DCTE 700 / 800
Research Methodology
Instructor Notes

Steve Terrell, Ed.D.
Professor
Nova Southeastern University

In the beginning…

All research starts out by identifying a problem, opportunity or area of inquiry about which we want to find a solution or gain better knowledge.

We can use the following general approaches to help us in our inquiry:

- Qualitative approaches
- Problem-solving approaches
- Quantitative approaches

What is Qualitative Research?

The collection of extensive narrative data on many variables over an extended period of time, in a naturalistic setting, in order to gain insights not possible using other types of research.

Gay and Airasian (2000)

Characteristics of Qualitative Research Strategies

- The researcher attempts to understand an issue from an insider's perspective
- The researcher is the primary instrument for data collection and analysis
- In many instances, qualitative research involves fieldwork
- The research usually uses an inductive approach

Examples of Qualitative Research

- Historical research – used to study, understand and explain past events as well as predict future events
- Grounded Theory – used to develop theories in an inductive manner
- Ethnography – study of cultural patterns and perspectives in a natural setting
- Case Study – an in-depth investigation of one individual, group, program, document, etc.

Problem Solving Approaches:

- Developmental – action oriented research where a new good, product or process is developed in order to address a problem or opportunity. Examples include new curriculums, computer programs and textbooks.

- Evaluative – action oriented research where new tools or techniques are evaluated to test their feasibility or efficacy. This approach can be used to evaluate the results of a development study.

What is Quantitative Research?

A research technique that places heavy emphasis on the collection and analysis of numeric data in order to explain or predict a phenomenon of interest. There are two types:

1. Descriptive research – used to study, understand and explain current events

2. Experimental research – used to investigate "cause and effect" relationships

This class focuses on quantitative research!

- While it is feasible to write a dissertation that is qualitative, developmental or evaluative in nature, our goal is to become familiar with quantitative research.

- Quantitative research is one of the building blocks of having a PhD.

- Many of the specific tools and procedures we will discuss are applicable to the other methodologies.

- A bibliography is attached showing texts dealing with all approaches.

We can learn more about a problem or opportunity by using:

- Past experience or experience of others
- Authority figures
- Both inductive and deductive reasoning
- The scientific method

The scientific method is:

- Orderly
- Methodical
- Used in both education and industry
- Similar to developmental methodologies
- The best way to ensure valid and reliable results when investigating your problem

Four steps in the scientific process:

- Recognize and identify a topic to be studied.
- Describe and execute procedures to collect information about the topic being studied.
- Analyze the collected data.
- State the results or implications based on analysis of the data.

Step 1: Recognize and Identify a Topic to Be Studied

(Gay and Airasian Task Two)

1. Introduction

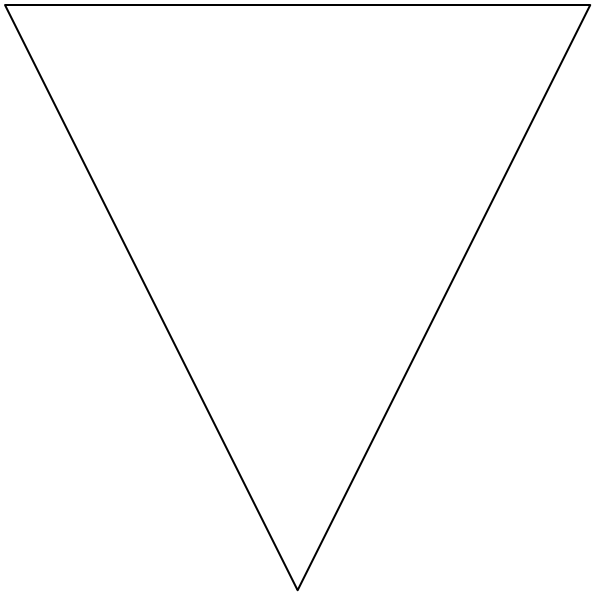
2. Actual Statement of the Problem

3. Review of Literature

4. Statement of the Hypothesis

Characteristics of a Good Problem

General Problem Area



Specific, manageable problem!

- It can be researched through the collection and analysis of data
- It has theoretical or practical significance
- The researcher is comfortable in terms of the skill, resources, time and knowledge needed to conduct the research.

Sources of problems

- Existent theory
- Personal experience
- The literature
- Replication of previous studies
- Problems in your own workplace or institution

An Example Introduction

Researchers (Smith, 1998) have noted that, although feedback is an integral part of the development of achievement motivation, elementary classroom teachers do not provide feedback in a manner that is conducive to the growth of such motivation. Given the relationship between feedback, motivation and achievement (Jones, 1999), it is imperative that educators attempt to develop and implement better methods of achievement oriented feedback.

An Example Problem Statement

Students have low levels of academic achievement. The purpose of this study is to investigate the effect of alternative methods of feedback on levels of student motivation and achievement.

Review of the Literature

- Provides a foundation for the acquisition of new knowledge
- Helps in identification, location and analysis of documents related to the problem
- Addresses problem area
- Determines what has already been done
- Helps determine research strategies, procedures and instrumentation to use
- Aids in interpretation of results

Steps in the ROL per Gay and Airasian

1. Identify key words associated with your problem
2. Identify sources of information
3. Evaluate your sources
4. Analyze, organize and report

Steps 1 thru 3: Identify key words, sources of information and evaluate your sources

- Consider primary and secondary sources
- The more research done,the less tangential your sources
- Use reliable sources
  - Journal articles and conference papers
  - Books
  - Search engines
  - Prior dissertations
- The watchword is “validity”

For Example:

- Keywords: achievement, motivation, feedback, metacognition
- Outline:
  - History of motivational research
    - Traditional Theories
  - Intrinsic and extrinsic motivation
    - Competence theorists
    - Cognitive theorists
    - Attribution theorists
  - Relationship between motivation and achievement
  - Research models

Putting it Together

- Outline, outline, outline!
- Synthesize!
- The ROL should flow – introduction, body, conclusion
- The most recent and related references should be last in a section
- **Review of Literature Example**

After You've Finished the ROL, Ask Yourself:

- Have I thoroughly covered my problem area?
- Have I provided a sound theoretical and conceptual framework for my study?
- Does the reader understand what I am saying?
- Am I making assumptions?
- Am I writing to my target audience?

What is a Hypothesis?

A statement expressing the researcher's beliefs about an event that has occurred or will occur. A well-stated hypothesis has four requirements:

1. It must be stated clearly and concisely.

2. It must be consistent with prior research or observations.

3. It must provide a reasonable explanation for the event that occurs.

4. It must be testable via the collection and analysis of data.

Two Terms We Need to Know

Independent variable – anything that we believe makes a difference in a behavior

Dependent Variable - the content area or construct about which we are collecting data

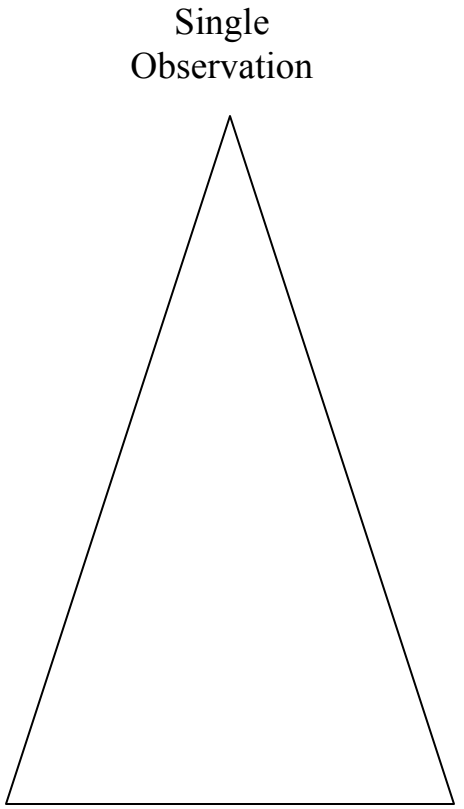
.

