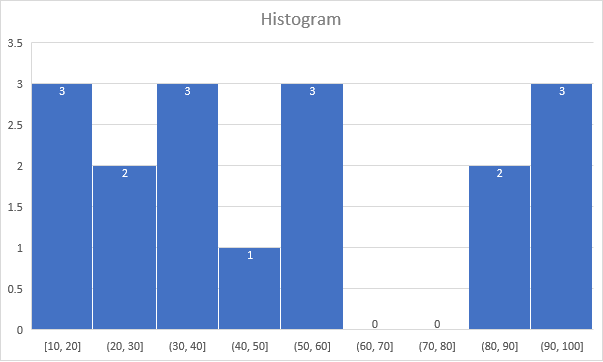
**Statistics Assignment (INEURON)**

**1) Plot a histogram,**

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



**Que 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.**

Sample mean: 520

Sample size: 25

Population standard deviation: 100

Confidence level: 80% (which translates to a z-score of 1.28)

2. Calculate the margin of error:

Margin of error = z \* σ / √n

= 1.28 \* 100 / √25

= 25.6

Lower bound = sample mean - margin of error

= 520 - 25.6

= 494.4

Upper bound = sample mean + margin of error

= 520 + 25.6

= 545.6

Therefore, the 80% confidence interval for the mean score on the CAT exam is (494.4, 545.6).

**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.**

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.**

Null Hypothesis (H0): The percentage of citizens in city ABC that owns a vehicle is 60% or less. This is what the car believes.

Alternative Hypothesis (Ha): The percentage of citizens in city ABC that owns a vehicle is more than 60%. This is what the sales manager is arguing.

Data:

Sample size: 250 residents

Number of vehicle owners: 170

Significance level: 10%

Analysis:

Proportion of vehicle owners in the sample: p̂ = 170 / 250 = 0.68 (68%)

Test Statistic: We can use a one-sample proportion z-test since we have a large sample size and are focusing on a single proportion.

Z = (p̂ - p₀) / √(p₀(1-p₀) / n)

Z = (0.68 - 0.60) / √(0.60(1-0.60) / 250) ≈ 1.54

Critical Value: At a 10% significance level, the critical z-score for a one-tailed test is 1.28.

Decision Rule: If the calculated z-score is greater than the critical value, we reject the null hypothesis. Otherwise, we fail to reject the null hypothesis.

Result:

The calculated z-score (1.54) is greater than the critical value (1.28). Therefore, at a 10% significance level, we reject the null hypothesis. This means there is enough evidence to suggest that the percentage of vehicle owners in city ABC is more than 60%, supporting the sales manager's claim.

**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**

The 99th percentile is 12.

* The 99th percentile represents the value below which 99% of the data points fall.
* In this case, 99% of the data points in the sorted list are less than or equal to 12.
* Therefore, 12 is the 99th percentile value.

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

**Draw the graph to represent the same.**

**Relationship between Mean, Median, and Mode in Skewed Data**

The relationship between mean, median, and mode in skewed data depends on the direction of the skew: left-skewed or right-skewed.

**Left-Skewed Data**:

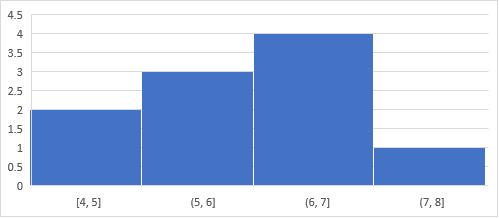
**Distribution**: Tail stretches out towards the left side.

**Relationship**: Mean < Median < Mode

**Explanation:** The long tail on the left pulls the mean down more than the median. The mode, representing the peak, lies furthest left.

Eg:

The histogram for the data: 4, 5, 6, 6, 6, 7, 7, 7, 7, 8 is not symmetrical. The right-hand side seems “chopped off” compared to the left side. A distribution of this type is called skewed to the left because it is pulled out to the left.



|  |  |
| --- | --- |
| **mean** | **6.3** |
| **median** | **6.5** |
| **mode** | **7** |

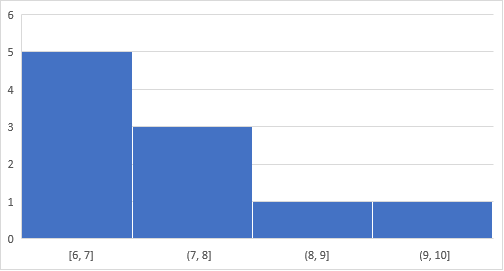
**Right-Skewed Data**:

**Distribution**: Tail stretches out towards the right side.

**Relationship**: Mean > Median > Mode

**Explanation**: The long tail on the right pulls the mean up more than the median. The mode, representing the peak, lies furthest right.

The histogram for the data: 6, 7, 7, 7, 7, 8, 8, 8, 9, 10 is also not symmetrical. It is skewed to the right.



|  |  |
| --- | --- |
| mean | 7.7 |
| median | 7.5 |
| mode | 7 |

To summarize, generally if the distribution of data is skewed to the left, the mean is less than the median, which is often less than the mode. If the distribution of data is skewed to the right, the mode is often less than the median, which is less than the mean.