Statistics problems

**Que 1) Plot a histogram,**

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

**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.**

* To construct an 80% confidence interval (CI) about the mean using the given information, you can follow these steps:
* Determine the critical value corresponding to the desired confidence level. For an 80% confidence level, the critical value can be found using a standard normal distribution or a t-distribution, depending on the sample size. Since the sample size is 25, we can use a t-distribution.
* Using a t-table or a statistical calculator, find the critical value for an 80% confidence level and 24 degrees of freedom. In this case, the critical value is approximately 1.711.
* Calculate the standard error (SE) of the sample mean using the formula:

SE = Population Standard Deviation / Square Root of Sample Size

= 100 / sqrt(25)

= 100 / 5

= 20

* Calculate the margin of error (ME) using the formula:

ME = Critical Value \* Standard Error

= 1.711 \* 20

≈ 34.22

* Calculate the lower and upper bounds of the confidence interval:

Lower Bound = Sample Mean - Margin of Error

= 520 - 34.22

≈ 485.78

* Upper Bound = Sample Mean + Margin of Error

= 520 + 34.22

≈ 554.22

* The 80% confidence interval about the mean is approximately (485.78, 554.22).
* Therefore, we can be 80% confident that the true population mean falls within the range of 485.78 to 554.22 based on the sample mean of 520 and a known population standard deviation of 100.

**Que 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.**

**To determine whether the car's belief that the percentage of citizens in city ABC who own a vehicle is 60% or less is valid, the sales manager conducted a hypothesis test. Let's analyze the situation using hypothesis testing.**

1. **State the null & alternate hypothesis.**
2. **At a 10% significance level, is there enough evidence to support the idea that vehicle owner in ABC city is 60% or less.**

Hypotheses:

* Null Hypothesis (H0): The percentage of citizens in city ABC who own a vehicle is 60% or less.
* Alternative Hypothesis (H1): The percentage of citizens in city ABC who own a vehicle is greater than 60%.
* Test Statistic:

In this case, we are dealing with proportions. We can use the Z-test as the sample size is relatively large (n > 30). The test statistic formula for the proportion is:

* Z = (p̂ - p₀) / √(p₀ \* (1 - p₀) / n)

where:

p̂ is the sample proportion (170/250 = 0.68 in this case)

p₀ is the hypothesized population proportion (0.6 in this case)

n is the sample size (250 in this case)

Calculation:

Using the given values, let's calculate the test statistic (Z):

Z = (0.68 - 0.6) / √(0.6 \* (1 - 0.6) / 250)

= 0.08 / √(0.6 \* 0.4 / 250)

≈ 0.08 / 0.03464

≈ 2.31

* Significance Level:

The significance level (α) is a predetermined threshold that helps us decide whether to reject the null hypothesis or not. Common choices for α are 0.05 (5%) and 0.01 (1%). Let's assume α = 0.05 for this test.

* Critical Value:

Since we are testing the alternative hypothesis that the percentage of citizens who own a vehicle is greater than 60%, we need to find the critical value for a right-tailed test at α = 0.05. Consulting the Z-table, we find that the critical value at α = 0.05 is approximately 1.645.

* Decision:

If the test statistic (Z = 2.31) is greater than the critical value (1.645), we reject the null hypothesis. Otherwise, we fail to reject the null hypothesis.

Since 2.31 > 1.645, we reject the null hypothesis.

Conclusion:

Based on the survey results and the hypothesis test, there is sufficient evidence to conclude that the percentage of citizens in city ABC who own a vehicle is greater than 60%.

**Que 4) What is the value of the 99 percentile 2,2,3,4,5,5,5,6,7,8,8,8,8,8,9,9,10,11,11,12**

To find the 99th percentile in a given set of data, we need to arrange the data in ascending order and then locate the value that corresponds to the 99th percentile. Let's 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

Since we are looking for the 99th percentile, we want to find the value below which 99% of the data falls. In other words, we are looking for the value that is greater than or equal to 99% of the data.

To calculate the position of the 99th percentile, we use the formula:

Position = (99/100) \* (n + 1)

where n is the total number of data points.

In this case, n = 20 (the number of data points).

Position = (99/100) \* (20 + 1) = 0.99 \* 21 = 20.79

The position is between the 20th and 21st values. To find the 99th percentile, we take the average of these two values:

99th percentile = (value at position 20 + value at position 21) / 2

99th percentile = (11 + 12) / 2 = 23 / 2 = 11.5

Therefore, the value of the 99th percentile in the given data set is 11.5.

**Que 5) In left & right-skewed data, what is the relationship between mean, median & mode? Draw the graph to represent the same.**

In left-skewed data (negative skew), the mean is typically less than the median, which is less than the mode. On the other hand, in right-skewed data (positive skew), the mean is usually greater than the median, which is greater than the mode. Let's explore this relationship further and visualize it graphically.

For left-skewed data:

Mean < Median < Mode

In a left-skewed distribution, the tail of the data extends to the left, causing the mean to be pulled in that direction. This results in the mean being smaller than the median. The mode, which represents the most frequently occurring value, is typically the smallest value in a left-skewed distribution.

Here is a visual representation of a left-skewed distribution:

markdown

| \*

| \*

| \*

| \*

| \*

| \*

| \*

| \*

|\*

---------------------------------

For right-skewed data:

Mode < Median < Mean

In a right-skewed distribution, the tail of the data extends to the right, pulling the mean in that direction. As a result, the mean is larger than the median. The mode, representing the most frequently occurring value, is typically the largest value in a right-skewed distribution.

Here is a visual representation of a right-skewed distribution:

-------------------------------------------------

| \*

| \*

| \*

| \*

| \*

It's important to note that the relationship between the mean, median, and mode can vary in different distributions, and the skewness is just one factor that influences their relative positions.