**Statistics**

* **Part -1 : Formulas**

**1) Frequency ( f ) =** Number of times a value occurs

**--------------------------------------------------------------------------------------------**

frequency

**2) Relative Frequency =** --------------------------------------

Total nu. of observations in the dataset

**--------------------------------------------------------------------------------------------**

Sum of the frequencies of all values

**3) Cumulative Frequency =** -------------------------------------------

including the current value

**--------------------------------------------------------------------------------------------**

Sum of all values

**4) Mean (Average) = ------------------------------------**

Number of all values

**--------------------------------------------------------------------------------------------**

N + 1

**5) Median = -----------------**

2

**6) Mode =** Highest peak of the distribution

**--------------------------------------------------------------------------------------------**

**7) Range =** Maximum – Minimum

**--------------------------------------------------------------------------------------------**

**8) Mean Deviation =** 1/N \*∑ (Xi - X̄)

**--------------------------------------------------------------------------------------------**

**9) Absolute Mean Deviation =** 1/N \*∑ |Xi - X̄|

**--------------------------------------------------------------------------------------------**

**10) Variance ( σ² ) =** 1/N \*∑ ( Xi - X̄) ²

**--------------------------------------------------------------------------------------------**

**11) Standard Deviation ( σ ) =** variance

**--------------------------------------------------------------------------------------------**

1 -(( X - μ)/σ)²

**12) Normal Distribution F(X)= -------** \* e

2√2πσ²

**--------------------------------------------------------------------------------------------**

**13) Covariance (x,y) =**  1/N \* ∑ ( Xi - X̄) ( Yi - Ȳ)

**--------------------------------------------------------------------------------------------**

**14) Correlation ( r ) =** Cov ( X , Y ) / σ X σ Y

**--------------------------------------------------------------------------------------------**

**15) Percentile =** ( N + 1 ) \* P / 100

**--------------------------------------------------------------------------------------------**

**16) Common outlier Boundaries :**

Lower Bound = Q1 – 1.5 \* IQR

Upper Bound = Q1 + 1.5 \* IQR

**--------------------------------------------------------------------------------------------**

**17) Extreme Outlier Boundaries :**

Lower Bound **=** Q1 – 3\* IQR

Upper Bound **=** Q1 + 3 \* IQR

**--------------------------------------------------------------------------------------------**

**X –** μ

**18) Standardization =**  ----------

σ

**--------------------------------------------------------------------------------------------**

X – min( X )

**17) Normalization = -----------------------**

Max (X) – Min (X)

**--------------------------------------------------------------------------------------------**

**18) Empirical Rule : ( 68 – 95 – 99.7 )**

68 % = μ-1σ **To** μ+1σ

95 % =μ-2σ **To** μ+2σ

99.7 % =μ-3σ **To** μ+3σ

**--------------------------------------------------------------------------------------------**

**19) Chebyshev Rule :** 1–1/K²

μ-1σ to μ+1σ (Not Valid)

75 % = μ-2σ **To** μ+2σ

90 % = μ-3σ to μ+3σ

**-------------------------------------------------------------------------------------------**

**Shrirang G.Kurhe**

**17/09/2024**