



**FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION - 2016
FOR RECRUITMENT TO POSTS IN BS-17
UNDER THE FEDERAL GOVERNMENT**

Roll Number

STATISTICS

TIME ALLOWED: THREE HOURS	PART-I (MCQS)	MAXIMUM MARKS = 20
PART-I(MCQS): MAXIMUM 30 MINUTES	PART-II	MAXIMUM MARKS = 80
NOTE: (i) Part-II is to be attempted on the separate Answer Book.		
(ii) Attempt ONLY FOUR questions from PART-II by selecting TWO questions from EACH SECTION. ALL questions carry EQUAL marks.		
(iii) All the parts (if any) of each Question must be attempted at one place instead of at different places.		
(iv) Candidate must write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper.		
(v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed.		
(vi) Extra attempt of any question or any part of the attempted question will not be considered.		
(vii) Use of Calculator is allowed.		

PART-II
SECTION-I

- Q. No. 2.** (a) Describe the importance of skewness and kurtosis. Also give the major characteristics of a frequency distribution. (8)
- (b) The fourth mean moment of a symmetrical distribution is 243. What would be the value of the standard deviation in order that the distribution may be mesokurtic? (6)
- (c) Find the arithmetic, geometric and harmonic means of the series: (6)
- 1, 2, 4, 8, 16,..., 2^n .

- Q. No. 3.** (a) Describe the importance of hypergeometric probability distributions. In reality where we can use it? (8)
- (b) If N becomes indefinitely large, the hypergeometric probability distribution tends to the binomial probability distribution. Prove it. (6)
- (c) A continuous r.v. X has the p.d.f., given by (6)
- $f(x) = w(2-x)(x-5)$, $2 \leq x \leq 5$, $= 0$, otherwise
- Find the value of w , mean and variance. What are the values of the mode and median of the distributions of X ?

- Q. No. 4.** (a) What are the assumptions made in a linear regression. Give the important uses of regression analysis? (8)
- (b) Differentiate between correlation and regression with two real life examples. (6)
- (c) The following measurements show the respective heights in inches of ten fathers (X) and their eldest sons (Y). (6)

X	67	63	66	71	69	65	62	70	61	72
Y	68	66	65	70	69	67	64	71	60	63

- (i) Find the regression line of Y on X .
- (ii) Estimate Y for the given X as 70 inches.
- (iii) Test the significance of the population regression coefficient β_{yx} .
- (iv) Calculate 95% confidence limits for β_{yx} .
- (v) Test the significance of the intercept of the lines Y on X .
- (vi) Find the regression line of X on Y .
- Q. No. 5.** (a) What are the steps involved in any nonparametric test of hypothesis. (8)
- (b) Describe briefly Kolmogorov-Smirnov test of goodness of fit in case of one and two samples. (6)
- (c) Following is a sequence of heads (H) and tails (T) in tossing of a coin 14 times, $HTTHHHTHTTHHTH$. Test whether the heads and tails occur in random order, [Given: For $\alpha = 5\%$, $r_L = 2$, $r_u = 12$]. (6)

STATISTICS

SECTION-II

- Q. No. 6.** (a) Describe the importance of sampling. Also discuss sampling with and without replacement. (8)
- (b) What steps should keep in mind to determine the sample size? (6)
- (c) A population consists of $N = 7$ numbers, 1, 1, 2, 3, 4, 4, 5. (6)

Draw all possible sample of size $n = 3$ without replacement from this population and find the sample properties of odd numbers in the samples. Construct the sampling distribution of sample proportion and verify:

$$(i) \quad E(p) = P \text{ and } (ii) \quad E(p) = \left(\frac{N-n}{N-1} \right) \frac{PQ}{n}.$$

- Q. No. 7.** (a) Differentiate between simple and composite hypothesis. Also discuss the role of "Test of significance". (8)
- (b) In a random sample of 500 men from Lahore city, 300 are found to be smokers. In one of 1000 men from Karachi city, 550 are smokers. Do the data indicate that the two cities are significantly different with respect to the prevalence of smoking among men? (6)
- (c) Test the hypothesis given in the following table that the variances of three populations are equal at $\alpha = 5\%$. (6)

Sample 1	4, 7, 6, 6
Sample 2	5, 1, 3, 5, 3, 4
Sample 3	3, 8, 6, 8, 9, 5

- Q. No. 8.** (a) What is meant by Analysis of Variance and degrees of freedom? What are the assumptions underlying a one-way analysis of variance? (8)
- (b) Given the following information: (6)

	Samples			
	1	2	3	4
Observations (n_i)	4	6	7	3
Sample means ($\bar{y}_i$)	58	57	43	42
Estimate of variance (s_i^2)	10	30.4	5.67	9

Construct an analysis of variance table and test the hypothesis that the population means are equal at $\alpha = 5\%$.

- (c) Describe the role of Pakistan Bureau of Statistics and NADRA. (6)



FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION - 2017
FOR RECRUITMENT TO POSTS IN BS-17
UNDER THE FEDERAL GOVERNMENT

Roll Number

STATISTICS

TIME ALLOWED: THREE HOURS	PART-I (MCQS)	MAXIMUM MARKS = 20
PART-I(MCQS): MAXIMUM 30 MINUTES	PART-II	MAXIMUM MARKS = 80
NOTE: (i) Part-II is to be attempted on the separate Answer Book. (ii) Attempt ONLY FOUR questions from PART-II by selecting TWO questions from EACH SECTION. ALL questions carry EQUAL marks. (iii) All the parts (if any) of each Question must be attempted at one place instead of at different places. (iv) Candidate must write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper. (v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed. (vi) Extra attempt of any question or any part of the attempted question will not be considered. (vii) Use of Calculator is allowed.		

PART-II
SECTION-I

- Q. No. 2.** (a) Suppose a set of observations has mean $\bar{X}$ and variance s^2 . What happens to the mean and variance if each score is divided by s and then $\bar{X}$ is subtracted from each quotient? Is the result same if you first subtract $\bar{X}$ from each score and then divide the difference by s ? (8)

- (b) Calculate (i) S. D (ii) Co-efficient of variation (12)
(iii) Co-efficient of Skewness and discuss the result from the following weekly wages of 300 workers.

Weekly wages	Below 5	Below 10	Below 15	Below 20	Below 25	Below 30	Below 35
No. of workers	7	78	120	189	205	275	300

- Q. No. 3.** (a) If the probability that an individual suffers a bad reaction of a given serum is 0.001, determine the probability that out of 2000 individuals (a) exactly 3 and (b) more than 2 individuals will suffer a bad reaction. Find the answer using both poisson and the binomial distributions. (8)

- (b) A random variable to assume the value 1 with probability p , with probability $q = 1 - p$. Prove that (a) $E(X) = p$ and (b) $E[(X - \bar{X})^2] = pq$. (6)

- (c) An experiment consists of drawing three cards in succession from a well-shuffled ordinary deck of cards. Let E_1 be the event "king" on the first draw, E_2 be the event "king" on the second draw, E_3 be the event "king" on the third draw. State in words the meaning of each of the following: (6)

- (i) $\text{pr}\{E_1 \bar{E}_2\}$ (ii) $\bar{E}_1 + \bar{E}_2$ (iii) $\bar{E}_1 \bar{E}_2 \bar{E}_3$
(iv) $\text{pr}\{E_1 + E_2\}$ (v) $\text{pr}\{\bar{E}_3 | E_1 \bar{E}_2\}$ (vi) $\text{pr}\{E_1 E_2 + \bar{E}_2 E_3\}$

- Q. No. 4.** (a) What is regression model? Write the assumptions made in a linear regression. Also proof the Comment on the unbiasedness of regression estimates. (8)

- (b) Calculate coefficient of correlation by the method of least squares for the following paired values of X and Y variables. Also verify that this value of r is same as that obtained by pearson's formula. (12)

X	10	12	13	17	18
Y	5	6	7	9	13

SECTION-II

- Q. No. 5.

(a)

Discuss time reversal test and explain the factor reversal test.

(8)

(b)

Per capita income of a person from 1980-81 to 1986-87 and the consumer price index with 1980-81 were as follows:

(12)

Year	Income per capita (Rs)	Index Nos.
1980-81	1627	100
1981-82	1851	103.5
1982-83	1993	103.4
1983-84	2290	109.4
1984-85	2494	110.9
1985-86	2735	113.8
1986-87	2970	115.6

Find real wages and real income indices.

- Q. No.6.

(a)

What is the difference between population census and vital statistics? What are the various uses of vital statistics for a country?

