

**ADVANCED  
MASS MEDIA RESEARCH**

**Prof. William Tillinghast  
San Jose State University  
AJEEP**

# **AJEEP**

## **Advanced Mass Media Research**

### **Course Description**

This course focuses on quantitative research and its ability to utilize probability theory to test for statistical significance of the differences obtained in research. These techniques, which are applicable to all of the social sciences, include three specific tests relevant to studying mass media: chi square analyses, t-tests, and correlation. They apply differently to different types of research. This course also focuses on preparing study results for publication.

### **Course Goals and Student Learning Objectives**

The course goal is that students demonstrate basic research statistical tests in order to assess the value of research results.

Student learning objectives focus on how students can:

1. Demonstrate how qualitative and quantitative research require different statistics
2. Demonstrate how to conduct the various tests: chi square, t-test, and correlation to determine statistical significance.
3. Prepare a version of a properly structured research report

### **Required Texts/Readings**

No text is required.

Students may receive handouts and may search the Internet

### **Assignments and Grading Policy**

Assignments will mainly focus on gathering of information from the Internet to support or rebut what the instructor says. There will two 20-point quizzes, one at the beginning of the third class and the other at the start of the fifth class. One out-of-class assignment will be worth 20 points. It will be a short paper topic agreed upon by both student and instructor. It will be assigned at conclusion of the third lecture. A final exam will be worth 40 points. Since the total number of points equal 100, a minimum grade of 60 points, or 60 percent is required to pass the course with a grade of D. The standard university policy dictates that 70 be required for a C; 80 for a B; and 90 for an A.

## Introduction to Media Research

<b>Class</b>	<b>Date</b>	<b>Topics, Readings, Assignments, Deadlines</b>
<b>1</b>		Topic: Quantitative research methods Assignment: Find media examples using content of Lecture #2
<b>2</b>		Topic: Content Analyses and Experiments Assignment: Find media examples utilizing content of Lecture #3
<b>3</b>		Topic: Surveys Assignment: Find newspaper/television coverage of surveys Quiz #1 at start of class
<b>4</b>		Topic: Chi Square Assignment: Find media examples that include examples of Lecture #4 content. Quiz #1 at start of class. (next page) Paper topic assigned
<b>5</b>		Topic: Experiments and t-tests. Assignment: Find media examples of content of Lecture #5 Quiz #2 at start of class (see next pages)
<b>6</b>		Topic: Data input, data analysis, and report writing Assignment: Find media examples of Lecture #6 content Topic: Qualitative v. Quantitative Methods.

Advanced Media Research – 1  
Quantitative Research Methods

STUDENTS: The three main things you should learn from this course are:

1. Why a sample is better than a census and why a sample can be the same size for a small town and a large city and yet have the same accuracy
2. When do you use a probability sample or a non-probability sample and what are the advantages and disadvantages
3. How two groups can have the same average yet be very different from each other

A census counts every person in the nation, the city, or whatever group you are interested in. A sample is a small part of the total group.

Believe it or not, many times the sample is more accurate than if had taken a census.

For example, suppose we wanted to know how Egyptians felt about the outcome of April Spring and the downfall of Hosni Mubarak and his regime. But instead of a civilian government, the military stayed in power.

The population of Egypt is about 80 million. Even Cairo is about seven million. How long would it take to reach every single Egyptian to get what they thought? And by the time you did get to the last person, events would have changed and people you talked to in the beginning would have changed their minds.

But a sample can be obtained almost overnight.

How can a sample reflect the entire group?

If you went home tonight for dinner, and your spouse or mother asked you to taste the soup to see if it needed more salt, how many teaspoon full would you have to taste? The answer is only one because the soup is the same throughout. The scientific word is homogenous. If there are many differences, it is called heterogeneity. And that is why the sample has to be drawn properly.

This is where the first question arises. Do you want the sample to reflect the overall population it came from, or are you more interested in study something in depth?

If you want to reflect the overall population you need a sample that is representative of the various groups in that population. You need to be able to generalize from the sample to the parent population. The basis of this selection is called E-P-S-E-M, the equal probability of selection method.

Every person in the overall group has the same chance of being selected.

