Data Analysis in Clinical Laboratories

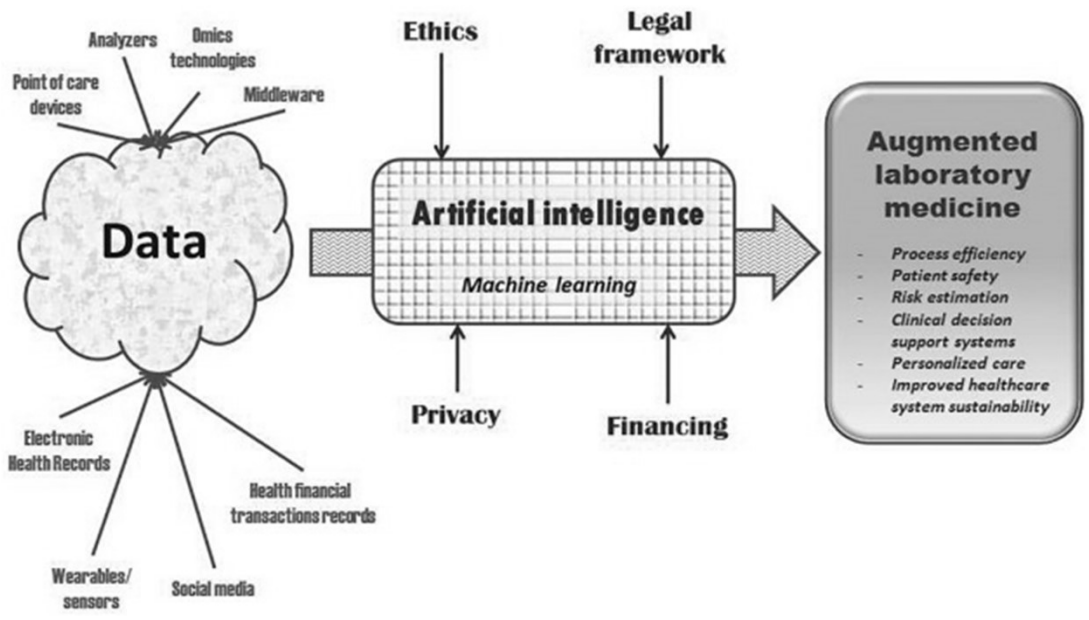
Dr. M Reza Bakhtiari, DCLS, PhD


MREZABAKHTIARI


dr.bakhtiari.academy

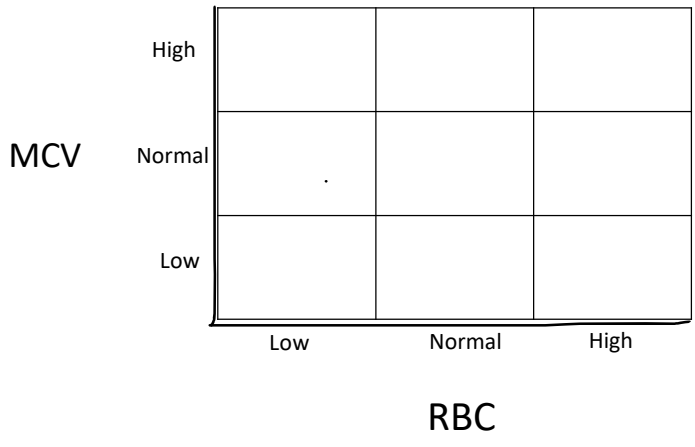
Lecturer Introduction

- **Mohammad Reza Bakhtiari**
  - Doctorate in Clinical Laboratory Sciences (DCLS)
  - PhD in Medical Biotechnology
  - Faculty Member to Biotech Dept. of IROST
  - Member to IFCC Committee on Standardization of Thyroid Function Tests (C-STFT)



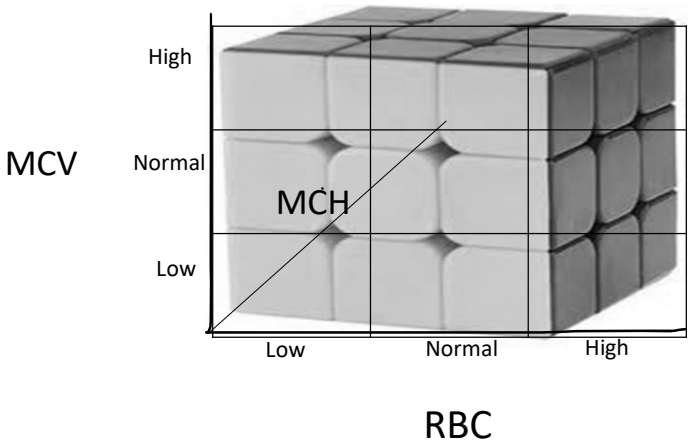
MREZABAKHTIARI dr.bakhtiari.academy 4/79

