BASIC STATISTICAL METHODS FOR QUALITY ASSURANCE(QA) AND QUALITY CONTROL (QC) SCHEMES IN WATER LABORATORIES

Presented by:

A. MANOHARAN

Scientist

Central Pollution Control Board

(Ministry of Environment & Forests)

Delhi – 110 032

India

E- MAIL: manocpcb@yahoo.com

STATISTICS: Definition

- Statistics is a science and in particular it is a branch of applied mathematics which is concerned with collection of numerical observation and interpreting them.
- The word Statistics is derived from the Latin word ***'status'***, Italian word '***State***' and German word '***Statik'*** meaning a political state.
- **R.A. Fisher**, who is mainly responsible for the development of the science of statistics defined **statistics as a branch of applied mathematics which deals with collection, classification and nterpretation of numerical facts.**
- **Ya-Lun-Chou defined " Statistics is a method of decision- making in the face of uncertainty on the basis of numerical data and calculated risks"**

FUNCTIONS OF STATISTICS

- Presents facts in a definite form 
- Simplifies mass of figures 
- Facilitates comparison 
- Helps in formulating and testing hypothesis 
- Helps in prediction 
- Helps in the formulation of suitable policies 

1. Range (R)

- Range (R) of a set of data is expressed simply the absolute difference between the maximum and minimum values.
- Range (R) = $X_{\mathbf{max.}} - X_{\mathbf{min.}}$
- Example 1
- For example, the range of the data, out of 25 observations given inTable 1 is as follows:
- Range (R) = 74 - 49 = 25
- Range is useful in assessing the extent of the spread of the data as first hand information.

2. Arithmetic Mean

- Arithmetic Mean, $\mathbf{X = \frac{X1 + X2 + X3\ldots Xn}{n}}$
  - Where X I, X2, X n are the individual values
  - n = No. Of observations

Example 2 (Refer **Table 1)**

- By computing the above formula the data of the table I, the arithmetic mean of all data (25 observations) is 61.4.
- Arithmetic mean is a common average and is useful in many statistical calculations.

3. Mean deviation

Mean deviation is the mean of the deviation of the set of values from the mean value. The mean deviation gives a [air idea about the degree of dispersion of the values and it is less significance in inferring the dispersion of the data as compared to standard deviation.

$$Mean \quad Deviation = \frac{\sum (X - \bar{X})}{n}$$

$$X = individual \ \ values$$

$$\bar{X} = Mean$$

$$n = No.of \ \ values$$

5. Variance (S^2)

Variance is the mean of the squares of the deviations from the mean value of the set *of* data.

$$Variance = \frac{\sum (X - \bar{X})^2}{n-1}$$

$X = individual\ \ values$

$\bar{X} = Mean$

$n = No.of\ \ values$

4. Standard deviation (SD)

Standard deviation is the measure of degree of dispersion. It is the most stable index of variability and most widely used in statistical calculations. It is the square root of the mean of the squared deviations. Otherwise, it is simply the square root of variations (S^2)

$$Sandard\ Deviation = \sqrt{\frac{\sum (X - \bar{X})^2}{n-1}}$$

Where

$$X = individual\ \ values$$

$$\bar{X} = Mean$$

$$n = No.of\ \ values$$

6. Co-efficient of variation (CV%) or Relative standard deviation(RSD)

Co-efficient (CV%) is the calculated by dividing the standard deviation by mean and multiply by 100 and it is expressed in percentage.

$$CV\ \% = \frac{SD}{Mean}\ X\ 100$$

Where

SD = Standard deviation (as calculated above)

