Basic Statistical Analysis and Interpretation

http://orcid.org/0000-0003-0006-8841
Dave E. Marcial
www.davemarcial.net
demarcial@su.edu.ph

Concepts and Variables
Structure of Statistical Analysis
Statistical Packages/Software

Why you need to use statistics in your research?

Why you need to use statistics in your research?

- ✓ measure things;
- ✓ examine relationships;
- ✓ make predictions;
- ✓ test hypotheses;
- ✓ construct concepts and develop theories;
- ✓ explore issues;
- ✓ explain activities or attitudes;
- ✓ describe what is happening;
- ✓ present information;
- ✓ make comparisons to find similarities and differences;
- ✓ draw conclusions about populations based only on sample results.

What is statistics?

- is a range of **procedures** for ***gathering***, ***organizing***, ***analyzing*** and ***presenting*** quantitative **data.**

'Data' is the term for facts that have been obtained and subsequently recorded, and, for statisticians, 'data' usually refers to quantitative data that are numbers

a scientific approach to analyzing numerical data.

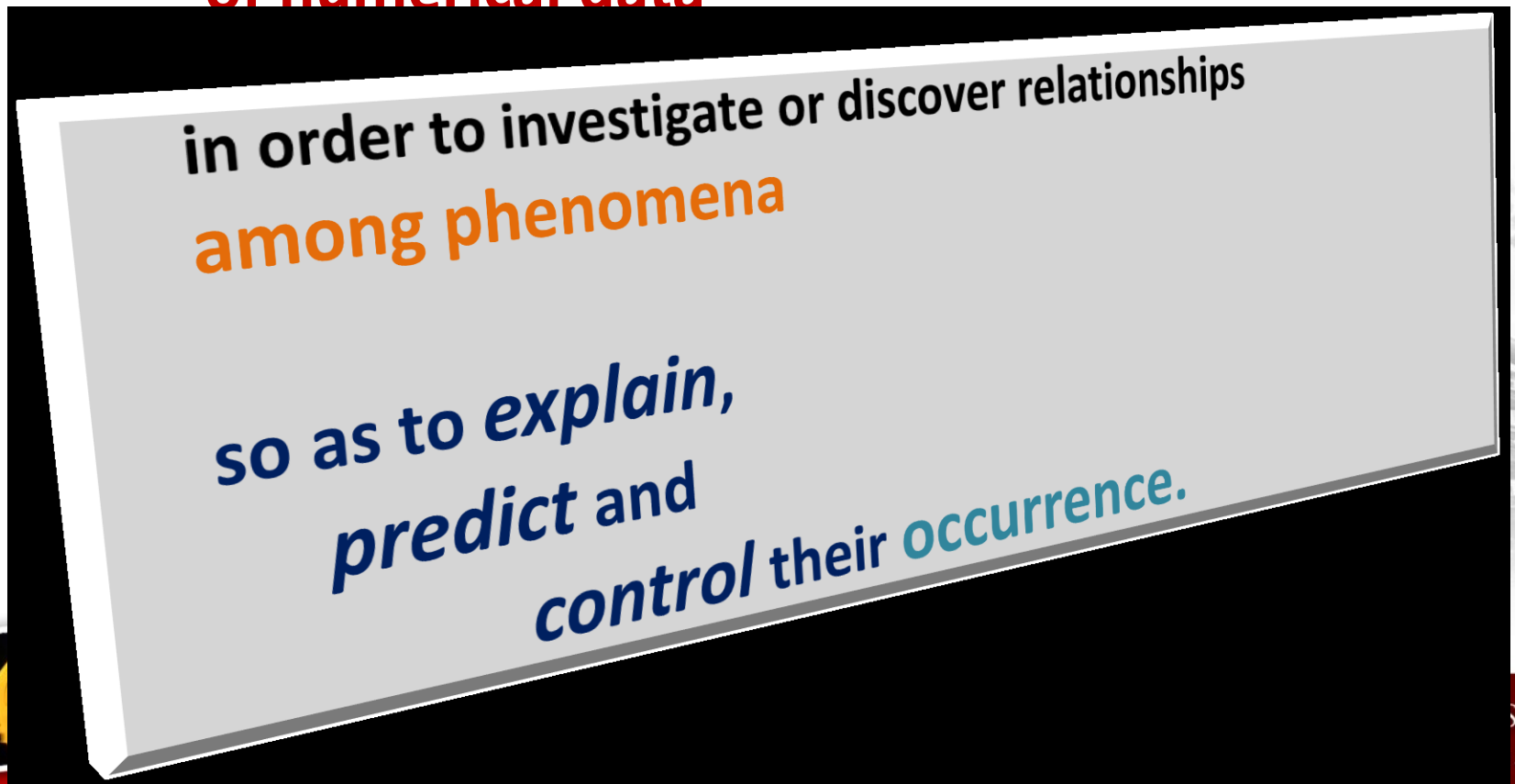
in order to enable us to maximize our interpretation, understanding and use

What is statistics?

is the **systematic**

collection and analysis

of numerical data



Objectives

Descriptive

To summarize and describe sets of observations

Inferential

To make an inference (determine significant differences, relationships between sets of observations)

Exploratory

Artificial classification of sets of observations

Variables

- is a concept that can take two or more values
- is the thing that is measured or counted; the thing of interest.

