Statistical Instruments and References Writing in Research

A. Yusuf, Ph.D.

Department of Arts and Social Sciences Education, Faculty of Education

University of Ilorin, Ilorin

What is Statistical Data Analysis? Data are not Information!

- Statistics is a set of methods that are used to collect, analyze, present, and interpret data.

- Statistical methods are used in a wide variety of occupations and help people identify, study, and solve many complex problems. In the education, business and economic world, these methods enable decision makers and managers to make informed and better decisions about uncertain situations.

Computer and Data Analysis

- Statistical software packages, include SPSS, Stata, Minitab, Mathlab, E- view

We do not need to know how to manually calculate?

- Computer software packages offers extensive data-handling capabilities and numerous statistical analysis routines that can analyze small to very large data statistics.

- The computer will assist in the summarization of data.

What we must Know

We must know (the 4Hs)

How appropriate is the statistical method for what research Design?

How to interprete?

How to make inferences?

How to predict?

Four Basic Steps Involved in Solving Problems

- Defining the problem

- Collecting the data

- Analyzing the data

- Reporting the results

Defining the Problem

- If we want to obtain accurate data, clear definition of the problem is necessary.

Collecting the Data

- Designing ways to collect data is an important job in statistical data analysis. Two important aspects of a statistical study are:

- **Population**: a set of all the elements of interest in a study

- **Sample**: a subset of the population

- Statistical inference refers to extending your knowledge obtained from a random sample from a population to the whole population. This is known in mathematics as an Inductive Reasoning.

What Type of Data to Collect?

- cross-sectional and time series data

- Categorical nominal-observations that can be coded

- Continuous -Observations that can be counted or measured

- Mixed matrix of categorical and continuous data

How to Collect Data?

- Observational study: In observational studies, no attempt is made to control or influence the variables of interest. For Example, a survey

- Experimental Studies: In experimental study, the variable of interest is identified. Then one or more factors in the study are controlled so that data can be obtained about how the factors influence the variables. For Example, true experiment or quasi experment
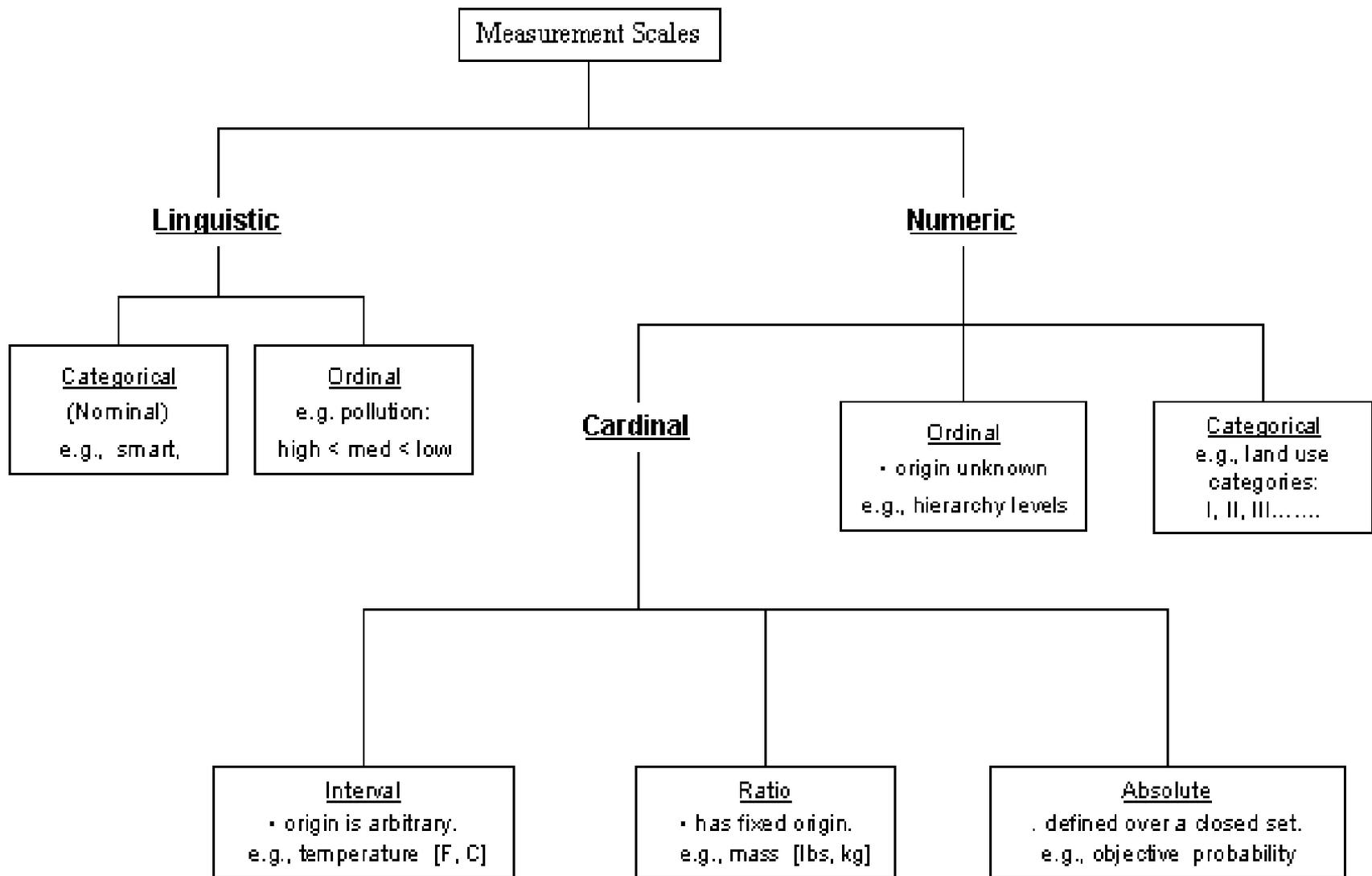
Analyzing the Data

