End Course Summative Assignment

Problem Statement: Write the Solutions to the Top 50 Interview Questions and

Explain any 5 Questions in a Video

Imagine you are a dedicated student aspiring to excel in job interviews. Your task is to write the solutions for any 50 interview questions out of 80 total questions presented to you. Additionally, create an engaging video where you thoroughly explain the answers to any five of these questions.

Your solutions should be concise, well-structured, and effective in showcasing your problem-solving skills. In the video, use a dynamic approach to clarify the chosen questions, ensuring your explanations are easily comprehensible for a broad audience.

Note:

1. Make a copy of this document and write your answers.
2. Include the Video Link here in your document before submitting.

GitHub Link - https://github.com/mishra2022/ALMA-Module2/blob/main/statistics.pdf

1. **What is a vector in mathematics?**

A vector is a mathematical quantity with both magnitude and direction.

The magnitude of vector u, denoted |u|, is it's length.

One way to express its direction is to give the angle it makes with a horizontal ray (that points to the right) that is parallel to the positive x-axis. This is called the vector's direction angle.

2. **How is a vector different from a scalar?**

A scalar quantity has only magnitude, but no direction. Vector quantity has both

magnitude and direction. Every scalar quantity is one-dimensional. Vector quantity can be one, two or three-dimensional.

3. **What are the different operations that can be performed on vectors?**

Different operations performed on vector operations that include,

- Addition of Two Vectors
- Subtraction of Two Vectors
- Multiplication of Vector with Scalar
- Product of Two Vectors
    - Dot Product
    - Cross-Product

4. How can vectors be multiplied by a scalar?

When a vector is multiplied by a scalar quantity, then the magnitude of the vector changes in accordance with the magnitude of the scalar but the direction of the vector remains unchanged.

5. What is the magnitude of a vector?

The magnitude of a vector is the length of the vector. The magnitude of the vector a is denoted as ‖a‖.

For a two-dimensional vector a=(a1,a2), the formula for its magnitude is

$$\|a\|=\sqrt{(a1+a2)}.$$

6. How can the direction of a vector be determined?

We know that the slope of a line that passes through the origin and a point (x, y) is y/x. We also know that if θ is the angle made by this line, then its slope is tan θ, i.e., tan θ = y/x. Hence, $\theta = \tan^{-1}(y/x)$. So, the direction of a vector (x, y) is found using the formula $\tan^{-1}(y/x)$

7. What is the difference between a square matrix and a rectangular matrix?

A square matrix has the same number of rows as columns. Square matrices are used for transformations.
A rectangular matrix is one where the number of rows or columns may not be the same.

8. What is a basis in linear algebra?

A set B of vectors in a vector space V is called a basis if every element of V may be written in a unique way as a finite linear combination of elements of B.

9. What is a linear transformation in linear algebra?

A linear transformation is a function from one vector space to another that respects the underlying (linear) structure of each vector space. A linear transformation is also known as a linear operator or map.

10. What is an Eigen vector in linear algebra?

The eigenvector is a vector that is associated with a set of linear equations. The eigenvector of a matrix is also known as a latent vector, proper vector, or characteristic vector. These are defined in the reference of a square matrix.

11. What is gradient in Machine learning?

A gradient simply measures the change in all weights with regard to the change in error.

12. What is back propagation in machine learning?

The practice of fine-tuning the weights of a neural net based on the error rate (i.e. loss) obtained in the previous epoch (i.e. iteration.)

13. What is concepts of derivatives in calculus?

Derivative is the rate of change of a function with respect to a variable.

14. How are the partial derivatives used in machine learning?

Partial derivatives and gradient vectors are used very often in machine learning algorithms for finding the minimum or maximum of a function. Gradient vectors are used in the training of neural networks, logistic regression, and many other classification and regression problems.

15. What is probability theory?

Probability theory studied random events and tells us about their occurrence. The two main approaches for studying probability theory are.

