

FEDERAL PUBLIC SERVICE COMMISSION

COMPETITIVE EXAMINATION FOR RECRUITMENT TO POSTS
IN BPS-17 UNDER THE FEDERAL GOVERNMENT, 2001.

STATISTICS

TIME ALLOWED: THREE HOURS

MAXIMUM MARKS: 100

NOTE: Attempt FIVE questions in all, including question No.8 which is COMPULSORY. All questions carry EQUAL marks.

1. (a) Explain classical, axiomatic and relative frequency definitions of Probability with one example in each case. Which definition you prefer in day to day problems solving in a Chaotic Situations.
- (b) Define law of total probability. Three facilities supply micro processors to a manufacturer of telemetry equipment. All are supposedly made to the same specifications. However, the manufacturer has for several years tested microprocessors, and records indicate following numerical facts:

Supply Facility	Fraction Defective	Fraction Supplied by
1	0.02	0.15
2	0.01	0.80
3	0.03	0.05

The director of manufacturing randomly selects a microprocessor, takes it to the test department, and finds that it is defective. If we let A be the event that an item is defective and B_i be the event that the item came from facility $i(i=1,2,3)$. Compute $P(B_i | A)$ $i=1,2,3$ and comment.

2. (a) The daily demand of Computer diskettes in a office follows the probability distribution:

X = x	0	1	2	3	4
P(X=x)	0.2	0.25	0.25	0.2	0.1

Compute $E(X)$ and $Var(X)$.

- (b) State Chebyshev's inequality. Estimate the demand interval such that the probability is at least $\frac{8}{9}$ that the demand will remain or lie in that interval.
3. Define binomial, Poisson and negative binomial random variables and find their mean and variance respectively. Comment on relation between mean and variance for each random variable.
4. (a) What do you understand by maximum likelihood estimation of parameter θ , if X follows pdf $f(x, \theta)$ and a random sample of size n is given on X. Discuss with an example.
- (b) Find maximum likelihood estimator of λ when r.v.x follows exponential distribution given by: $f(x, \lambda) = \lambda e^{-\lambda x}$, $x > 0$.
5. (a) Define regression line of Y on X and regression line of X on Y. How regression coefficients are related with correlation between X and Y.
- (b) For regression line of Y on X is $y = \alpha + \beta x + \epsilon$. Give complete procedure for testing $H_0: \beta=0$ where ϵ follows $N(0, \sigma^2)$ and σ^2 is not known.

6. In a large city of Pakistan, we are interested to study socio-economic conditions of the citizens. It is known that lower middle class, middle class, higher middle class and affluent people are living in the city. Discuss the sampling technique which is most suitable in such situations. How the sample size will be determined?
7. In an organization married and unmarried individuals are working and we are interested to study the following null and alternative hypothesis using statistical methods. A sample of 500 employees was selected and following results in the tabular form are obtained under the Hypothesis:

H_0 : Absentee behaviour is independent of marital status.

H_1 : Absentee behaviour is dependent of marital status.

Marital Status vs Absentee Rate				
Marital Status ↓	Absentee Rate →			Row Total
	Zero	1-5	over 5	
Single	84	82	34	200
Married	50	64	36	150
Divorced	50	34	16	100
Widow	16	20	14	50
Column Total	200	200	100	500

Test H_0 against H_1 as stated above and write conclusion.