Types of Hypotheses

- Source = Inductive or deductive
- Type = Research or statistical
- Direction = Directional or non-directional
- These three types can be combined in that you can have, for example, an inductive, directional hypothesis.

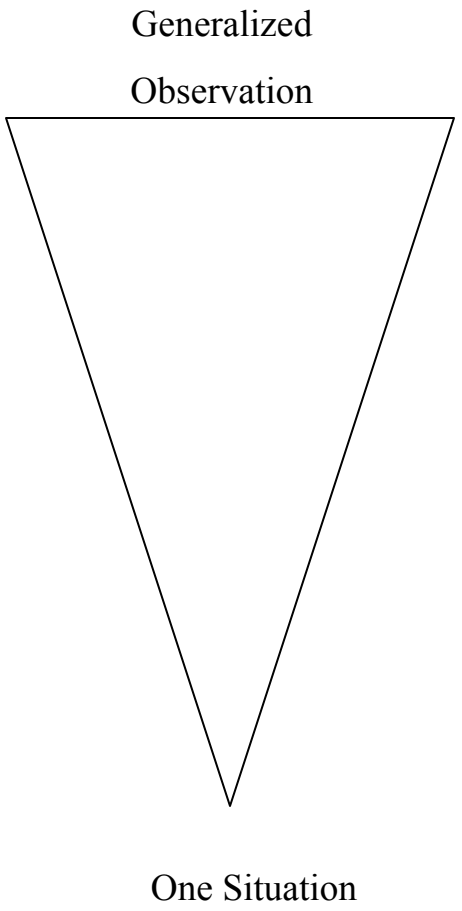
Inductive Hypotheses

- Inductive hypotheses are developed by taking a single observation and trying to generalize it
- Bob is short and Bob is smart, therefore all short people are smart
- Primarily used in qualitative research



Deductive Hypotheses

- Deductive hypotheses are the opposite in that they take an generalization and make it applicable to a give situation
- All research classes are difficult and this is a research class, so this research class is difficult
- More the realm of quantitative research



Stating Your Hypothesis

- Research (alternate) hypothesis:
  - Students receiving CAI will score significantly higher on the SAT than students receiving lectures.
- Statistical (null) hypothesis:
  - There will be no significant difference (different due to reasons other than chance) in SAT scores between children who get CAI and those that receive lectures.

Directional Hypotheses:

- A hypothesis can be directional which means that it contains a > or < sign. This is also called a "one-tailed" test. For example:
  - Students sleeping less than eight hours per night will have significantly lower levels of achievement than students getting eight or more hours of sleep.
  - Students receiving graphical report cards will have significantly higher levels of intrinsic motivation than students receiving traditional report cards.

Non-directional Hypotheses:

- A non-directional hypothesis has an inferred "not equal" condition. This is called a "two-tailed" test. For example:
  - There will be a significant difference in motivational levels of students getting daily report cards and those getting monthly report cards.

Testing Your Hypotheses

The purpose of your study will be to test your hypothesis. In doing so, you will either "reject" or "fail to reject" your null hypothesis.

Rejecting your null hypothesis simply means that the null hypothesis is not true – in other words, there is a significant difference between the groups you have measured. This is the same as saying that the research hypothesis is true.

Failing to reject your null means that no significant difference exists between the groups being measured; in other words, the null hypothesis is true.
This is the same as saying that the research hypothesis is not true.

Some “hypothetical” closing thoughts…

- The hypothesis must be a logical extension of the review of literature.
- Don’t state your hypothesis and then look for a problem to match it!
- If you conduct a good study, it’s not important whether or not you reject your null hypothesis. What is important is that your results are valid.

Step 2: Describe and execute procedures to collect information about the topic being studied.

- Gay and Airasian
  - Task 3: Create a Brief Research Plan
  - Task 4: Describe Your Sample
  - Task 5: Selection of Instrumentation
  - Task 6: Combination of Tasks 3 thru 5

The Validity of a Study is Our Most Important Concern

- Internal Validity
  - The degree to which changes in the dependent variable are affected by the manipulated independent variable. Maintaining high internal validity means controlling for all other independent variables other than the one(s) being studied

- External Validity
  - The degree to which the results of a study can be generalized to the "real world". Factors that negatively affect external validity also negatively affect the generalizability of the results

The Validity of a Study Depends on:

- The sampling methodology used
- Instrument reliability and validity
- The design of the study

What is Sampling?

Sampling is the selection of a group of persons for a study (i.e., the sample) that is representative of the entire population (i.e., the sample frame). The purpose of the sample is to gather information that is generalizable to the population.

Random Sampling

In research we are generally unable to use the entire population because of size, inconvenience, cost, etc. With this in mind, we try to select a random sample that accurately reflects the population from whence it was drawn. This homogeneity is important since the researcher wants to make inferences based on statistics calculated about the sample. Concerns arising from sampling include reliability, validity, systematic error and random error.

Sampling Definitions

- Parameters
  - Numeric values computed about all of the values in a population. For example, the average age of all students in a high school.
- Statistics
  - Numeric values computed about a sample drawn from a population. For example, if we randomly select 100 students from a high school and compute the average age, it is called a “statistic”.

More Definitions!

- Sampling Error
  - Error introduced into a sample due to some fault of the researcher

- Sampling Bias
  - Error that is expected due to the very nature of a sampling process. We can statistically control for this.

Even More Definitions!!

- Generalizability
  - The degree to which results of a study can be generalized to other populations

Steps in Sampling

- Identify the population (the sample frame)
- Determine the sample size needed
- Select the sample using most appropriate sampling method
- Avoid sampling bias!

Sampling Rules of Thumb

- The larger the sample, the more generalizable
- Sample sizes (Gay & Airasian, 2000)
  - Descriptive: 10% for small populations, smaller percentages for larger populations
  - Correlational: at least 15 subjects per group
  - Causal-comparative: at least 30 subjects per group
  - Experimental studies: 25 subjects per group
- Randomization is the key word
- Statistical formulas exist for exact sample size

Probabilistic Sampling

- Random sampling
  - All subjects have an equal chance
  - Best representative sample
- Stratified sampling
  - Selection is made from subgroups
  - Sampling repeated for each subgroup
- Cluster sampling
  - Clusters or groups from the population are sampled
- Systematic sampling
  - Every nth member of the population is selected
  - Is random if list of population is random

Non-probabilistic Sampling Methods

- Convenience sampling
  - Data is collected from a readily available group
- Quota sampling
  - Researchers given exact characteristics and quotas of persons to be interviewed
- Purposive sampling
  - Researcher selects sample based on experience

Problems with Non-probabilistic Sampling

- Cause sampling bias
- Create generalizability problems
- Generally is the fault of the researcher
- Must be detailed in final report as delimitation

Two Rules for Instrumentation!

- Validity is your primary concern!
- Reliability is a secondary concern!

Validity: Does Your Instrument Measure What It's Supposed to Measure?

- **Achievement Tests**
  - Purchased or developed
- **Personality Inventories**
  - Non-projective inventories (e.g., Myers-Briggs)
  - Attitude scales
  - Interest inventories
  - Projective inventories (e.g.,Rorschach, Draw an Animal)
- **Aptitude Tests**
  - General
  - Specific
  - Readiness

Three Measures of Validity

1. Construct validity
2. Criterion related validity
3. Content validity

Construct Validity

- The degree to which an instrument actually measures whether or not an underlying construct is being measured.

- Does a personality test actually measure personality?

Criterion Related Validity

- Concurrent validity
  - Degree to which scores on one test are correlated with scores on another test administered at the same time.

- Predictive validity
  - Degree to which scores on one test predicts scores on a test administered in the future.