Some CBC Data

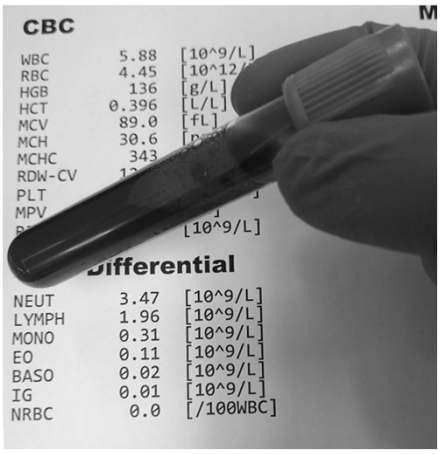


MREZABAKHTIARI dr.bakhtiari.academy 5/79

Some CBC Data



CBC + CMP Data

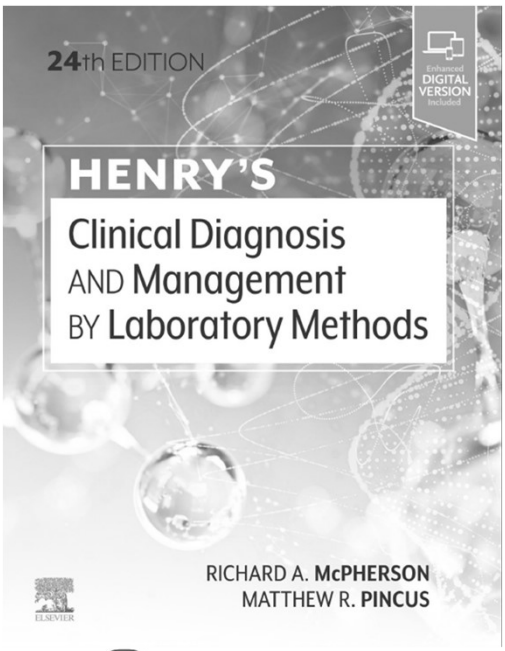


	HPβCD-Naringenin	Naringenin	Control
Na^+	139.0±1.0	138.0±1.0	134.3±2.5
K^+	5.9±0.6	5.3±0.3	4.4±0.7
Cl^-	102.0±1.0	102.0±2.0	89.7±3.9
Glu	168.0±14.4	157.0±6.1	181.2±26.5
Ca^{2+}	9.3±0.7	9.7±0.2	10.2±0.5
BUN	16.7±3.1	19.3±5.1	14.0±3.6
Cre	0.2±0.0	0.3±0.0	0.4±0.1
ALP	147.3±84.6	135.3±37.5	311.7±64.6
ALT	56.7±11.1	61.3±20.3	59.8±30.4
AST	307.3±107.8	248.3±45.5	205.7±124.0
TBIL	0.3±0.0	0.3±0.1	0.4±0.0
ALB	1.6±0.2	1.6±0.3	2.1±0.1
TP	5.6±0.5	5.6±0.5	6.2±0.3

Glu, Glucose; BUN, blood urea nitrogen; Cre, Creatinine; ALP, alkaline phosphatase; ALT, alanine transaminase; AST, aspartate transaminase; TBIL, total bilirubin; ALB, albumin; TP, total protein.
doi:10.1371/journal.pone.0018033.t002

CBC + CMP Data

References



CHAPTER 10

LABORATORY STATISTICS

Richard A. McPherson

KEY POINTS

- For statistical analyses, nominal variables can take on only a limited number of values (or categories), whereas continuous variables are used to report quantitative data.

accuracy also are based on these principles. For data that are not continuous but take on only two or a few discrete values (e.g., positive or negative), the analysis might consist of counting the number in each category and looking at the proportions of all values by category.

Comparison of data typically asks the question whether one group is

CHAPTER 12

CLINICAL LABORATORY INFORMATICS

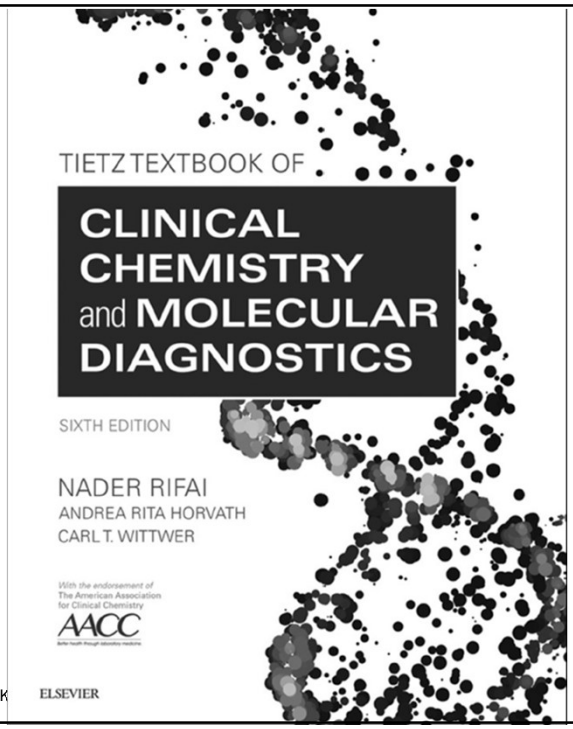
Ulysses G.J. Balis

KEY POINTS

- The practice of Pathology Informatics/Clinical Laboratory Informatics is central to all aspects of data stewardship in the clinical laboratory.
- Fundamental knowledge of database technology and database principles is critical to understanding key aspects of laboratory information system (LIS) operation.
- The LIS is an integral element of the larger enterprise-wide portfolio of information technology solutions that may be rendered as either

Informatics Association's "Clinical informatics subspecialty delineation of practice" document (Silverman et al., 2019), The material contained within this chapter constitutes a reasonable contemporary survey of the areas of expertise and topics that pathology informaticists might typically encounter during their daily activities. In topic areas in which substantial additional domain knowledge is available but exceeds the level of coverage provided by this chapter, appropriate references are provided, along with mention of the significantly expanded knowledge base that can be perused and assimilated as required.

12/79



13/79

2

Statistical Methodologies in Laboratory Medicine: Analytical and Clinical Evaluation of Laboratory Tests

Kristian Linnet, Karel G.M. Moons, and James C. Boyd

ABSTRACT

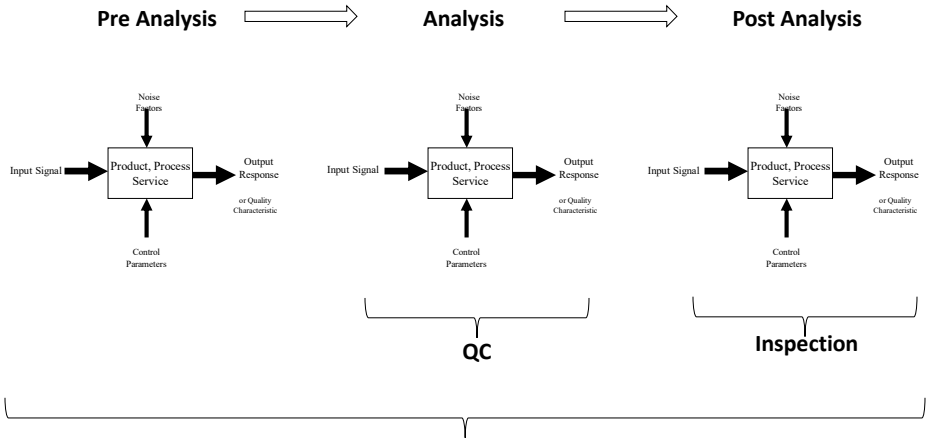
Background

The careful selection and evaluation of laboratory tests are key steps in the process of implementing new measurement procedures in the laboratory for clinical use. Method evaluation in the clinical laboratory is complex and in most countries is a regulated process guided by various professional recommendations and quality standards on best laboratory practice.