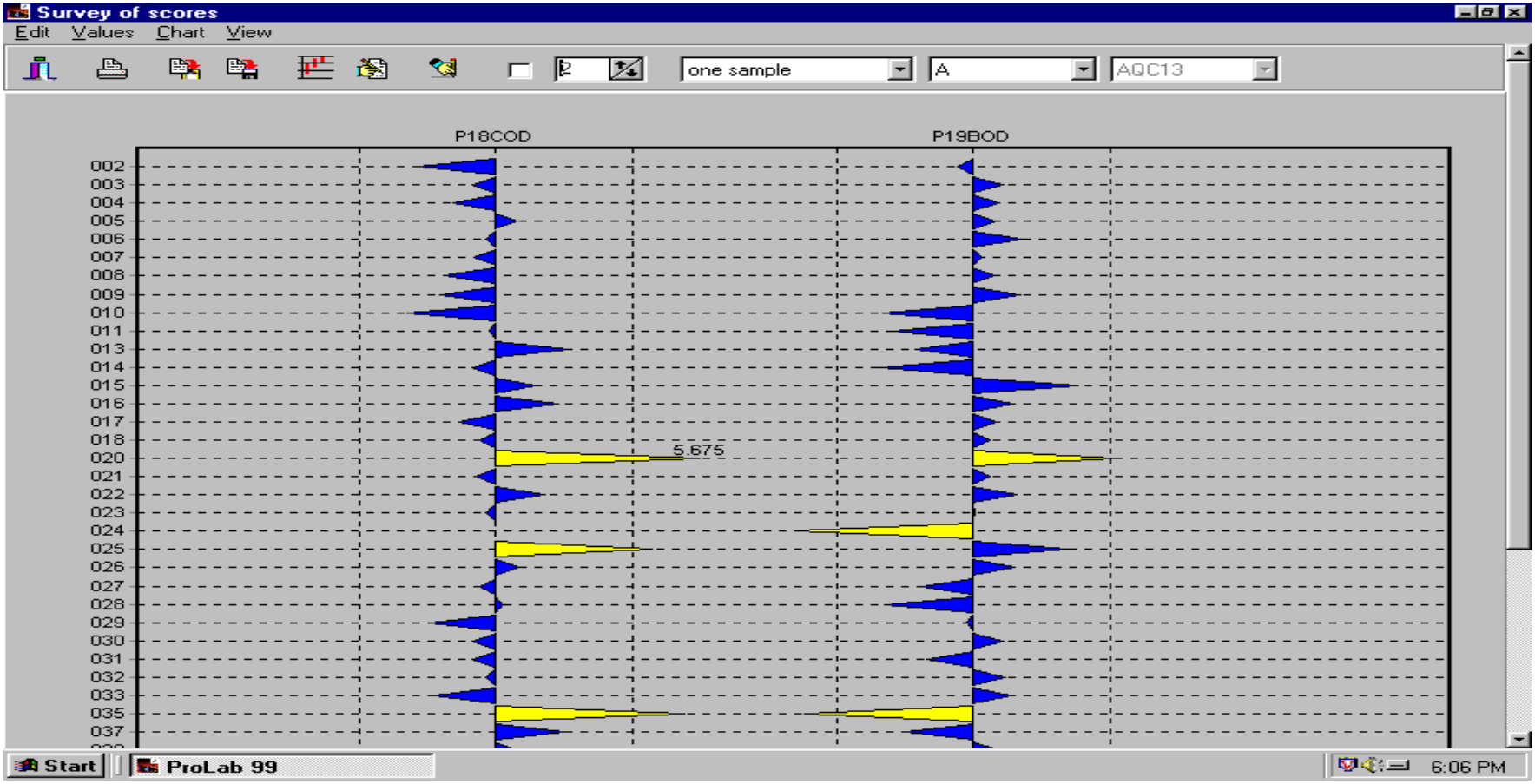
C V % or RSD

CV% is used to compare the variability of two sets of data. Suppose if the two means of 2 sets of data are same, one can differentiate the variability by looking into their standard deviations. But if the two means & ,units and their standard deviations are different from each other, One can compare the variability by looking into the differences in CV% values.

	Lab- B	Lab - C
Mean	**70.8**	**61.0**
S.D	**2.17**	**4.67**
C V %	**3.06 %**	**7.64 %**

From the above data, one can infere that the laboratory B (CV% = 3.06%) showed more variability than laboratory A (CV% = 2.29) though the mean and standard deviation are differ.

COD & BOD PARAMETERS
z- scores



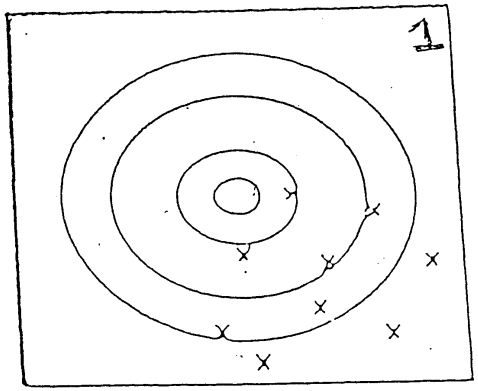
What is Z- score ?