Content Validity

- Content validity is the degree to which a specific content area is measured. Concerns include:
  - Item Validity: are all of the items representative of the domain to be measured?
  - Sampling Validity: has the whole domain been covered? Are there items outside of the domain?
  - Established by expert judgment

Measures of Reliability

- **Reliability is the consistency with which an instrument measures the construct or content area it is intended to measure**
  - Stability (test / re-test)
  - Equivalence (alternate forms)
  - Equivalence and Stability combined
  - Internal consistency
  - Scorer-rater reliability

Stability

- Often called “test / re-test”
- Measures whether scores on one administration of the test are equivalent to scores on another administration
- One test administered at two different times
- Reliability coefficient with a range of 0.00 to 1.00 is calculated
- Has problems with measurement error

Equivalence

- Often called “alternate forms”
- Used to determine if two forms of same test are equivalent
- One group, two administrations over time
- High correlation means high equivalence
- Used when alternate forms are desired for pre / post test
- Problems exist with measurement error

Equivalence and Stability

- Best approach to reliability
- Two forms of test are administered to one group at two different times
- Reduces problems with measurement error

Internal Consistency

- Focuses on the degree to which the individual items are correlated with each other and is thus often called homogeneity.
  - Split-half
    - Used with dichotomous tests
  - Kuder-Richardson 20 / 21
    - Improvement on split-half
  - Cronbach's Alpha
    - Only used with instruments with more than two scores (e.g., Likert Scales)

Scorer / Rater Reliability

- Used with subjective, open-ended test, observations, etc.
- Used to ensure that ratings are reliable across raters
- Used to ensure ratings are consistent for a given rater

Instrument Selection Guidelines

- It's almost always easier to find an instrument than to develop one.
- There are a variety of instruments for a single purpose. Find the one that best suits your purposes – don't pick the first one you find!
- Consider scoring time, costs, professional needs, etc.

If you develop…

- Validity is the primary consideration
- Reliability is the second consideration
- Use an expert panel to aid in the pre-validation of your instrument
- Allow ample time to pre-test the instrument to assess validity and reliability

If you do not develop your own

- Those you have access to – ensure validity and reliability
- Buro's Mental Measurements Yearbook
- Tests in Print (Previous Buro's)
- Publishers and distributors
- ETS collection
- Professional journals

Study Designs

1. True Experimental Studies
2. Quasi-experimental Studies
3. Pre-experimental Studies
4. Descriptive Studies

Results Are Generalizable to the Degree They are Valid

- Internal Validity
  - The degree to which changes in the dependent variable are affected by the manipulated independent variable. Maintaining high internal validity means controlling for all other independent variables other than the one(s) being studied

- External Validity
  - The degree to which the results of a study can be generalized to the "real world". Factors that negatively affect external validity also negatively affect the generalizability of the results

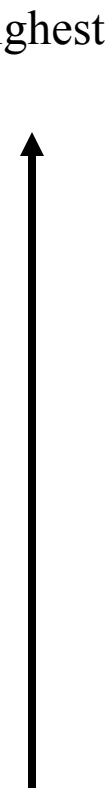
Validity and the Type of Study:

True Experimental Studies

Quasi-experimental Studies

Pre-experimental Studies

Descriptive Studies



Descriptive Studies

- Used to understand and explain current events
- There are many, many different types of descriptive studies
- Sample selection and data collection are critical
- Since new data is collected it is often necessary to develop new data collection tools (instrumentation)
- Data may be collected via self-report or observation
- Cause and effect is not stringently tested but can be implied (limited validity)

Types of Descriptive Studies

- Surveys
  - Used to examine current status of something. Common types include questionnaire and interviews.

- Relationship Studies
  - These studies explain how events or things are related to one another. Common types are correlation studies and causal-comparative studies.

- Developmental Research Studies
  - These studies look at progress, change or development over a period of time and include trend analysis and longitudinal studies. These should not be confused with developmental problem solving where a good or service is developed.

Survey Studies: Questionnaire Study

- Permits very large sample – usually mailed
- Consideration must be given to construction of instrument, selection of subjects, validation of instrument, cover letter and return rate
- Aim for 70% return rate
- Follow-up is very important

Survey Studies: Interview Study

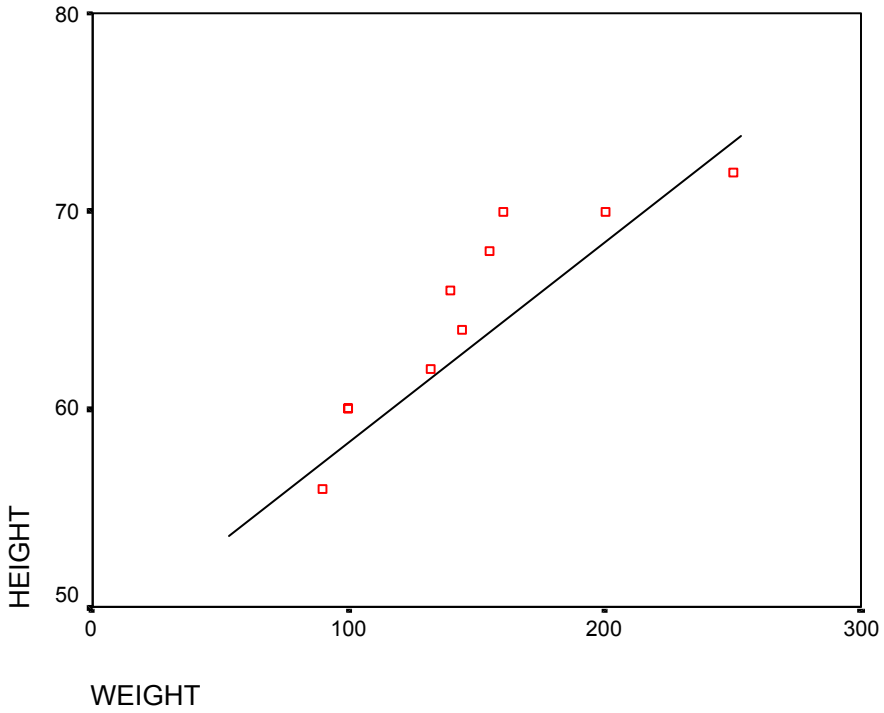
- Interviews may be structured, semi-structured or informal
- Validity is associated with interview type
- If used, considerable concern should be given to the interview guide
- Mechanical or manual recording is allowed

Relationship Studies: Correlation Study

- Data are collected for two or more variables to determine if a relationship exists between the variables
- The relationship is expressed as a correlation coefficient
- Coefficients range from –1.00 to +1.00
  - A negative coefficient implies an inverse relationship
  - A positive coefficient implies a positive relationship
  - A coefficient of zero indicates no relationship
- Correlations do not demonstrate causality but can be used for prediction or inference

Positive Correlation

Height	Weight
60.00	100.00
66.00	140.00
56.00	90.00
70.00	160.00
68.00	155.00
64.00	144.00
62.00	132.00
70.00	200.00
60.00	100.00
72.00	250.00