If you don't have any need to generalize, you can choose a non-generalizable method.

There are essentially four non-generalizable methods: availability, volunteer, purposive, snowball, and quota. There are also four generalizable methods: simple random, systematic random, stratified random. The latter are known as probability selection methods, and the first group is called non-probability methods.

A basic description of the types is as follows:

**Simple Random** requires a population list and a blind selection process. An example would be drawing names out of a hat.

**Systematic Random** uses a random start number by every kth number. For example, if we wanted to draw a sample of 300 from a company with 3,000 employees. That means we want one out of ten workers. Our start number has to be a number under 10 chosen at random. If that number is six, than our sample from the company list are employee number six, number 16, number 26, number 36, and number 46 and so on.

**Stratified Random** is used when you want to ensure that certain groups have proportional representation. Suppose in the above group of 3,000 employees, only 500 were female. That is 16.6 percent. Since our sample size will be 300, we want to ensure that 16.6 percent of the 300 are women. That number is 300 times .166 equals 50. The computer must be instructed to draw 50 women and 450 men.

Multi-Cluster Random draws sample in states, such as selecting so many cities in Afghanistan to be included, and then so many different sections in the selected cities, and then specific house addresses in those areas. It is used when the first population is very large.

The non-probability samples do not require that way of choosing subjects when there is no intention of generalizing to the parent population. Those selection methods are:

**Availability** sampling is just asking whoever is nearby, your friends, neighbors, fellow workers. It is usually done more to test the questionnaire for clearness than for any other reason.

**Volunteer** samples are self-selective, that is the researchers does not choose the subjects, the subjects choose to participate. Television news in the United States often did this. The anchor would start the newscast stating that that night's poll was on gun control. If you favored gun control, call a certain phone number, and if you were against gun control, call another number. At the end of the newscast, the anchor would report the results, such as 3,456 in favor of gun control and 2,033 against gun control. But it is not news. It is not enough accurate. People interested call in. Others do not.

**Purposive** sampling includes only individuals who have something in common. Suppose we wanted to know how Afghan citizens living in the Bay Area felt about the U.S. military in Afghanistan, we would focus on just finding and interviewing those citizens.

**Quota** sampling is sort of like Stratified Random sampling. You need the get a sample that reflects your population categories, but it is left to the interviewer to find them.

**Snowball** sampling uses the first subject to give the researcher names of other people with the same research interest, say marijuana smokers. When you interview those people, they give you the names of more marijuana smokers.

Sometimes you just need to find out if something works. If you film a story differently do people like it? You don't need a large group and you are not trying to generalize to some larger population.

Another way of looking at the difference between qualitative and quantitative research is that qualitative research can tell you why, and quantitative research can tell you how much. You want to know the demographics of your viewing audience. You need a sample to generalize to the overall viewing population.

The next question you need to ask yourself is how big a sample you need. Normally qualitative sampling only needs a few people or a few situations. Quantitative research usually uses probability sample, and usually at least 200 subjects or more.

Other than time and expense, sample accuracy may be a consideration. You need to know how accurate is the sample, or in other words, does it really represent the parent population?

There is an easy formula to determine the accuracy of any sample. Divide the number "one" by the square root of the sample size. The result is the percentage range of accuracy for your findings. If you have 200 subjects, then you need obtain the square root, which are 14.14. One divided by 14.14 is 0.07. That is seven percent.

It is used this way. Suppose you asked 200 randomly drawn members of your TV viewing audience if they approved of the way your new female anchor dressed. The answer was 54 percent said yes and 46 percent said no. Does that mean that most of your audience likes the new dress? No, it does not. The correct answer for approval is 54+/- 7 percent or, between 47 and 61 percent.

If you want more accuracy, that is less spread in the range of numbers, you will have to increase the size of your sample. A sample of 400 subjects has a square root of 20. Divide one by 20 and your accuracy range is five percent. A sample size of



1,600 has a square root of 40. Divide one by 40 and your accuracy range is two-and-one-half percent.