(8)

(b)

The population and its distribution by sex and number of births in a tehsil in 1991 and survival rates are given below:

(12)

Group	Population	Males	Females	Male birth	Female birth	Total birth	Survival rate
15-19	11832	6145	5687	65	60	125	0.91
20-24	10538	5214	5324	144	132	276	0.90
25-29	9375	4655	4720	135	127	262	0.84
30-34	7843	3910	3933	82	81	163	0.87
35-39	7270	3600	3670	62	56	118	0.85
40-44	6315	3290	3025	12	15	27	0.83
45-49	5394	2793	2601	3	3	6	0.82
Total	58567	29607	28960	503	474	977	

From the given data, calculate

- (i) General fertility rate

(ii) Age specific fertility rate

(iii) Total fertility rate

(iv) Gross production rate

(v) Net reproduction rate

Assuming no mortality.

- Q. No. 7.

(a)

Given the following data obtained from a completely randomized design with four treatments; analyse the given data and draw conclusion about the equality of treatment effects.

(8)

Treatments			
T ₁	T ₂	T ₃	T ₄
20.9	23.7	13.2	5.8
12.4			6.1
10.1			4.8
4.2			1.5

- (b)

Give Statistical model for completely randomized design with one observation per unit.

(6)
- (c)

What is the role of randomization in the process of the experimentation? Discuss and define experimental error. What factors are responsible for determining the number of replications?

(6)

- Q.No.8.

(a)

A random sample of 100 workers in a farm took an average of 14 minutes to complete a task. A random sample of 150 workers in another large farm took an average of 11 minutes to complete the task. Can it be assumed at 1% level of significance that the average time taken by the workers in the two farms is same, if the S.D. of the workers of first farm and second farm are 2 minutes and 3 minutes respectively?

(8)
- (b)

Describe the general procedure for testing a hypothesis about a difference between population mean, when sample size is large.

(6)
- (c)

Distinguish between:

(6)

(i) Sampling error and non-sampling error

(ii) Sampling with replacement and without replacement

(iii) Probability and non-probability Sampling



**FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION-2018
FOR RECRUITMENT TO POSTS IN BS-17
UNDER THE FEDERAL GOVERNMENT**

Roll Number

STATISTICS

TIME ALLOWED: THREE HOURS	PART-I (MCQS)	MAXIMUM MARKS = 20
PART-I(MCQS): MAXIMUM 30 MINUTES	PART-II	MAXIMUM MARKS = 80
NOTE: (i) Part-II is to be attempted on the separate Answer Book.		
(ii) Attempt ONLY FOUR questions from PART-II by selecting TWO questions from EACH SECTION. ALL questions carry EQUAL marks.		
(iii) All the parts (if any) of each Question must be attempted at one place instead of at different places.		
(iv) Candidate must write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper.		
(v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed.		
(vi) Extra attempt of any question or any part of the attempted question will not be considered.		
(vii) Use of Calculator is allowed.		
(viii) Use of statistical table is allowed.		

**PART-II
SECTION – A**

- Q. 2. (a)** The mean and standard deviation of a sample of 20 observations were found to be 75 and 2.5 respectively. On checking the original figures, it was discovered that one observation which was actually 68, was copied down as 86. Find the correct mean and standard deviation. (5)
- (b)** Given below is the distribution of monthly income (to the nearest rupee) of 100 employees working in each of two banks A and B. Compare the variations in salaries of employees by calculating coefficient of variation in each case? (9)
- | Income Rs. 000 | 35 – 49 | 50 – 64 | 65 – 79 | 80 – 94 | 95 – 109 | 110 – 124 |
|----------------|---------|---------|---------|---------|----------|-----------|
| f_A | 10 | 15 | 35 | 20 | 10 | 10 |
| f_B | 15 | 20 | 40 | 15 | 7 | 3 |
- (c)** What can you say of the skewness in each of the following cases? (6) (20)
- (i) Mean = Rs. 39000 and Mode = Rs. 45000
- (ii) The median is 79.39, which the two quartiles are 50.25 and 95.00.
- (iii) Mean = Rs. 39000 and Median = Rs. 38500
- Q. 3. (a)** A box contains 12 red and 20 white rose flowers. Flowers are picked up at random one by one without replacement. What is the probability that: (6)
- (i) The third one is red given that the first 2 are white?
- (ii) There are 3 red and 2 white flowers in the first five picked up?
- (b)** Two hundred and fifty passengers have made reservations for an airplane flight. If the probability that a passenger who has a reservation will not show up is 0.005, what is the probability that exactly five will not show up? (8)
- (c)** A soft drink machine is regulated so that it discharges an average of 250 milliliters per glass. If the amount of drink is normally distributed with a standard deviation equal to 15 millimeters, then: (6) (20)
- (i) What fraction of the glasses will contain more than 280 milliliters?
- (ii) What is the probability that a glass contains between 240 milliliters to 260 milliliters.
- Q. 4. (a)** What is a multiple regression? Explain the basic differences between simple regression and multiple regression. (5)
- (b)** Differentiate between multiple correlation and partial correlation. (5)
- (c)** Find the coefficient of correlation between persons employed and cloth manufactured in a textile mill. Interpret the result (10) (20)
- | | | | | | | | | |
|---------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| Persons employed | 139 | 219 | 140 | 213 | 180 | 215 | 249 | 280 |
| Cloth Manufactured (000 meters) | 28 | 55 | 26 | 49 | 42 | 54 | 51 | 60 |

- Q. 5. (a) What is meant by an experimental design? Describe in brief the basic principles of experimental designs. (8)
- (b) A consumer agency wanted to find out if the mean time it takes for each of three brands of medicines to provide relief from a headache is the same. The 1st drug was administered to six randomly selected patients, the 2nd to four randomly selected patients and the 3rd to five randomly selected patients. The following table gives the time (in minutes) taken by each patient to get relief from a headache after taking the medicine. (12) (20)

Drug 1	Drug 2	Drug 3
25	15	44
38	21	39
42	19	54
65	25	58
47		73
52		

At 5% significance level, will you conclude that mean time taken to provide relief from a headache is the same for each of the three drugs?

- Q. 6. (a) The management of a super market wanted to investigate if the percentages of men and women who prefer to buy local brand products over the imported products are different. A sample of 900 men shoppers at the company’s super markets showed 400 of them prefer to buy local brand products over imported brand. Another sample of 1200 women shoppers showed that 420 of them prefer to buy local brand products over the imported brand. (6)
- (i) Construct a 99% confidence interval for the difference between the two proportions.
- (ii) At 5% significance level, can you conclude that the difference between all male customers at this market is less than that by all female customers?
- (b) An auto manufacturing company wants to estimate the variance of kilometers per litre for its one of the auto model. A random sample of 25 cars of this model showed that the variance of kilometres per litre for all such cars are approximately normal. Test at the 5% significance level whether the sample result indicates that the population variance is different from 1.05. (7)
- (c) A researcher wanted to study the relationship between gender and owning expensive cell phones. A sample of 1000 young people were taken and information given are in the following table. (7) (20)

Status cell phones	Male	Female	Total
Own	300	400	700
Do not own	200	100	300
Total	500	500	1000

At the 5% level of significance, can we conclude that gender and owning an expensive cell phone are related for all young people?

- Q. 7. (a) Define and distinguish between: (6)
- (i) Target and Sampled Populations
- (ii) Simple Random Sampling and Stratified Random Sampling
- (iii) Sample Distribution and Sampling Distribution
- (b) Explain with examples the properties of a good point estimator. (6)
- (c) The heights of a large number of shrubs of the same kind produced for sale by a horticultural nursery are normally distributed with mean 1.14 m and standard deviation 0.25 m. Fifty samples, each consisting of 100 shrubs are selected. In how many of these samples would you expect to have the mean sample being to be (i) greater than 1.16 m; (ii) between 1.13 m and 1.18 m? (8) (20)

- Q. 8. (a) What are the various statistical organizations in Pakistan? Discuss the main functions of Statistics Division and NADRA. (10)
- (b) Calculate age-specific fertility rates, total fertility rate, gross-reproduction rate and net reproduction rate from the following data, assuming sex-ratio at birth to be 106.18 per cent (10) (20)

Age-group (years)	Female Population (000)	Registered Births	Probability of Survival
15 – 19	1,424	27,639	0.9645
20 – 24	1,531	226,817	0.9607
25 – 29	1,653	280,506	0.9554
30 – 34	1,658	194,526	0.9489
35 – 39	1,741	113,966	0.9416
40 – 44	1,669	32,363	0.9324
45 – 49	1,561	2,215	0.9201



FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION-2019
FOR RECRUITMENT TO POSTS IN BS-17
UNDER THE FEDERAL GOVERNMENT

Roll Number

STATISTICS

TIME ALLOWED: THREE HOURS	PART-I (MCQS)	MAXIMUM MARKS = 20
PART-I(MCQS): MAXIMUM 30 MINUTES	PART-II	MAXIMUM MARKS = 80
NOTE: (i) Part-II is to be attempted on the separate Answer Book. (ii) Attempt ONLY FOUR questions from PART-II by selecting TWO questions from EACH SECTION. ALL questions carry EQUAL marks. (iii) All the parts (if any) of each Question must be attempted at one place instead of at different places. (iv) Write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper. (v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed. (vi) Extra attempt of any question or any part of the question will not be considered. (vii) Use of Calculator is allowed.		