COMPULSORY QUESTION

8. Select the correct answer by writing (a), (b), (c) or (d) (for each part of the question in the answer book. Don't reproduce questions:
- (1) Statistic is used for:
- (a) Subject statistics (b) a number
(c) random number (d) (b) and/or (c).
(e) None of these.
- (2) If a random variable X is measurable then the probability $P(X = 0)$ is
- (a) 1 or 0 (b) less than 1
(c) 1 or less than 1 (d) zero.
(e) None of these.
- (3) If $A_1, A_2 \subseteq S$ and S is sample space then $P(A_1 \cup A_2 | S) \leq P(A_1) + P(A_2)$ if:
- (a) $A_1 \subset A_2$ (b) $A_1 \subseteq A_2$
(c) $A_1 \wedge A_2 \neq \Phi$ (d) $A_1 \wedge A_2 = \Phi$
(e) None of these.
- (4) If a frequency distribution is normal then:
- (a) $\beta_1 = 3, \beta_2 = 0$ (b) $\beta_1 = 0, \beta_2 = 3$
(c) $\beta_1 = 1, \beta_2 = 2$ (d) None of these.
- (5) If $\alpha = \text{Prob}(\text{Reject } H_0 | H_0 \text{ true}), \beta = \text{Prob}(\text{Accept } H_0 | H_0 \text{ true})$:
- (a) If α increases then β decreases.
(b) If α increases then β remains unchanged.
(c) If α decreases then β decreases.
(d) If α increases then β increases.
(e) None of these.

- (6) In binomial distribution:
- Number of successes are fixed.
 - Number of successes are random.
 - Number of trials are random.
 - Number of trials and successes are random.
 - None of these
- (7) If events A and B are not mutually exclusive then:
- $P(A|B) = 0$
 - $P(A|B) = P(A)$
 - $P(A|B) \cdot P(B) = P(A \cap B)$
 - $P(A|B) = P(B)$
 - None of these.
- (8) The joint density function of X_1 and X_2 is given by $f(x_1, x_2) = \frac{1}{500}$; $0 \leq x < 0.25$; $0 \leq x \leq 2000$ then:
- X_1 and X_2 independent
 - X_1 and X_2 are not independent
 - X_1 depends on X_2
 - X_2 depends on X_1
 - None of these.
- (9) If X be a r.v. with pdf $p(x) = k \cdot q^{x-1}$, $x = 1, 2, \dots$ and k is constant then
- $k = 1$
 - $k = \frac{1}{p}$
 - $k = p$
 - $k = q$
 - None of these.
- (10) For a negative binomial distribution, if $p = 2$ then for $r = 50^{\text{th}}$ success needs on average:
- 50 trials
 - 100 trials
 - 150 trials
 - 200 trials
 - None of these.
- (11) For normal distribution, pdf $f(x; \mu, \sigma^2)$ is:
- $f(x + \mu, \sigma) = f(x - \mu; \sigma)$
 - $f(x + \mu, \sigma) = f(-x + \mu; \sigma)$
 - $f(x + 2\mu, \sigma) = f(-x + \mu; \sigma)$
 - $f(2x + \mu, \sigma) = f(2x - \mu; \sigma)$
 - None of these.
- (12) Equality of two population means is tested by:
- Z-test with $\sigma_1^2 = \sigma_2^2$ is known.
 - t-test with $\sigma_1^2 = \sigma_2^2$ is known.
 - chi-square test.
 - None of these.
- (13) If $n \rightarrow \infty$ and p is fixed then binomial probabilities can be computed using:
- normal (np, npq)
 - Poisson (np)
 - hypergeometric
 - χ^2 - distribution
 - None of these.

(14) If x has a binomial distribution with parameters p and n then $\frac{x}{n}$ has

variance:

- (a) npq (b) n^2pq
 (c) $\frac{pq}{n}$ (d) $\frac{pq}{n^2}$
 (e) None of these.

(15) If x is $n(\mu, \sigma^2)$ then $Z = \frac{x - \mu}{2\sigma}$ is :

- (a) $n(0,1)$ (b) $n(0, \frac{1}{4})$
 (c) $n(0, 2\sigma)$ (d) $n(1, \sigma^2)$
 (e) None of these.

(16) $Y_i = \alpha + \beta x_i + \epsilon_i, i = 1, 2, \dots, n$ if $H_0: \beta = 0$ is:

- (a) rejected then there is linear relationship between x & y .
 (b) accepted then there is linear relationship between x & y .
 (c) rejected then there is no linear relationship between x & y .
 (d) accepted then there is no linear relationship between x & y .
 (e) None of these.

