STA301-STATIC
FINAL TERM SHORT NOTES
Prepared by: JUNAID MALIK

❖ Chebychev's Inequality:

If X is a random variable having mean μ and variance σ2 > 0, and k is any positive constant, then the probability that a value of X falls within k standard deviations of the mean is at least.

❖ CONTINUOUS RANDOM VARIABLE:

A random variable X is defined to be continuous if it can assume every possible value in an interval [a, b], a < b, where a and b may be –∞ and +∞ respectively.

❖ JOINT DISTRIBUTIONS:

The distribution of two or more random variables which are observed simultaneously when an experiment is performed is called their JOINT distribution.

❖ BIVARIATE PROBABILITY FUNCTION:

The joint or bivariate probability distribution consisting of all pairs of values (xi, yj).

❖ MARGINAL PROBABILITY FUNCTIONS:

The point to be understood here is that, from the joint probability function for (X, Y), we can obtain the INDIVIDUAL probability function of X and Y. Such individual probability functions are called MARGINAL probability functions.

❖ CONDITIONAL PROBABILITY FUNCTION:

Let X and Y be two discrete r.v.'s with joint probability function f(x, y). Then the conditional probability function for X given Y = y, denoted as f(x|y).

❖ INDEPENDENCE:

Two discrete r.v.'s X and Y are said to be statistically independent, if and only if, for all possible pairs of values (xi, yj) the joint probability function f(x, y) can be expressed as the *product* of the two marginal probability functions.

AL-JUNAID INSTITUTE GROUP

❖ COVARIANCE OF TWO RANDOM VARIABLES:

The covariance of two r.v.'s X and Y is a numerical measure of the extent to which their values tend to increase or decrease *together*. It is denoted by σXY or Cov (X, Y), and is defined as the expected value of the product.

❖ BINOMIAL DISTRIBUTION:

The binomial distribution is a very important discrete probability distribution. It was discovered by James Bernoulli about the year 1700.

❖ PROPERTIES OF A BINOMIAL EXPERIMENT:

- Every trial results in a success or a failure.
- The successive trials are independent.
- The probability of success, p, remains constant from trial to trial.
- The number of trials, n, is fixed in advanced.

❖ Binomial probability distribution:

A probability distribution showing the probability of x successes in n trials of a binomial experiment.

❖ Binomial probability function:

The function used to compute probabilities in a binomial experiment.

❖ Binomial experiment:

A probability experiment having the following four properties: consists of n identical trials, two outcomes (success and failure) are possible on each trial, probability of success does not change from trial to trail, and the trials are independent.

❖ PROPERTIES OF A BINOMIAL EXPERIMENT:

- Every item selected will either be defective (i.e. success) or not defective (i.e. failure)
- Every item drawn is independent of every other item
- The probability of obtaining a defective item i.e. 7% is the same (constant) for all items. (This probability figure is according to relative frequency definition of probability.

❖ **Hyper Geometric Probability function:**

The function used to compute the probability of x successes in n trials when the trials are dependent.

❖ **PROPERTIES OF HYPERGEOMETRIC EXPERIMENT:**

- The outcomes of each trial may be classified into one of two categories, success and failure.
- The probability of success changes on each trial.
- The successive trials are not independent.
- The experiment is repeated a fixed number of times.

❖ **Hyper Geometric Distribution:**

There are many experiments in which the condition of independence is violated and the probability of success does not remain constant for all trials. Such experiments are called hyper geometric experiments.

❖ **PROPERTIES OF THE HYPERGEOMETRIC DISTRIBUTION:**

- If N becomes indefinitely large, the hyper geometric probability distribution tends to the BINOMIAL probability distribution.
- The above property will be best understood with reference to the following important points:
- There are two ways of drawing a sample from a population, sampling with replacement, and sampling without replacement.
- Also, a sample can be drawn from either a finite population or an infinite population.

❖ **Poisson Probability Distribution:**

A probability distribution showing the probability of x occurrences of an event over a specified interval of time or space.

❖ **Poisson Probability Function:**

The function used to compute Poisson probabilities.

❖ PROPERTIES OF POISSON DISTRIBUTION:

- It is a limiting approximation to the binomial distribution, when p, the probability of success is very small but n, the number of trials is so large that the product np = μ is of a moderate size;
- a distribution in its own right by considering a POISSON PROCESS where events occur randomly over a specified interval of time or space or length.