PART – II
SECTION – I

- Q. 2. (a)** What is meant by a frequency distribution? Describe briefly the main steps in the preparation of a frequency table from raw data. (6)
- (b)** A man travels from A to B at average speed of 30 miles per hour and returns from B to A along the same route at an average speed of 60 miles per hour. Find the average speed of the entire journey. (6)
- (c)** Define mean-deviation and its co-efficient. Discuss its advantages and uses. (8) **(20)**
Estimate the mean deviation from the arithmetic mean of the following set of examination marks.

Marks	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79
No. of students	2	3	8	24	27	40	11	5

- Q. 3. (a)** Define mutually exclusive events. State and prove the theorem of addition of probabilities concerning mutually exclusive events. (6)
- (b)** Show that the multiplication law $P(A \cap B) = P(A/B)P(B)$, established for two events, may be generalized to three events as follows; (6)
$$P(A \cap B \cap C) = P(A/B \cap C) P(B/C) P(C)$$
- (c)** There are three coins, identical in appearance, one of which is ideal and the other two biased with probabilities $1/3$ and $2/3$ respectively for a head. One coin is taken at random and tossed twice. If a head appears both the times, what is the probability that the ideal coin was chosen? (8) **(20)**

- Q. 4. (a)** (i) Explain briefly how the principle of least squares is used to find a regression line based on a sample of size n . Illustrate on a rough sketch the distance whose squares are minimized, taking care to distinguish the dependent and independent variables. (6)
- (ii) Find the least square estimates of parameters in a simple linear regression model $Y_i = \alpha + \beta X_i + e_i$ where e_i 's are distributed independently with mean zero and constant variance.
- (iii) What are the properties of least square regression line?
- (b)** The following means, standard deviations and correlations are found for (6)
 X_1 = Seed-hay crops in owts. Per acre
 X_2 = Spring rainfall in inches
 X_3 = Accumulated temperature above 42° F in spring in a certain district in England during 20 years.

$$\begin{aligned}\bar{X}_1 &= 28.02, & S_1 &= 4.42, & r_{12} &= 0.80, \\ \bar{X}_2 &= 4.91, & S_2 &= 1.10, & r_{13} &= -0.40, \\ \bar{X}_3 &= 594, & S_3 &= 85, & r_{12} &= -0.56,\end{aligned}$$

Find the partial correlation and the regression equation for hay-crop on spring rainfall and accumulated temperature.

STATISTICS

- (c) What do you understand by nonparametric tests? Why such tests are also called distribution-free tests? Give the advantages and disadvantages of nonparametric tests over parametric tests. Describe the Wilcoxon signed-rank test for one sample. How does it differ from the sign test? (8) (20)

SECTION-II

- Q. 5. (a) Explain what you understand by the probability sampling and non probability sampling. What are their relative advantages and disadvantages? (6)
- (b) What is a sampling distribution? Describe the properties of the sampling distribution of the means. (6)
- (c) A finite population consists of the numbers 2, 4 and 6. Form a sampling distribution of sample mean, when random samples of size 4 is drawn with replacement. Also verify its properties. (8) (20)
- Q. 6. (a) Under what condition is the sampling distribution of $\frac{s_1^2}{s_2^2}$ an F-distribution? Explain the relationship between the F and t distributions, between the F and Chi-Square distributions. (6)
- (b) The proportion of families buying milk from company A in a certain city is believed to be $p=0.6$. If a random sample of 10 families shows that 3 or less buy milk from company A, we shall reject the hypothesis that $p=0.6$ in favour of the alternative $p<0.6$. Evaluate α if $p=0.6$, evaluate β for the alternatives $p=0.3$, $p=0.4$ and $p=0.5$. (6)
- (c) Define a Chi-square random variable and its density function. Discuss the important properties of Chi-square distribution. Show that the Chi-square distribution tends to normal distribution for large degrees of freedom. (8) (20)
- Q. 7. (a) Describe the Randomized Complete Block Design, its model and analysis. What are its advantages and disadvantages? (6)
- (b) Compare Randomized Complete Block experiments with Completely Randomized experiments, comparing their respective advantages and relative efficiency, with illustrations. (6)
- (c) Three varieties A, B and C of a crop are tested in a randomized block design with four replications, the layout being given below. The plot yields in pounds are also indicated therein. Analyze the experimental yields and state your conclusions. (8) (20)

Replications	1	A 32.1	C 34.2	B 31.7
	2	C 30.7	A 17.0	B 32.7
	3	A 40.8	B 25.3	C 48.2
	4	B 47.9	C 59.6	A 26.8

- Q. 8. (a) Define gross and net production rates. Explain how would you compute the net production rate and what interpretations can be made if it is 1, less than 1 or greater than 1. (6)
- (b) Explain with suitable illustrations the object of standardizing various vital statistics relating to births, deaths and marriages. (6)
- (c) Compute the gross and net reproduction rates for the following data: (8) (20)

Age-group (years)	Female Population (000)	Female births	Probability of survival
15-19	1558	18900	0.914
20-24	1112	71100	0.899
25-29	1595	96900	0.884
30-34	1629	64200	0.868
35-39	1627	34900	0.852
40-44	1522	10800	0.834
45-49	1401	800	0.813



FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION-2020
FOR RECRUITMENT TO POSTS IN BS-17
UNDER THE FEDERAL GOVERNMENT

Roll Number

STATISTICS

TIME ALLOWED: THREE HOURS	PART-I (MCQS)	MAXIMUM MARKS = 20
PART-I(MCQS): MAXIMUM 30 MINUTES	PART-II	MAXIMUM MARKS = 80
NOTE: (i) Part-II is to be attempted on the separate Answer Book. (ii) Attempt ONLY FOUR questions from PART-II by selecting TWO questions from EACH SECTION. ALL questions carry EQUAL marks. (iii) All the parts (if any) of each Question must be attempted at one place instead of at different places. (iv) Candidate must write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper. (v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed. (vi) Extra attempt of any question or any part of the attempted question will not be considered. (vii) Use of Calculator is allowed. (viii) Use of statistical table is allowed.		

PART-II
SECTION – A

- Q. 2. (a)** Two bags A and B contain red and blue marbles. Bag A contains 7 red and 8 blue marbles. Bag B contains 9 red and 7 blue marbles. One bag is selected randomly and one marble is drawn. If the drawn marble is red then what is the probability that this drawn marble is from bag A? (10)
- (b)** For the given set of observations showing weekly sale of a specific type of refrigerators. (10) (20)

35, 56, 43, 21, 43, 56, 78, 12, 56, 47, 76, 23, 52

- (i) Find mean and standard deviation.
(ii) Find and describe $\bar{x} \pm 2(sd)$

- Q. 3.** Raw material used in the production of a synthetic fiber is stored in a place that has no humidity control. Measurements of the relative humidity (y) and moisture control (x), on ten days, are given below. Fit a Least Square model: (20)
- $y = \beta_1 + \beta_2 x$. Further find and explain coefficient of determination.

Humidity	46	53	37	42	34	29	60	48	41	48
Moisture	12	14	11	13	10	8	17	12	10	15

- Q. 4.** Three teaching methods were implemented to a homogenous group of school level students. Groups of students, selected randomly, were taught with a particular method and their scores were recorded as given below: (20)

Method A 94, 88, 91, 74, 87, 97
Method B 85, 82, 79, 84, 61, 72, 80
Method C 89, 67, 72, 76, 69

Use Kruskal-Wallis Test, at 5% level of significance to test the hypothesis that, on average there is no significant difference between the average score of these teaching methods.