(17) The mean square error of an estimator $\hat{\theta}$ of θ is:

- (a) $V(\hat{\theta})$ if $\hat{\theta}$ is biased estimator of θ .
 (b) $V(\hat{\theta})$ if $\hat{\theta}$ is unbiased estimator of θ .
 (c) $V(\hat{\theta})$ if $\hat{\theta}$ is unbiased or biased estimator of θ .
 (d) None of these.

(18) If $\hat{\theta}_1$ estimates θ with $V(\hat{\theta}_1)$ and $\hat{\theta}_2$ estimates θ with $V(\hat{\theta}_2)$ then:

- (a) $\hat{\theta}_1$ is better than $\hat{\theta}_2$ if $V(\hat{\theta}_1) < V(\hat{\theta}_2)$
 (b) $\hat{\theta}_1$ is better than $\hat{\theta}_2$ if $V(\hat{\theta}_1) > V(\hat{\theta}_2)$
 (c) $\hat{\theta}_1$ is unbiased and minimum variance estimator.
 (d) $\hat{\theta}_1$ is biased and minimum variance estimator.

(19) Sample correlation coefficient between x & y is γ then:

- (a) $|\gamma| < 1$ (b) $|\gamma| > 1$
 (c) $1 < |\gamma| \neq 0$ (d) None of these.

(20) The variance of sampling distribution of mean is:

- (a) σ_p^2 (b) $n\sigma^2$
 (c) $n^2\sigma^2$ (d) $\frac{\sigma^2}{n}$

if $v(x_i) = \sigma^2$ and $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$.

COMPETITIVE EXAMINATION FOR RECRUITMENT TO POSTS
IN PBS-17, UNDER THE FEDERAL GOVERNMENT, 2002

STATISTICS

TIME ALLOWED: THREE HOURS

MAXIMUM MARKS: 100

NOTE: Attempt **FIVE** questions in all, including **QUESTION NO. 8** which is **COMPULSORY**. All questions carry **EQUAL** marks. Statistical Table will be provided.

1. (a) Define:
 - i) Marginal and conditional probabilities
 - ii) Distribution function and probability density function
 - iii) Stochastic independence.
- (b) The following probability values are given:
 $\Pr\{A_2|A_1\} = 0.3$ $\Pr\{A_1\} = 0.6$ $\Pr\{A_2\} = 0.4$
 $\Pr\{A_1|A_2\} = 0.4$ $\Pr\{A_3|A_1 \text{ and } A_2\} = 0.3$
 Use the general multiplication law to find
 i) $\Pr\{A_1 \text{ and } A_2\}$ ii) $\Pr\{A_2 \text{ and } A_3\}$ iii) $\Pr\{A_1 \text{ and } A_2 \text{ and } A_3\}$
(10+10)
2. (a) How the Poisson distribution differ form the Binomial and hypergeometric distributions?
 (b) Find the Variance of negative Binomial distribution and Normal distribution.
(8+12)
3. (a) An examination is given to determine whether veterans of different IQs performed better. The scores obtained are shown in following table. Determine at the 0.05 significance level whether there is a difference in scores due to difference in
 (i) veterans and (ii) IQ.

Table	TEST SCORE		
	High IQ	Medium IQ	Low IQ
Veteran	90	81	74
Nonveteran	85	78	70

- (b) What is the underlying logic for the method of maximum likelihood as an estimating procedure?
- (c) How does a point estimator differ from a confidence interval estimator?
(10+5+5)
- Q.4 a) Find the coefficient of correlation between X and Y, also fit the regression line Y on X.

X	5	7	6	3	4
Y	10	8	3	9	5

- b) For each of the following situations, indicate whether a correlation analysis, a regression analysis, or both would be appropriate. In each case, give the reasons for your choice.
 - i) To choose advertising media, an agency account executive is investigating the relationship between a woman's age and her annual expenditures on a client firm's cosmetics.
 - ii) A trucker wishes to establish a decision rule that will enable him to determine when to inspect or replace his tires, based on the number of kilometer driven.