Content

with comparison of assays in detail, including using difference plots and regression analysis, the focus is on quantification of the (added) diagnostic value of laboratory assays or tests. First, the evaluation of tests in isolation is outlined, which corresponds to simple diagnostic scenarios, when only a single test result is decisive (eg, in the screening context). Subsequently, the chapter addresses the more common clinical situation in which a laboratory assay or test is considered as part of a diagnostic workup and thus a test's added value is at issue. This involves use of receiver

Quality Management
(A System Approach)



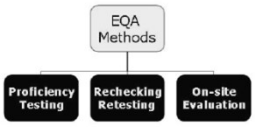
ISO 15189: Clinics & BioStatistics

5.5 Examination processes

5.5.2 Biological reference intervals or clinical decision values

- The laboratory shall define the biological reference intervals **or** clinical decision values, document the basis for the reference intervals or decision values and communicate this information to users.
- When a particular biological reference interval or decision value is no longer relevant for the population served, appropriate changes shall be made and communicated to the users.
- When the laboratory changes an examination procedure or pre-examination procedure, the laboratory shall review associated reference intervals and clinical decision values, as applicable.

ISO 15189: Clinics & BioStatistics



5.6.3 Interlaboratory comparisons

5.6.3.1 Participation

- The laboratory shall participate in an inter-laboratory comparison programme(s) (such as an external quality assessment programme or proficiency testing programme) appropriate to the examination and interpretations of examination results. The laboratory shall monitor the results of the inter-laboratory comparison programme(s) and participate in the implementation of corrective actions when predetermined performance criteria are not fulfilled.
- NOTE The laboratory should participate in inter-laboratory comparison programmes that substantially fulfil the relevant requirements of ISO/IEC 17043.
- The laboratory shall establish a documented procedure for inter-laboratory comparison participation that includes defined responsibilities and instructions for participation, and any performance criteria that differ from the criteria used in the inter-laboratory comparison programme.
- Inter-laboratory comparison programme(s) chosen by the laboratory shall, as far as possible, provide clinically relevant challenges that mimic patient samples and have the effect of checking the entire examination process, including pre-examination procedures, and post-examination procedures, where possible.

Lean Six Sigma
(the best available method to reduce process variability)

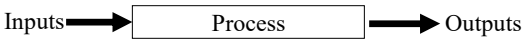
The main steps:

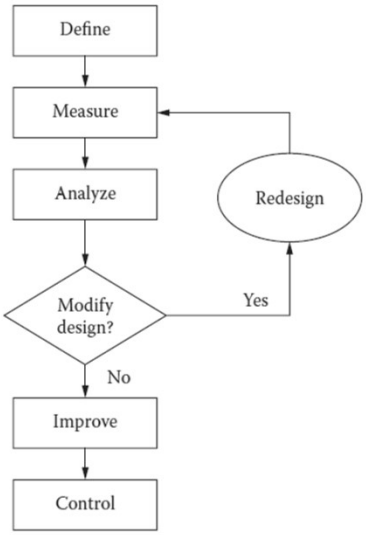
1. Identify the needs of the customer.
2. Translate these needs into the process expert's language through quality function deployment (QFD).
3. Omit redundant processes (Lean)
4. Make improvements through the DMAIC process [y = f(x)]
5. Hold the gains through statistical process control (SPC).
6. Provide customer satisfaction.

DMAIC Steps

DMAIC:

1. **Define:** Understand the project output *Y* and how to measure it.
2. **Measure:** Priority-wise, determine potential *X*s and measure *X*s and *Y*.
3. **Analyze:** Determine *X*–*Y* relationships and, after verification, quantify important *X*s.
4. **Improve:** Devise solutions to optimize *X*s to improve *Y*.
5. **Control:** Control and monitor important *X*s and the output *Y* over time.