SECTION-B

- Q. 5. (a)** Draw all possible samples of size 3, without replacement, from the population, 6, 12, 3, 9, 15, and 21. Find sample means and prove the following relationships, using usual notations. (15)

(i) $E(\bar{x}) = \mu$ and (ii) $V(\bar{x}) = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)$

- (b)** Define Stratified Random sampling method, identify situations where this type of sampling is beneficial. Give an example. (5) (20)

STATISTICS

- Q. 6. (a) To compare the effectiveness of two medicines M1 and M2, for headache, a study was conducted. Samples from a homogeneous group of headache patients were selected randomly and administered M1 (six patients) and M2 (8 patients) selected randomly. Recovery times (in minutes) of Patients were recorded as follows: (15)
- Medicine M1

12 9 8 11 10 9

Medicine M2

5 11 7 6 8 6 5 4
- Could it be concluded at 5% level of significance that, on average, medicine M2 is better than M1?
- (b) While testing hypothesis one may commit errors when we make decisions. State and explain such errors, supporting by real life example. (5) (20)
- Q.7. (a) The following Latin Square layout displays the scores secured by nine college students. Students are of different ethnic background and various professional interests. (15)

Professional Interest	Ethnic Background		
	X	Y	Z
Law	A 75	B 86	C 69
Medicine	B 95	C 79	A 86
Engineering	C 70	A 83	B 93

In this table A, B, and C are the three instructors. Analyze and test following hypotheses, Use $\alpha = 0.05$.

- (i) Having a different instructor has no effect on the scores.

(ii) Differences in ethnic backgrounds have no effect on the scores

(iii) Differences in professional interest have no effect on the scores.
- (b) Describe the role of Multiple Comparison tests in Analysis of Variance. Name few Multiple Comparison tests and explain one method. (5) (20)
- Q.8. (a) Explain the terms Demography and Vital statistics. List few sources of demographic data both locally and globally. (5)
- (b) Using the information given in the following table. Calculate Total Fertility Rate (TFR) and Gross Fertility Rate (GFR). (10)

Age (years)	Women Population	No of births to women
15-19	84790	343
20-24	70010	14541
25-29	72660	16736
30-34	75920	12218
35-39	75100	756
40-44	71620	82
45-49	66660	45

- (c) Differentiate between Rates and Ratios. Explain Crude Death Rate and Specific Death Rate. (5) (20)



FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION-2021 FOR RECRUITMENT
TO POSTS IN BS-17 UNDER THE FEDERAL GOVERNMENT

Roll Number

STATISTICS

TIME ALLOWED: THREE HOURS	PART-I (MCQS)	MAXIMUM MARKS = 20
PART-I(MCQS): MAXIMUM 30 MINUTES	PART-II	MAXIMUM MARKS = 80
NOTE: (i) Part-II is to be attempted on the separate Answer Book. (ii) Attempt ONLY FOUR questions from PART-II by selecting TWO questions from EACH SECTION. ALL questions carry EQUAL marks. (iii) All the parts (if any) of each Question must be attempted at one place instead of at different places. (iv) Candidate must write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper. (v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed. (vi) Extra attempt of any question or any part of the attempted question will not be considered. (vii) Use of Calculator is allowed. (viii) Use of statistical table is allowed.		

PART-II
SECTION – A

- Q.2. (a)** Describe the measures of central tendency. What is the purpose served by it? What are its desirable qualities? (08)
- (b)** A man gets rise of 10% in salary at the end of his first year of service and further rises of 20% and 25% respectively at the end of second and third years. The rise in each year being calculated on his salary at the beginning of the year. What annual percentage increases in this equivalent? (06)
- The following has been obtained from two different frequency distributions of
- (c)** Weights (lb) having 125 and 200 observations after making the substitution. (06) (20)
- $$X = 16 + 5u, Y = 20 + 2v,$$
- (i) $\sum fu = -46, \sum fu^2 = 306, \sum fu^3 = -242, \sum fu^4 = 1962.$
- (ii) $\sum fv = 21, \sum fv^2 = 1265, \sum fv^3 = -627, \sum fv^4 = 14169.$
- Find (i) Which of the distributions is more consistent? (ii) Which of the distribution is negatively skewed? (iii) Which of the distribution is mesokurtic?
- Q.3. (a)** Describe the linear regression model. Explain the assumptions underlying the linear regression model. (08)
- (b)** An instructor of mathematics wished to determine the relationship of grades on a final examination to grades on two quizzes given during the semester. Calling X_1, X_2 and X_3 the grades of a student on the first quiz, second quiz and final examination respectively. He made the following computations for a total of 120 students. (06)
- $$\bar{X}_1 = 6.8, S_1 = 1.0, r_{12} = 0.60, \bar{X}_2 = 7.0, S_2 = 0.8, r_{13} = 0.70, \bar{X}_3 = 74, S_3 = 9.0, r_{23} = 0.65.$$
- (i) Find the least-squares regression equation of X_3 on X_1 and X_2 .
- (ii) Estimate the final grades of two students who scored respectively (I) 9 and 7, (II) 4 and 8, on the two quizzes. (iii) Compute $R_{3.12}$.
- (c)** Why non-parametric tests are important? Describe the run test in detail and give its advantages and disadvantages. (06) (20)
- Q.4. (a)** Describe the random experiment and give its properties. Also discuss the mutually exclusive events with two real life examples. (08)
- (b)** An urn contains four balls which are known to be either: (i) all white or (ii) two white and two black. A ball is drawn at random and is found to be white. What is the probability that all the balls are white? (06)
- (c)** Describe the hypergeometric distribution and give its properties. In real life under what situations we can use it? (06) (20)

STATISTICS

SECTION-B

- Q.5. (a) Describe the probability and non-probability sampling methods. Also differentiate between multiphase and multistage sampling. (08)
- (b) A sample of size $n=3$ is selected by using without replacement sampling from a population having $N=5$ units whose values are 0, 3, 6, 9 and 12. (i) Find the sampling distribution of the sample mean $(\bar{X})$. (ii) Calculate the mean and standard deviation of $\bar{X}$, and verify that
$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right).$$
 (06)
- (c) A population consists of 5 observations 1, 2, 3, 4, 5. Draw all possible samples of size 2 without replacement. Find the mean of the sampling distribution of the variances. Compare it with the variance of the population. (06) (20)
- Q.6. (a) Differentiate between simple and composite hypothesis. Also describe the importance of type-I and type-II errors. (08)
- (b) In a poll of college students in a large state university, 300 of 400 students living in dormitories approved a certain course of action, whereas 200 of 300 students not living in dormitories approved it. Estimate the difference in proportions favoring the course and compare 95% confidence interval for it. (06)
- (c) A manufacturer of house-dresses sent out advertising by mail. He sent samples of material to each of 2 groups of 1000 women. For one group, he enclosed a white return envelope and for the other group, a blue envelope. He received orders from 10% to 30% respectively. Do the data indicates that the colour of the envelope has an effect on the sales? Use 5% level of significance. (06) (20)

- Q.7. (a) Describe the point estimation. Also discuss the properties of good point estimator with two real life examples. (08)
- (b) Describe the chi-square test for goodness-of-fit. Also discuss its procedure for analysis. (06)
- (c) The following table shows the number of male and female births in 800 families having 4 children. (06) (20)

Number of male births	0	1	2	3	4
Families	32	178	290	236	64

Test whether the data are consistent with the hypothesis that the binomial law holds and that the chance of a male birth is equal to that of a female birth i.e. $p = q = 1/2$.

- Q.8. (a) Describe the importance of vital statistics with its uses and shortcomings. (08)
- (b) From the following data, find the gross and net-reproduction rates by assuming sex-ratio at birth to be 105.2%. (06)

Age group (Years)	Female Population	Registered births	Survival among females out of 1000
15-19	8981	1835	634
20-24	5875	2616	602
25-29	3613	2563	568
30-34	3380	1062	530
35-39	3345	558	488
40-44	3248	37	444

- (c) Describe the “Nation Income”, “Gross and Domestic Product” and “Saving and Wealth” in terms of country’s economy. (06) (20)



FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION-2022 FOR RECRUITMENT
TO POSTS IN BS-17 UNDER THE FEDERAL GOVERNMENT

Roll Number

STATISTICS

TIME ALLOWED: THREE HOURS

PART-I(MCQS): MAXIMUM 30 MINUTES

PART-I (MCQS)

PART-II

MAXIMUM MARKS = 20

MAXIMUM MARKS = 80

- NOTE: (i) Part-II is to be attempted on the separate Answer Book.**
(ii) Attempt ONLY FOUR questions from PART-II by selecting TWO questions from EACH SECTION. ALL questions carry EQUAL marks.
(iii) All the parts (if any) of each Question must be attempted at one place instead of at different places.
(iv) Candidate must write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper.
(v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed.
(vi) Extra attempt of any question or any part of the attempted question will not be considered.
(vii) Use of Calculator is allowed.
(viii) Use of statistical table is allowed.

PART-II
SECTION – A

Q. 2. (a) What can you say of the skewness in each of the following cases? (09)

- i) The median is 60 while the two quartiles are 40 and 80.
ii) Mean = 140 and Mode = 140.

The first three moments about 16 are respectively -0.35, 2.09 and -1.93.

