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.

1. What is a vector in mathematics?

A vector is **a quantity or phenomenon that has two independent properties: magnitude and direction**. The term also denotes the mathematical or geometrical representation of such a quantity. Examples of vectors in nature are velocity, momentum, force, electromagnetic fields, and weight.

2. How is a vector different from a scalar?

Scalars are quantities that only have magnitude (or size), while **vectors have both magnitude and direction**. Explore some examples of scalars and vectors, including distance, displacement, speed, and velocity.

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

Operations on vectors are the vectors that are performed especially on vector quantities, such quantities have both magnitude and direction, and operating them with

normal rules of mathematics is not possible. So we have to use various 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?

Multiplying vectors can be done in two forms namely dot product and cross product. If a vector is multiplied by a scalar it means that the magnitude of a vector is multiplied by a number.

5. What is the magnitude of a vector?

Magnitude of a vector is defined as the length of the vector. As the magnitude of the vector denotes the length of the vector it is always positive. For any vector A its magnitude is represented as |A|. Suppose a vector is defined as xi + yj then its magnitude is defined as the square root of the sum of squares of the individual terms. The magnitude of the vector represents the length of the vector i.e. the value or impact the vector has.

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

The direction of a vector is the angle made by the vector with the horizontal axis, that is, the X-axis. **The direction of a vector is given by the counterclockwise rotation of the angle of the vector about its tail due east**.

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

A square matrix is a matrix that contains the same number of rows and the same number of columns. **If a matrix is not a square matrix, then it is known as a**

rectangular matrix. We can also say that the matrices which have different numbers of rows and columns are called rectangular matrices.

8. What is a basis in linear algebra?

In mathematics, **a set B of vectors in a vector space V** is called a basis (pl. : bases) if every element of V may be written in a unique way as a finite linear combination of elements of B. The coefficients of this linear combination are referred to as components or coordinates of the vector with respect to B.

9. What is a linear transformation in linear algebra?

A linear transformation is **also known as a linear operator or map**. The range of the transformation may be the same as the domain, and when that happens, the transformation is known as an endomorphism or, if invertible, an automorphism. The two vector spaces must have the same underlying field.

10. What is an eigenvector in linear algebra?

In linear algebra, an eigenvector, or characteristic vector of a linear transformation is **a non-zero vector that changes at most by a scalar factor when that linear transformation is applied to it**. The corresponding eigenvalue, denoted by the symbol λ, is the factor by which the eigenvector is scaled.

11. What is the gradient in machine learning?

Gradient descent is an optimization algorithm that's used when training a machine learning model and is based on a convex function and tweaks its parameters iteratively to minimize a given function to its local minimum (that is, **slope = 0**)

12. What is backpropagation in machine learning?

Backpropagation is **a supervised learning algorithm used for training artificial neural networks**. It involves computing gradients of the loss function with respect to the model's parameters and updating the parameters accordingly.

13. What is the concept of a derivative in calculus?

Differential calculus is **a procedure for finding the exact derivative directly from the for- mula of the function, without having to use graphical methods**. In practise we

use a few rules that tell us how to find the derivative of almost any function that we are likely to encounter.

14. How are partial derivatives used in machine learning?

How are 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, **a branch of mathematics concerned with the analysis of random phenomena**. The outcome of a random event cannot be determined before it occurs, but it may be any one of several possible outcomes. The actual outcome is considered to be determined by chance.

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 **the probability of an event given that another event has occurred**. It is calculated as the probability of both events happening divided by the probability of the event that has already occurred.

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

Bayes' theorem can be used **to calculate the probability of a disease given a positive test result, also known as the post-test probability**. This is particularly useful in medical testing where false positives are common. The theorem formula is P(A|B) = [P(B|A)*P(A)] / P(B).

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

A Random Variable is different from the variable in algebra as **it has whole set of values and it can take any of those randomly**. Variable used in algebra cannot have more than a single value at a time. Then X could be 0, 1, 2 or 3 randomly where each of them might have a different probability.1

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 Central Limit Theorem **allows us to do so based on just ONE set of random samples if we were to take random samples over and over again**. Moreover, the property of the Central Limit Theorem tells us, as the sample size increases, our estimate of the population mean will be more precise with a smaller variability.

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 measure of central tendency is a single value that describes(represents) the central position within the dataset. Three most common measures of central tendency are **Mean, Median, and Mode**

To find central tendency, **calculate the mean by summing all values and dividing by the total number, find the median by locating the middle value, or determine the mode as the most commonly occurring value**.

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

Percentiles are used **to compare and interpret data**. For example, an observation at the 50th percentile would be greater than 50 percent of the other obeservations in the set. Quartiles divide data into quarters.

25. How do you detect and treat outliers in a dataset?

Box plots are a graphical depiction of numerical data through their quantiles. It is a very simple but effective way to visualize outliers. Think about the lower and upper whiskers as the boundaries of the data distribution. Any data points that show above or below the whiskers, can be considered outliers or anomalous.

One method is to **remove outliers as a means of trimming the data set**. Another method involves replacing the values of outliers or reducing the influence of outliers through outlier weight adjustments. The third method is used to estimate the values of outliers using robust techniques.

26. How do you use the central limit theorem to approximate a discrete probability distribution?

The central limit theorem **can be applied to sums of discrete random variables** as well as continuous random variables.

