3. (a) Box A contains 5 green and 7 red balls .Box B contains 3 green ,3 red and 6 yellow balls. A box is selected at random and a ball is drawn at random from it. What is the probability that ball drawn is green,

(b) Let X and Y be two discrete random variables with the following joint p.d

Y\X	2	4
1	0.10	0.15
3	0.20	0.30
5	0.10	0.15

Show that E(XY)=E(X)E(Y)

(10+10=20 Marks)

Q.No.4. (a) Let X be random variable with p.d.f.

$f(x) = 2(x-1) \quad 1<x<2$

Find the $E(2X-1)$ and $E(X^2)$.

(b) Let X and Y have the joint p.d.f described as follows.

(x,y)	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)
f(x,y)	2/15	4/15	3/15	1/15	1/15	4/15

and f(x,y) is equal to zero elsewhere . Find the two marginal p.d.f 's and the correlation coefficient.

(10+10=20 Marks)

Q.No.5. (a) A certain event is believed to follow the Binomial distribution. In 1024 samples of 5, the result was observed once 405 times and twice 270 times . Find p and q .

(b) A random sampling of 4 members of a 150 members club has shown that 3 prefer no smoking in the clubhouse dining room. What is the probability that this will occur if in fact only 20% of members prefer no smoking in the dining room. Find this probability assuming that the sample was obtained under

i) sampling without replacement, and

ii)sampling with replacement.

Compare the two answers. **(10+10=20 Marks)**

Q.No.6. (a) For a normal distribution prove that mean=median=mode=μ.

(b) A random variable X is normally distributed with mean μ=166 and σ = 20 .Find

i)P(170<X<200) ii) P(148<X<172)

(10+10=20 Marks)

Q.No.7 (a) If n pairs of values of two variables a and b are given, whereas each variable is ranked in order(1 to n),show that the coefficient of correlation between ranks is given by
$r_s = 1 - (6\sum d^2 / (n(n^2-1))$

(b) Obtain the product moment coefficient of correlation between following values

a	7.4	9.0	11.0	2.5	4.6	6.4
b	8.5	6.1	2.4	6.7	12.6	3.3

Rank the values and hence find a rank correlation between the two sets

(6+14=20 Marks)

Q.No.8 (a) Given the following pairs of values of X and Y.

X	0	1	2	3	4
Y	10	17	28	41	62

Fit a suitable curve.

(b) The following data represent concomitant values of three variables.

X_1	32	18	32	16	42	48
X_2	3	2	3	1	4	6
X_3	2	4	2	5	3	9

Calculate all the multiple correlation coefficients.

(10+10=20 Marks)

PUNJAB PUBLIC SERVICE COMMISSION
COMBINED COMPETITIVE EXAMINATION
FOR RECRUITMENT TO THE POSTS OF
PROVINCIAL MANAGEMENT SERVICE, ETC -2021
CASE NO. 3C2022

SUBJECT: STATISTICS (PAPER-II)

TIME ALLOWED: THREE HOURS **MAXIMUM MARKS: 100**

NOTE:

i. All the parts (if any) of each Question must be attempted at one place instead of at different places.
ii. Write Q. No. in the Answer Book in accordance with Q. No. in the Q. Paper.
iii. No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed.
iv. Extra attempt of any question or any part of the question will not be considered.

NOTE: Attempt any FIVE questions in ALL. Calculator is allowed. (Not Programmable).

Q No. 1 a) Describe briefly the difference between
i) Probability and Non-probability sampling.
ii) Stratified and Cluster sampling

b) A local community is stratified in four blocks. If we wish to select a stratified random sample of size n = 40 by proportional allocation on the basis of number of houses in each block.

Block	A	B	C	D
No. of households	144	162	198	216

Calculate the sample size allocated to each block.

c) A population consists of 2, 4, 4, 4, 6, 8 and 10.
i) Draw all possible samples of size $n = 2$ without replacement.
ii) Calculate the mean of each sample and verify that

$$\mu_{\bar{x}} = \mu \text{ and } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

iii) Between what two values would you expect *at least* $\frac{8}{9}$ of the sample means to fall?

