STATISTICS

FEDERAL PUBLIC SERVICE COMMISSION
COMPETITIVE EXAMINATION FOR RECRUITMENT TO POSTS IN BPS-17 UNDER THE FEDERAL GOVERNMENT, 2009

S.No.	
R.No.	

STATISTICS

TIME ALLOWED:	**(PART-I) 30 MINUTES**	**MAXIMUM MARKS:20**
	(PART-II) 2 HOURS & 30 MINUTES	**MAXIMUM MARKS:80**

NOTE: **(i)** **First** attempt **PART-I (MCQ)** on separate **Answer Sheet** which shall be taken back after **30 minutes.**
(ii) **Overwriting/cutting of the options/answers will not be given credit.**
(iii) **Statistical Table will be provided if required**.
(iv) **Use of Scientific Calculator is allowed.**

PART – I (MCQs)

(COMPULSORY)

Q.1. Select the best option/answer and fill in the appropriate box on the Answer Sheet. (20)

(i) The probability of event given the event B is P(A/B) is equal to P(A) if B is:
(a) any event in sample S (b) sample space S
(c) $A \subset B$ (d) B is dependent on A

(ii) If an event $A = (A\cap B_1) \cup (A\cap B_2) \cup \ldots (A\cap B_n)$ and sample space $S = B_1\cup B_2\cup \ldots B_n$ and $B_i\cap B_j = \phi$, i # j, i, j = 1,2, …, n then:
(a) P (A)=1 (b) $P(A)=\sum_{i=1}^{n} P(B_i)$
(c) $P(A)=\sum_{r=1}^{n} P(A|B_i)P(B_i)$ (d) $P(A)=\sum_{i=1}^{n} P(B_i|A)$

(iii) A family has two children, then the probability of the event that atleast one of them is a boy is:
(a) $\frac{1}{4}$ (b) $\frac{1}{2}$
(c) $\frac{1}{3}$ (d) $\frac{3}{4}$

(iv) The value of $\sum_{k=0}^{n}\binom{n}{k}$ is:
(a) n^k (b) k^n (c) 2^k (d) 2^n

(v) A student is attempting to log on internet with 0.5 chance of successful attempt in each trial. The average number of attempts required to log on successfully is:
(a) 1 (b) 2 (c) 3 (d) 4

(vi) The mean of binomial random variable, with parameter probability of success is twice the probability of failure in a single trail then:
(a) greater than $\frac{2}{9}$. n (b) less than twice the variance (c) greater than $\frac{2n^2}{3}$ (d) none of these

(vii) If x follows normal distribution with polf exp $- \barwedge x^2$ then its mean and variance are:
(a) Π, Π (b) $\barwedge, \barwedge/2$ (c) $o, \frac{1}{2\barwedge}$ (d) 0, 1

(viii) Let X be the number of patients arriving at OPD on any day in a hospital according to poison distribution with the probability of at least one arrival in a day is $1-\bar{e}^{2}$. Then average number of arrivals of patients per day is:
(a) 8 (b) 4 (c) 2 (d) 1

(ix) To test the hypothesis $H_o: \mu_1 = \mu_2 = \mu_3$ at $\propto = 0.05$, then one can use:
(a) Regression Analysis (b) Analysis of Variance (c) z-test (d) t-test

STATISTICS

(x) Suppose we have random sample of size n from normal population with mean μ and variance σ^2, then maximum likelihood estimate of σ^2, when $\hat{\mu} = x$, is:

(a) $\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ (b) $\frac{1}{n}\sum_{i=1}^{n}(x - \mu)^2$ (c) $\frac{1}{n-1}\sum_{i=1}^{n}(x - \bar{x})^2$ (d) $\frac{1}{n}\sum_{i=1}^{n}(x - \mu)^2$

(xi) The probability of accepting a hypothesis when it is false is 0.2 then the probability of rejecting this hypothesis when it is false is:

(a) 0.95 (b) 0.9 (c) 0.85 (d) 0.8

(xii) If (x_1, y_1), ….. (x_n, y_n) set of n observations on Variable X = Hours studied, random variable Y = test score and Y = a + bx is the least square line that approximates the regression of test scores on the number of hours studied is given by Y = 21.819 + 3.471 X. If the desired test score is at least 60 then hours of studied should be at least:

(a) none (b) at most 10 (c) at least 10 (d) at least 11

(xiii) The inter arrival time between two messages in a communication/service system follows negative exponential distribution $2\ \bar{e}^{2x}, x > 0,$ then average inter arrival time between two messages is:

(a) 1 (b) 2 (c) $\frac{1}{2}$ (d) $\frac{1}{4}$

(xiv) In random sampling with replacement, the probability that all n specified units of a sampling fram are selected in n draws, with population size is N, is:

(a) $\frac{1}{n^N}$ (b) $\frac{1}{N^n}$ (c) $\frac{1}{n}$ (d) $\frac{1}{\binom{N}{n}}$

(xv) For population with heterogeneous groups, the suitable sampling scheme is:

(a) Simple Random Sampling (b) Systematic Sampling
(c) Cluster Sampling (d) Stratified Sampling

(xvi) The variance of x, y, z, u, v objects is:

(a) 5 (b) $\sqrt{5}$ (c) 1 (d) none of these

(xvii) For a size of size n from $N(\mu, \tilde{o}^2), \tilde{o}^2$ is unknown, H_o: $\mu = \mu_o$ against H_1: $\mu \# \mu_o$ then:

