FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION FOR
RECRUITMENT TO POSTS IN BS-17
UNDER THE FEDERAL GOVERNMENT, 2015

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 FIVE** 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) 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.

PART-II

Q. No. 2. **(a)** Discuss probability and its significance in social, economic and political problems. **(05)**

(b) Sam is going to assemble a computer by himself. He has the choice of chips from three brands, a hard drive from five, memory from two, and an accessory bundle from six local stores. How many different ways can Sam order the parts? **(05)**

(c) According to a Consumer Digest (July/August, 1996), the probable location of a personal computer (PC) in the home is as follows: **(06)**
> Adult Bedroom: 0.03
> Child Bedroom: 0.15
> Other bedroom: 0.14
> Office or den: 0.40
> Other rooms: 0.28
(i) What is the probability that a PC is in bedroom?
(ii) What is the probability that a PC is not in a bedroom?

Q. No. 3. **(a)** On average, 2 traffic accidents per month occur at a certain intersection. What is the probability that in any given month at this intersection? **(08**
(i) exactly 6 accidents will occur?
(ii) fewer than 4 accidents will occur?

(b) A soft-drink machine is regulated so that it discharges an average of 100 millilitres per cup. If the amount of drink is normally distributed with a standard deviation equal to 10 millilitres, **(08)**
(i) How many cups will probably overflow if 130-milliter cups are used for the next 1000 drinks?
(ii) Below what value do we get the smallest 30% of the drinks?

Q. No. 4. **(a)** Differentiate between the following: **(06)**
(i) Null and Alternative hypothesis
(ii) One and two sided tests
(iii) Rejection and Acceptance region

(b) The variable X, is the carbon monoxide concentration in air, and ten measurements are obtained as follows: **(10)**
10.25, 10.37, 10.66, 10.47, 10.56, 10.22, 10.44, 10.38, 10.63, 10.40 mg/m^3.
(i) Test whether the mean concentration of carbon monoxide in air is 10.00 mg/m^3.
(ii) Calculate the power of the above test if the mean concentration of carbon monoxide in alternative hypothesis is assumed to be 10.75mg/m^3.

STATISTICS

Q. No. 5. **(a)** How do you test the equality of variances of two normal populations? **(04)**

(b) Differentiate between simple and composite hypothesis. **(04)**

(c) Given the statistics of two samples drawn from two normal populations **(08)**
$N(\mu_1, \sigma^2)$ *and* $N(\mu_2, \sigma^2_2)$ as,

$$n_1 = 6, \qquad \bar{x}_1 = 25, \qquad s_1^2 = 36$$
$$n_2 = 8, \qquad \bar{x}_2 = 20, \qquad s_2^2 = 25$$

Test $H_0 : \mu_1 = \mu_2 \; vs \; \mu_1 \neq \mu_2$ under two situations (i) $\sigma_1^2 = \sigma_2^2$ (ii) $\sigma_1^2 \neq \sigma_2^2$

Q. No. 6. **(a)** How can the parameters β_0 and β_1 be estimated in simple linear regression model? **(04)**

(b) Delineate the properties of regression coefficients. **(04)**

(c) The following measurements show the respective height in inches of ten fathers **(08)**
and their eldest sons.

Father(X):	67	63	66	71	69	65	62	70	61	72
Son(Y):	68	66	65	70	69	67	64	71	60	63

(i) Find the regression line of son's height on father's height.
(ii) Estimate the height of son for the given height of father as 70 inches.

Q. No. 7. **(a)** Discuss the properties of correlation coefficient. **(06)**

(b) Suppose that an advertising campaign for a new product is conducted in 10 test **(10)**
cities. The intensity of the advertising x, measured as the number of exposures
per evening of prime-time television, is varied across cities; the awareness
percentage y is found by survey after the ad campaign:

x:	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
y:	10.1	10.3	10.4	21.7	36.7	51.5	67.0	68.5	68.2	69.3

(i) Calculate and interpret the correlation coefficient r_{yx}
(ii) Does the relation appear linear you? Does it appear to be increasing?

Q. No. 8. **(a)** What is the need of sampling as compared to complete enumeration? **(04)**

(b) Highlight the advantages of stratified random sampling. **(04)**

(c) For a population with elements 5, 7, 8, 10, 13, draw all possible samples of **(08)**
size 3 without replacement and compute the sampling distribution of the
sample means.

Q. No. 9. Write note on any FOUR of the following: **(04 each) (16)**
(a) Applications of Poisson distribution

(b) Role of Statistics in social, law and economics

(c) Statistical importance of Normal distribution

(d) Conditional expectation

(e) Maximum likelihood estimation of the mean of a normal population

(f) Applications of t-distribution

****

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)**
$f(x) = w(2-x)(x-5), 2 \le x \le 5 = 0,$ 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 coin14 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) $E(p) = P$ and   (ii) $E(p) = \left(\dfrac{N-n}{N-1}\right)\dfrac{PQ}{n}$.

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)**

****

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
(iii) Co-efficient of Skewness and discuss the result from the following weekly wages of 300 workers. **(12)**

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			5.8
	23.7	13.2	
12.4			6.1
	14.4	10.2	
10.1			4.8
	9.0	5.1	
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)**

STATISTICS

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

SECTION-B

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 administrated 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
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 $P(A \wedge B)=P(A/B)P(B)$, established for two events, may be generalized to three events as follows; (6)
$$P(A \wedge B \wedge C)=P(A/B \wedge C) \, P(B/C) \, P(C)$$

(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.
$$\bar{X}_1 = 28.02, \quad S_1 = 4.42, \quad r_{12} = 0.80,$$
$$\bar{X}_2 = 4.91, \quad S_2 = 1.10, \quad r_{13} = -0.40,$$
$$\bar{X}_3 = 594, \quad S_3 = 85, \quad r_{12} = -0.56,$$
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 (15) 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:

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 (5) **(20)** explain such errors, supporting by real life example.

Q.7. (a) The following Latin Square layout displays the scores secured by nine college students. (15) Students are of different ethnic background and various professional interests.

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 (5) **(20)** Multiple Comparison tests and explain one method.

Q.8. (a) Explain the terms Demography and Vital statistics. List few sources of demographic (5) data both locally and globally.

(b) Using the information given in the following table. Calculate Total Fertility Rate (TFR) (10) and Gross Fertility Rate (GFR).

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 (5) **(20)** Rate.

****