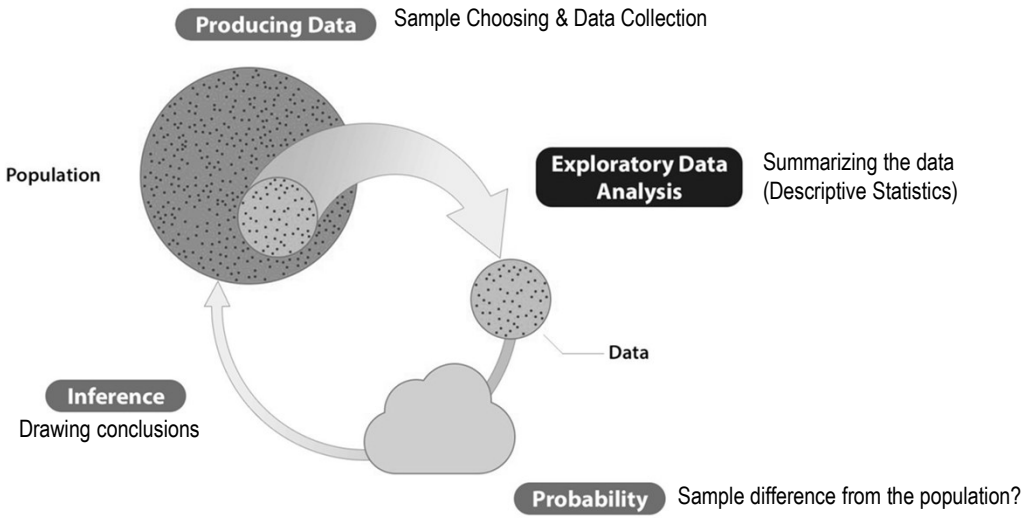


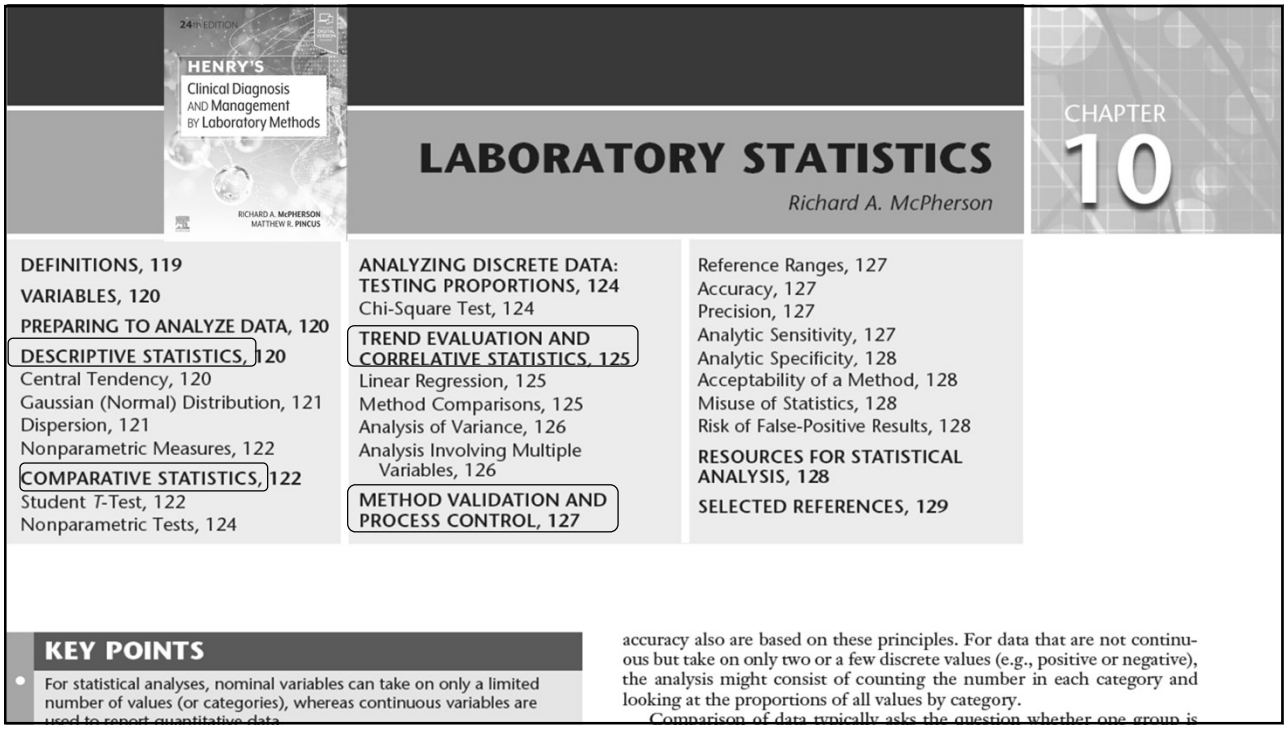
20/79

21

BioStatistics

"The Big Picture," the four-step process that encompasses statistics

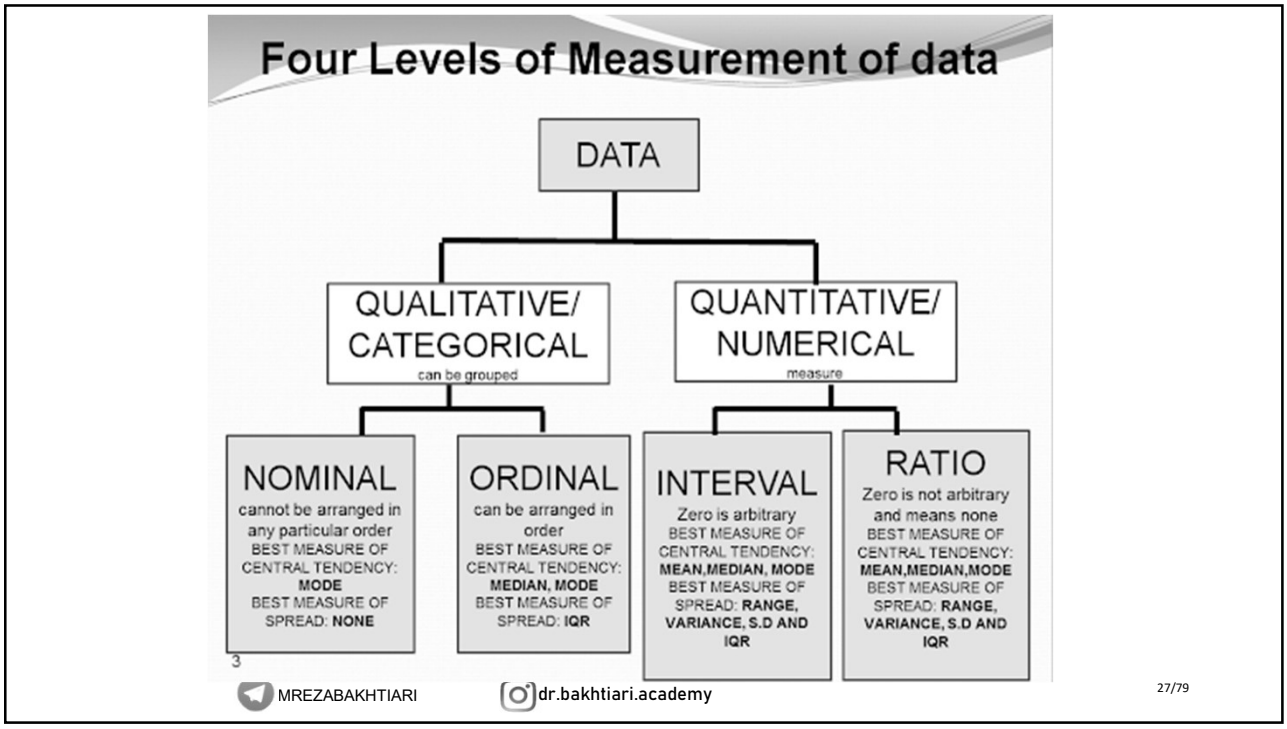




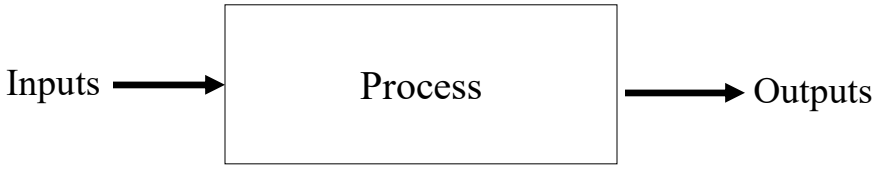
Variables

- *Variables:* The things that we measure, count, or otherwise delineate are termed *variables* because the values they can assume vary.
- ***Variables* scales:**
  1. *Nominal scale* : the variable can take on only a limited number of values, usually called categories (or characters), e.g. gender (male or female) and risk factors (e.g., smoker or nonsmoker).
  2. *Ordinal scale:* the variable takes on specific values that have some inherent order such as magnitude but without equivalent distances between categories (e.g., trace amount, 1+, 2+, etc., of protein in urine).
  3. *Interval scale:* the variable takes on values in a quantitative range with defined differences between points.
- It is conventional to treat most numeric laboratory measurements as continuous variables, even though they may be reported as discrete values (e.g., glucose values of 123 or 124 mg/dL, but not 123.857… mg/dL).