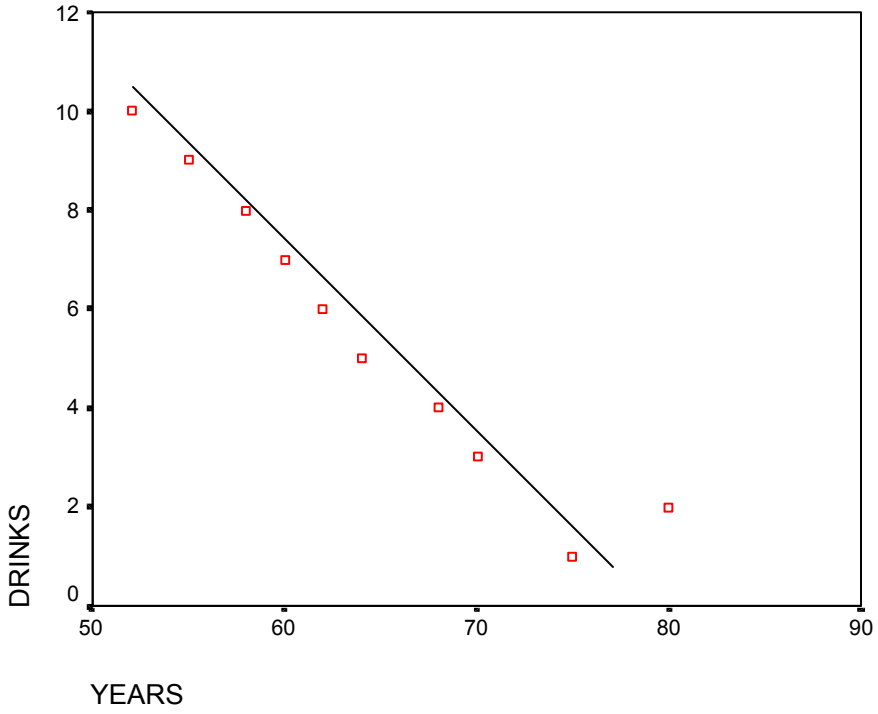
Correlations

		HEIGHT	WEIGHT
Pearson Correlation	HEIGHT	1.000	.904**
	WEIGHT	.904**	1.000
Sig. (2-tailed)	HEIGHT	.	.000
	WEIGHT	.000	.
N	HEIGHT	10	10
	WEIGHT	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

Negative Correlation

Drinks Daily	Life Expectancy
1.00	75.00
2.00	80.00
3.00	70.00
4.00	68.00
5.00	64.00
6.00	62.00
7.00	60.00
8.00	58.00
9.00	55.00
10.00	52.00



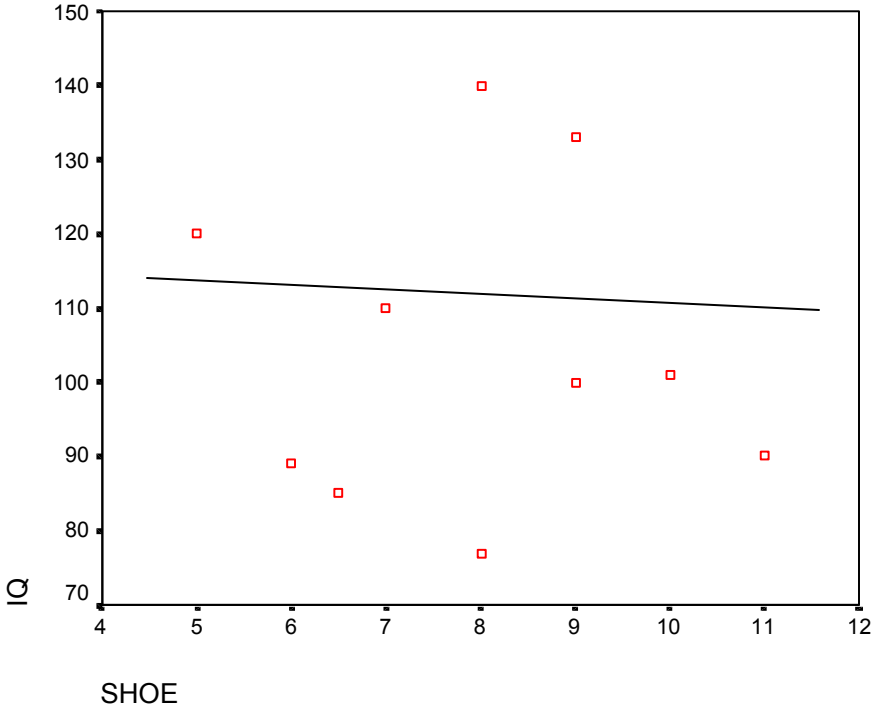
Correlations

		DRINKS	YEARS
Pearson Correlation	DRINKS	1.000	-.968**
	YEARS	-.968**	1.000
Sig. (2-tailed)	DRINKS	.	.000
	YEARS	.000	.
N	DRINKS	10	10
	YEARS	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

Zero Correlation

Shoe Size	IQ
7.00	110.00
6.50	85.00
8.00	140.00
9.00	100.00
11.00	90.00
8.00	77.00
5.00	120.00
6.00	89.00
9.00	133.00
10.00	101.00



Correlations

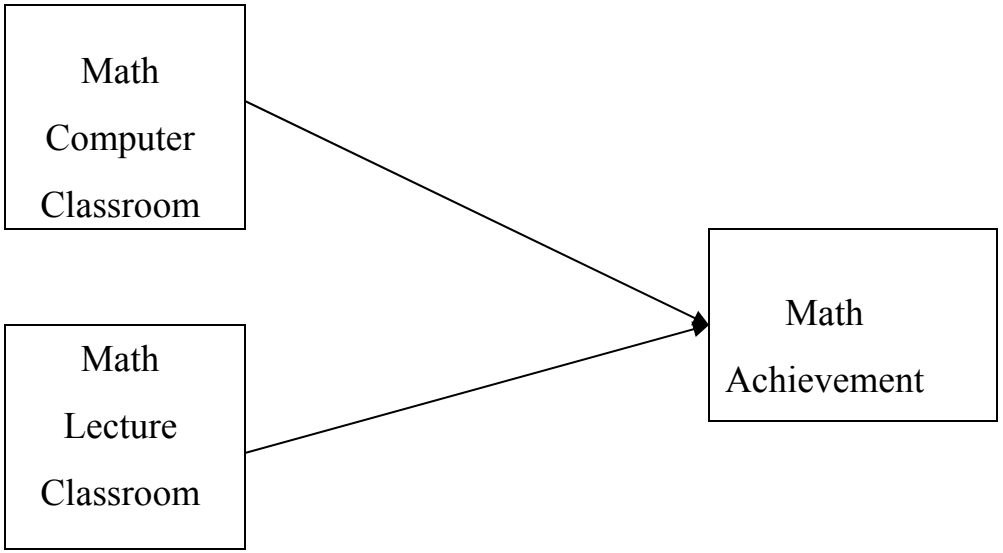
		IQ	SHOE
Pearson Correlation	IQ	1.000	-.052
	SHOE	-.052	1.000
Sig. (2-tailed)	IQ	.	.887
	SHOE	.887	.
N	IQ	10	10
	SHOE	10	10

Relationship Studies: Causal-Comparative Study

Two or more different levels of an existing independent variable are measured using a dependent variable. For example, the achievement (dependent variable) of students in self-contained classrooms (level of an independent variable called "classroom type") is compared against the achievement of students in open classrooms (another level of the same independent variable "classroom type").

Because neither treatment or group membership is randomly assigned, this does not test cause and effect but inferences can be made. The best inferences are made when the groups being measured are approximately equal on all factors that might affect the dependent variable other than the one in question.

Causal-Comparative Studies



No Random Assignment of Treatment to Groups – the classes were existent

No Random Assignment of Students to Groups – the students were already assigned to the classes

Be Careful! As will be seen, these designs look very similar to some experimental designs!

Types of Experimental Studies

- Pre-experimental designs
- Quasi-experimental designs
- True experimental designs
- The “higher” we go, the closer we get to cause and effect

The Experimental Designs and Threats to Validity

- The experimental designs are set up to control for threats to the internal and external validity of these types of studies.

- Each of these designs controls for these threats to one degree or another. The lower level designs (pre-experimental and quasi-experimental) have fewer controls than the true experimental designs.