STATISTICS

- iii) A government agency wishes to identify which field offices of various sizes (based on numbers of employees) do not conform to the prevailing pattern of working days lost due to illness.
 - iv) A research firm conducts attitude surveys in two stages. The first stage identifies coincident factors, such as age and income. The second stage is more detailed and involves a separate study to predict the values of one variable using the known values of other variables associated with it in the initial stage. (10 + 10)
5. (a) What is systematic sampling? What are its advantages and drawbacks?
(b) Under what conditions can cluster sampling be more efficient than other types of random sampling designs?
(c) How can stratification increase efficiency? (7+7+6)
6. What is the role of statistics in solving the following problems:
(a) Law and crime
(b) Public health
(c) Socio-political inequality.
7. Write short notes on the followings:
(a) Multiple and Partial Correlation.
(b) Testing of hypothesis.
(c) Relationship between T and F distributions.
(d) Importance of normal distribution.

COMPULSORY QUESTION

- Q.8 Read the following statements carefully and decide which one is true or false(T/F). Do not reproduce the question.
- (1) Sampling error and bias can appear in data at the same time.
 - (2) Random sampling means that no system is used in the sampling process.
 - (3) Subjective probabilities of events are those based on observations of past events.
 - (4) Conditional probability is always subjective probability.
 - (5) The Poisson distribution is skewed to the right.
 - (6) In the binomial distribution the probability of success, p , remains the same from trial to trial.
 - (7) The F distribution is symmetrical
 - (8) An analysis of variance is a useful tool for proving or disproving a null hypothesis about several means.
 - (9) There are two main types of chi-square tests- tests of goodness of fit and tests of badness fit.
 - (10) A regression equation must be computed if a correlation coefficient is to be useful.
 - (11) The higher the coefficient of correlation the lower the standard error of estimate.
 - (12) If r is negative in a correlation analysis then we know that Y decreases as X increases.
 - (13) If $b_{11.2}$ is negative then $r_{11.2}$ will be negative
 - (14) $R_{123} = r_{12.3} + r_{11.2}$
 - (15) The hypergeometric distribution does not require independence between successive trials.
 - (16) Two sample tests must be two-sided tests as well.
 - (17) A random variable assumed only numerical values.
 - (18) A census, if not conducted carefully, may result in sampling bias.
 - (19) Confidence intervals are not true probability intervals.
 - (20) Joint events are those that are not mutually exclusive.

STATISTICSTIME ALLOWED: THREE HOURSMAXIMUM MARKS: 100

NOTE: Attempt FIVE questions in all, including QUESTION NO. 1 which is **COMPULSORY**. All questions carry **EQUAL** marks. Statistical tables will be provided if required.

COMPULSORY QUESTION

a. 1. Write only the correct answer in the Answer Book. Do not reproduce the question. Cutting or overwriting is not allowed.