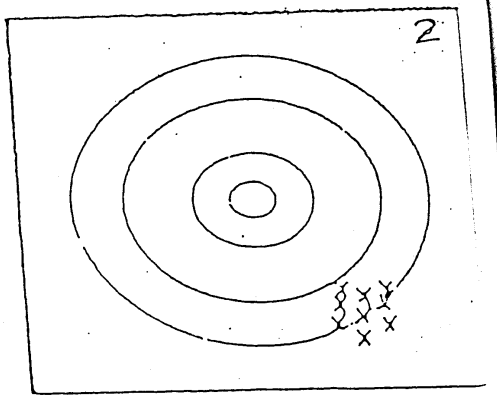
$$\text{Classical Z- Score} = \frac{\text{Result- Reference value}}{\text{Standard deviation}}$$

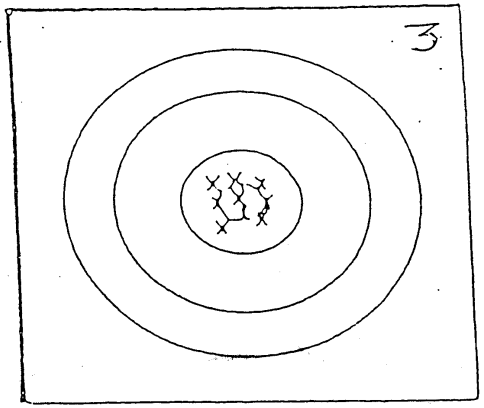
$$\text{Robust Z- Score} = \frac{\textbf{Result- Reference value}}{\textbf{Normalised IQR}}$$

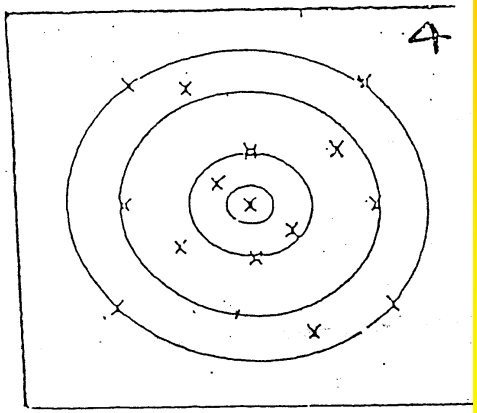
NIQR= IQR x 0.7413

AQC - TARGET MODEL
TYPES OF ERRORS









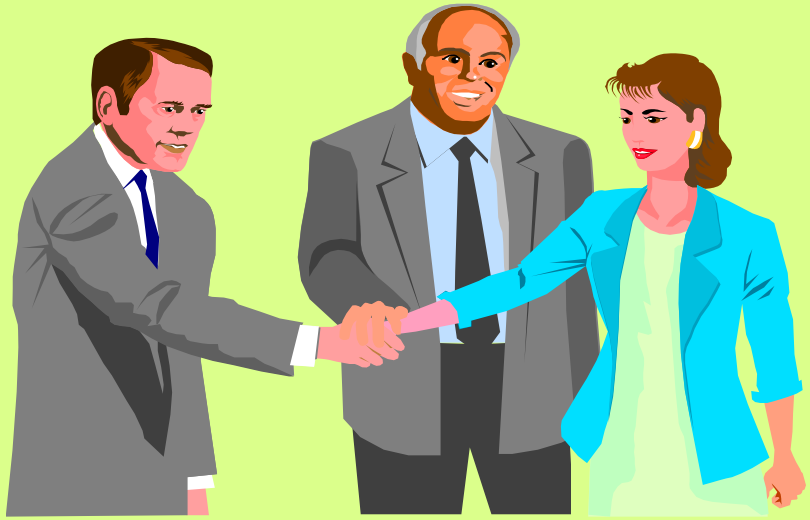
	ACCURACY	PRECISION
1.	POOR	POOR
2.	POOR	GOOD
3.	GOOD	GOOD
4	GOOD	POOR

CENTRAL POLLUTION CONTROL BOARD - DELHI

LABORATORY DIVISION - WATER LABORATORY

AQC/WATER: SHEWART'S ANALYTICAL QUALITY CONTROL (AQC) CHART

SAMPLE NAME:	PRE PERIOD	CONTROL PERIOD
PARAMETER:		UCL
		UWL
Conc. of Quality Control Sample:		
	MEAN =	MEAN
METHOD:		LWL
MATRIX:		LCL
Measured Conc. :		
Date:		
Remarks:		



THANK YOU