❖ POISSON PROCESS:

It may be defined as a physical process governed at least in part by some random mechanism.

❖ NORMAL DISTRIBUTION:

A continuous random variable is said to be normally distributed with mean μ and standard deviation if its probability density function is given by (Formula of Normal Distribution)

❖ THE STANDARD NORMAL DISTRIBUTION:

A normal distribution whose mean is zero and whose standard deviation is 1 is known as the standard normal distribution.

❖ THE PROCESS OF STANDARDIZATION:

In other words, the standardization formula converts our normal distribution to the one whose mean is 0 and whose standard deviation is equal to 1.

❖ SAMPLING DISTRIBUTION:

The probability distribution of any statistic (such as the mean, the standard deviation, the proportion of successes in a sample, etc.) is known as its sampling distribution.

❖ CENTRAL LIMIT THEOREM:

The theorem states that:
"If a variable X from a population has mean μ and finite variance σ2, then the sampling distribution of the sample mean−X approaches a

normal distribution with mean μ and variance σ2/n as the sample size n approaches infinity."

❖ Sampling Distribution:

A probability distribution consisting of all possible values of a sample statistic.

❖ SAMPLING DISTRIBUTION OF THE SAMPLE PROPORTION:

In this regard, the first point to be noted is that, whenever the elements of a population can be classified into two categories, technically called "success" and "failure", we may be interested in the proportion of "successes" in the population.

❖ POINT ESTIMATION:

Point estimation of a population parameter provides as an estimate a single value calculated from the sample that is likely to be close in magnitude to the unknown parameter.

❖ UNBIASEDNESS:

An estimator is defined to be unbiased if the statistic used as an estimator has its expected value equal to the true value of the population parameter being estimated.

❖ CONSISTENCY:

An estimator θˆ is said to be a consistent estimator of the parameter θ if, for any arbitrarily small positive quantity e.

❖ EFFICIENCY:

An unbiased estimator is defined to be efficient if the variance of its sampling distribution is smaller than that of the sampling distribution of any other unbiased estimator of the same parameter.

❖ METHODS OF POINTESTIMATION:

- The Method of Moments
- The Method of Least Squares
- The Method of Maximum Likelihood

These methods give estimates which may differ as the methods are based on different theories of estimation.

❖ THE METHOD OF LEAST SQUARES:

The method of Least Squares, which is due to Gauss (1777-1855) and Markov (1856-1922), is based on the theory of linear estimation. It is regarded as one of the important methods of point estimation.

❖ METHOD OF MAXIMUM LIKELIHOOD:

This method was introduced in 1922 by Sir Ronald A. Fisher (1890-1962).The mathematical technique of finding Maximum Likelihood Estimators is a bit *advanced*, and involves the concept of the Likelihood Function.

❖ HYPOTHESIS-TESTING:

It is a procedure which enables us to decide on the basis of information obtained from sample data whether to accept or reject a statement or an assumption about the value of a population parameter.

❖ NULL HYPOTHESIS:

A null hypothesis, generally denoted by the symbol H0, is any hypothesis which is to be tested for possible rejection or nullification under the assumption that it is true.

❖ ALTERNATIVE HYPOTHESIS:

An alternative hypothesis is any other hypothesis which we are willing to accept when the null hypothesis H0 is rejected. It is customarily denoted by H1 or HA.

❖ TYPE-I AND TYPE-II ERRORS:

On the basis of sample information, we may reject a null hypothesis H0, when it is, in fact, true or we may accept a null hypothesis H0, when it is actually false. The probability of making a Type I error is conventionally denoted by α and that of committing a Type II error is indicated by β.

❖ TEST-STATISTIC:

A statistic (i.e. a function of the sample data not containing any parameters), which provides a basis for testing a null hypothesis, is called a test-statistic.

❖ PROPERTIES OF STUDENT'S t-DISTRIBUTION:

i) The t-distribution is bell-shaped and symmetric about the value t = 0, ranging from $-\infty$ to ∞.

ii) The number of degrees of freedom determines the shape of the t-distribution.

❖ PROPERTIES OF F-DISTRIBUTION:

1. The F-distribution is a continuous distribution ranging from zero to plus infinity.

2. The curve of the F-distribution is positively skewed.

❖ ANALYSIS OF VARIANCE (ANOVA):

It is a procedure which enables us to test the hypothesis of equality of several population means.