Threats to Internal Validity

- History
  - Events during study affect dependent variable
- Maturation
  - Subjects change physically, intellectually and emotionally over time
- Testing
  - Effects of pre-tests or practice tests on post-testing
- Instrumentation
  - The lack of reliability in a single test instrument or the instrument is changed from pre-test to post-test

Threats to Internal Validity - Continued

- Statistical Regression
  - Extremely high and low scorers tend to regress to mean
- Differential Selection of Participants
  - Different characteristics of groups affect dependent variable differently
- Mortality
  - Subjects drop out, for whatever reason
- Selection Interactions
  - Maturation rates
  - History
  - Instrumentation

Threats to External Validity

- Pre-test – treatment interaction
  - Pre-test sensitizes subjects to part of treatment thereby changing the manner in which the subject reacts
- Selection – treatment interaction
  - Non-random or volunteer subjects affect generalizability
- Multiple treatment interference
  - When subjects receive multiple treatments, effect of one interferes with the other
- Specificity of variables
  - Poor specification of variables leading to no ability to generalize

Threats to External Validity - Continued

- Treatment diffusion
  - Groups interaction thereby weakening their specific treatment
- Experimenter effects
  - Conscious or unconscious actions of researchers or teachers that affect treatment
- Reactive Effects
  - Being in the study negatively affects the participant. Good examples are the Hawthorne effect and the John Henry effect

Some notes on nomenclature…

- We will use "shorthand" to describe the experimental designs
- X will always be a level of the independent variable
- O will always be an point where data is collected about the dependent variable
- R will indicate that random assignment has taken place
- Numbers will be used to identify treatments to different groups

Pre-experimental designs

One-shot case study

X – O

One-group pre-test / post-test

O – X – O

Static group comparison

X1 – O

X2 – O

At least two groups

One group is randomly chosen to receive different treatment

The One-Shot Case Study

Hypothesis:

Children receiving tutoring will do better in math.

Annotation: X - O

1. One independent variable (type of instruction) with one level
2. This intervention is new to the subjects involved
3. One dependent variable, achievement
4. You do not know level of ability prior to study
5. You do not know if changes occurred. If they did, you would not be able to attribute them to tutoring.

The One-Group Pre-Test / Post-Test Design

Hypothesis:

Children receiving tutoring will do better in math

Annotation: O - X – O

1. One independent variable (instruction) with one level
2. This intervention is new to the subjects involved
3. One dependent variable, achievement
4. You know level of ability prior to study
5. You cannot reliably ascertain cause of any change

Static Group Comparison

Hypothesis:

Children receiving tutoring will do better in math than students not receiving tutoring.

Annotation: X1 – O

X2 – O

1. One independent variable (instruction) with two levels, X1 and X2. You can randomly assignment treatment.
2. One dependent variable, achievement
3. You do not know level of ability prior to study but you can compare the two groups.
4. If there is a difference, it might have been pre-existing. Because of that, you don't know if changes are due to intervention.

Quasi-experimental designs

Non-Equivalent Control Group Design

O – X1 – O

O – X2 – O

Example: existing classrooms

Time-Series Design

OOOOO – X – OOOOOO

Multiple measures before and after intervention

Counter-Balanced Design

X1 – O – X2 – O – X3 – O

X2 – O – X3 – O – X1 – O

X3 – O – X1 – O – X2 - O

All groups receive all treatments but in different order.

Non-Equivalent Control Group Design

Hypothesis:

Children receiving tutoring will do better in math than students not receiving tutoring.

Annotation: O – X1 – O

O – X2 – O

1. One independent variable, class, with two levels, tutoring and other. The treatment is randomly assigned to one of the groups.
2. One dependent variable, achievement.
3. The groups are pre-existing, therefore they might represent different populations.
4. Because of that, any post-test differences might be due to pre-test differences.
5. Most widely used in social-science research.

Time-Series Design

Hypothesis:

Student motivation will be higher after the use of graphical report cards.

Annotation: OOOOO – X – OOOOOO

1. Multiple measures before and after intervention
2. Elaboration of one-group pre-test / post-test design.
3. In this case, multiple measures would be taken prior to the intervention to ensure that motivation was not naturally increasing. Multiple post-test measures would be taken to ensure that any changes after the intervention are lasting.

Counter-Balanced Design

Hypothesis:

There will be a significant difference in levels of achievement between students receiving lecture, students receiving CAI and students using a combination of both.

Annotation: X1 – O – X2 – O – X3 – O

X2 – O – X3 – O – X1 – O

X3 – O – X1 – O – X2 – O

1. All groups receive all treatments but in different order.
2. Used with existing groups when pre-testing is not available.
3. Controls for more threats to validity than the static-group comparison.
4. Number of groups must be equal to number of treatments.

Experimental designs

Pre-Test / Post-Test Control Group Design

R – O – X1 – O

R – O – X2 – O

Post-Test Only Control Group Design

R – X1 – O

R – X2 – O

Solomon Four-Group Design

R X1 – O

R X2 – O

R – O – X1 – O

R – O – X2 – O

Pre-Test / Post-Test Control Group Design

Hypothesis:

Students receiving tutoring will have higher achievement in math than students that do not receive tutoring.

Annotation: R – O – X1 – O

R – O – X2 – O

1. Subjects are randomly assigned to groups.
2. Treatment is randomly assigned to group.
3. Controls for all threats to validity.
4. Some argue that pre-test leads to pre-test / post-test sensitivity.

Post-Test Only Control Group Design

Hypothesis:

Students receiving tutoring will have higher achievement in math than students that do not receive tutoring.

Annotation: R – X1 – O

R – X2 – O

1. Subjects are randomly assigned to groups.
2. Treatment is randomly assigned to group.
3. Controls for all threats to validity.
4. Some argue that need for pre-test leads to validity issues.

Solomon Four Group Design

Hypothesis:

Students receiving tutoring will have higher achievement in math than students that do not receive tutoring.

Annotation:

R – X1 – O

R – X2 – O

R – X1 – O

R – X2 – O

1. Subjects are randomly assigned to groups.
2. Treatment is randomly assigned to group.
3. Half of each group is pre-tested.
4. Addresses concerns of pre-test / post-test and post-test only designs.

Experimental Designs and Internal Validity

	History	Mature	Testing	Instruments	Regression	Selection	Mortality	Selection Interaction
One-Shot Case Study	No	No	N/A	N/A	N/A	N/A	No	N/A
One-Group Pre/Post	No	No	No	No	No	N/A	Yes	N/A
Static Group	Yes	No	N/A	N/A	N/A	No	No	No
Pre/Post Control Grp	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Post Only Control Grp	Yes	Yes	N/A	N/A	N/A	Yes	No	Yes
Solomon Four Group	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-equiv Control Grp	Yes	Yes	Yes	Yes	No	Yes	Yes	No
Time Series	No	Yes	Yes	No	Yes	N/A	Yes	N/A
Counter-balanced	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

Experimental Designs and External Validity

	Pre-Test / Post-Test Interaction	Multiple Treatment Interaction
One-Shot Case Study	N/A	N/A
One-Group Pre-Post	No	N/A
Static Group	N/A	N/A
Pre/Post Control Group	No	N/A
Post-Test Only Control Group	N/A	N/A
Solomon Four-Group	Yes	N/A
Non-equivalent Control Group	No	N/A
Time Series	No	N/A
Counter-Balanced	No	No

Procedures

- Step by step plan for what you are going to do
- It should be so precise someone could pick up your work and continue or replicate it
- Example of Procedures Section

Step 3: Analyze the collected data.

- Gay and Airasian Task 7 – First part
- Terrell – entire book

Five Steps to Statistics

- State the hypothesis that you are investigating
- Identify the independent variable and it's levels (there may be more than one independent variable)
- Identify and statistically "describe" the dependent variables (there may be more than one dependent variable)
- Select and run appropriate statistical test
- Interpret the hypothesis in light of the results

What is the Independent Variable?

Children using CAI will have higher achievement than children receiving lectures.

In this hypothesis, the independent variable is the type of instruction received. It has two levels (or factors): CAI instruction and lecture instruction. An independent variable can, theoretically, have any number of levels although more than four or five levels is unusual. This is called a manipulated independent variable in that we are actively assigning members to each group.

What is the Independent Variable?

High school seniors will have a significantly higher number of absences than children in the other grades

In this hypothesis, the independent variable is the high school year in which the student falls. There are four levels: freshman, sophomore, junior and senior. This is a latent independent variable in that we are not putting children into a specific class, we are measuring them in the class into which they naturally fall.