1. Theoretical Probability
2. Experimental Probability

16. What are the primary components of probability theory?

The probability space associated with a random experiment is determined by three components: the outcome space Ω whose element ω is an outcome of the experiment, a collection of events F whose elements are subsets of Ω, and a probability measure IP assigned to the elements in F.

17. What is conditional probability, and how is it calculated?

Conditional probability is defined as the likelihood of an event or outcome occurring, based on the occurrence of a previous event or outcome.

$P(A|B) = P(A \cap B) / P(B)$

Where,

$P(A \cap B)$ represents the probability of both events A and B occurring simultaneously, and
$P(B)$ represents the probability of event B occurring.

18. What is Bayes theorem, and how is it used?

The general statement of Bayes' theorem is "The conditional probability of an event A, given the occurrence of another event B, is equal to the product of the event of B, given A and the probability of A divided by the probability of event B."

$P(A|B) = P(B|A)P(A) / P(B)$

where,
$P(A)$ and $P(B)$ are the probabilities of events A and B
$P(A|B)$ is the probability of event A when event B happens
$P(B|A)$ is the probability of event B when A happens

Example: Bayes' theorem defines the accuracy of the medical test by taking into account how likely a person is to have a disease and what is the overall accuracy of the test.

19. What is random variable, and how is it different from a regular variable?

20. What is the law of large numbers, and how does it relate to probability theory?

The law of large numbers, in probability and statistics, states that as a sample size grows, its mean gets closer to the average of the whole population.

This is due to the sample being more representative of the population as the sample become larger.

21. What is the central limit theorem, and how is it used?

The CLT is a statistical theory that states that - if you take a sufficiently large sample size from a population with a finite level of variance, the mean of all samples from that population will be roughly equal to the population mean.

22. What is the difference between discrete and continuous probability distributions?

Discrete distribution is a probability distribution where the random variable can only take on a finite or countable number of values.
In contrast, continuous distribution refers to a probability distribution where the random variable can take on any value within a certain range or interval.

23. What are some common measures of central tendency, and how are they calculated?

The mean, median and mode are the three commonly used measures of central tendency.
The mean (average) of a data set is found by adding all numbers in the data set and then dividing by the number of values in the set.
The median is the middle value when a data set is ordered from least to greatest.

24. What is the purpose of using percentiles and quartiles in data summarization?

The mode is the number that occurs most often in a data set.
Quartiles and percentiles are measures of variation, which describes how spread out the data is.
Quartiles and percentiles are both types of quantiles.
Percentiles are values that separate the data into 100 equal parts.
For example, The 95th percentile separates the lowest 95% of the values from the top 5%.

25. How do you detect and treat outliers in a dataset?
26. How do you use theorem to approximate a discrete probability distributions?
27. How do you test the goodness of fit of a discrete of a discrete probability distributions?

There are multiple types of goodness-of-fit tests, but the most common is the chi-square test.
The chi-square test determines if a relationship exists between categorical data.
The Kolmogorov-Smirnov test determines whether a sample comes from a specific distribution of a population.

28. What is a joint probability distribution?

A joint probability distribution represents a probability distribution for two or

more random variables. Instead of events being labelled A and B, the condition is to use X and Y as given below. f(x,y) = P(X = x, Y = y) The main purpose of this is to look for a relationship between two variables.

29. How do you calculate the joint probability distribution?

Probabilities are combined using multiplication, therefore the joint probability of independent events is calculated as the probability of event A multiplied by the probability of event B. This can be stated formally as follows: Joint Probability: P(A and B) = P(A) * P(B).

30. What is difference between a joint probability distribution and probability distributions?

A joint probability distribution represents a probability distribution for two or more random variables. Instead of events being labelled A and B, the condition is to use X and Y as given below. This table can be used to find the probabilities of events. Example: Find the probability of X = 3 and Y = 3.

31. What is the covariance of a joint probability distribution?

Covariance is a statistical measure of the relationship between two random variables. It is used to determine how much two variables change together and to what extent they are related**.**

32. How do you determine if two random variables are independent based on their joint probability distributions?

Two random variables are independent if they convey no information about each other and, as a consequence, receiving information about one of the two does not change our assessment of the probability distribution of the other.