(a) t test with $n-1$ d.f. at $\propto = 0.05$ will be used
(b) F test with n d.f. at $\propto = 0.05$ will be used
(c) t test with $n-1$ d.f. at $\propto = 0.025$ will be used
(d) X^2 test with $n-1$ d.f. at $\propto = 0.025$ will be used

(xviii) Height of date trees, say, follow N(8,4) then the third moment about mean is:

(a) 3×64 (b) 4×256 (c) 0×4 (d) none of these

(xix) In a sample of size n, x are girls with variance V(x) and n–x are boys with variance is::

(a) V(x) (b) V(x) + n^2 (c) V(x) – n^2 (d) none of these

(xx) If two random variables are independent then correlation or covariance zero. If correlation or covariance between two variables X and Y is zero then:

(a) X and Y are independent of one another (b) X and Y may be independent on one another
(c) X and Y may be mutually exclusive (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) Extra attempt of any question or any part of the attempted question will not be considered.

Q.2. (a) Explain the concept of conditional probabilities using daily life events. Also justify the common formula of conditional probability of an event A given B is $P(A|B) = \frac{P(A\ and\ B)}{P(B)}$. **(4)**

(b) In answering a question on a MCQ test a student either knows the answer or guesses. Let p be the probability that student knows the answer and 1–p the probability that he/she guesses. Assume that a student who guesses at the answer will be correct with probability $\frac{1}{n}$, where n is the number of MC alternatives. What is the conditional probability that a student knew the answer to a question given that she answered it correctly? **(8)**

(c) A Laboratory blood test in 95% effective in detecting a certain disease when it is, infact, present but also yields a “false positive” result for 1% of the healthy persons tested. If 0.5 percent of the population actually has the disease, what is the probability a person has the disease given that his test result is positive. **(8)**

Q.3. If X is the amount of money (in Hundreds of rupees) that a salesman spends on gasoline during a day and Y is the corresponding amount of money (in Hundred of rupees) for which he/she is reimbursed, the joint density of these two random variables is given by $f(x, y) = K.\frac{20-x}{x}$, for $10 < x < 20, \frac{x}{2} < y < x$ and o else where. Find

(a) K (b) $f_x(x)$ (c) $f_{Y|X}(y|x=12)$

(d) the probability that the salesman will be reimbursed at least 8 units of money when spending 12 units of money. **(4×5)**

Q.4. In a certain city three T.V. channels are available. During prime time on Saturday nights Channel 1 has 50% of the viewing audience, Channel 2 has 25% of the viewing audience and Channel 3 has remaining percent of the viewing audience:

(a) Compute the probability that among 10 T.V. viewers in that city, randomly chosen on a Saturday night, 50% watching Channel 1, 30% watching Channel 2 and 20% watching Channel 3. **(10)**

(b) Calculate the average number of viewers watching Channel 1, Channel 2, Channel 3 out of 10 randomly selected. **(10)**

Q.5. The best yardstick to measure the social and moral maturity of a society is the state of its children. In a recent report titled ‘The State of Pakistan Children 2007’ the infant mortality rate at 84 per 1000 live births, under-five mortality rate is 125 per 1000 and 38% of children under five being under-weight.

(a) Construct 95% C.I. for infant mortality rate. **(5)**

(b) Construct 95% C.I. for under-five mortality rate. **(5)**

(c) Construct 95% C.I. for children under-five being under weight. **(5)**

(d) Write a brief report in the light of inferences made in (a), (b) and (c) so that non-technical person can understand. **(5)**

Q.6. (a) Define Chi-square Goodness-of-fit test with a simple example. **(8)**

(b) Mendalian theory indicates that the shape and colour of certain variety of pea ought to be grouped into 4 groups, “round and yellow,” “round and green,” “angular and yellow” and “angular and green,” according to ratio 9/3/3/1. For a sample of size n = 556 peas, the following results were obtained: Round and Yellow 315, Round and green 108, Angular and yellow 101 and Angular and green 32. Test H_o: $p_1 = \frac{9}{16}$, $p_2 = \frac{1}{3} p_1$ $p_3 = p_2$ and $p_4 = 1 - p_1 - p_2 - p_3$. **(12)**

Q.7. (a) To learn good programming techniques, two courses: C++ and C-Sharp are taught by an I.T Department of a University. The success of each course is evaluated by the scores achieved by the students in the Departments Programmers Test. Nine students using course C++ achieved an average test score of 89.6 with a sample variance of 12.96. Seven students using course C-Sharp got an average score of 81.9 with a sample variance of 161.29. Assuming all test scores are normally distributed, test 1+0: $\mu_x = \mu_y$ against H_o: $\mu_x > \mu_y$ at $\propto = 0.01$. **(8)**

(b) (i) Explain a test statistic which test the hypothesis on difference of two variances on normal populations. **(6)**

(ii) Consider part (a) of Q.7. At $\propto = 0.05$, whether it is reasonable to assume that the variance is the same for two courses mentioned above. **(6)**

Q.8. (a) Explain Systematic Sampling with an example. Compare this method of sampling with simple random sampling. **(6)**

(b) Describe the relationship of systematic sampling with Cluster Sampling **(6)**

(c) Write notes on the following terms: **(8)**

(i) Maximum Likelihood Estimation (ii) Least Squares Estimation of Regression Coefficient.

(iii) Census and Registration (iv) Bayes Theorem

****