- (b) Discuss the various measures or quantities by which the characteristics of frequency distributions are measured and compared. (06)**
(c) Differentiate between descriptive and inferential statistics. (05) (20)

Q. 3. The following data give information on the ages (in years) and the number of breakdowns during the past month for a sample of seven machines at a large company. (20)

Age	12	7	2	8	13	9	4
Number of breakdowns	10	5	1	4	12	7	2

- i) Taking age as independent variable and the number of breakdowns as a dependent variable what is your hypothesis about the sign of B in the regression line.
ii) Compute SS_{xx} , SS_{yy} , and SS_{xy} .
iii) Find the least squares regression line by choosing appropriate dependent and independent variables based on your answer in part i.
iv) Calculate r and r^2 and explain what they mean.
v) Predict the number of breakdowns for age equal to 10 years.
vi) Compute the standard deviation of errors.
vii) Construct a 90% confidence interval for B.
viii) Test the hypothesis $H_0 : \beta = 0$ against $\beta \neq 0$.

Q. 4. (a) A company has 800 employees. Twenty percent of the employees have college degrees but half of these people are in non-management positions. Thirty percent of the non-degree people are in management positions. (10)

- i) How many managers does the company have?
ii) What is the conditional probability of being a college graduate is a manager?
iii) Find the probability of randomly selecting someone who is either a college graduate or a manager.

STATISTICS

- (b) A soft drink machine is regulated so that it discharges an average of 200 milliliters per cup. If the amount of drink is normally distributed with a standard deviation equal to 15 millimeters. (10) (20)
- i) What is the probability that a cup contains between 191 and 209 milliliters?
 - ii) How many cups will likely overflow if 230 milliliters cups are used for the next 1000 drinks?
 - iii) Below what value do we get the smallest 25% of the drinks?

SECTION-B

- Q. 5. (a) A study of time spent on housework found that men who are employed spend an average of 8.2 hours per week doing housework (Americans’ Use of Time Project, University of Maryland, American Demographics, November 1998). Assume that the amount of time spend on housework per week by all employed men in the United States is normally distributed with a mean of 8.2 hours and a standard deviation of 2.1 hours. (10)
- i) What is the probability that a randomly selected employed man spends 9 or more hours per week on housework?
 - ii) What is the probability that a randomly selected employed man spends 7 hours or less on housework per week?
- (b) A population consists of $N = 6$ numbers 0, 1, 5, 6, 8 and 11. Draw all possible samples of size $n = 3$, without replacement, from the population and find the sample means. Construct the sampling distribution of sample means and verify that: (10) (20)

$$\mu_{\hat{x}} = \mu \text{ and } Var(\hat{X}) = \frac{\sigma^2}{n} \left(\frac{N - n}{N - 1} \right)$$

- Q. 6. (a) A candidate for mayor in a large city believes that he appeals more of the women voters than the men voters. He hires the services of a poll-taking organization, and they find that 62 out of 100 women interviewed support the candidate, and 69 out of 150 men support him. Calculate 99% confidence interval for the difference between two proportions. (10)
- (b) Explain the six steps of Tests of Hypothesis. (10) (20)

- Q. 7. (a) Given the following ANOVA for a RCB design (12)

S.V	d.f.	SS	MS
Blocks	8	2.1074	
treatments	3	2.1006	
Error	24	2.6249	

- i) Complete the analysis; fill in the mean squares
 - ii) Compute the standard error for a treatment and for the difference between two treatment means.
 - iii) Treatment means are 1.464, 1.195, 1.325 and 1.662. What mean or means do you suspect might represent different population by applying LSD test.
- (b) Define the relative efficiency. How would you calculate the relative efficiency of LS design relative to RCB design with rows as blocks? (08) (20)

- Q.8. (a) Construct the following weighted aggregative price index numbers for 1990 and 1991 from the given data: (12)
- i) Laspeyres’ index
 - ii) Paasche’s index
 - iii) Fisher’s “Ideal” index

Commodity	Prices (Rs. per 40 kg)			Quantities (tons)		
	1986 (base)	1990	1991	1986	1990	1991
A	64	75	80	270	276	290
B	40	45	41	124	118	144
C	18	21	20	130	121	137
D	58	68	56	185	267	355

- (b) Describe the importance of vital statistics. Also state its uses and limitations. (08) (20)



FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION-2023 FOR RECRUITMENT
TO POSTS IN BS-17 UNDER THE FEDERAL GOVERNMENT

Roll Number

STATISTICS

TIME ALLOWED: THREE HOURS PART-I(MCQS): MAXIMUM 30 MINUTES	PART-I (MCQS) PART-II	MAXIMUM MARKS = 20 MAXIMUM MARKS = 80
NOTE: (i) Part-II is to be attempted on the separate Answer Book. (ii) Attempt ONLY FOUR questions from PART-II by selecting TWO questions from EACH SECTION. ALL questions carry EQUAL marks. (iii) All the parts (if any) of each Question must be attempted at one place instead of at different places. (iv) Candidate must write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper. (v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed. (vi) Extra attempt of any question or any part of the attempted question will not be considered. (vii) Use of Calculator is allowed. (viii) Use of statistical table is allowed.		

PART-II
SECTION – A

- Q. 2. (a) Calculate the Q1, median and Q3 from the following distribution of weight of containers in Kg, (10)
and comment on the symmetry of distribution.

Weight (Kg)	18- 26	27- 35	36- 44	45- 53	54- 62	63- 71	72- 80
# of Containers	13	20	39	40	25	6	12

- (b) What is frequency distribution? Discuss briefly the steps involve in construction of frequency (05)
distribution.
- (c) The first three moments of a distribution about the value 2 of the variable are 1, 16 and -40. (05) (20)
Show that the mean is 3, the variance 15 and m_3 is -86. Also show that the first three moments
about $x=0$ are 3, 24 and 76. (5)

- Q. 3. (a) From the following data, determine the linear regression equations of X_1 on X_3 and of (10)
 X_2 on X_3 .

X_1	07	12	14	17	20
X_2	04	07	08	09	12
X_3	01	02	04	05	08

Find the deviations of observed values of X_1 from the regression, i.e., $X_{1.3}$. Repeat the same of X_2 , i.e., obtain $X_{2.3}$. Determine the simple correlation co-efficient between the two sets of deviations $X_{1.3}$ and $X_{2.3}$.

- (b) What is meant by: (05)
(i) Regression (ii) Regresand (iii) Regressor (iv) Regressor co-efficient
- (c) Describe the Properties of the correlation co-efficient? (05) (20)
- Q. 4. (a) Derive the Poisson distribution as the limiting form of the binomial distribution, stating clearly (08)
the assumptions you make.
- (b) Enumerate all the possible (i) combinations and (ii) permutations of 3 letters chosen from the (06)
four letters A, B, C, and D.
- (c) A box contains 4 bad and 6 good tubes. Two tubes are drawn together at random. One of them (06) (20)
is tested and found to be good. What is the probability that other one is good?

STATISTICS

SECTION-B

- Q. 5. (a) Draw all possible random sample of size $n_1=2$ with replacement from a finite population consisting of 4, 6, 8. Similarly draw all possible random samples $n_2=2$ with replacement from another finite population consisting of 1, 2, 3. (12)
- (i) Find the possible difference between the sample means of the two populations.
 - (ii) Construct the sampling distributions of $\bar{x}_1 - \bar{x}_2$ and compute its mean and variance.
 - (iii) Verify that $\mu_{\bar{x}_1 - \bar{x}_2} = \mu_1 - \mu_2$ and $\sigma_{\bar{x}_1 - \bar{x}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$
- (b) Explain sampling and non-sampling errors. What method would you suggest to control each type of error? (08) (20)

- Q. 6. (a) The two samples A and B detailed below, were taken from normal populations of standard deviation 0.8. Test whether the difference of means is significant (12)

A	10.5,	11.6,	12.7,	12.9,	13.5,	13.6,	14.8
B	11.3,	12.4,	12.4,	13.9,	14.2,	14.7,	14.9, 15.6

- (b) Explain with examples the difference between: (08) (20)
- (i) Null and Alternative hypothesis
 - (ii) Simple and composite hypothesis
 - (iii) Type I-error and Type II-error
 - (iv) Critical and non-critical region.