❖ EXPERIMENTAL DESIGN:

By an experimental design, we mean a plan used to collect the data relevant to the problem under study in such a way as to provide a basis for valid and objective inference about the stated problem. The plan usually includes:

- The selection of treatments, whose effects are to be studied,
- The specification of the experimental layout, and
- The assignment of treatments to the experimental units.

❖ SYSTEMATIC AND RANDOMIZED DESIGNS:

In this course, we will be discussing only the randomized designs, and, in this regard, it should be noted that for the randomized designs, the analysis of the collected data is carried out through the technique known as Analysis of Variance.

❖ THE COMPLETELY RANDOMIZED DESIGN (CR DESIGN):

A completely randomized (CR) design, which is the simplest type of the basic designs, may be defined as a design in which the treatments are

assigned to experimental units completely at random, i.e. the randomization is done without any restrictions.

❖ THE RANDOMIZED COMPLETE BLOCK DESIGN (RCB DESIGN):

A randomized complete block (RCB) design is the one in which

- The experimental material (which is not homogeneous overall) is divided into groups or blocks in such a manner that the experimental units within a particular block are relatively homogeneous.
- Each block contains a complete set of treatments, i.e., it constitutes a replication of treatments.
- The treatments are allocated at random to the experimental units within each block, which means the randomization is restricted. (A new randomization is made for every block.)The object of this type of arrangement is to bring the variability of the experimental material under control.

❖ Linear Square Method:

The method used to develop the estimated regression equation. It minimizes the sum of squared residuals (the deviations between the observed values of the dependent variable, yi, and the estimated values of the dependent variable, yi)

❖ Level of Significance:

Level of significance of a test is the probability used as a standard for rejecting null hypothesis Ho when Ho is assumed to be true. The level of significance acts as a basis for determining the critical region of the test.

❖ THE LEAST SIGNIFICANT DIFFERENCE (LSD) TEST:

According to this procedure, we compute the smallest difference that would be judged significant, and compare the absolute values of all differences of means with it. This smallest difference is called the least significant difference or LSD.

❖ PROPERTIES OF THE CHI-SQUARE DISTRIBUTION:

The Chi-Square ($\chi 2$) distribution has the following properties:

- It is a continuous distribution ranging from 0 to $+\infty$ .The number of degrees of freedom determines the shape of the chi-square

distribution. (Thus, there is a different chi-square distribution for each number of degrees of freedom.

- As such, it is a whole family of distributions.)
- The curve of a chi-square distribution is positively skewed. The skewness decreases as well as increases.

❖ ASSUMPTIONS OF THE CHI-SQUARE TEST OF GOODNESS OF FIT:

While applying the chi-square test of goodness of fit, certain requirements must be satisfied, three of which are as follows:

- The total number of observations (i.e. the sample size) should be at least 50.
- The expected number ei in any of the categories should not be less than 5. (So, when the expected frequency ei in any category is less than 5, we may combine this category with one or more of the other categories to get ei ≥ 5.)
- The observations in the sample or the frequencies of the categories should be independent.

❖ DEGREES OF FREEDOM:

As you will recall, when discussing the t-distribution, the chi-square distribution, and the F-distribution, it was conveyed to you that the parameters that exists in the equations of those distributions are known as degrees of freedom.

❖ P value:

The p-value is a property of the data, and it indicates "how improbable" the obtained result really is.

❖ LATEST STATISTICAL DEFINITION:

Statistics is a science of decision making for governing the state affairs. It collects analyzes, manages, monitors, interprets, evaluates and validates information. Statistics is Information Science and Information Science is Statistics. It is an applicable science as its tools are applied to all sciences including humanities and social sciences.

Question: **What is meant by A true zero point in a ratio scale?**

Answer: A value of zero does not mean the total absence of the variable under study. e.g A 0-degree temperature does not mean that there is no temperature.

Question: What are the fractals?

Answer: A fractal is that point below which a stated fraction (or decimal equivalence) of the values lie. A fractal and a percentile is same.

Question: What the effect will occur in the result of mean, median, mode, variance, and standers deviation if we add or subtract a number from the data?