1. If $P(A)=0.2$, $P(B)=0.3$ and $P(A \cap B)=0.06$ then
 - a) A and B are mutually exclusive events.
 - b) A and B are exhaustive events.
 - c) $P(A \cap B)$ represents the probability that either event A or B will occur.
 - d) $P(A \cap B)$ represents the joint probability that both events A and B will occur.
2. Which one of the following statements is incorrect?
 - a) If two events A and B are statistically independent, then $P(A \cap B)=P(A)P(B)$.
 - b) If $P(A \cap B)=P(A)P(B)$, then events A and B are statistically independent.
 - c) If A and B are statistically independent then $P(B/A)=P(B)$.
 - d) If A and B are statistically independent, then $P(B/A)=P(A/B)$.
3. The relative frequency of Mathematics majors at a university with 8000 students is 0.015. The frequency of the total number of Mathematics majors in the university is
 - a) 150
 - b) 9
 - c) 15
 - d) 120
4. A manufacturing process produces 10% defective articles. If 2 articles are drawn from this process, the probability that they will both be good is:
 - a) 0.01
 - b) 0.91
 - c) 0.81
 - d) 0.18
5. According to Chebysev's inequality, at least _____ percent of the probability distribution is included within two standard deviations from the mean:
 - a) 75
 - b) 25
 - c) 89
 - d) 11
6. Which of the following values cannot be the probability of an event:
 - a) 0.78
 - b) 0.00
 - c) 1.25
 - d) 1.00
7. A normal probability distribution is symmetric around the value 3, which of the following statements is false?
 - a) The expected value of that distribution is 3.
 - b) The standard deviation of that distribution can also be 3.
 - c) Both of the above are true.
 - d) None of the above is true.
8. Which of the following events can't occur?
 - a) A poisson random variable takes on the value 0.
 - b) A binomial random variable takes on the value -1.
 - c) A normal random variable takes on the value 0.
 - d) All of the above.
9. In which one of the situations below are statistics useless?
 - a) Predicting an unborn baby's height.
 - b) Helping the President prepare the budget.
 - c) Determining whether there are human beings in another solar system.
 - d) Deciding which medicine is better for AIDS patients.
10. The bias of non responses in an interview survey may be corrected by:
 - a) More accurate measurements.
 - b) Improved interviewing techniques.
 - c) Either of the above.
 - d) None of the above.
11. Random sampling is preferable to judgment sampling because:
 - a) Measures of precision of estimates of population values can be derived in random sampling.
 - b) Random sampling costs less.
 - c) A random sample is easy to obtain.
 - d) None of the above.

contd. P/2

STATISTICS:

12. A sample observation can contain _____
a) Systematic error b) Random error
c) Both of the above d) None of the above
13. The basic concept in confidence interval estimation is
a) the standard score b) the sampling distribution
c) the number of degrees of freedom d) the finite population correction
14. Which of the following statements is false for the normal distribution?
a) the sample mean is an unbiased estimator of the μ .
b) the sample mean is an efficient estimator of the μ .
c) the sample mean is a consistent estimator of the μ .
d) at least one of the above is false.
15. The coefficient of determination r^2 is equal to 1 minus the ratio of _____
a) the standard error of estimate to the standard deviation of y.
b) unexplained variation to total variation.
c) explained variation to total variation.
d) Variation around the mean of Y to variation around the regression line.
16. A coefficient of correlation of $r=-1.5$ indicates that _____
a) the regression line passes through the origin of the scatter diagram.
b) there is a strong inverse relationship between y and x.
c) the regression line slopes from upper left to down right on the scatter diagram.
d) an error has been made.
17. Which of the following is referred to as homoscedasticity?
a) Population variations of y values for all x values are equal.
b) y is normally distributed.
c) the y values are independent of one another.
d) y values given x are normally distributed.
18. Positive auto correlation leads to _____
a) low values of R^2 . b) strong multicollinearity.
c) low observed values of the Durbin - Watson statistics. d) none of the above.
19. The test to use in evaluating the overall significance of a fitted regression model is the _____
a) t test b) chi-square test
c) test for auto correlation d) F test
20. Which of the following statements is true?
a) A price index can be used as an inflation barometer.
b) An un-weighted average of relatives index uses quantity data.
c) Values can never be appropriate weights.
d) All of the above.
- Q.2. a) A police department in a small city consists of 10 officers. If the department policy is to have 5 of the officers patrolling the streets. 2 of the officers working full time at the station and 3 of the officers on reserve at the station, how many different divisions of the 10 officers into the 3 groups are possible.
b) A retail establishment accepts either the American Express or the VISA credit card. A total of 24 percent of its customers carry an American Express card, 61 percent carry a VISA card, and 11 percent carry both. What percentage of its customers carry a credit card that the establishment will accept?
c) At a certain stage of a criminal investigation the inspector in charge is 60 percent convinced of the guilt of a certain suspect. Suppose now that a new piece of evidence that shows that the criminal has a certain characteristics is uncovered. If 20 percent of the population possesses this characteristics, how certain the guilt of the suspect should the inspector now be if it turns out that the suspect has this characteristics?
- Q.3. a) Show that hyper geometric probabilities add up to unity.
b) Show that the binomial distribution expressed in standard measure tends to the standard normal distribution for n tendency to infinity.