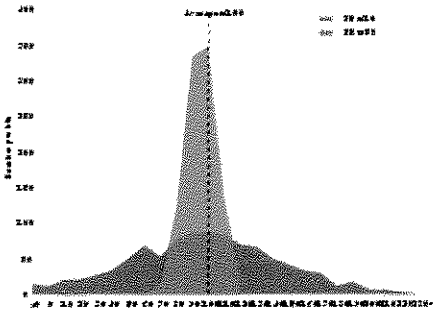
But what if you want to draw samples from different size cities, such as the three largest cities in Afghanistan? Their populations are: Kabul, 3,300,000; Herat, 450,000; and Kandahar, 400,000. Kabul is almost eight times as big as the other two. It would seem that its sample should also be eight times larger.

The answer is no. The accuracy range is the accuracy of the sample, not the parent population. Therefore, the sample sizes can be the same.

The accuracy range is also not completely a sure thing. Mathematically, we are 95 percent certain, that if we continued to use different samples of that population we would get the same results 95 of 100 times.

Often when we look at quantitative data we need to know how close or how far apart the scores are. For example suppose two school teachers gave their students the same exam and the average score, the mean, was 75 in each class. Does that mean the students were equally the same? No.

If you look closely at the picture below, you will first notice that the two groups, the red and the blue, are very different. The red is much higher but not very wide. The blue is very wide but not very tall. But the longer you look at it; you notice the dotted black line up through the blue and the red to the very top.



That dotted black line represents the mean, the average of 75 for each class. By looking at the chart, you can see that most of the students in the red group scored very close together while the students in the blue group scored very low as well as very high.

What you need to know is the standard deviation of the scores. The red class above has a very small standard deviation. The blue class has a very high standard deviation. Standard deviation is an indicator of the amount of variation all of the scores are from the mean.

This “dispersion” of scores totals the individual score differences from the mean, then takes the square root and doubles it. The resulting number is both added and subtracted to a score to indicate either the spread of the numbers, or to indicate that you are 95 percent confident that the real average score in the population will fall in that sample range of scores.

## Advanced Media Research – 2 Content Analysis and Correlation

STUDENTS: The main things you should learn from this lecture are:

1. In what respects are Content Analyses superior to other research forms
2. Correlation is relational but not necessarily causal.
3. Why correlations can be both positive and negative
4. Why visuals of the data are a must in correlational research

Researchers like content analysis for several reasons. Unlike people, your subject matter can't cancel appointments at the last minute. Your subject sits on the shelf until it is convenient for you. Your subject is anything you can see, smell, touch or sense in any way that you can measure. Your subject may be newspaper articles, recorded TV programs, graffiti on walls, music, or many other things.

And the subject cannot lie to you. It does not know it is being observed and measured. So, it gives you accuracy, if not truth. It is WYSIWYG, what you see, is what you get.

Content Analysis can be defined as a method of studying and analyzing communication in a systematic, objective and quantitative manner for the purpose of examining the content of recorded information.

Media Content Analysis is a research technique to examine observable text and broadcasts through quantitative, systematic, repeatable and objective measures.

Basically, it is counting what you observe.

Content analysis is not an impressionistic summary of media content; an examination of one theme of a selected text; or an analysis of foreign language media using software translation. Content is:

1. Print—newspapers, magazines, books, etc.
2. Other writing – the web, ads, posters, etc.
3. Broadcast – radio, television, movies, etc.
4. Other recordings – music, photos, drawings
5. Live situations -- speeches, plays, interviews
6. Observations – gestures, products, rooms

Why do Content Analysis? In order to:

1. Find links between causes (ad content) and effects (consumer behavior)
2. Evaluate and improve programming

Before you conduct a content analysis:

1. Formulate your research questions or hypotheses
2. Define the universe
3. Select a sample from the population
4. Select a unit of analysis
5. Construct the categories to be analyzed
6. Establish a quantification system
7. Train coders and conduct a pilot study
8. Code the content according to definitions
9. Analyze the data
10. Interpret the results

The advantages of content analysis that it . . .

1. Can find central aspects of social interaction
2. Can utilize both quantitative and qualitative
3. Can provide historical, over-time trends
4. Is unobtrusive so avoids subject structuring responses

The disadvantages of content analysis are that it . . .