33. What is the relationship between the correlation coefficient and the covariance of a joint probability distributions?

The correlation coefficient normalizes the covariance by dividing by the geometric mean of the total variances for the two random variables.

34. What is sampling in statistics, and why is it important?

In statistics, a sample is an analytic subset of a larger population. The use of samples allows researchers to conduct their studies with more manageable data and in a timely manner.

35. What are the different sampling methods commonly used in statistical inference?

There are several types of sampling methods that can be used to collect data. These include:
Random sampling: selecting a person to interview or site to measure, at random.
Stratified sampling: dividing the population into subgroups and then selecting a sample from each subgroup.
Systematic sampling: collecting data in an ordered or regular way, eg every 5 metres or every fifth person.
Non-random sampling: involves non-random selection based on convenience or other criteria, allowing you to easily collect data.
Capture recapture: a method used to estimate the size of a population by capturing and marking a sample of individuals, then recapturing them later and counting how many are marked

36. What is the central limit theorem, and why is it important in statistical inference?
Refer to 21.
37. What is the diffrence between parameter estimation and hypothesis testing ?

In statistics, estimation and hypothesis testing are two fundamental concepts that are used to make inferences about a population based on a sample.
Estimation is the process of determining the value of an unknown population parameter based on a sample statistic. It involves calculating a point estimate or an interval estimate for the population parameter. A point estimate is a single value that is used to estimate the population parameter, while an interval estimate is a range of values that is likely to contain the population parameter with a certain level of confidence 1.
Hypothesis testing is a statistical method used to determine whether a hypothesis about a population parameter is supported by the sample data. It involves setting up two hypotheses, the null hypothesis and the alternative hypothesis, and using the sample data to determine which hypothesis is more likely to be true. The null hypothesis is the hypothesis that there is no significant difference between a specified population parameter and a hypothesized value, while the alternative hypothesis is the hypothesis that there is a significant difference between the population parameter and the hypothesized value
In summary, estimation is used to determine the value of an unknown population parameter based on a sample statistic, while hypothesis testing is used to determine whether a hypothesis about a population parameter is supported by the sample data.

38. What is the p-value in hypothesis testing?

The p value, or probability value, tells us how likely it is that your data could have occurred under the null hypothesis. It does this by calculating the likelihood of your test statistic, which is the number calculated by a statistical test using your data.

39. What is confidence internal estimation?

A confidence interval is the mean of your estimate plus and minus the variation in that estimate. This is the range of values you expect your estimate to fall between if you redo your test, within a certain level of confidence

40. What are Type I and Type II errors in hypothesis testing?

Type-I Error	Type-II Error
It is a false positive conclusion.	It is a false negative conclusion.
The consequences of this error is lower ...	The consequences of this error is compar ...
It is a situation when we reject null hy ...	It is a situation where we accept null h ...
Here we fail to accept null hypothesis.	Here we fail to reject null hypothesis.

41. What is the difference between correlation and causation?

A **correlation** between variables, however, does not automatically mean that the change in one variable is the cause of the change in the values of the other variable.

Causation indicates that one event is the result of the occurrence of the other event; i.e. there is a causal relationship between the two events.

42. How is the confidence interval defined in statistics?

A confidence interval is the mean of your estimate plus and minus the variation in that estimate. This is the range of values you expect your estimate to fall between if you redo your test, within a certain level of confidence. Confidence, in statistics, is another way to describe probability.

43. What is the confidence level represent in a confidence interval?

Confidence, in statistics, is another way to describe probability. For example, if you construct a confidence interval with a 95% confidence level, you are confident that 95 out of 100 times the estimate will fall between the upper and lower values specified by the confidence interval.

