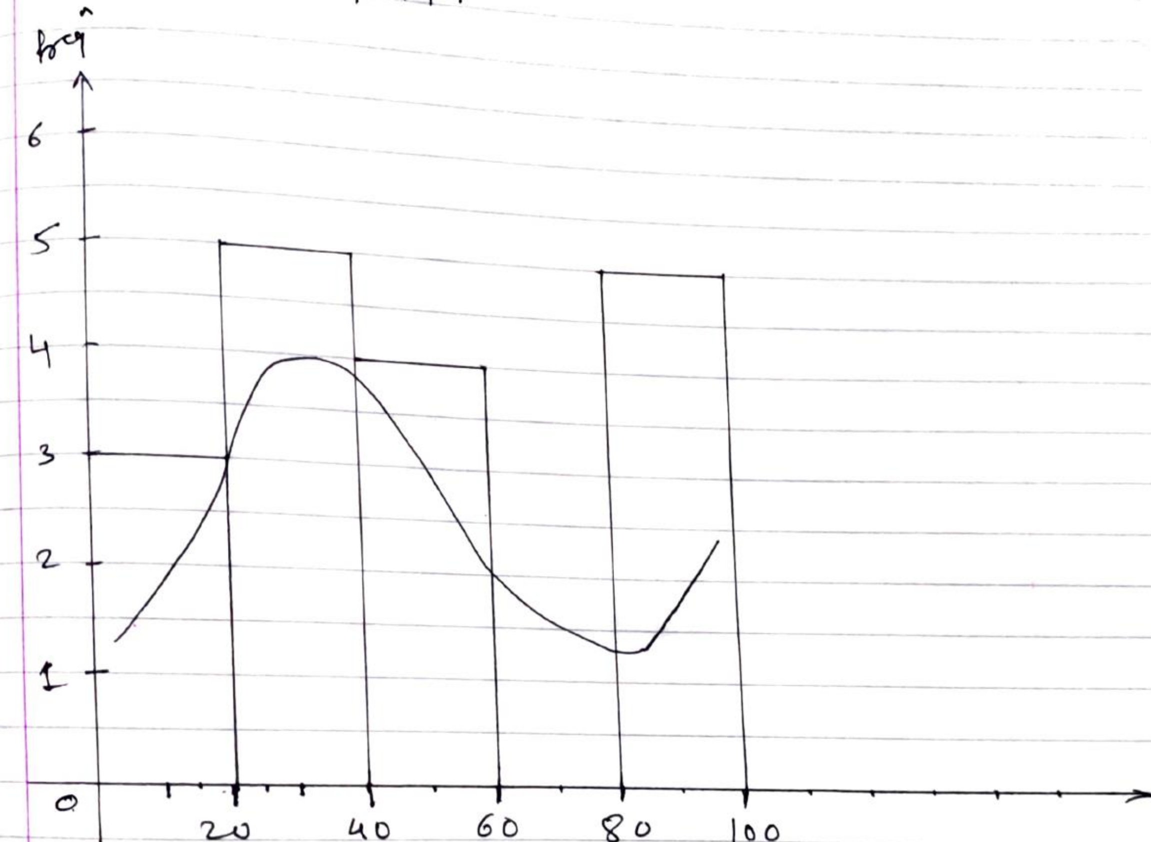
1) 1) Plot a histogram,

10, 13, 18, 22, 27, 32, 38, 40, 45, 51, 56, 57, 88, 90, 92, 94, 99

ANS:-

Assignment :

Eg: 10, 13, 18, 22, 27, 32, 38, 40, 45, 51, 56, 57, 88, 90, 92, 94, 99.



~~Bin~~

Given : Bins = 5

Bin size = 20.

2) In a quant test of the CAT Exam, the population standard deviation is known to be 100. A sample of 25 tests taken has a mean of 520. Construct an 80% CI about the mean.

ANS :- To construct an 80% confidence interval about the mean, we need to follow these steps:

Determine the level of significance (alpha). Since the confidence level is 80%, the level of significance is 1 - 0.80 = 0.20.

Find the critical value. Since the sample size is 25, we can use a t-distribution with degrees of freedom = 24 (df=n-1). From a t-table with 24 degrees of freedom and a significance level of 0.10 (0.20/2), the critical t-value is approximately 1.711.

Calculate the margin of error. The margin of error (E) is given by: E = critical value * (population standard deviation / sqrt(sample size)). Plugging in the values, we get: E = 1.711 * (100 / sqrt(25)) = 68.44.

Calculate the confidence interval. The confidence interval is given by: mean ± margin of error. Plugging in the values, we get: 520 ± 68.44, which gives us the interval (451.56, 588.44).

Therefore, we can say with 80% confidence that the true population mean of the quant test scores is between 451.56 and 588.44.

3) A car believes that the percentage of citizens in city ABC that owns a vehicle is 60% or less. A sales manager disagrees with this. He conducted a hypothesis testing surveying 250 residents & found that 170 residents responded yes to owning a vehicle.