—Ex:

- Sex

(male, female);

- marital status (single, married, divorced, widowed)

Variables

Independent Variables

* causes changes in another

Dependent Variables

* a variable that is affected or explained by another variable

Ex:

- "family status and scholastic achievement"
- Independent: family status
- Dependent: scholastic

Variables

Discrete

* measurement uses whole units or numbers, with no possible values between adjacent units
* counted not measured

Ex: family size: 2, 4, 7

Continuous

* are measured, not counted
* measurement uses smaller increments of units

Ex: height, distance, time, age, temperature etc

Variables

Discrete

* if sample size is < 40, the data set is not normally distributed (non-parametric test)

Continuous

* has the tendency to assume a normal distribution (parametric tests)

The type of data set is one of the determinants in choosing the appropriate analysis.

Levels of Measurement

-Young / old -Single / married / widowed -Nationality -Type of shoes -Skin color -Type of music	-Size (smallest, small, big, biggest) -Quality (poor, good, very good, excellent)	temperature -Calendar time -Attitude scales -IQ scores	-Number of family members -Weight -Length -Distance -Number of books

Non-parametric **Parametric**

Levels of Measurement

NOMINAL

- **simplest, lowest, most primitive type**
- **involves classification of events into categories that must be distinct, one-dimensional, mutually exclusive and exhaustive; and the resulting scales are "naming" scales**
- Characteristics:
  - It involves nominal categories & is essentially a qualitative and a non-mathematical measurement
  - It names and classifies data into categories
  - It doesn't have a zero point
  - It cannot be ordered in a continuum of low-high
  - It produces nominal or categorical data
  - It assumes no equal units of measurement
  - It assumes the principle of equivalence

-Male / female
-Black / white
-Young / old
-Single / married / widowed
-Nationality
-Type of shoes
-Skin color
-Type of music

Levels of Measurement

ORDINAL

- involves not only categorizing elements into groups but also ordering of data and ranking of variables in a continuum ranging according to magnitude, that is, from the lowest to the highest point
  - Characteristic:
    - It refers to ranks based on a clear order of magnitude of low and high signifying that some elements have more value than others
    - The numbers have actual mathematical meaning as well as having identification properties
    - It is essentially a quantitative measurement
    - It shows a relative order of magnitude

-Status (low, middle, high)

-Size (smallest, small, big, biggest)

-Quality (poor, good, very good, excellent)

Non-parametric

Levels of Measurement

INTERVAL

- Provides information about the distance between the values, and contains equal intervals, ordering subjects into one of them
  - Characteristic:
    - It includes equal units
    - It is essentially quantitative measurement
    - It specifies the numerical distance between the categories
    - It does not have a true zero point

-Degrees of temperature

-Calendar time

-Attitude scales

-IQ scores

Parametric

Levels of Measurement

RATIO

- includes the other three forms offer, plus the option of an absolute true zero as its lowest value, which in essence indicates absence of the variable in question.
- Allows the researcher to make statements about proportions and ratios, that is, to relate one value to stimulus

-Interval level with 0

-Number of family members

-Weight

-Length

-Number of books

Parametric

Important items to consider in choosing a particular analysis

1. The problem or the specific objective

If the problem requires for the data to be summarized and described → **Descriptive Statistics**

If the problem requires for an inference to be made → **Inferential Statistics**

If the problem requires for data to be classified or pattern determined → **Exploratory Statistics**

Important items to consider in choosing a particular analysis

2. The type of data set

⇒Discrete Data (counts, ranks)

Non-Parametric Tests

⇒Continuous Data (ratio, interval)

Parametric Tests

Important items to consider in choosing a particular analysis

3. Number of Variables

There are different tests for 2 variables and > 2 variables

Important items to consider in choosing a particular analysis

4. The population where the samples were taken

⇒Dependent Population

data of variables to be compared were taken from the same population (e.g. before and after experiment measurements)

⇒Independent Population

data of variables to be compared were taken from two separate and distinct population

B. Structure of Statistical Analysis