- Q. 7. (a) The following data represent the result of 3 questions obtained by 3 students in three subjects: (12)

Students	Subjects		
	English	Mathematics	Statistics
1	13	23	22
	18	20	23
	15	16	20
2	21	20	20
	16	14	15
	24	24	22
3	18	17	19
	15	13	21
	12	16	18

Perform an analysis of variance upon these data and test the hypothesis that:

- (i) The subjects are of equal difficulty.
 - (ii) The students are of equal ability, and
 - (iii) The students and subjects do not interact.
- (b) Discuss why using multiple two-sample t-tests is not an appropriate alternative of analysis of variance? (08) (20)
- Q.8. (a) Given the population 1, 1, 1, 3, 4, 5, 6, 6, 6, and 7. Find (10)
- (i) The probability that a random sample of size 36 selected with replacement will yield sample mean between 3.26 and 4.74
 - (ii) The mean and standard deviation for the sampling distribution of means for a sample size of 4 selected at random without replacement. Between what values would you expect at least $\frac{3}{4}$ of the sample mean to fall?
- (b) Explain sampling and non-sampling errors. What methods would you suggest to control each type of error? (05)
- (c) Explain with examples the following properties of a point estimator: (05) (20)
- (i) Unbiasedness, (ii) Consistency, and (iii) Efficiency.
- *****



FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION-2024 FOR RECRUITMENT
TO POSTS IN BS-17 UNDER THE FEDERAL GOVERNMENT
STATISTICS



TIME ALLOWED: THREE HOURS	(PART-I MCQs) MAXIMUM MARKS: 20
PART-I (MCQs) : MAXIMUM 30 MINUTES	(PART-II) MAXIMUM MARKS: 80
NOTE: (i) First attempt PART-I (MCQs) on separate OMR Answer Sheet which shall be taken back after 30 minutes. (ii) Overwriting/cutting of the options/answers will not be given credit. (iii) There is no negative marking. All MCQs must be attempted.	

PART-I (MCQs) (COMPULSORY)

Q.1. (i) Select the best option/answer and fill in the appropriate Box on the OMR Answer Sheet. (20x1=20)
(ii) Answers given anywhere else, other than OMR Answer Sheet, will not be considered.

- A study based on complete enumeration is known as:**
(A) Sample survey (B) Pilot survey (C) Census survey (D) None of these
- If the estimated value of an item is 50 and its actual value is 60, the relative error is:**
(A) -20 (B) 0.16 (C) 21.2 (D) 0.20
- Statistical results are:**
(A) Absolutely correct (B) Not true (C) True on average (D) Universally true
- The value 43,572.6 approximated to the thousandth place by adding figure is:**
(A) 43,500 (B) 43,000 (C) 44,000 (D) 44,600
- In a grouped data, the number of classes preferred are:**
(A) Minimum possible (B) Adequate (C) Maximum possible (D) Arbitrarily chosen number
- The headings of the rows given in the first column of a table are called:**
(A) Stubs (B) Captions (C) Titles (D) Prefatory notes
- With the help of ogive curve, one can determine:**
(A) Median (B) Deciles (C) Percentiles (D) All of these
- If a constant value is subtracted from each observation of a set, the mean of the set is:**
(A) Increased by 50 (B) Decreased by 50 (C) Is not affected (D) Zero
- Extreme value have no effect on:**
(A) Average (B) Median (C) Geometric mean (D) Harmonic mean
- In case of weighted mean, the accuracy or utility of the mean:**
(A) Decreases (B) Increases (C) Is unaffected (D) None of these
- The individual probabilities of occurrence of two events A and B are known, the probability of occurrence of both the events together will be:**
(A) Increased (B) Decreased (C) One (D) Zero
- For a Bernoulli distribution with probability p of a success and q of a failure, the relation between mean and variance that holds is:**
(A) Mean < variance (B) Mean > variance (C) Mean = variance (D) Mean ≤ variance
- A random sample of 17 items from a heap of machine parts gives a mean of 42 and S.D.= 6.25. The value of statistic t to test the hypothesis that the population mean = 38 is:**
(A) 2.64 (B) 6.6 (C) 2.56 (D) None of these
- The shape of Chi-square distribution curve with 1 or 2 degree of freedom is:**
(A) A parabola (B) A hyperbola (C) J-shaped curve (D) A bell shaped curve
- F-distribution curve in respect of tails is:**
(A) Negative skew (B) Positive skew (C) Symmetrical (D) All of these
- 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 of these
- If a discrete random variable takes on four values -1, 0, 3, 4 with probabilities 1/6, k, 1/4 and 1-6k, where k is a constant, then the value of k is:**
(A) 1/3 (B) 2/9 (C) 1/12 (D) 5/24
- If a random variable X has mean 3 and S.D.= 5, then the variance of the variable Y = 2X - 5 is:**
(A) 25 (B) 45 (C) 100 (D) 50
- Let X has a random variable U(0,1), then the variable y = -2 log X follows:**
(A) Log normal distribution (B) Gamma distribution
(C) Chi-square distribution (D) Exponential distribution
- If all observations in a set of observations are same, the variance of the set of values is:**
(A) Zero (B) One (C) Infinity (D) Not possible to calculate

- NOTE:**
- (i) Part-II is to be attempted on the separate Answer Book.
 - (ii) Attempt **FOUR** questions in all by selecting **TWO** Questions each from **SECTION-A** and **SECTION-B**.
 - (iii) **ALL** questions carry **EQUAL** marks.
 - (iv) All the parts (if any) of each Question must be attempted at one place instead of at different places.
 - (v) Write Q. No. in the Answer Book in accordance with Q. No. in the Q. Paper.
 - (vi) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed.
 - (vii) Extra attempt of any question or any part of the question will not be considered.
 - (viii) Use of Calculator is allowed.

SECTION-A

- Q. No.2. (a)** What methods you employed in the collection of statistical data when the field of inquiry is (i) small, (ii) fairly large and (iii) very large, if you are to pay due regard to accuracy, labor, cost and time. (10)
- (b)** If $\log_e x$ is normally distributed with $\mu = 1$ and $\sigma^2 = 4$, find $P(1 < x < 2)$, given $\log_e 2 = 0.693$. (10) (20)

- Q. No.3.** In a certain experiment to compare two types of sheep food A and B, the following results of increase in weights were observed:

Sheep No.	1	2	3	4	5	6	7	8
Food A	49	53	51	52	47	50	52	53
Food B	52	55	52	53	50	54	54	53

- (i)** Assuming that the two samples of sheep are independent, can we conclude that food B is better than food A? (10)
- (ii)** Examine the case when the same set of eight sheep was used in both the foods. (10) (20)
- Q. No.4. (a)** An urn contains 10 white and 3 black balls. Another urn contains 3 white and 5 black balls. Two balls are transferred from first urn and placed in the second and then one ball is taken from the latter. What is the probability that it is a white ball? (10)
- (b)** Differentiate between Poisson distribution and Poisson process. Give examples in each case. (10) (20)

- Q. No.5. (a)** Given the following population distribution:

x	1	2	3	4
f(x)	1/7	3/7	2/7	1/7

Find the sampling distribution of the mean if a sample of three numbers is taken without replacement. How does the variance of the sampling distribution compare with the population variance? (10)

- (b)** A random sample of size $n = 100$ is taken from a population having a mean of 20 and a standard deviation of 5. The shape of the population distribution is unknown.
- (i)** What can you say about the sampling distribution of the sample mean? (10)
 - (ii)** Find the probability that the sample mean exceeds 20.75. (10) (20)

SECTION-B

- Q. No.6. (a)** A confidence interval is constructed from a random sample of size $n = 50$, for the mean yield of a normal population which has $\sigma = 21$ tons. The limits for the interval are 866.11 and 875.89 tons. What confidence co-efficient was used? (10)
- (b)** How will you construct confidence interval for difference of means of a normal population? (10) (20)

Q. No. 7. (a) Describe the main steps involved in the construction of index numbers of wholesale prices. **(10)**

(b) The prices and quantities of three commodities during 2002 and 2004 are given below: **(10) (20)**

Commodity	Prices (Rs. Per 40 Kilogram)		Quantities Produced (Kilograms)	
	2002	2004	2002	2004
A	3.95	4.25	9,675	10,436
B	34.80	38.04	78	83
C	61.56	59.70	118	116

Compute the Marshall-Edgeworth and the Walsh's price index numbers for 2004, using 2002 as the base period.

Q. No. 8. (a) Describe a Randomized Complete Block Design experiment and give its layout. **(10)**

(b) In a Randomized Complete Block Design, in each of four blocks I, II, III, IV, four varieties of wheat A, B, C, D are grown in the layout given below and the yields are also indicated therein: **(10) (20)**

I	B 27	A 17	C 15	D 25
II	A 28	D 22	B 26	C 16
III	D 14	C 11	A 22	B 25
IV	C 18	B 18	D 19	A 17

- (i)** Perform the analysis of variance to test at 0.05 significance level, the difference in the yields of varieties and in blocks.
- (ii)** What would be the result if no blocking had been done, i.e., if we consider it as a completely randomized design?



FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION – 2025
FOR RECRUITMENT TO POSTS IN BS-17
UNDER THE FEDERAL GOVERNMENT

Roll Number

Statistics

TIME ALLOWED: THREE HOURS	PART-I (MCQS)	MAXIMUM MARKS = 20
PART-I(MCQS): MAXIMUM 30 MINUTES	PART-II	MAXIMUM MARKS = 80
NOTE: (i) First attempted Part-I (MCQS) on the separate OMR Answer Book which shall be taken back after 30 minutes. (ii) Overwriting/cutting of the options/answers will not be given credit. (iii) There is no negative marking . All MCQs must be attempted.		

PART-I (MCQs) (COMPULSORY)

- Q.1.** (i) Select the best option/answer and fill in the appropriate Box on the **OMR Answer Sheet. (20 × 1 = 20)**
(ii) Answers given anywhere else, other than the OMR Answer Sheet, will not be considered.

1. Statistics deals with:

- (A) Individuals
- (B) Particular facts
- (C) Isolated Items
- (D) Aggregative facts

2. The branch of Statistics that deals with procedure and methodology for obtaining valid conclusion is called:

- (A) Descriptive
- (B) Advance
- (C) Inferential
- (D) All of these

3. Sum of the absolute deviation is least when deviation is taken from:

- (A) Mean
- (B) Median
- (C) Mode
- (D) Geometric Mean

4. In t-distribution, which one is true?

- (A) Mean = Median = Mode
 - (B) Mean > Median > Mode
 - (C) Mean < Median < Mode
 - (D) All of these
-

5. What is the probability of a sure event?

- (A) 1
 - (B) 0
 - (C) 0.5
 - (D) 0.2
-

6. Which formula represents the probability of the complement of event A?

- (A) $1 + P(A)$
 - (B) $1 - P(A)$
 - (C) $P(A)$
 - (D) $P(A) - 1$
-

7. In regression analysis, the variable that is being predicted is:

- (A) Dependent variable
 - (B) Independent variable
 - (C) Intervening variable
 - (D) None of these
-

8. Paired t-test is applicable when the observations in the two samples are:

- (A) Equal in number
 - (B) Paired
 - (C) Correlation
 - (D) All of these
-

9. The degree to which numerical data tend to spread about an average is called:

- (A) The Dispersion
 - (B) Regression
 - (C) Correlation
 - (D) None of these
-

10. The types of estimates are:

- (A) Point estimate
- (B) Interval estimates
- (C) Estimation of confidence region
- (D) All of these

11. Ranking scale also includes the properties of _____ scale.

- (A) Nominal
 - (B) Interval
 - (C) Ratio
 - (D) All of these
-

12. The difference between a statistic and the parameter is called:

- (A) Sampling error
 - (B) Random error
 - (C) Non-random error
 - (D) Probability
-

13. The standard deviation of any sampling distribution is called:

- (A) Standard error
 - (B) Non-sampling error
 - (C) Type-I error
 - (D) Type-II error
-

14. A survey conducted by a sampling design is called:

- (A) Sample survey
 - (B) Population survey
 - (C) Systematic survey
 - (D) None of these
-

15. The sum of the frequencies of the frequency distribution of a statistic is equal to:

- (A) Sample size
 - (B) Population size
 - (C) Possible samples
 - (D) Sum of X values
-

16. Sampling error can be reduced by:

- (A) Non-probability sampling
 - (B) Increasing the population
 - (C) Decreasing the sample size
 - (D) Increasing the sample size
-

17. The range of test statistic t is:

- (A) 0 to ∞
- (B) 0 to 1

- (C) $-\infty$ to $+\infty$
(D) -1 to +1
-

18. The probability associated with committing a Type-I error is:

- (A) β
(B) α
(C) $1 - \beta$
(D) $1 - \alpha$
-

19. The degree of freedom for paired t-test based on n pairs of observations is:

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

20. Experimental error is due to:

- (A) Experimenter's mistakes
(B) Extraneous factors
(C) Variation in treatment effects
(D) None of these

PART-II

NOTE:

- (i) **Part-II** is to be attempted on the separate **Answer Book**.
- (ii) Attempt **ONLY FOUR** questions from **PART-II**. **ALL** questions carry **EQUAL** marks.
- (iii) All the parts (if any) of each Question must be attempted at one place instead of at different places.
- (iv) Write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper.
- (v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed.
- (vi) Extra attempt of any question or any part of the question will not be considered.
- (vii) **Use of calculator is allowed.**

SECTION-A

- Q. No.2.** (a) What is the purpose of frequency distribution and what are its desirable qualities? For a certain frequency distribution, the mean was 40.5 and the median 36. Find the mode approximately using the formula connecting the three. **(10)**
- (b) Define Statistics. Discuss the importance of study of statistics by giving examples. How it can help the extension of scientific knowledge? **(10) (20)**
- Q. No.3.** (a) Explain what is meant by the skewness of the distribution, and define a suitable measure **(10)**

of Skewness. For given p.d.f.

$$f(x) = k(x-x^2), 0 \leq x \leq 1$$

Find the Skewness and discuss.

- (b) Define the normal distribution and obtain its mean and variance. Show that for the normal distribution, the mean, mode and median are the same. (10) (20)

- Q. No.4.** (a) If X has a binomial distribution, then show that $E(X) = np$, $\text{Var}(X) = npq$. Derive the m.g.f of the binomial distribution and explain its uses? (10)

- (b) What are the main characteristics of the Poisson distribution? Explain with the help of examples. Also give its properties, applications and relationship with other distributions. (10) (20)

- Q. No. 5.** (a) From the following set of values: (10)

Y	6.5	5.3	8.6	1.2	4.2	2.9	1.1	3.0
X	3.2	2.7	4.5	1.0	2.0	1.7	0.6	1.9

- 1) Compute the residuals and verify that they add to zero and draw the conclusion about the results.
- 2) Compute the standard error of estimate, $S_{y.x}$
- (b) Under what conditions the correlation among the random variables exists. Also describe the properties of correlation coefficient. Calculate the correlation co-efficient for a sample of 20 pairs of observations, given that
Mean of X = 2, Mean of Y = 8, $\sum X^2 = 180$, $\sum Y^2 = 1424$ and $\sum XY = 404$
Also interpret its results. (10) (20)

SECTION-B

- Q. No.6.** (a) What is finite-correction factor? When is it appropriately used in sampling applications and when can it, without the great undesirable consequences, be ignored? (10)

- (b) Given the population 2, 4, 8, 8, 10, 10. (10) (20)

- 1) How many samples of size $n = 2$ can be drawn without replacement from this population?
- 2) Compute and tabulate the sampling distribution of the mean for samples of size $n = 2$.

- Q. No.7.** (a) Explain what is meant by: (10)

- (i) a statistical hypothesis, (ii) test-statistic, (iii) test of significance,
(iv) level of significance, (v) type-I error and type-II error

- (b) The weights of 4 persons before they stopped smoking and 5 weeks after they stopped smoking are as follows: (10) (20)

Person	1	2	3	4
Before	148	176	153	116
After	154	176	151	121

Use the t-test for paired observations to test the hypothesis at the 0.05 level of significance, that giving up smoking has no effect on a person's weight.

- Q. No.8.** (a) Explain the procedure of randomization in a completely randomized design where we have 3 varieties of wheat and 18 experimental plots available. Also explain what is explained by significant F value in an experiment. (10)

- (b) Four varieties of wheat were tried in a randomized complete block design in four replications. Yield in kilogram per plot is shown in the table given below. Test the hypothesis that there is no difference in the means of four varieties. $\alpha = 0.05$. **(10) (20)**

Replicates	Varieties			
	V_1	V_2	V_3	V_4
I	2	5	4	1
II	2	3	3	1
III	4	6	6	2
IV	1	4	2	3



FEDERAL PUBLIC SERVICE COMMISSION

COMPETITIVE EXAMINATION – 2026
FOR RECRUITMENT TO POSTS IN BS-17
UNDER THE FEDERAL GOVERNMENT

Roll Number

STATISTICS

TIME ALLOWED: THREE HOURS	(PART-I MCQs)	MAXIMUM MARKS = 20
PART-I (MCQs): MAXIMUM 30 MINUTES	(PART-II)	MAXIMUM MARKS = 80
NOTE: (i) First attempt PART-I (MCQs) on separate OMR Answer Sheet which shall be taken back after 30 minutes. (ii) Overwriting/cutting of the options/answers will not be given credit. (iii) There is no negative marking . All MCQs must be attempted.		

PART-I (MCQs) (COMPULSORY)

- Q.1.** (i) Select the best option/answer and fill in the appropriate Box on the **OMR Answer Sheet**. ($20 \times 1 = 20$)
(ii) Answers given anywhere else, other than OMR Answer Sheet, will not be considered.

