

**Sunandan Divatia School of Science**

**MSc Applied Statistics and Analytics**

**Syllabus of Entrance Test**

**Basic Statistics** Measure of Central Tendency, Dispersion, Skewness and Kurtosis, Correlation and Regression, Index Numbers.

**Probability Theory and Distribution:** Classical (Mathematical) and Empirical Definitions of Probability and their properties. Conditional Probability, Baye's theorem and its applications, Standard Univariate Discrete Distributions, Standard Univariate Continuous Distributions, Sampling Distribution: t, F, Z Distribution and their applications.

**Statistical Inference:** Point and Interval Estimation, Testing of Hypothesis, Stochastic Processes, Markov Chains.

**Sampling Theory:** Simple Random Sampling With and Without Replacement (SRSWR/SRSWOR), Stratified Sampling, Systematic Sampling, Cluster Sampling, Two-Stage Sampling.

**Design of Experiments:** Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD), Factorial Experiment  $2^2$ ,  $3^2$ .

**Operations Research:** Linear Programming Problem (LPP) (solution to the LPP using Graphical Method, Simplex Method, Big M Method and Two-Phase Method), Duality in LPP, Queuing Theory.

**Exploratory Data Analysis:** Time Series Analysis, Multivariate Analysis (Principle Component Analysis (PCA), Factor Analysis, Discriminant Analysis).

**Sectorial Application of Statistics:** Life Science (Bioassays, Clinical Trials), Marketing Analytics (Principal of Consumer Behavior and Marketing Strategy, Key Aspects of Marketing Process), Financial Risk Analytics (Risk Management Process, Types of Risk, Risk Measurement and Analysis).

## Sample Questions

- 1) For two populations  $N(\mu_1, \sigma^2)$  and  $N(\mu_2, \sigma^2)$  with  $\sigma^2$  unknown, the test statistic for testing  $H_0: \mu_1 = \mu_2$  based on small samples with usual notations is
- (a)  $t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})}}$                       (b)  $t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$
- (c)  $t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\sigma^2(\frac{1}{n_1} + \frac{1}{n_2})}}$                       (d) any of the above
- 2) The probability of guessing at least 8 out of 10 answers correctly, on a true-false examination, is given by
- (a) 7/64    (b) 7/128  
(c) 45/1024                                        (d) 175/1024
- 3) Which of the following is not a necessary assumption underlying the use of the Analysis of variance technique?
- (a) The samples are independent and randomly selected.  
(b) The populations are normally distributed.  
(c) The variances of the populations are the same.  
(d) The means of the populations are equal.
- 4) Acceptance sampling plans are preferable due to
- (a) the economy in the inspection                      (b) protection to perishable items  
(c) increased efficiency in the inspection of items                      (d) all of the above
- 5) Which of the following is not a sampling technique?
- (a) Piggy backing                      (b) Snowball sampling                      (c) Lurk and grab  
(d) None of them -- they are all techniques used in sampling for qualitative studies
- 6) If  $T = t(X_1, X_2, \dots, X_n)$  is a sufficient statistics for a parameter  $\theta$  and an unique MLE  $\hat{\theta}$  of  $\theta$  exists, then
- (a)  $\hat{\theta} = t(X_1, X_2, \dots, X_n)$                       (b)  $\hat{\theta}$  is a function of t  
(c)  $\hat{\theta}$  is independent of t                      (d) none of the above
- 7) If  $\alpha=0.10$  for a particular test, then we are saying that
- (a) 10% is our minimum standard for acceptable probability.  
(b) 10% is the risk we take of rejecting a hypothesis that is true.  
(c) 10% is the risk we take of accepting a hypothesis that is false.  
(d) None of the above.
- 8) The purpose of an *x-bar* chart is to determine whether there has been a
- (a) Gain or loss in uniformity.  
(b) Change in the percent defective in a sample.  
(c) Change in the central tendency of the process output.

- (d) Change in the number of defects in a sample.
- 9) The acronym DMAIC stands for
- (a) Develop, Metrics, Analyze, Implement, Contain
  - (b) Define, Measure, Analyze, Improve, Control
  - (c) Define, Metrics, Analysis, Implement, Control
  - (d) None of the above
- 10) Which of the following is an unsupervised learning method?
- (a) Logistic Regression
  - (b) K means Clustering
  - (c) Discriminant Analysis
  - (d) Random Forest