Write sl
i)
ii)
According
by an org
the curri
distribution
i)

a)
b)

c)

Q.5. a)

i
i
ii
iv
v
vi
vii

Q.6. a)
b)

Q.7. a)
b)

c)

Q.8. Write
a)
b)
c)
d)
e)

Contd.... P/3

- a) Write short notes on the following:
- Multinomial distribution
 - Random variable and its expectations
- b) According to a survey of Association of women head basketball coaches conducted by an organization the mean base salary of these coaches is \$44961. Assume that the current base salary of all these coaches have an approximately normal distribution with a mean of \$44961 and a standard deviation of \$6255.
- What is the prob. That the base salary of a randomly selected women's head is between \$34000 and \$50000.
 - What is the prob. that the base salary of a randomly selected women's heads is \$42000 and lower.
 - What is the prob. that the base salary of a randomly selected women's heads is \$39000 and higher.
 - The lowest paid 5% of women head earn what salary?
- c) An athlete finds that in a high jump he can clear a height of 1.68m in once in five attempts and a height of 1.52 m nine times out of ten attempts. Assuming the heights he can clear in various jumps form a normal distribution, estimate the mean and standard deviation of the distribution.

- Q.5. a) A random sample of eight auto drivers insured with a company and having similar auto insurance policies was selected. The following table lists their driving experience (in years) and the monthly auto insurance premium (in dollars) paid by them:

Driving Experience (yrs)	5	2	12	9	15	6	25	16
Monthly Insurance Premium (\$)	64	87	50	71	44	56	42	60

- Do you think experience depends on the monthly insurance or the monthly insurance depends on experience?
- With experience as an independent variable and the insurance premium sold as a dependent variable, what is your hypothesis about the sign of B in the regression model?
- Construct a scatter diagram for these data. Does the scatter diagram exhibit a liner relationship between the two variables?
- Find the least squares regression line. Is the sign of b the same as the one you hypothesized for B in part b?
- Give a brief interpretation of the values of the y-intercept and slope calculated in above.
- Computer r and r^2 and explain what they mean.
- Predict the insurance premium for 11 years of experience.
- Compute the standard deviation of errors.

- Q.6. a) Distinguish between the stratified and cluster sampling by giving at least one example.
- b) What role does statistics play in the economics and social problems solving.

- Q.7. a) If (X,Y) has a bivariate normal distributions then find the marginal distribution of Y.
- b) The two random variables have the join p.d.f.

$$f(x,y) = 24y(1-x) \quad 0 \leq x \leq 1$$

$$\text{find } E(x/y=y) \quad E(y/x=x), \quad 0 \leq y \leq x$$

Var (x/y=y) and correlation coefficient between x and y.

- c) Define lognormal distribution and explain its connection with normal distribution. Find its mean.

- Q.8. Write short notes on the following:

- Axiomatic approach to probability.
- Factorial moments and factorial moment generating function.
- Multiple and partial correlation.
- Maximum likelihood estimator.
- Analysis of variance and its assumptions.

www.maxpapers.com

FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION FOR RECRUITMENT TO POSTS
IN BPS-17, UNDER THE FEDERAL GOVERNMENT, 2004

STATISTICS

TIME ALLOWED: 3 HOURS

MAXIMUM MARKS: 100

NOTE: Attempt **FIVE** questions in all, including **QUESTION NO. 8** which is **COMPULSORY**. All question carry equal marks.