The central limit theorem states that under very weak conditions (almost all probability distributions you will see will satisfy them) the sum of n i.i.d. random variables, Sn , will converge, appropriately normalised to a standard Normal distribution as $n \to \infty$ $n \to \infty$.

27. How do you test the goodness of fit of a discrete probability distribution?

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

29. How do you calculate the joint probability distribution?

We consider the two events to be A and B. A is the occurrence of five on the first roll, whereas B is the occurrence of five on the second roll. The probability of event B,

P(B)=16. Therefore the Joint Probability is, **P(A∩B)=P(A)×P(B)**=16×16=136.

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

Joint probability is the probability of two events occurring simultaneously. Marginal probability is the probability of an event irrespective of the outcome of another variable. Conditional probability is the probability of one event occurring in the presence of a second event

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

The covariance of X and Y for the given joint probability distribution is: **X/Y**. **0**.

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

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

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

Sampling is **a process in statistical analysis where researchers take a predetermined number of observations from a larger population**. Sampling allows researchers to conduct studies about a large group by using a small portion of the population.

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

There are five types of sampling: **Random, Systematic, Convenience, Cluster, and Stratified**. Random sampling is analogous to putting everyone's name into a hat and drawing out several names.

36. What is the central limit theorem, and why is it important in statistical inference?

The Central Limit Theorem (CLT) is a fundamental concept in statistics and probability theory. It states that when independent random variables are added together, their sum

tends to follow a normal distribution (also known as a Gaussian distribution) regardless of the original distribution of the individual variables.

The central limit theorem is useful when analyzing large data sets because **it allows one to assume that the sampling distribution of the mean will be normally-distributed in most cases**. This allows for easier statistical analysis and inference.

37. What is the difference between parameter estimation and hypothesis testing?

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

The P value is defined as **the probability under the assumption of no effect or no difference (null hypothesis), of obtaining a result equal to or more extreme than what was actually observed**. The P stands for probability and measures how likely it is that any observed difference between groups is due to chance.

39. What is confidence interval estimation?

For both continuous and dichotomous variables, the confidence interval estimate (CI) is **a range of likely values for the population parameter based on: the point estimate, e.g., the sample mean**. **the investigator's desired level of confidence** (most commonly 95%, but any level between 0-100% can be selected)

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

For both continuous and dichotomous variables, the confidence interval estimate (CI) is **a range of likely values for the population parameter based on: the point estimate, e.g., the sample mean**. **the investigator's desired level of confidence** (most commonly 95%, but any level between 0-100% can be selected)

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 a confidence interval defined in statistics?

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

44. What is hypothesis testing in statistics?

Hypothesis testing is **a form of statistical inference that uses data from a sample to draw conclusions about a population parameter or a population probability distribution**. First, a tentative assumption is made about the parameter or distribution.

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

Null hypothesis testing is a formal approach to **deciding whether a statistical relationship in a sample reflects a real relationship in the population or is just due to chance**.

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

The main difference between one-tailed and two-tailed tests is that **one-tailed tests will only have one critical region whereas two-tailed tests will have two critical regions**. If we require a 100(1−α) 100 (1 − α) % confidence interval we have to make some adjustments when using a two-tailed test.

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

A proper experimental design **serves as a road map to the study methods, helping readers to understand more clearly how the data were obtained and, therefore, assisting them in properly analyzing the results**.

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

True experiments have four elements: **manipulation, control , random assignment, and random selection**. The most important of these elements are manipulation and control. Manipulation means that something is purposefully changed by the researcher in the environment.

49. How can sample size determination affect 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 potential sources of bias in experiment design?

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

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

53. How are optimization algorithms with calculus used in training deep learning models?

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?

A right-skewed distribution is longer on the right side of its peak than on its left. Left skew (also called negative skew). A left-skewed distribution is longer on the left side of its peak than on its right. Zero skew. It is symmetrical and its left and right sides are mirror images.

57. What is Bessel’s correction?

Bessel's Correction is a correction applied while calculating the sample variance and sample standard deviation where the denominator is (N-1) instead of N, where N is the sample size or the number of observations in the sample. This is done in order to correct the bias in the estimation of population variance (and standard deviations).

58. What is kurtosis?

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

The probability of obtaining a sum of 5 on the first roll and a sum of 8 on the second roll is (1/36) * (1/36) = **1/1296**. You can get a sum of 5 by rolling (1,4), (2,3), (3,2), (4,1), and a total of 8 by rolling (2,6), (3,5), (4,4), (5,3), (6,2).

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

Descriptive statistics summarize and describe data, while inferential statistics draw conclusions and make predictions about populations based on sample data.

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 a person sleeps and the amount of productive work he does. What can be inferred from this?

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

Degrees of freedom, often represented by v or df, is **the number of independent pieces of information used to calculate a statistic**. It's calculated as the sample size minus the number of restrictions.

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?

Statistical power is positively correlated with the sample size, which means that given the level of the other factors viz. alpha and minimum detectable difference, a larger sample size gives greater power.

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

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

The confidence interval is based on the margin of error. There are three factors that determine the size of the confidence interval for a given confidence level. These are: **sample size, percentage and population size**. The larger your sample, the more sure you can be that their answers truly reflect the population.

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?

Increasing the sample size decreases the width of confidence intervals, because it decreases the standard error.

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?