(4+4+12 = 20 Marks)

Q No. 2 a) Describe the following
i) Point estimator
ii) Confidence level
iii) Maximum likelihood Estimation

b) If X_1, X_2 and X_3 are a random sample from a normal population with the mean μ and the variance σ^2, what is the relative efficiency of the estimator $T_1 = \frac{X_1 + 2X_2 + X_3}{4}$ with respect to $T_2 = \bar{X}$?

c) There is a proposal under consideration to build an overhead bridge in a locality. As a part of feasibility and acceptability of this proposal, a poll is taken among the residents of the city and its suburbs. If 2400 of 4000 city residents favour the proposal and 1500 of 2000 suburban residents favour it, find a 90% confidence interval for the true difference between the proportion of city and suburban residents who favour the proposal to construct the overhead bridge. **(6+6+8 = 20 Marks)**

Q No. 3 a) Define Type-I error and Type-II error.

b) A sample of 25 observations from a normal population with $\sigma = 3$, is selected at random. Test the hypothesis $H_0 : \mu = 67$ against $H_1 : \mu > 67$ at 5% level of significance.

c) Given two random samples of size $n_1 = 9$ and $n_2 = 16$, from two independent normal populations, with $\bar{x}_1 = 75$, $\bar{x}_2 = 60$, $s_1 = 13.61$ and $s_2 = 12.5$, test the hypothesis at the 10% level of significance that $\mu_1 = \mu_2$ against the alternative that $\mu_1 > \mu_2$. Assume that the populations have equal variances.

(2+8+10 = 20 Marks)

Q No. 4 a) In the context of analysis of variance, define the following:
i) Main effects ii) Interaction effect

b) Determinations of yields of a process with four treatments are given:

	Treatments			
	1	2	3	4
Yields	11	6	8	14
	4	4	6	27
	4	3	4	8
	5	6	11	18

i. Test the hypothesis that no differences exist among the four treatments at $\alpha = 0.05$.
ii. Apply Least Significant Difference test to identify the pairwise significant differences at 5% level of significance.

(4+16 = 20 Marks)

Q No. 5 a) The following is percentage distribution by income level and ownership of a random sample of 400 families in the city of Lahore.

	Monthly Income		
	Less than Rs.60,000	Rs.60,000 to Rs.100,000	More than Rs.100,000
Home Owner	5%	25%	20%
Renter	15%	25%	10%

Test the hypothesis that the home ownership is independent of the family income level, using 1% level of significance.

b) Given the two samples below, test the null hypothesis that the population medians are equal against the alternative that $M_1 < M_2$, at $\alpha = 0.05$ by applying the Wilcoxon rank-sum test.

Sample 1	26,	25,	38,	33,	42,	40,	44,	26,	25,	43,	35,	48,	37,		
Sample 2	44,	30,	34,	47,	35,	46,	35,	47,	48,	34,	32,	42,	43,	49,	46,	47

(10+10 = 20 Marks)

Q No. 6 a) Compute the consumer price index number for 2020 with 2015 as base for the following data. Use as weights (i) quantities consumed in the base year (ii) the values in the base year.

Article	Quantity	Price (Rs.)	
	2015	2015	2020
Food	50 kg	180	265
Cloth	30 metre	260	280
Electricity	75 units	25	30
Rent	1 room	3000	3750
Miscellaneous	34 units	50	70

b) Given the following data.

Year	Quarters			
	I	II	III	IV
2015	112	125	129	110
2016	119	132	147	115
2017	120	142	150	118
2018	128	151	162	125

i) Fit a linear trend to the annual averages.
ii) Calculate quarterly trend values from the trend equation obtained in part (i).

(8+12 = 20 Marks)

Q No. 7 a) Describe the functions of Pakistan Bureau of Statistics.
b) Calculate the crude death rate and the standardized death rate for the data:

Age (years)	District A				Standard Population ('000)	
	Population		Number of Deaths			
	Males	Females	Males	Females	Males	Females
0 – 14	2,110	2,010	30	27	99	93
5 – 14	3,340	3,230	6	8	109	102
15 – 34	7,320	7,310	16	20	177	180
35 – 59	7,960	8,750	70	57	121	122
60 & over	3,240	4,280	196	230	34	41

(8+12 = 20 Marks)

Q No. 8 a) The following data were computed from personal records of a manufacturing firm
X: number of years of service
Y: weekly wage rate

$n = 23, \sum X = 2433, \sum X^2 = 281019, \sum Y = 4245,$

$\sum Y^2 = 841786$ and $\sum XY = 482788.$

i. Fit a least squares regression line $Y = \alpha + \beta X + \varepsilon$

ii. Test the hypothesis $H_0 : \beta = 0.$

b) The price of rice (X) and price of wheat (Y) at 243 shops are recorded with the results:

$\sum X = 5442.2, \sum X^2 = 122155.04, \sum Y = 4019.6,$

$\sum Y^2 = 66588.92$ and $\sum XY = 90113.83.$

i. Test the hypothesis $H_0 : \rho = 0.$

ii. Calculate 95% confidence interval for the true correlation coefficient between X and Y.

(10+10 = 20 Marks)