Answer: mean is not effected by change of origin (addition or subtraction of any no) and change of scale(multiplication or division of any no) but variance is effected by change of origin and scale. With change of origin it will remain unchanged i-e v(X+Y)=v(X)+v(Y) or v(X-Y)=v(X)+v(Y) and when one number is multiplied or divided then it will become double. i-e v(5X)=25v(X) or v(1/5x)=1/25v(X)

Question: What is practical importance's of median and quartile or in which cases these commodities are used?

Answer: Median is one of the measure of central location. It is good to use the median when a frequency distribution involves "open end" classes like those of income and prices. In a highly skewed distribution, median is an appropriate average to use as it is not affected by extreme values. It can be located when the values are not capable of quantitative measurement. While quartiles are used when the same nature of data is to be dealt with but they are used to divide the data into four equal parts.

Question: What is the important of relation between Arithmetic, Geometric and harmonic?

Answer: Relation between arithmetic mean, geometric mean and harmonic mean is given below: Arithmetic Mean > Geometric Mean >Harmonic Mean

I.e. for a data arithmetic mean is greater than geometric mean and harmonic mean. And geometric mean is greater than harmonic mean.

Question: State what is Grouped and Row data?

Answer: Grouped data The data presented in the form of frequency distribution is also known as grouped data. Raw data Data that have not been processed in any manner. It often refers to uncompressed text that is not stored in any priority format. It may also refer to recently captured data that may have been placed into a database structure, but not yet processed.

Question: How will decide number of classes and class interval for the given data?

Answer: There are no hard and fast rules for deciding on the number of classes which actually depends on the size of data. Statistical experience tells us that no less than 5 and no more than 20 classes are generally used. Use of too many classes will defeat the purpose of condensation and too few will result in too much loss of information. Deciding on the number of classes does not depend on the value of range. In the given example no. of classes was chosen 8. It is chosen with respect the size of data. It is not decided after seeing the value of range which is 1.38 in this example. To find class interval 'h' we should first find the range and divide it by number of classes

Question: Define the Mean Deviation.

Answer: The mean deviation is used to characterize the dispersion among the measures in a given population. To calculate the mean deviation of a set of scores it is first necessary to compute their average (mean or median) and then specify the distance between each score and that mean without regard to whether the score is above or below (negative and positive) the mean. The mean deviation is defined as the mean of these absolute values.

Question: What is meant by variability?

Answer: Variability is the spread or dispersion in a set of data. Consider the following sets of data. 9, 9, 9, 9, 9, 9, 9, 9, 9, 9 10, 6, 2, 8, 4, 14, 16, 12 13, 10, 7, 6, 21, 3, 7, 5 All these three sets of data have same mean (9) but they

are different in variability. First set of values has no dispersion and there is greater variability is third data set as compared to second set of data as its values are more spread away as compared to the values of second set of data.

Question: What is EQUALLY LIKELY EVENTS?

Answer: The two events are said to be equally likely if they have the same chance of occurring. For example, in our coin-tossing experiment, the two events, heads and tails, are equally likely. Both have the same chances of occurring. There is 50% chance for occurring both events.

Question: What is the difference between dependent and independent event?

Answer: Independent and Dependent Events By independent we mean that the first event does not affect the probability of the second event. Coin tosses are independent. They cannot affect each other's probabilities; the probability of each toss is independent of a previous toss and will always be 1/2. Separate drawings from a deck of cards are independent events if you put the cards back. An example of a dependent event, one in which the probability of the second event is affected by the first, is drawing a card from a deck but not returning it. By not returning the card, you've decreased the number of cards in the deck by 1, and you've decreased the number of whatever kind of card you drew. If you draw an ace of spades, there are 1 fewer aces and 1 fewer spades. This affects our simple probability: (number of favorable outcomes)/ (total number of outcomes. This type of probability is formulated as follows: If A and B are not independent, then the probability of A and B is P(A and B) = P(A) × P(B|A) where P(B|A) is the conditional probability of B given A. Example If someone draws a card at random from a deck and then, without replacing the first card, draws a second card, what is the probability that both cards will be aces? Solution Event A is that the first card is an ace. Since 4 of the 52 cards are aces, P(A) = 4/52 = 1/13. Given that the first card is an ace, what is the probability that the second card will be an ace as well? Of the 51 remaining cards, 3 are aces. Therefore, p(B|A) = 3/51 = 1/17, and the probability of A and B is 1/13 × 1/17 = 1/221.

Question: Explain the Conditional Probability.