What is the Dependent Variable?

Children receiving graphical feedback will have significantly higher levels of motivation than children receiving traditional feedback.

The dependent variable is the construct titled “motivation”. It would be measured using a specialized instrument designed specifically for the construct.

What is the Dependent Variable?

High school seniors will have a significantly higher number of absences than children in the other grades.

In this hypothesis, the dependent variable is “number of absences”. It would be measured by counting the number of times a given student was absent during the period under consideration.

Before we can describe our dependent variable, we must understand data types

- Discrete data
  - Nominal (categorical) level data
  - Ordinal (rank) level data
- Continuous (quantitative) data
  - Interval level data
  - Ratio level data

Discrete Data

- Nominal (categorical) – frequencies of occurrence are counted.
  - Gender
  - Class or group
- Ordinal (rank) – Numbers represent rank order of observation.
  - Class standing

Continuous (Quantitative) Data

- Interval data – numbers are assigned on a scale where intervals are constant throughout the scale. There is no absolute zero.
  - Intelligence
  - Achievement
  - Aptitude

- Ratio data – numbers are assigned so that statements of ratio can be made. Numbers represent units from absolute zero.
  - Weight
  - Elapsed Time
  - Distance

Descriptive Statistics

- Measures of Central Tendency
  - Mean, Median, Mode
- Measures of Variation
  - Range, Standard Deviation, Variance
- Measures of Relative Standing
  - z scores, t-scores, percentiles
- Graphical presentations
  - Histograms, pie charts, box-and-whisker plots

Descriptive Statistics

- 1.00
- 1.00
- 2.00
- 2.00
- 2.00
- 3.00
- 4.00
- 5.00
- 7.00
- 8.00
- 9.00

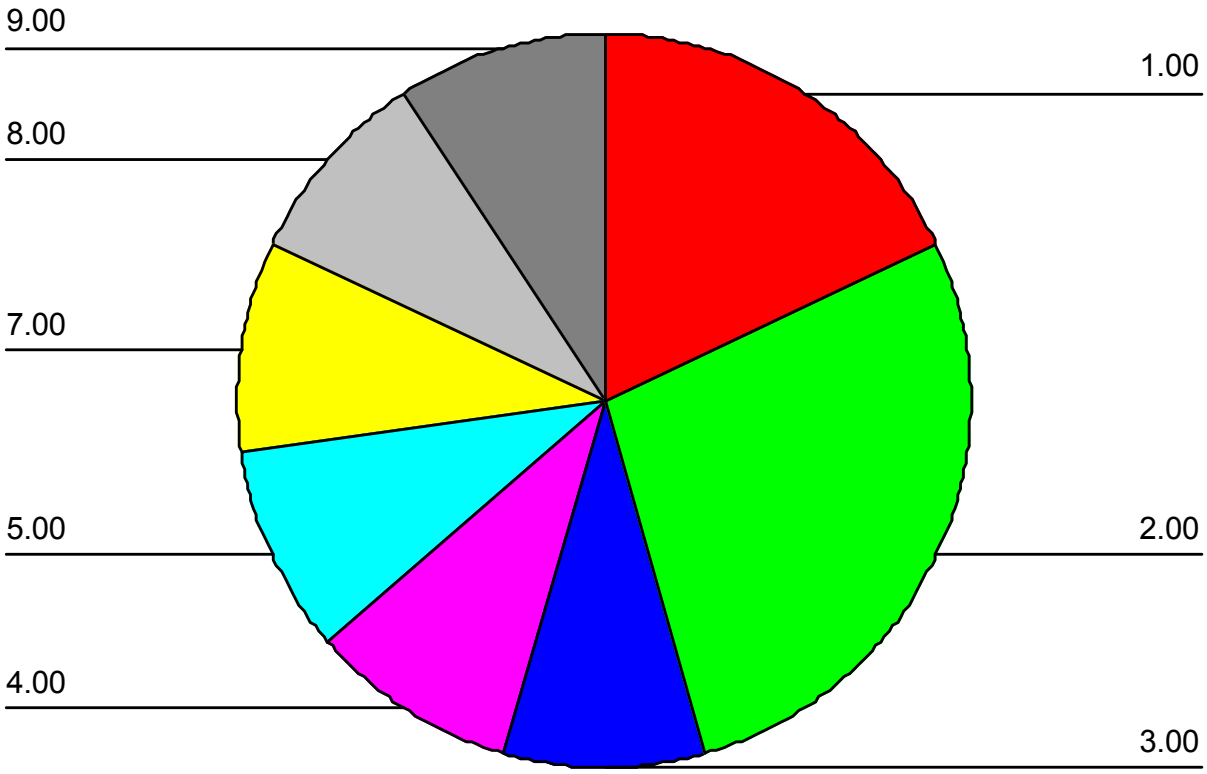
Descriptives

			Statistic	Std. Error
EXAMPLE	Mean		4.0000	.86340
	95% Confidence Interval for Mean	Lower Bound	2.0762	
		Upper Bound	5.9238	
	5% Trimmed Mean		3.8889	
	Median		3.0000	
	Variance		8.200	
	Std. Deviation		2.86356	
	Minimum		1.00	
	Maximum		9.00	
	Range		8.00	
	Interquartile Range		5.0000	
	Skewness		.718	.661
	Kurtosis		-.966	1.279

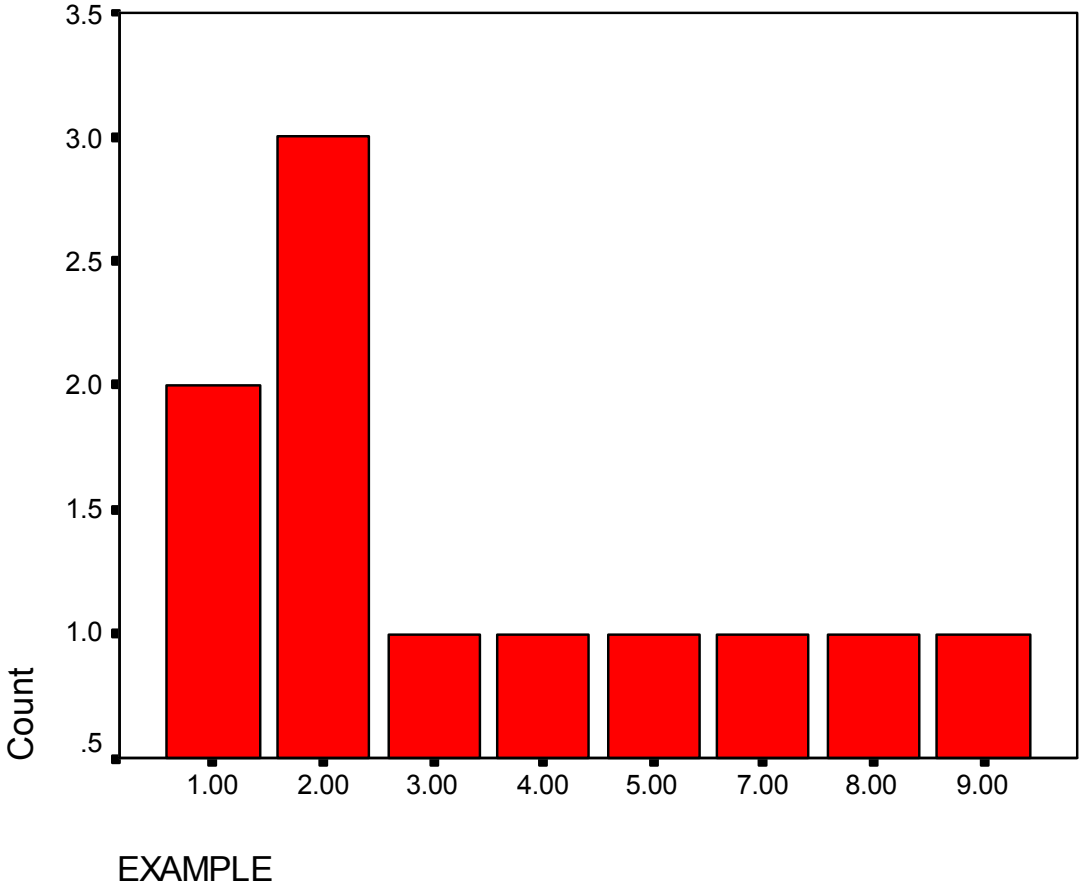
z Scores Show Relative Distance From the Mean in Terms of the Standard Deviation