1.
 - a) Explain the association between randomness and probability. (5)
 - b) Give the limitations of classical and relative frequency definitions of probability. (5)
 - c) Can two mutually exclusive events be independent? Give your answer with proof. (5)
 - d) If A, B, and C are mutually independent events then show that $P(A + B + C) = 1 - P(A')P(B')P(C')$ (5)

2.
 - a) A survey is being conducted to determine the public opinion concerning the construction of a dam to control flooding in the ABC Valley. Fifteen residents of the area are to be randomly selected and surveyed. If, in fact, 80% of the people living in the area oppose the dam, what is the probability that a majority of those surveyed will be in the favour of the dam? (10)
 - b) Show that mean, median and mode, all have the same value for Normal distribution. Also give the applications of Normal distributions. (10)

3.
 - a) Suppose, we want to compare the mean protein intake of all people with income below the poverty level to that of all people with incomes above the poverty level. The data given below, display the protein intakes, in grams, of a 24-hours period for independent random sample of 10 people with income below the poverty level and 8 people with income above the poverty level. Test, at 5% level of significance, whether the intake of rich people is higher than that of poor?

Below Poverty Level:	51.4,	76.7,	73.7,	66.2,	65.5,	49.7,	65.8,	62.1,
	75.8,	62.0						
Above Poverty Level:	86.0,	59.7,	68.6,	98.6,	87.7,	69.0,	80.2,	78.1

 (10)
 - b) Explain the desirable properties of a good point estimator. (5)
 - c) What is power of a test? Also define confidence coefficient. (5)

4.
 - a) Explain the coefficient of determination and write its interpretation. (5)
 - b) Explain, with example, the need of partial correlation. (5)
 - c) Test (at 10% level of significance) whether there is any correlation between X_1 and X_2 , where

X_1 :	12	16	18	18	25	23	23	30	45	47
X_2 :	7	10	18	12	15	13	22	25	20	35

 X_1 = Additive oil (in gallons), X_2 = Mileage (in ,000 miles) (10)

5.
 - a) Explain the general procedures to take a stratified and systematic samples. (10)
 - b) What is the difference between; (5)
 - (i) precision & reliability
 - (ii) stratum & cluster

STATISTICS

- c) Differentiate between sampling plan and sampling frame. Also write a situation of investigation where there is no way other than taking a sample. (5)
6. a) How can you utilize the subject Statistics to evaluate the health problem of the community? (10)
- b) What is the logic behind ANOVA for testing equality of several means? Also give the assumptions for ANOVA. Hence also mention that which test is used for overall fit of a regression model? (10)
7. Write short note on the following:
- a) Baye's Theorem
- b) Negative binomial distribution
- c) Regression analysis
- d) Disadvantages of sampling (20)

COMPULSORY QUESTION

8. Write only the most appropriate option in your Answer Book, about the following statements. Do not reproduce the question. Cutting or overwriting is not allowed. (20)

- (1) The standard deviation of a sample statistic is called
- a) Sampling bias b) Standard error
- c) Sampling error d) Sample deviation of statistic
- (2) Which one of the following sample allocation procedure can be used when no information other than the total number of units in the stratum is given
- a) Equal allocation b) Proportional allocation
- c) Neyman allocation d) Optimum allocation
- (3) In Poisson distribution
- a) mean > variance b) mean < variance
- c) mean = variance d) mean \leq variance
- (4) Binomial distribution approaches to Poisson when
- a) n is very large b) np remains constant
- c) p is very small d) all a,b, & c
- (5) In regression analysis, the scatter of data about fitted line is measured by
- a) standard deviation b) with the help of slope
- c) with the help of intercept d) standard error of estimate
- (6) In a positive skewed distribution
- a) mean = median = mode b) mean < median < mode
- c) mean > median > mode d) none of these
- (7) In testing of hypotheses, which of the following statement is always true for α & β
- a) $\alpha + \beta = 1$ b) $\alpha + \beta \neq 1$
- c) $\alpha > \beta$ d) $\alpha < \beta$
- (8) Testing of hypotheses is a phase to check the validity of
- a) population parameter b) sample estimate
- c) population d) all a,b, & c

FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION FOR RECRUITMENT TO POSTS
IN BPS-17, UNDER THE FEDERAL GOVERNMENT, 2005

STATISTICS

TIME ALLOWED: 3 HOURS

MAXIMUM MARKS: 100

NOTE: Attempt FIVE questions in all, including QUESTION NO. 8 which is COMPULSORY. All question carry equal marks.