Answer: Conditional Probability In many situations, once more information becomes available; we are able to revise our estimates for the probability of further outcomes or events happening. For example, suppose you go out for lunch at the same place and time every Friday and you are served lunch within 15 minutes with probability 0.9. However, given that you notice that the restaurant is exceptionally busy, the probability of being served lunch within 15 minutes may reduce to 0.7. This is the conditional probability of being served lunch within 15 minutes given that the restaurant is exceptionally busy. The usual notation for "event A occurs given that event B has occurred" is "A | B" (A given B). The symbol | is a vertical line and does not imply division. P(A | B) denotes the probability that event A will occur given that event B has occurred already. A rule that can be used to determine a conditional probability from unconditional probabilities is: P(A|B) = P(AÇB)/P(B) Where: P(A | B) = the (conditional) probability that event A will occur given that event B has occurred already P(AÇB) = the (unconditional) probability that event A and event B both occur P(B) = the (unconditional) probability that event B occurs

Question: What is the difference between mutually exclusive events & exhaustive events with the help of an example?

Answer: If only one of two or more events can occur, the events are called mutually exclusive events. For example, in our coin-tossing experiment, the two events, heads and tails, are mutually exclusive: if one occurs, the other cannot occur. When a set of events for an experiment includes every possible outcome, the set is said to be collectively exhaustive. Thus, heads and tails are a collectively exhaustive set of events for our coin-tossing experiment. One requirement we place on probability numbers is that the sum of the profanities for a collectively exhaustive set of mutually exclusive events be equal to 1.

Question: Explain the concept of "Central limit theorem".

Answer: The central limit theorem states that given a distribution with a mean μ and variance s2, the sampling distribution of the mean approaches a

normal distribution with a mean (μ) and a variance s2/n as 'n', the sample size, increases. The amazing thing about the central limit theorem is that no matter what the shape of the original distribution, the sampling distribution of the mean approaches a normal distribution. Furthermore, for most distributions, a normal distribution is approached very quickly as 'n' increases.

Question: Explain "P -value"? with examples.

Answer: Each statistical test has an associated null hypothesis, the p-value is the probability that your sample could have been drawn from the population(s) being tested (or that a more improbable sample could be drawn) given the assumption that the null hypothesis is true. A p-value of .05, for example, indicates that you would have only a 5% chance of drawing the sample being tested if the null hypothesis was actually true.

Question: What is the difference between average and Central tendency?

Answer: A single value used to represent the distribution is called average. Most commonly used averages are Mean, Median and Mode. And measures of dispersion are used to measure how the data are dispersed about the average. For example, it is quite possible that two or more sets of data may have the same average (mean, median and mode) but their individual observations may differ considerably from the average. Thus the value of central tendency does not describe data. So we therefore need some additional information concerning with how the how the data are dispersed about the average. There are several measures of dispersion, the most common being the range, quartile deviation, mean deviation and standard deviation.

Question: What is the coefficient of variation?

Answer: Co-efficient of variation is used to compare the variability and to check the consistency of two or more series. It is most commonly used relative measure of dispersion. Symbolically, the coefficient of variation, denoted by C.V., is given by C.V = [Standard deviation / Arithmetic mean] × 100 It is used as a criterion of consistent performance; the smaller the coefficient of variation, the more consistent is the performance It is also used

as the criterion of variability; the larger the coefficient of variation, the more variability in the data.

Question: Define mean deviation from median.

Answer: Mean deviation from median deviation is defined as the average of the deviations of the values from median; the deviations are taken without considering algebraic signs. The median deviation of a set of n values X1, X2, Xn, denoted by M.D., is given by M.D = summation | X – median | /n Where | X – median | indicate the absolute deviations of the observations from the median of a sample

Question: Which line is best fitted line, how we can judge?

Answer: According to the principal of least squares, the best-fitting line to a set of points is the one for which the sum of the squares of the vertical distances between the points and the line is minimum

Question: What are Covariance & Correlation?

Answer: The Covariance of two r.v.'s X and Y is a numerical measure of the extent to which their values tend to increase or decrease together. The correlation is used to describe the degree to which one variable is linearly related to another. Often, correlation is used in conjunction with regression to measure how well the regression line explains the variation of the dependent variable; Y. Correlation can also be used by itself, however, to measure the degree of association between two variables. Statisticians have developed measure for describing the correlation between two variables: the coefficient of correlation.