- exploratory methods and confirmatory methods.

- **Exploratory methods** are used to discover what the data seems to be saying by using simple arithmetic and easy-to-draw pictures to summarize data.

- **Confirmatory methods** use ideas from probability theory in the attempt to answer specific questions. Probability is important in decision making because it provides a mechanism for measuring, expressing, and analyzing the uncertainties associated with future events.

Reporting the Results

- The results may be reported in the form of a table, a graph or a set of percentages.

- Because only a small collection (sample) has been examined and not an entire population, the reported results must reflect the uncertainty through the use of probability statements and intervals of values.

Type of Data and Levels of Measurement

- remember the French word NOIR for colour black
- Nominal
- Ordinal
- Interval
- Ratio

```
                        ┌─────────────────────┐
                        │  Measurement Scales │
                        └─────────────────────┘
```

Measurement Scales

- **Linguistic**
  - Categorical (Nominal) — e.g., smart,
  - Ordinal — e.g. pollution: high < med < low

- **Numeric**
  - **Cardinal**
    - Interval — · origin is arbitrary. e.g., temperature [F, C]
    - Ratio — · has fixed origin. e.g., mass [lbs, kg]
    - Absolute — . defined over a closed set. e.g., objective probability
  - Ordinal — · origin unknown — e.g., hierarchy levels
  - Categorical — e.g., land use categories: I, II, III.......

Measurement Scales

Types of Statistics

- There are two types of statistics
- descriptive
- inferential

Descriptive Statistics

- "Descriptive Statistics involve measuring data using graphs, tables, and basic descriptions of numbers such as averages or means. These universally accepted descriptions of numbers are called parameters" (Coolidge, 2000, p. 5-6).

- It describes a sample's characteristics through the measures of central tendency, variability, and relationship.

Inferential Statistics

- Making conclusions about the population (a large group of data) from the sample's characteristics (a small group of data).

- A general formula in using inferential statistics is Fly IDAIR--identify the problem, design the statistical test, apply the method, infer from the test, and reporting the results.

FAMILIES OF TESTS

- parametric and nonparametric
- Parametric tests are based upon the assumption that the data are sampled from a Gaussian distribution.
- Nonparametric tests are based upon the assumption that the data are sampled from a Non-gaussian distribution.
- Tests that do not make assumptions about the population distribution are referred to as nonparametric- tests.

Goal	Type of Data	
	Measurement (from Gaussian Population)	**Rank, Score, or Measurement (from Non-Gaussian Population)**
Describe one group	Mean, SD	Median, interquartile range
Compare one group to a hypothetical value	One-sample t test	Wilcoxon test or Chi Square
Compare two unpaired groups	Unpaired t test	Mann-Whitney test
Compare two paired groups	Paired t test	Wilcoxon test
Compare three or more unmatched groups	One-way ANOVA	Kruskal-Wallis test
Compare three or more matched groups	Repeated-measures ANOVA	Friedman test
Quantify association between two variables	Pearson correlation	Spearman correlation Keddar
Predict value from another measured variable	Simple linear regression or Nonlinear regression	Nonparametric regression**
Predict value from several measured or binomial variables	Multiple linear regression* or Multiple nonlinear	

CHOOSING BETWEEN PARAMETRIC AND NONPARAMETRIC TESTS

Parametric	Nonparametric
...atistical technique used for group ...mparison when the characteristic ...ing studied (e.g., learning outcomes) ... normally distributed in the ...pulation, sample was randomly ...lected, and data being analyzed are ...erval or ratio (e.g., test scores).	Statistical techniques used for group comparison the characteristic being studied is not normally distributed in the population, sample size is small and not randomly selected, and data being analyzed are ordinal (rank) or nominal (categories).
...est: for dependent samples (same ...oup) and independent samples (two ...fferent groups) T test is used to test ...e statistical significance of mean ...fferences of one or two groups. ...alysis of Variance (ANOVA): ...NOVA is similar to t test. But used ...hen you compare more than two ...oups or have more than one ...dependent variable. ...alysis of Covariance (ANCOVA): ...milar to ANOVA but for controlling ...e influence of an IV that may vary ...tween groups before the treatment is ...plemented. ...ultivariate Analysis of Variance ...IANOVA): multivariate ANOVA. ...sed when there is more than one ...sponse variable.	**Wilcoxon matches pairs test (t-equivalent):** used with dependent samples and ordinal data. McNemar's test, used in certain 2 × 2 tables with pairing **Mann-Whitney U Test (t-equivalent):** used with two independent samples and ordinal data. **Mann—Whitney—Wilcoxon** (MWW) or Wilcoxon rank-sum test) is a non-parametric statistical hypothesis test for assessing whether one of two samples of independent observations tends to have larger values than the other. It is one of the most well-known non-parametric significance tests. **Friedman Two-Way Analysis of Variance:** used with more than two dependent samples and ordinal data. **Kruskal-Wallis One Way Analysis of Variance:** used with more than two independent samples and ordinal data. **Chi-Square (for categorical data):** used to test the statistical independence of two variables (e.g., gender and learning styles).