Example	z-score
1.00	-1.04765
1.00	-1.04765
2.00	-0.69843
2.00	-0.69843
2.00	-0.69843
3.00	-0.34922
4.00	0.00000
5.00	0.34922
7.00	0.69843
8.00	1.39636
9.00	1.74608

Pie Chart

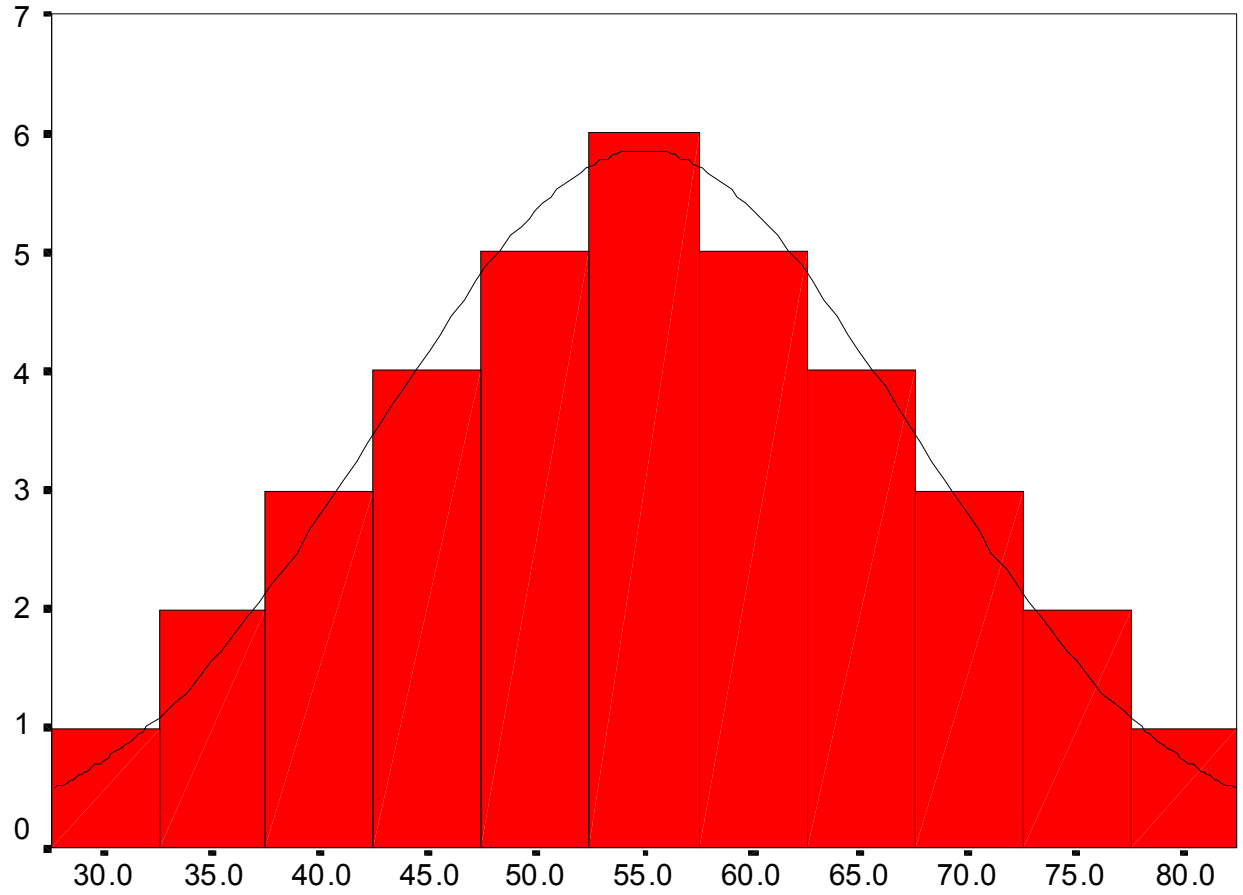


Histogram



What is a Normal Distribution?

- Only applies to quantitative data
- Data that are in a "mound shaped" distribution
- Meet the guidelines of the Empirical Rule
- The shape doesn't have to be perfect
- Statistical procedures allow for skewed or kurtotic distributions
- Data can be mathematically manipulated in order to help normality of distribution



Mean = 55 Standard Deviation = 10 Variance = 100

Empirical Rule says that about 68% of values lie with +/- 1 standard deviation from the mean, 95% within +/- 2 standard deviations from the mean and nearly all (99.7%) within +/- 3 standard deviations from the mean.

Inferential (Decision Making) Statistics

- Discrete data use non-parametric statistics
- Continuous data generally use parametric statistics unless data are not normally distributed
- Continuous data may use non-parametric statistics if distribution is not normal

Common Parametric Statistics

- Independent sample *t*-test
  - 1 independent variable with 2 levels that are independent of one another
  - 1 dependent variable with continuous data
- Dependent sample *t*-test
  - 1 independent variable with 2 levels that are related to one another
  - 1 dependent variable with continuous data

Common Parametric Statistics

- One-way analysis of variance (ANOVA)
  - 1 independent variable with 3 or more levels
  - 1 dependent variable with continuous data
- N-way analysis of variance (factorial ANOVA)
  - More than 1 independent variable with 2 or more levels
  - 1 dependent variable with continuous data
- Multivariate analysis of variance (MANOVA)
  - 1 or more independent variables with 2 or more levels
  - More than 1 dependent variable with continuous data

Common Non-Parametric Statistics

- Chi-square
  - Used with nominal data to determine if cell sizes are statistically equivalent
- Mann-Whitney U Test
  - Equivalent to *t*-test when data are not normally distributed or data are ordinal in nature
- Kruskal-Wallis H Test
  - Equivalent to ANOVA when data are not normally distributed or data are ordinal in nature

Definitions

- *alpha* value – the percentage of risk we're willing to take when statistically evaluating the null hypothesis. This is generally set to .05 in social science research.

- *p* value – value computed from statistical procedure indicating probability that groups being evaluated came from same population (e.g., probability that groups being evaluated are statistically equivalent)

The Heart of Statistical Decision Making

- Significantly different – when data being compared is different to the degree that it cannot be attributed to chance

  - If $p < alpha$, reject the null hypothesis
  - If $p = alpha$, fail to reject the null hypothesis
  - If $p > alpha$, fail to reject the null hypothesis

Possible Errors

- Type I error – rejecting the null hypothesis when the groups being evaluated are statistically equivalent. Also referred to as the *alpha* value.

- Type II error – failing to reject the null hypothesis when the groups being evaluated are actually significantly different. Sometimes referred to as the *beta* value.

- Both of these errors are possible given the sampling process

An Example of Hypothesis Testing

- Hypothesis: Children receiving daily report cards will have higher averages than children receiving monthly report cards
- Descriptive statistics
  - Mean of daily report cards = 88
  - Mean of monthly report cards = 86
- They are different, but are they significantly different?

Example continued

- Establish *alpha* level of .05
- Use independent sample *t*-test to analyze data from both groups
- Output of *t*-test indicates *p* value of .04
- Determine if null should be rejected:
  - If $p > alpha$ then fail to reject null hypothesis
  - If $p = alpha$ then fail to reject null hypothesis
  - If a < *alpha* then reject null hypothesis

What would you do?