www.SarzaminDownload.com


Four Levels of Measurement of data
DATA
QUALITATIVE/ CATEGORICAL
can be grouped
QUANTITATIVE/ NUMERICAL
measure
NOMINAL
cannot be arranged in any particular order
BEST MEASURE OF CENTRAL TENDENCY: MODE
BEST MEASURE OF SPREAD: NONE
ORDINAL
can be arranged in order
BEST MEASURE OF CENTRAL TENDENCY: MEDIAN, MODE
BEST MEASURE OF SPREAD: IQR
INTERVAL
Zero is arbitrary
BEST MEASURE OF CENTRAL TENDENCY: MEAN,MEDIAN, MODE
BEST MEASURE OF SPREAD: RANGE, VARIANCE, S.D AND IQR
RATIO
Zero is not arbitrary and means none
BEST MEASURE OF CENTRAL TENDENCY: MEAN,MEDIAN,MODE
BEST MEASURE OF SPREAD: RANGE, VARIANCE, S.D AND IQR
3
MREZABAKHTIARI
dr.bakhtiari.academy
27/79

Variability



x → f (x) → y

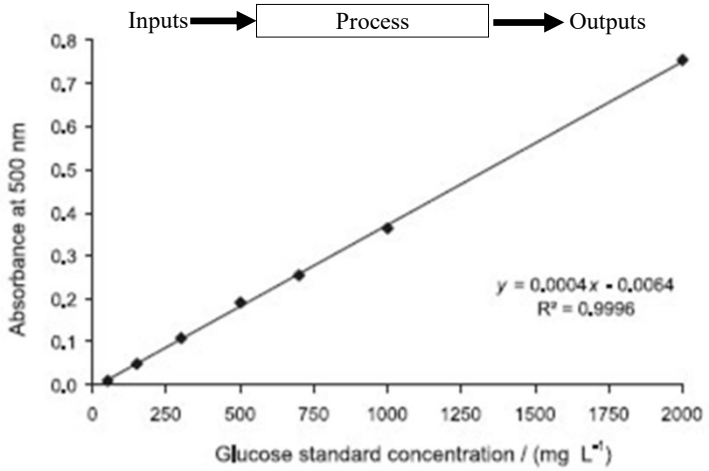
$$y = f(x)$$

Functions & Sets

$$y = f(x) = m.x + b$$

$$y = f(x) = \{(x_1,y_1), (x_2,y_2), (x_3,y_3), \ldots.\}$$

Functions & Sets (Calibration Curves)



$y = f(x)$

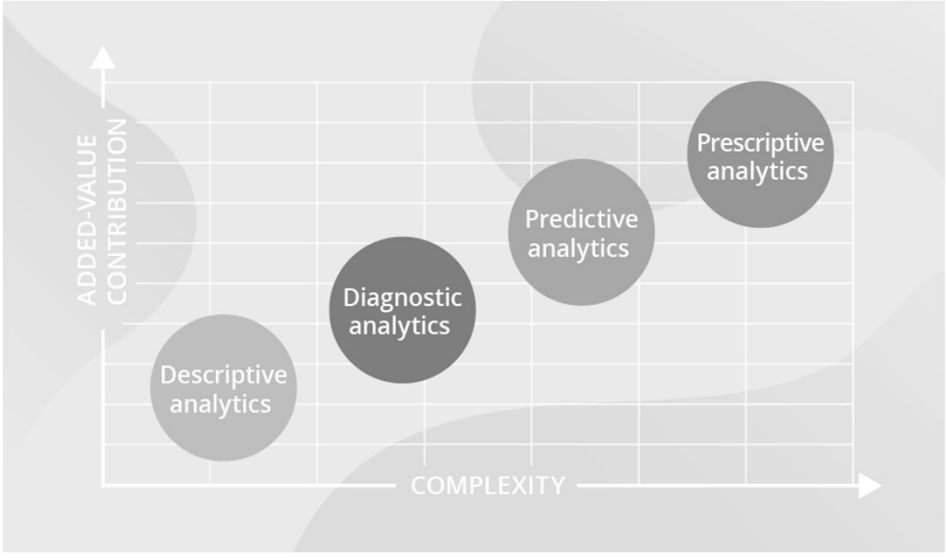
$y = f(x) = m.x + b$

$y = f(x) = \{(x_1,y_1), (x_2,y_2), (x_3,y_3), \ldots.\}$

Data Analysis

ScienceSoft

4 Types of Data Analytics to Improve Decision-Making



CHAPTER 10

LABORATORY STATISTICS

Richard A. McPherson

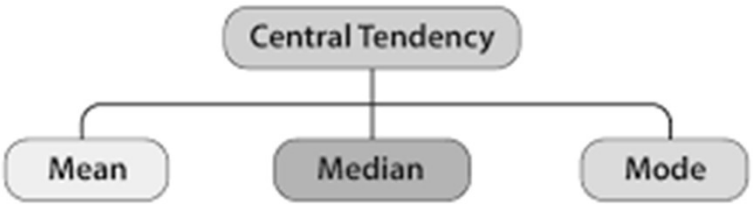
KEY POINTS

- For statistical analyses, nominal variables can take on only a limited number of values (or categories), whereas continuous variables are used to report quantitative data

accuracy also are based on these principles. For data that are not continuous but take on only two or a few discrete values (e.g., positive or negative), the analysis might consist of counting the number in each category and looking at the proportions of all values by category.