The Meaning and Interpretation of P-values (what the data say?)

P-value	Interpretation
P< 0.01	very strong evidence against H0
0.01£ P < 0.05	moderate evidence against H0
0.05£ P < 0.10	suggestive evidence against H0
0.10£ P	little or no real evidence against H0

Statistics for questions of description.

- For a question of description, you will use statistical measures of central tendency and dispersion. For normally distributed "bell curve" data measured on an ordinal, interval, or ratio level of measurement, you will report mean, median, mode, range, and standard deviation. For non-normal data, you will report median, mode, and range, but for data reported on a nominal level of measurement, you will report only mode.

Select statistics for questions of difference

- . For a question of difference, you will use statistics regarding the significance of group differences. For normally-distributed "bell curve" data measured on interval or ratio levels of measurement, you will conduct a t-test for a comparison of two groups, or an analysis of variance / ANOVA for a comparison of three or more groups. For non-normal data, with data reported on a nominal level of measurement, you would use nonparametric statistics like chi square; with data reported on an interval level, you would use nonparametric statistics such as the Mann-Whitney U test.

Statistics for questions of association

- For normally-distributed "bell curve" data measured on interval or ratio levels of measurement, you will calculate a Pearson product moment correlation, r.

- For non-normal data, with data reported on a nominal level of measurement, you will calculate a contingency coefficient, C. For non-normal data reported on an ordinal level of measurement, you will calculate a Spearman or Kendall rank correlation, r-sub-s or tau.

Select statistics for questions of prediction.

- For normally-distributed "bell curve" data, measured on interval or ratio levels of measurement, you will calculate a regression or multiple regression equation. You may also try doing this for data measured on an ordinal level of measurement, but results must be interpreted cautiously.

What are some of the Tests Used on Categorical Data?

- Chi-squared test-to determine the relationship between variables
- Fisher' s test-to compare two unpaired groups McNemar's test-to analyze a matched case-control study
- Kappa-to quantify interafter agreement
- Wilcoxon test- to compare one group to a hypothetical value
- Kruskal-Wallis test-to compare three or more groups Spearman test-to compare association between variables

What are some of the Tests Used on Continuous Data?

- T test- to compare one or two groups ANOVA-to compare three or more groups

- ANCOVA-

- MANOVA-

- MANOCOVA-

- Pearson test-to compare the association between variables

How to Draw Conclusions from Data?

- Use of graphical presentations Use of statistical analyses Sharing data among colleagues and receiving constructive feedback Critically analyzing data and results

How to Present Research Findings?

- Tables-matrix of rows and columns representing variables

- Figures-visual organization of data/observations -pictures -pie charts -line charts -bar charts -flow charts -organizational charts -cartogram charts -Gantt charts -scatter plot charts

How to Present your Paper According to the APA style?

- APA stands for American Psychological Association APA is the most commonly used style in the Social Sciences. APA provides useful guidelines on: style in text citations and references

APA Style of writing

- APA Style Third person should be used.

- Active voice should be used.

- Writing should be clear, concise and plain.

- General format should be: paper size-8.5 x 11 1â€• margins on all sides double spaced written in 12 point font using Times New Roman short title and page number on the upper right hand corner of every page

GENERAL TEMPLATE

- General Format of APA Paper
- Title
- Abstract
- Introduction
- Methodology
- Results
- Discussion of Findings
- Conclusions and Recommendations
- References

APA Referencing Style

- Last name of author, followed by initials Last names of authors arranged alphabetically in the reference list

- Capitalize the first letter of the first word in titles and subtitles

- Capitalize all major words in journal titles Italicize titles of books and journals

- APA In-text Citation Authors last name and year of publication placed in parenthesis next to where the information is cited Three or more authors, use all three when first cited then followed by the first author et al., year for subsequent in-text citations Six or more authors, use the last name of the first author followed by et al., year of publication

How to reference material

- King, M. (2000). *Wrestling with the angel: A life of Janet Frame.* Auckland, New Zealand: Viking.

- Treviño, L. K., & Nelson, K. A. (2007). *Managing business ethics: Straight talk about how to do it right*. Hoboken, NJ: Wiley.

THANK YOU

A. Yusuf, Ph.D.

THE END