44. What is hypothesis testing in statistics?

A statistical hypothesis test is a method of statistical inference used to decide

whether the data at hand sufficiently support a particular hypothesis. Hypothesis testing allows us to make probabilistic statements about population parameters

45. What is the purpose of a null hypothesis in hypothesis testing?

The null hypothesis is a kind of hypothesis which explains the population parameter whose purpose is to test the validity of the given experimental data. This hypothesis is either rejected or not rejected based on the viability of the given population or sample.

46. What is difference between one-tailed and two-tailed test?

One-tailed tests allow for the possibility of an effect in one direction. Two-tailed tests test for the possibility of an effect in two directions—positive and negative.

47. What is experiment design, and why is it important?

Design of experiments, also called experimental design, is a structured and organized way of conducting and analyzing controlled tests to evaluate the factors that are affecting a response variable.

48. What are key elements to consider when designing an experiment?

Key elements of a well-designed experiment

- Ensure your experiment is unbiased.
- Make sure your experiment is adequately powered.
- Consider the range of applicability of your experiment.
- Simplify your experiment.
- Indicate the uncertainty in your results.
- References and further reading.

49. How can sample size determination affect of experiment design?

The smaller the difference or effect size you expect, the larger the sample size that will be required. Sample size increases proportionally to the variance. The larger the uncertainty of the outcome measurement (variability of a result), the larger the sample size must be

50. What are some strategies to mitigate potentials source of bias in experiment design?

Best practices for minimizing bias in experimental procedures, including: blinding; systematic random sampling; inclusion of positive and negative controls; and methods of quality control for reliability and reproducibility.

51. What is the geometric interpretation of the dot product?

The dot product has the following geometric interpretation: Let α be the angle between A and B. Then $A \cdot B = |A| \cdot |B| \cdot \cos(\alpha)$. A number of important properties of the dot product should be noted. Most of these are obvious consequences, either of the definition or of the above geometric formula or both.

52. What is the geometric interpretation of the cross product?

The cross product a × b is defined as a vector c that is perpendicular (orthogonal) to both a and b, with a direction given by the right-hand rule and a magnitude equal to the area of the parallelogram that the vectors span.

53. How optimization algorithm with calculus used in deep learning models?

Multivariate Calculus: This is used for advanced optimization algorithms, such as stochastic gradient descent, which are used in deep learning algorithms to optimize models with a large number of parameters. Optimization Theory: This is used for designing algorithms that find the best solution to a problem.

54. What are observational and experimental data in statistics?

55. How are confidence tests and hypothesis tests similar? How are they different?

56. What is the left-skewed distribution and the right-skewed distribution?

57. What is Bessel's correction?

58. What is kurtosis?

59. What is the probability of throwing two fair dice when the sum is 5 and 8?

60. What is the difference between Descriptive and Inferential Statistics?

61. Imagine that Jeremy took part in an examination. The test has a mean score of 160, and it has a standard deviation of 15. If Jeremy's z-score is 1.20, what would be his score on the test?

62. In an observation, there is a high correlation between the time person sleeps and amount of productive work he does. What can be inferred from this?

63. What is the meaning of degrees of freedom (DF) in statistics?

64. If there is a 30 percent probability that you will see a supercar in any 20-minute time interval, what is the probability that you see at least one supercar in the period of an hour (60 minutes)?

65. What is the empirical rule in Statistics?

66. What is the relationship between sample size and power in hypothesis testing?

67. Can you perform hypothesis testing with non-parametric methods?

68. What factors affect the width of a confidence interval?

69. How does increasing the confidence level affect the width of a confidence interval?

70. Can a confidence interval be used to make a definitive statement about a specific individual in the population?

71. How does sample size influence the width of a confidence interval?

72. What is the relationship between the margin of error and confidence interval?

73. Can two confidence intervals with different widths have the same confidence level?

74. What is a Sampling Error and how can it be reduced?

75. What is a Chi-Square test?

76. What is a t-test?

77. What is the ANOVA test?

78. How is hypothesis testing utilised in A/B testing for marketing campaigns?

79. What is the difference between one-tailed and two tailed t-tests?

80. What is an inlier?