Comparison of data typically asks the question whether one group is

CENTRAL TENDENCY



Measures of Central Tendency

- *Mean (Arithmetic)*

$$Mean = \frac{sum\ of\ all\ values}{total\ number\ of\ values}$$

$$Mean = \bar{x} = (x_1 + x_2 + \ldots x_n) \div n = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- *Median*

$$Median = middle\ value\ (when\ the\ data\ are\ arranged\ in\ order)$$

To find the median number:

1. Put all the numbers in numerical order.
2. If there is an odd number of results, the median is the middle number.
3. If there is an even number of results, the median will be the mean of the two central numbers

- *Mode*

$$Mode = most\ common\ value$$

Geometric Mean

Another measure of central tendency is the *geometric mean*, which has the feature of minimizing the influence from extreme values in a distribution. The geometric mean is calculated as the nth root of the product of all n values from a population, or:

$$Geometric\ mean = \sqrt[n]{x_1 \times x_2 \times \cdots x_n}$$

The following transformation is more convenient for calculating the geometric mean:

$$\log Geometric\ mean = \sum_{i=1}^{n} \frac{\log x_i}{n}$$

BiModal Distribution

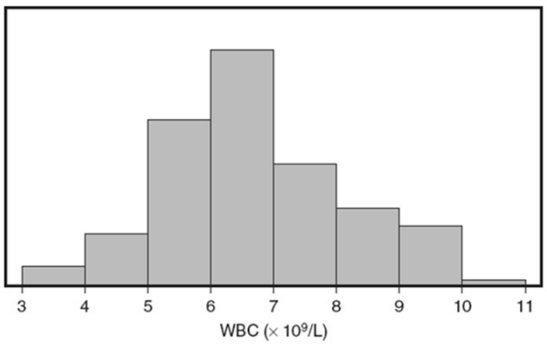
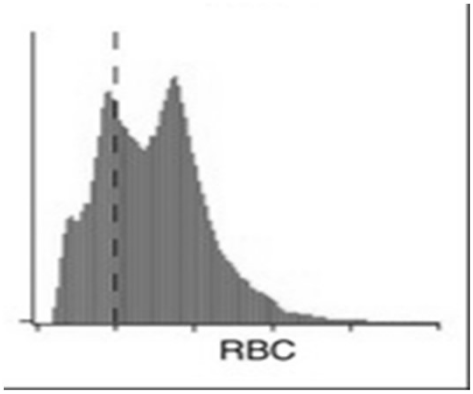
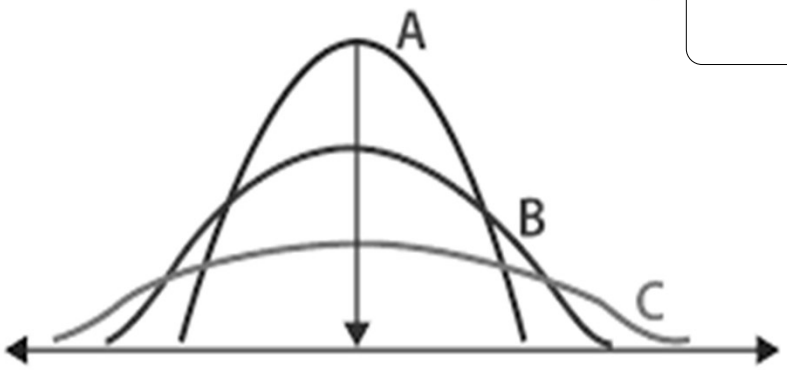


Figure 9-2 Distribution of white blood cells (WBCs) in the blood of 85 healthy individuals.

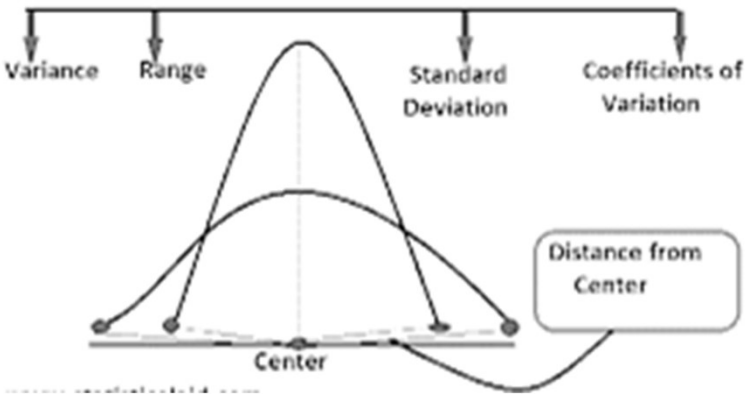


DISPERSION AND MEASURES OF DISPERSION

BYJU'S



Measures of Dispersion



Measures of Dispersion

- *Range*

$$\text{Range}(X) = \text{Max}(X) - \text{Min}(X)$$

- *Variance*

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

- *Standard Deviation*

$$SD = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

- *Coefficient of Variation*

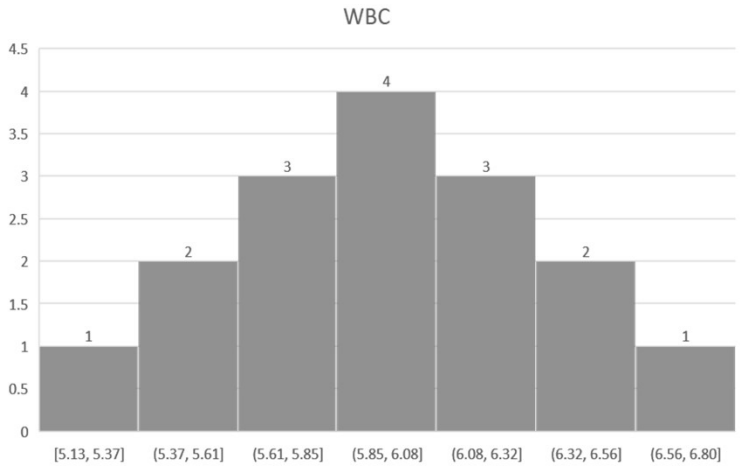
$$CV\ (\%) = \left(\frac{Standard\ deviation}{Mean}\right) \times 100$$