ANS :- To test the sales manager's hypothesis, we need to follow these steps:

State the null and alternative hypotheses. The null hypothesis is that the percentage of citizens in city ABC that owns a vehicle is 60% or more. The alternative hypothesis is that the percentage of citizens in city ABC that owns a vehicle is less than 60%.

Determine the level of significance (alpha). Let's assume alpha is 0.05.

Choose the appropriate statistical test. Since we are testing a proportion and have a sample size greater than 30, we can use the z-test for proportions.

Calculate the test statistic. The test statistic is calculated as (sample proportion - hypothesized proportion) / standard error. The sample proportion is 170/250 = 0.68. The hypothesized proportion is 0.60. The standard error is calculated as sqrt(p*(1-p)/n), where p is the hypothesized proportion and n is the sample size. Plugging in the values, we get: standard error = sqrt(0.60*(1-0.60)/250) = 0.045. Therefore, the test statistic is (0.68-0.60)/0.045 = 1.78.

Determine the p-value. The p-value is the probability of observing a test statistic as extreme as the one calculated, assuming the null hypothesis is true. From a standard normal distribution table, the p-value for a test statistic of 1.78 is approximately 0.037.

Make a decision. Since the p-value (0.037) is less than the level of significance (0.05), we reject the null hypothesis. Therefore, there is sufficient evidence to conclude that the percentage of citizens in city ABC that owns a vehicle is less than 60%.

In summary, the sales manager's hypothesis is supported by the survey results, which suggest that the percentage of citizens in city ABC that owns a vehicle is less than 60%.

3) A car believes that the percentage of citizens in city ABC that owns a vehicle is 60% or less. A sales manager disagrees with this. He conducted a hypothesis testing surveying 250 residents & found that 170 residents responded yes to owning a vehicle.

A] State the null & alternate hypothesis.

B] At a 10% significance level, is there enough evidence to support the idea that vehicle owner in ABC city is 60% or less.

ANS:- Null hypothesis: The percentage of citizens in city ABC that owns a vehicle is 60% or more.

Alternative hypothesis: The percentage of citizens in city ABC that owns a vehicle is less than 60%.

To test whether there is enough evidence to support the idea that vehicle owners in ABC city is 60% or less, we can use a one-tailed hypothesis test with a significance level of 10%.

Using the sample data provided, the sample proportion of residents who own a vehicle is 170/250 = 0.68.

Next, we need to calculate the test statistic, which follows a standard normal distribution under the null hypothesis. The test statistic is calculated as:

z = (sample proportion - hypothesized proportion) / sqrt((hypothesized proportion * (1 - hypothesized proportion)) / n)

Plugging in the values, we get:

z = (0.68 - 0.6) / sqrt((0.6 * 0.4) / 250) = 2.26

The corresponding p-value for a one-tailed test with a test statistic of 2.26 is approximately 0.012.

Since the p-value is less than the significance level of 0.10, we can reject the null hypothesis and conclude that there is sufficient evidence to support the idea that the percentage of citizens in city ABC that owns a vehicle is less than 60%.

Therefore, the sales manager's hypothesis is supported by the survey results, which suggest that the percentage of citizens in city ABC that owns a vehicle is less than 60%.

4) What is the value of the 99 percentiles?

2,2,3,4,5,5,5,5,6,7,8,8,8,8,8,8,9,9,10,11,11,12

ANS: - To find the 99th percentile, we need to arrange the data in ascending order:

2, 2, 3, 4, 5, 5, 5, 6, 7, 8, 8, 8, 8, 8, 9, 9, 10, 11, 11, 12

The 99th percentile is the value that separates the highest 1% of the data from the rest. To find this value, we first need to calculate the rank of the 99th percentile. The rank is calculated as:

Rank = (99/100) * n

where n is the total number of observations.

In this case, n = 20, so the rank of the 99th percentile is:

Rank = (99/100) * 20 = 19.8

Since the rank is not a whole number, we need to use interpolation to estimate the value of the 99th percentile. The formula for interpolation is:

Value = L + (R - L) * f

where L is the value at the lower rank, R is the value at the higher rank, and f is the fractional part of the rank.

The lower rank is 19 and the higher rank is 20. The fractional part of the rank is 0.8. So, the value of the 99th percentile is:

Value = 11 + (12 - 11) * 0.8 = 11.8

Therefore, the value of the 99th percentile is 11.8

5) In left & right-skewed data, what is the relationship between mean, median & mode?

Draw the graph to represent the same.

ANS: - In left-skewed data, the mean is typically less than the median, which is less than the mode. This is because the left-skewed data have a long tail on the left side of the distribution, which pulls the mean to the left of the median, which is also pulled to the left of the mode.

In right-skewed data, the mean is typically greater than the median, which is greater than the mode. This is because right-skewed data have a long tail on the right side of the distribution, which pulls the mean to the right of the median, which is also pulled to the right of the mode.

However, in symmetrical data, the mean, median, and mode are all equal. In this case, the distribution is perfectly balanced around the central point.