- Hypothesis: Children using a combination of lecture and CAI will do better than children receiving either lecture alone or CAI alone
- What is the IV? What are the levels?
- What is the DV? What type of data?
- Which of the statistical tools would you use?
- If *alpha* is .05 and *p* is .06, what would you do?

What would you do?

- Hypothesis: Weekly sessions with the guidance counselor will lead to higher levels of student intrinsic motivation
- What is the IV? What are the levels?
- What is the DV? What type of data?
- Which of the statistical tools would you use?
- If *alpha* is .05 and *p* is .03, what would you do?

What would you do?

- Hypothesis: High school students that get eight or more hours of sleep will have significantly different levels of achievement than high school students receiving less than eight hours of sleep
- What is the IV? What are the levels?
- What is the DV? What type of data?
- Which of the statistical tools would you use?
- If the achievement mean of the first group is 89 and the achievement mean of the second group is 91 and *alpha* is .05 with *p* equal to .08, what would you do?

Step 4: State the results or implications based on analysis of the data.

- Gay and Airasian Task 7 – second part
- Summarize the study
- State results in terms of hypotheses - discuss results
- Conclusions
- Implications
- Recommendations

Three Stages of the Dissertation

- You try to solve all of the world's problems
- You try to solve some of the world's problems
- You solve one problem and graduate

Aldenderfer, M.S., & Blashfield, R.K. (1984). Cluster Analysis. Newbury Park, CA: Sage Publications.

Arminger, G., Clogg, C.C., & Sobel, M.E. (1995). Handbook of Statistical Modeling for the Social and Behavioral Sciences. New York: Plenum Press.

Box, G.E.P., Hunter, W.G., & Hunter, J.S. (1978). Statistics for Experimenters. New York: John Wiley and Sons.

Bray, J.H., & Maxwell, S.E. (1985). Multivariate Analysis of Variance. Newbury Park, CA: Sage Publications.

Byrkit, D.R. (1987). Statistics Today - A Comprehensive Introduction. Menlo Park, CA: Benjamin/Cummins.

Campbell, D.T., & Stanley, J.C. (1966). Experimental and Quasi-Experimental Designs for Research. Boston, MA: Houghton Mifflin.

Conover, W.J. (1980). Practical Nonparametric Statistics. New York: John Wiley and Sons.

Converse, J.M., & Presser, S. (1986). Survey Questions - Handcrafting the Standardized Questionnaire. Newbury Park, CA: Sage Publications.

Creswell, J.W. (1995). Research Design - Qualitative & Quantitative Approaches. Thousand Oaks: Sage Publications.

DeFina, A.A. (1992). Portfolio Assessment - Getting Started. New York: Scholastic.

DeVellis, R.F. (1991). Scale Development: Theory and Applications. Newbury Park: Sage Publications.

Dunteman, G.H. (1989). Principal Components Analysis. Newbury Park, CA: Sage Publications.

Edwards, A.J. (1984). An Introduction to Linear Regression and Correlation. New York: W.H. Freeman.

Fitz-Gibbon, C.T., & Morris, L.L. (1987). How to Design a Program Evaluation. Newbury Park: Sage Publications.

Fowler, F.J. (1988). Survey Research Methods. Newbury Park: Sage Publications.

Freund, J.E., & Williams, F.J. (1966). Dictionary / Outline of Basic Statistics. New York: McGraw-Hill.

Gay, L.R. (1985). Educational Evaluation and Measurement - Competencies for Analysis and Application. New York:Merrill.

Gay, L.R. and Airasian, P. (2000). Educational Research: Competencies for Analysis and Application.

New York: Merrill.

Gordon, S.P., & Gordon, F.S. (1994). Contemporary Statistics - A Computer Approach. New York: McGraw-Hill.

Gronlund, N.E. (1993). How to Make Achievement Tests and Assessments. Boston, MA: Allyn and Bacon.

Gunst, R.F., & Mason, R.L. (1991). How to Construct Fractional Factorial Experiments. Milwaukee: ASQC.

Healy, M.J.R. (1986). Matrices for Statistics. Oxford: Clarendon Press.

Hopkins, C.D., & Antes, R.L. (1990). Educational Research - A Structure for Inquiry. Itasca, IL: Peacock.

Huck, S. & Cormier, W. (1996). Reading Statistics and Research. New York: Harper Collins.

Issac, S. & Michael, W. (1997). Handbook in Research and Evaluation. San Diego: Edits.

Kachigan, S.K. (1991). Multivariate Statistical Analysis - A Conceptual Introduction. New York: Radius Press.

Kalton, G. (1983). Introduction to Survey Sampling. Newbury Park, CA: Sage Publications.

Kerlinger, F. (1986). Foundations of Behavioral Research. New York: Holt, Rinehart and Winston.

Kim, J., & Mueller, C.W. (1978a). Introduction to Factor Analysis - What It Is and How To Do It. Newbury Park, CA:Sage Publications.

Kim, J., & Mueller, C.W. (1978b). Factor Analysis - Statistical Methods and Practical Issues. Newbury Park, CA: Sage Publications.

Klecka, W.R. (1980). Discriminant Analysis. Newbury Park, CA: Sage Publications.

Kume, H. (1985). Statistical Methods for Quality Improvement. Tokyo: AOTS.

Leedy, P.D. (1993). Practical Research - Planning and Design. New York: Macmillan.

Lehman, R. (1991). Statistics and Research Design in the Behavioral Sciences. Pacific Grove, CA: Brooks/Cole Publishing.

Lewis-Beck, M.S. (1980). Applied Regression - An Introduction. Newbury Park, CA: Sage Publications.

Long, J.S. (1983a). Covariance Structure Models - An Introduction to LISREL. Newbury Park, CA: Sage Publications.

Long, J.S. (1983b). Confirmatory Factor Analysis. Newbury Park, CA: Sage Publications.

Long, T.J., Convey, J.J., & Chwalek, A.R. (1985). Completing Dissertations in the Behavioral Sciences and Education.San Francisco: Jossey-Bass.

Mandel, J. (1964). The Statistical Analysis of Experimental Data. New York: Dover.

Merriam, S. (1998). Qualitative Research and Case Study Applications in Education. Jossey-Bass.

Neter, J., Wasserman, W., & Kunter, M.H. (1990). Applied Linear Statistical Model. Homewood, IL: Richard D. Irwin.

Newmark, C.S. (1985). Major Psychological Assessment Instruments. Boston: Allyn and Bacon.

Nunnally, J.C., & Bernstein, I.H. (1994). Psychometric Theory. New York: McGraw-Hill.

Oppenheim, A.N. (1992). Questionnaire Design, Interviewing and Attitude Measurement. London: Pinter Publishers.

Payne, D.A. (1992). Measuring and Evaluating Educational Outcomes. New York: Macmillan Publishing.

Phillips, J.L. (1992). How to Think About Statistics. New York: W.H. Freeman and Company.

Shapiro, S. (1990). How to Test Normality and Other Distributional Assumptions. Milwaukee: ASQC.

Sincich, T. (1993). Statistics By Example. New York: MacMillan.

Stevens, J. (1992). Applied Multivariate Statistics for the Social Sciences. Hillsdale, NJ: Lawrence Erlbaum.

Stevens, J.W. (1990). Intermediate Statistics - A Modern Approach. Hillsdale, NJ: Lawrence Erlbaum Associates.

Strauss, A.L. (1994). Qualitative Analysis for Social Scientists. Cambridge: Cambridge University Press.

Wildt, A.R., & Ahtola, O.T. (1978). Analysis of Covariance. Newbury Park, CA: Sage Publications.

Wolcott, H.F. (1990). Writing Up Qualitative Research. Newbury Park, CA: Sage Publications.

Wolf, F.M. (1986). Meta-Analysis - Quantitative Methods for Research Synthesis. Newbury Park, CA: Sage Publications.

Yin, R.K. (1994). Case Study Research - Design and Methods. Newbury Park: Sage Publications.