Statistical Software Programs

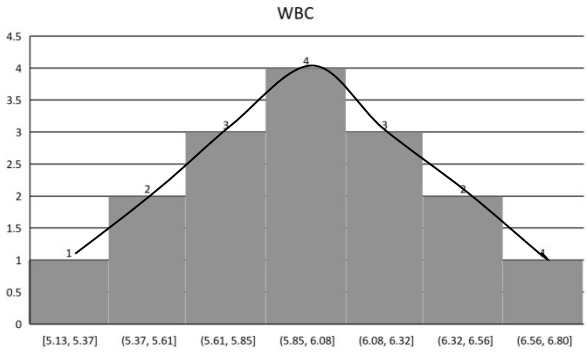
- *Open-source Programs*
  - DataMelt
  - DAP
  - JASP
  - www.westgard.com/ …..
- *Proprietary Programs*
  - *SPSS*
  - *SAS*
  - *GLIM*
  - *MiniTab*
  - *MedCalc*
  - *Excel*
  - …..
- *Online Programs*
  - www.socscistatistics.com/

WBC Example

	WBC	
1	5.13	1
2	5.40	2
3	5.40	
4	5.70	3
5	5.70	
6	5.70	
7	6.00	4
8	6.00	
9	6.00	
10	6.00	
11	6.13	3
12	6.13	
13	6.13	
14	6.40	2
15	6.40	
16	6.80	1



Normal (Gaussian) Distribution



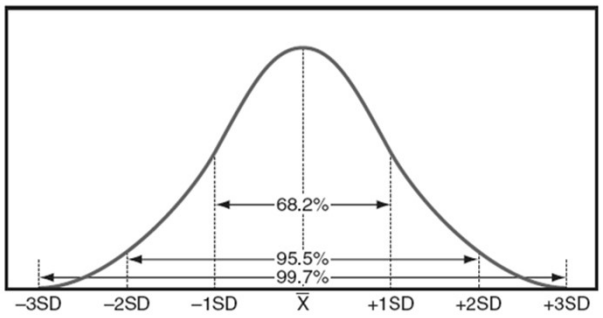
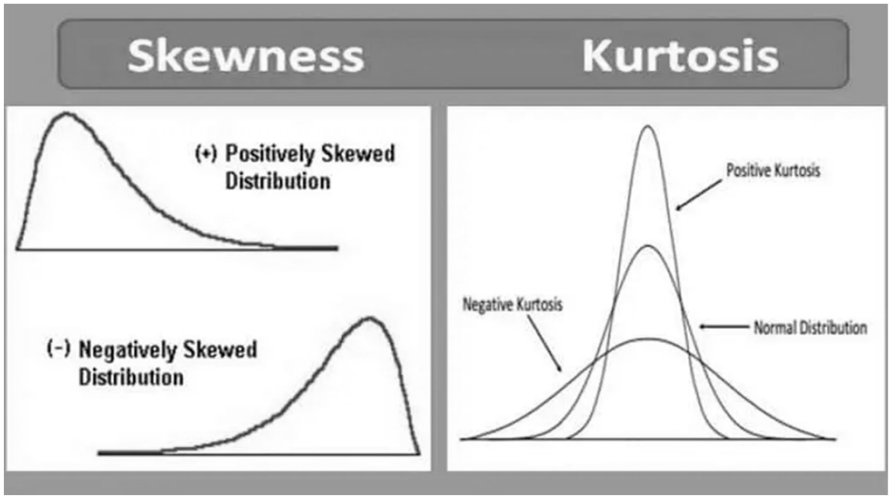
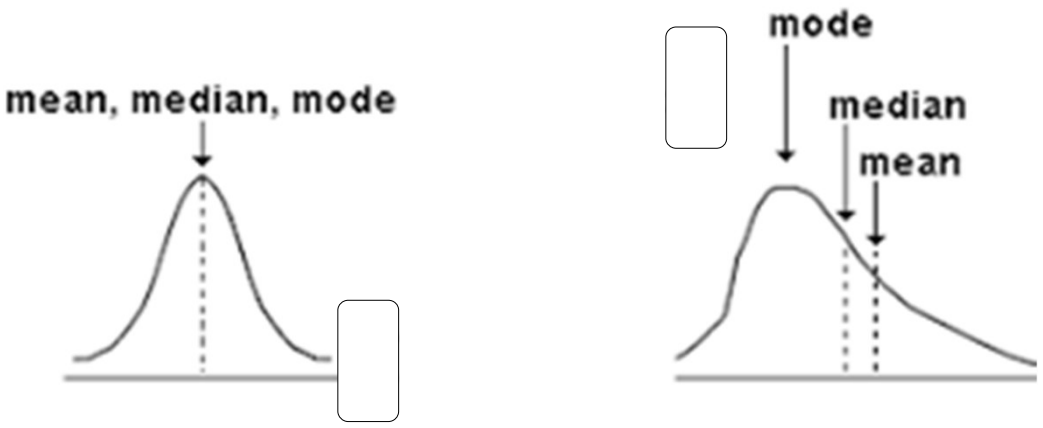


Figure 9-1 Idealized Gaussian (normal) distribution showing areas under the curve corresponding to mean ±1, 2, and 3 standard deviations (SD).

Divergence in Data Distribution (Skewness vs Kurtosis)



Normal vs Non-Gaussian Distribution



Non-Gaussian Distribution

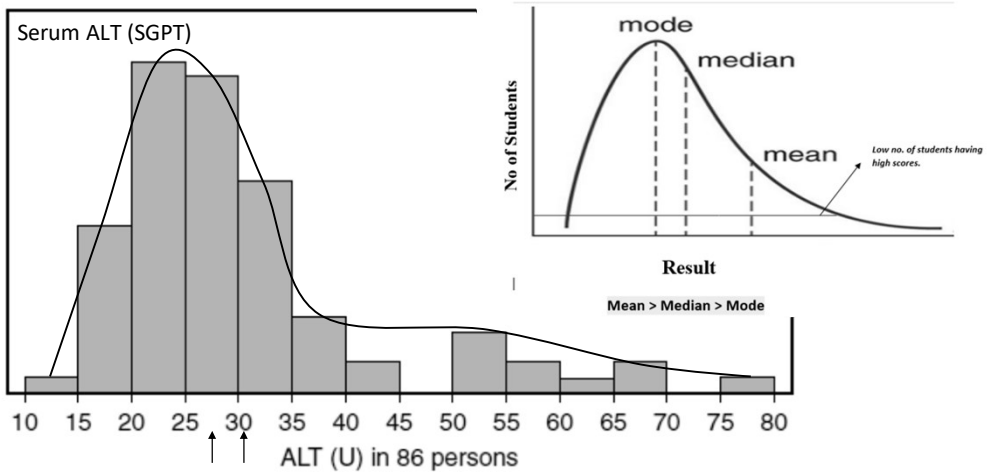


Figure 10.3 Distribution of alanine aminotransferase (ALT) in the serum of 86 healthy individuals.