1. a) Write the applications of probability theory in decision making. (10)
- b) A study of credit card fraud was conducted by a crime researcher. According to collected data, it was found that 243 cases were of stolen cards, 85 of counterfeit card, 52 of mail orders and 46 of both stolen card and mail orders. Find the probability that a randomly selected case
 (i) belongs to counterfeit card
 (ii) belongs to stolen card or mail order fraud (10)
2. a) According to the Labour Department, 15% of adult workers have a vocational diploma but they did not attend any institution. If 8 adults are randomly selected find the probability that
 (i) at least half of them have diploma but they did not attend any institution.
 (ii) there is no such case that the adult has diploma but he did not attend any institution. (10)
- b) The serum cholesterol levels in men aged 18-24 are normally distributed with a mean of 178.1 and a standard deviation of 40.7 (all units are in mg/100ml). If a man aged 21 is selected at random, find the probability that his serum cholesterol level is between 100 and 200. Find the probability that for a sample of 10 men aged 21-24 will have average serum cholesterol level greater than 200. (10)
3. a) Differentiate between
 (i) One-tailed test and two-tailed test
 (ii) Interval and point estimation
 (iii) Simple and composite hypotheses (10)
- b) In an insurance study of pedestrian deaths in a country, monthly fatalities are analyzed for two different time periods. Sample data from the both time periods are summarized by the following statistics:
 $n_1 = 12, \bar{x}_1 = 46.42, s_1 = 11.07; \quad n_2 = 15, \bar{x}_2 = 51.06, s_2 = 10.39$
 At 0.05 level of significance, test the claim that both time periods have the same mean. (10)
4. a) What is the difference between a correlation problem and a regression problem? Also give the situation where we have to use partial correlation rather than simple correlation. (10)
- b) Randomly selected subjects ride a bicycle at 5.5mi/h for one minute. Their weights (in pounds) are given with the number of calories used. Find the correlation coefficient between the both quantities.
 Weights: 167 191 112 129 140 173 119
 Calories used: 4.23 4.69 3.21 3.47 3.72 4.45 3.36 (10)
5. a) Where do we use Stratified Random Sampling? Write a general procedure to draw a stratified random sample. Also differentiate between stratum and cluster. (10)

- (9) With a lower level of significance, the probability of rejecting a true null hypothesis
- | | |
|-----------------|------------------|
| a) Remains same | b) Increases |
| c) Decreases | d) None of these |
- (10) To test of goodness of fit, we use
- | | |
|--------------------|-----------|
| a) Chi-Square test | b) t-test |
| c) F-test | d) Z-test |
- (B) Write "T" for "True" or "F" for "False" in your Answer Book, as the case may be about the following statements:
- | | |
|---|-------|
| (1) The coefficient of Kurtosis measures the spreadness of a distribution. | (10) |
| (2) The index number for a base year is always zero. | (T/F) |
| (3) The probability of type-I error is also referred as confidence coefficient. | (T/F) |
| (4) Probability always deals with uncertainty. | (T/F) |
| (5) Hypergeometric distribution has two parameters. | (T/F) |
| (6) Equality of several variances can be tested by ANOVA. | (T/F) |
| (7) Coefficient of determination can also be determined by squaring of correlation coefficient. | (T/F) |
| (8) Poisson distribution has only one parameter that is variance. | (T/F) |
| (9) Random variable is always a real-valued function. | (T/F) |
| (10) In systematic sampling, sampling interval can be non-integer. | (T/F) |

(End)