|  |
| --- |
| Both standard deviation and mean deviation are used to measure of dispersion. Write the calculating formula of them. Why do most researcher use standard deviation to measure dispersion of data? |
| SUMMIT, a Power company need information about customer usage to obtain accurate forecasts of demands. Investigators from Summit Power determined energy consumption (BTUs) during a particular period for a sample of 90 electric heater homes.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Class | 1 – <3 | 3 – <5 | 5 – <7 | 7 – <9 | 9 – <11 | | Frequency | 1 | 1 | 11 | 21 | 25 |  |  |  |  |  |  | | --- | --- | --- | --- | --- | | Class | 11 – 13 | 13 – <15 | 15 – <17 | 17 – <19 | | Frequency | 17 | 9 | 4 | 1 |     Using mid-point of the class interval calculate mean, standard deviation. Is this the actual measures of location and dispersion of sample data? -Justify your answer. |
| Draw a histogram of the data using relative frequency on the vertical scale, and comment on its features. |
| Why do most of the researchers use standard deviation as a measure of dispersion? |

|  |
| --- |
| The main two branches of statistics are descriptive statistics and inferential statistics- give examples of the use of these. |
| Calculate mean, median, mode, and standard deviation, from the data given below and comments on the result.   |  |  |  |  | | --- | --- | --- | --- | | 5 | 2 | 7 | 5 | | 8 | 4 | 3 | 10 | | 8 | 8 | 4 | 5 | |
| What do you mean by 3rd quartile of data? Suppose 90th percentile of a dataset is 7.5 . what will be your explanation? |
| Suppose the quartiles of a data set are 20, 45, 50. What can be guessed from these quartiles about the data? |
|  |
| What do you mean by probability and explain different types of probability with example |
| A box contains **two** 40-W bulbs, **three** 60-W bulbs, and **four** 75-W bulbs. If two bulbs are selected one by one in random order,   1. Write the sample space of this experiment. 2. What is the probability that both the bulbs are rated 75 W? |
| Two balls are drawn at random in succession without replacement from an urn containing 3 red balls and 5 black balls and 2 white balls.   1. Find the probabilities of all possible outcomes. 2. If we denote the event having red ball as A and having black or white ball as B, assume that ball in the first draw is ‘not red’ and it’s probability is p(B) = 7/10, what is the probability of p(White / B). |
| What is the difference between Bernoulli and Binomial distribution? What are we assume before using Binomial distribution? |
| Write an application of using binomial probability distribution? |
| Consider whether the next person buying a computer at a university book store buys a laptop or a desktop model. Let X=1 if the customer purchases a laptop computer and X= 0 if the customer purchases a desktop computer. Let P(X=1)=0.6 and P(X=0)=0.4. What is the probability that among the next 5 customers 2 will buy laptop? |
| Among all continuous probability distribution, why Standard Normal Distribution is most useful distribution? Write some of its properties. |
| Why does exponential distribution called a memoryless distribution? |
| A machine that produces ball bearings has initially been set so that the true average diameter of the bearings it produces is .500 inch. A bearing is acceptable if its diameter is within .004 inch of this target value. Suppose, however, that the setting has changed during the course of production, so that the bearings have normally distributed diameters with mean value .499 in. and standard deviation .002 in. What percentage of the bearings produced will not be acceptable? |
|  |
| When do we use hypothesis? Define null hypothesis, critical value, and p-value. |
| 1. What are the steps of testing a hypothesis? |
| Low-back pain (LBP) is a serious health problem in many industrial settings. After a survey among the workers in a factory, it was reported that summery of Isodynamic Evaluation of Trunk Muscles and Low-Back Pain are as follows:   |  |  |  |  | | --- | --- | --- | --- | |  | Density of Trunk Muscle | | | | Condition | Sample | Mean | S.D. | | LBP | 28 | 91.5 | 5.5 | | No LBP | 31 | 88.3 | 7.8 |   Does it appear that true average muscle density are higher among the workers who do not have LBP?   1. Formulate the hypothesis to test the belief. 2. If the test statistic value is 1.97, what will be your comment? Evaluate p-value from the printed table you carrying. |
|  |
| Why do we use correlation between two variable? What does a regression model reveal actually? |
| When do we use rank correlation? |
| A study to assess the capability of subsurface flow wetland systems to remove biochemical oxygen demand (BOD) and various other chemical constituents resulted in the accompanying data on X = BOD mass loading (kg/ha/d) and Y = BOD mass removal (kg/ha/d):   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | X | 3 | 8 | 10 | 11 | 13 | 16 | 27 | 30 | 35 | 37 | 38 | 44 | 103 | 42 | | Y | 4 | 7 | 8 | 8 | 10 | 11 | 16 | 26 | 21 | 9 | 31 | 30 | 75 | 30 |   Construct a scatter plot of the data, and comment on any interesting features. Also construct simple linear regression model to predict BOD mass removal using BOD mass loading. |