CHAPTER 10

LABORATORY STATISTICS

Richard A. McPherson

KEY POINTS

- For statistical analyses, nominal variables can take on only a limited number of values (or categories), whereas continuous variables are used to report quantitative data.

accuracy also are based on these principles. For data that are not continuous but take on only two or a few discrete values (e.g., positive or negative), the analysis might consist of counting the number in each category and looking at the proportions of all values by category.

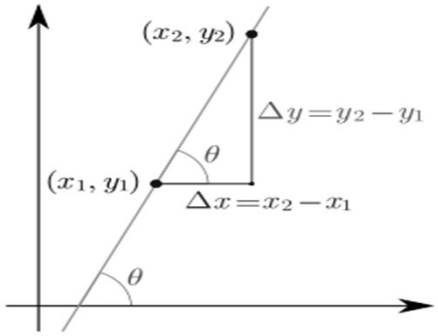
Comparison of data typically asks the question whether one group is

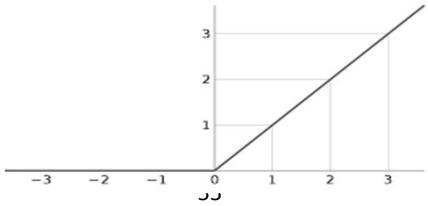
Straight Line Equation

$$y = f(x) = m.x + b$$

$$m = \frac{\text{change in } y}{\text{change in } x} = \frac{\Delta y}{\Delta x}$$

$$m = \text{tg}\ \theta$$





y = f(x) = 1.x + 0

y = x

CORRELATIVE STATISTICS & TREND EVALUATION

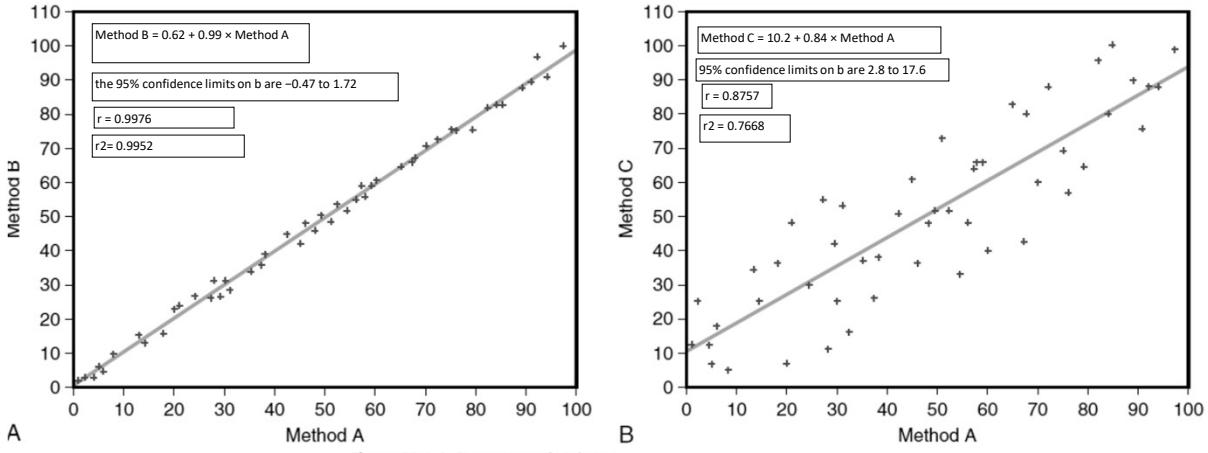
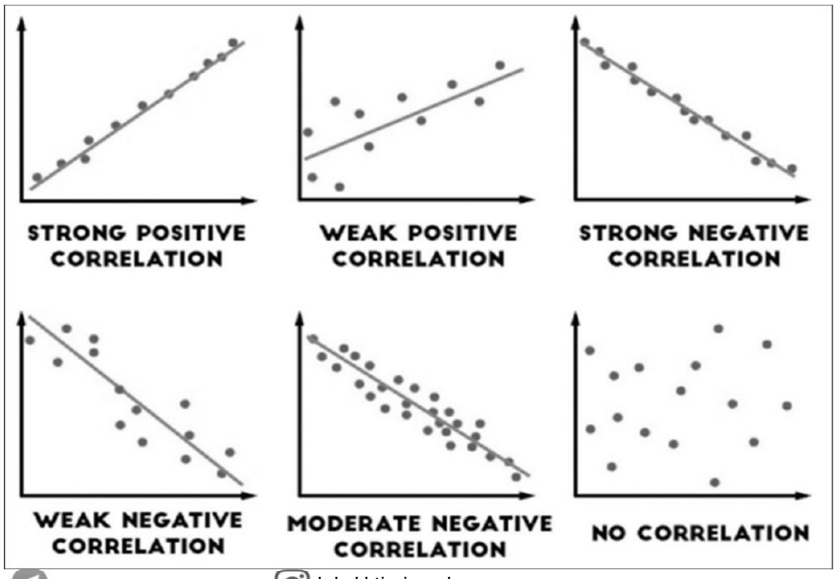


Figure 10.6 A, Regression analysis between method A and method B; strong correlation. **B,** Regression analysis between method A and method C; weaker correlation.

CORRELATIVE STATISTICS & TREND EVALUATION



58/79

Good Luck !

- **حضرت حافظ:**

آسایش دو گیتی تفسیر این دو حرف است
با دوستان مروت با دشمنان مدارا

موفق باشید !

bakhtiari.09@gmail.com