- Which of the following best describes what statistics deals with?
(A) Individuals (B) Particular facts (C) Isolated Items (D) Aggregative facts
- Five-point Likert scales (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree) are frequently used to measure motivations and attitudes. A Likert scale is a:
(A) Discrete variable (B) Ordinal variable (C) Categorical variable (D) All of these
- A numerical value used as a summary measure for a sample, such as sample mean, is known as a:
(A) Parameter (B) Estimator (C) Estimate (D) Statistic
- The general purchasing power of the currency of a country is determined by:
(A) Retail price index (B) Wholesale price index (C) Composite index (D) Value index
- If a data set is transformed by adding or subtracting a constant c to every observation, which measure remains unchanged?
(A) Mean (B) Median (C) Range (D) Standard deviation
- The weighted arithmetic mean of the first n natural numbers whose weights are equal to the corresponding numbers is:
(A) $(n+1)(2n+1)/6$ (B) $n(n+1)(2n+1)/6$ (C) $(2n+1)/3$ (D) $(2n+1)/6$
- In a random experiment of n trials, with fixed probability of success, the probability to observe success at k -th trial is modeled by the following distribution:
(A) Binomial (B) Poisson (C) Uniform (D) Normal
- If A and B are two events, the probability of occurrence of either A or B is given as:
(A) $P(A) + P(B)$ (B) $P(A \cup B)$ (C) $P(A \cap B)$ (D) None of these
- The diagonal matrix $D = \begin{vmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{vmatrix}$ the determinant of D^{-1} :
(A) $1/8$ (B) $4/8$ (C) $3/8$ (D) $6/8$
- If X is an exponential random variable with CDF $F(x) = 1 - e^{-2x}$ ($x \geq 0$), the mean of the random variable X is:
(A) 2 (B) 0.5 (C) -2 (D) -0.5
- Significance of partial regression coefficients can simultaneously be tested by:
(A) t-test (B) z-test (C) Chi-square test (D) F-test
- Which of the following is useful for computing bounds on probability when only the mean and standard deviation of the variable are known?
(A) Markov inequality (B) pdf of X (C) Chebyshev's inequality (D) Central limit theorem

13. The sample standard deviation of the first five natural numbers (1, 2, 3, 4, 5) is:

- (A) 1.414214 (B) 1.581139 (C) 2 (D) 4

14. The CDF of the random variable X is defined by:

$$F(x) = \begin{cases} 0 & x < 2 \\ K(x-2) & 2 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$

What is the value of K ?

- (A) 1/4 (B) 1/2 (C) 2 (D) None of these

15. Given a system of linear equations: $x_1 + x_3 = 3$; $x_1 - x_2 - x_3 = 1$; $-x_1 + x_2 = 4$. This non-homogeneous system has:

- (A) No solution (B) Infinite solutions (C) Unique solution (D) All of these

16. The probability associated with committing Type-II error is:

- (A) β (B) α (C) $1 - \beta$ (D) $1 - \alpha$

17. The following are the scores a professor gave on the first test in a statistics class: 52, 90, 88, 61, 75, 82, 75, 83, 88, and 86. What was the median score on this test?

- (A) 82 (B) 82.5 (C) 78.5 (D) 52

18. Which of the following is NOT a property of a consistent estimator?

- (A) As $n \rightarrow \infty$, bias $\rightarrow 0$ (B) It converges in probability to the true parameter
(C) As $n \rightarrow \infty$, variance $\rightarrow 0$ (D) It always attains the Cramér-Rao lower bound

19. In a two-way ANOVA with factors A (3 levels) and B (4 levels), the degrees of freedom for the interaction term are:

- (A) $2 \times 3 = 6$ (B) $(3-1)(4-1) = 6$ (C) $3 + 4 - 2 = 5$ (D) $12 - 1 = 11$

20. A country's infant mortality rate (IMR) is 45 deaths per 1,000 live births. If the number of live births in the year was 250,000, how many infant deaths occurred?

- (A) 13,750 (B) 12,500 (C) 11,250 (D) 10,000

PART-II

NOTE:

- (i) **Part-II** is to be attempted on the separate **Answer Book**.
- (ii) Attempt **FOUR** questions in all by selecting **TWO** Questions each from **SECTION**. **ALL** questions carry **EQUAL** marks.
- (iii) All the parts (if any) of each Question must be attempted at one place instead of at different places.
- (iv) Write Q. No. in the Answer Book in accordance with Q. No. in the Q.Paper.
- (v) No Page/Space be left blank between the answers. All the blank pages of Answer Book must be crossed.
- (vi) Extra attempt of any question or any part of the question will not be considered.
- (vii) **Use of Calculator is allowed.**

SECTION-A

- Q. No. 2.** (a) Define Type I and Type II errors in hypothesis testing. Explain the trade-off between these errors. (10)
(b) The first four moments of a frequency distribution about the origin are $\mu'_1 = -0.55$, $\mu'_2 = 4.46$, $\mu'_3 = -0.43$ and $\mu'_4 = 68.52$. Find skewness and kurtosis. Also comment on the shape of the data distribution. (10) (20)
- Q. No. 3.** (a) Explain the concept of exchangeability in probability. How does it differ from independence? (10)
(b) Define skewness of a distribution. What are the different measures of skewness? (10) (20)
(i) In a frequency distribution, the coefficient of skewness based upon quartiles is 0.6. If the sum of the upper and lower quartiles is 100 and median is 38, find the values of Q_1 and Q_3 .
(ii) Karl Pearson's coefficient of skewness of a distribution is 0.32. If the standard deviation is 6.5 and mean is 29.6, determine the mode and the median of the distribution.

- Q. No. 4.** (a) Explain the difference between Parametric and Non-Parametric tests. Provide examples of each in the context of ANOVA. (10)
- (b) A secretary makes 2 errors per page, on average. What is the probability that on the next page he or she will make: (10) (20)
- (i) 4 or more errors?
- (ii) No error?

- Q. No. 5.** (a) Define conditional probability, with examples. (10)
- (b) In a small town only three newspapers, A, B, and C, are available for the readers. Suppose that 60% of readers subscribe to newspaper A, 40% subscribe to newspaper B, and 30% to newspaper C. Suppose also that 20% subscribe to both A and B, 10% subscribe to both A and C, 20% subscribe to both B and C, and 5% subscribe to all three newspapers. (10) (20)
- (i) Construct Venn Diagram to present the above situation.
- (ii) What percentage of newspaper readers subscribe none of the three newspapers?
- (iii) What percentage of newspaper readers subscribe at least one of the three newspapers?

SECTION-B

- Q. No. 6.** (a) Critically evaluate the role of probability theory in statistical inference. (8)
- (b) One of the larger components of the cost of living is the amount spent on housing. Housing cost includes rent (for tenants), mortgage payments (for home owners), heating, electricity and water. An Economist undertook a five years study to determine how housing costs have changed. Five years ago, he took a random sample of 200 households and recorded the percentage of total income spent on housing. This year he took a sample of 100 households. (12) (20)
- The statistics from the two samples are:

$$\bar{x}_1 = 32.42; s_1 = 6.08; \bar{x}_2 = 33.72; s_2 = 6.75$$

Test the hypothesis at 5% level, to determine whether the Economist can infer that housing cost as a percent of total income has increased over the last five years.

- Q. No. 7.** (a) Suppose you want a confidence interval for a Normal Mean with small sample size. Explain how you would construct it. (10)
- (b) The population and death of two towns, according to age groups are given below: (10) (20)

Age group	Town A		Town B	
	Population	No. of deaths	Population	No. of deaths
Under 10	100	40	1000	300
11 – 25	1500	150	500	40
26 – 50	850	108	400	24
Above 50	50	37	100	6

- (i) Calculate crude death rate for town A.
- (ii) Calculate SDR of town A taking the population of town B as a standard population.

- Q. No. 8.** (a) Critique the use of p-values in statistical hypothesis testing. Discuss alternatives for interpreting statistical significance. (10)
- (b) Suppose in a field experiment there are 4 treatments (T1, T2, T3, and T4) and 4 blocks (B1, B2, B3, and B4). Consider the field is partitioned into 4 plots (blocks), each of these blocks is homogeneous. Assign each treatment randomly to each plot (block) by fulfilling the requirements of randomized complete block design conditions (requirements), that you have to complete the layout of the experiment according to the RCBD. What will be the ANOVA table (a general form of RCBD ANOVA table) for this experiment? (10) (20)

Blocks	T1	T2	T3	T4
Block I	2	5	4	1
Block II	2	3	3	1
Block III	4	6	6	2
Block IV	1	4	2	3

Perform the analysis of variance to test at the 0.05 level of significance.