1. Can be extremely time consuming
2. Often devoid of theoretical base
3. Is simplistic, often just word count
4. Often disregards content that produced text

## 5. Can be difficult to computerize

When you count things in content analysis, the level of measure is usually ratio level, meaning that there is a possible zero for an answer such that what you are counting does not occur sometimes, but occurs once, twice, three times, or more or other occasions.

For example, suppose one of your reporters thinks that you don't cover your provincial officials properly. He says you seem to give them all about the same amount of coverage regardless of their importance. Furthermore, he says you can do a statistical test to find out. You can correlate the number of inches of coverage each of your officials get with the salaries, on the assumption that those who earn more are more important, and probably should receive more news coverage, more inches of coverage.

You can conduct a Pearson correlation which measures to what degree two variables, such as column inches and salary vary together, that is how much does one variable, newspaper column inches go up as salary increases.

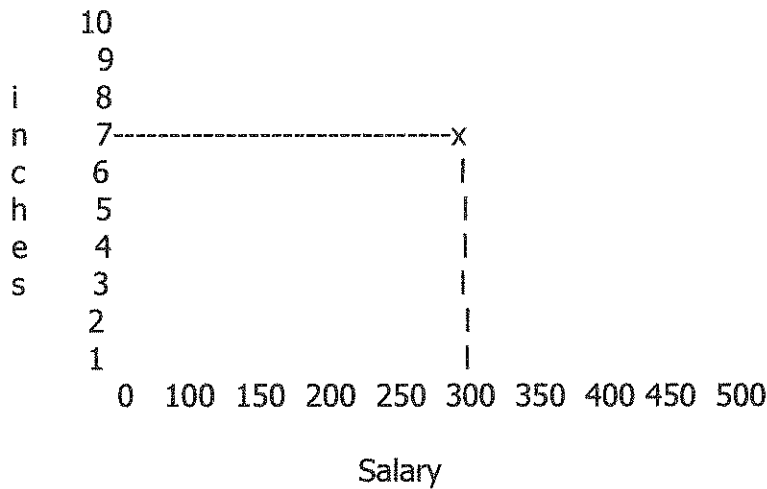
Think about this for a bit. We will come back to it later.

When summer weather gets too hot, people wear as few clothes as possible. And, when the temperature gets near the freezing point, people bundle up with several layers of clothes. The relationship is correlational, that as one variable changes, another variable also changes.

If both change in the same direction, it is called a positive. But if one goes up as another goes down, it is a negative or what can be called an inverse correlation. The

two clothing situations are both negative or inverse, one variable increasing and the other decreasing.

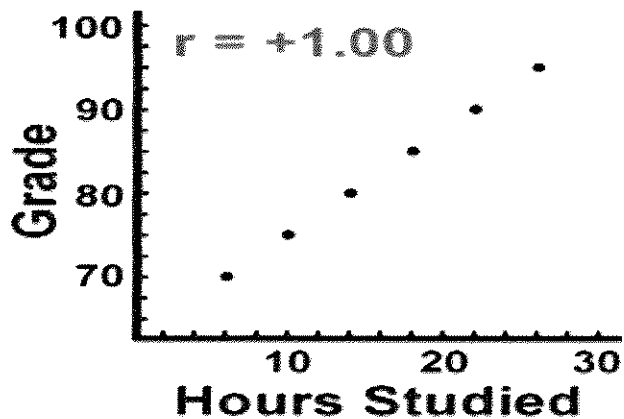
Correlation is based on how able a computer, or any mathematical calculation, can two scores on one thing, such as column inches and salary and put them on a straight line. If we construct a graph where the horizon line is salary and the vertical line is inches, it would like a bit like this, but larger:



Each person or thing measured in correlation has two scores, salary and inches. A dot is placed as the point where the vertical line of salary crosses the horizontal line of inches. For example, if an individual earned \$300 per week and averages seven inches of newspaper coverage, his position would be at the x on the above table. The computer will draw a straight line between all the x's in the table getting as close to each x as is possible. It then mathematically calculates how close it is to each x. The shorter the distance, the higher the correlation, meaning that if salary increases, so does inches of coverage.

All correlations range between +1.00 through zero to -1.00. Any other number is the result of a mathematical error. If a correlation was perfect, it would look like this.

**Perfect positive** (all points fall on a straight line)

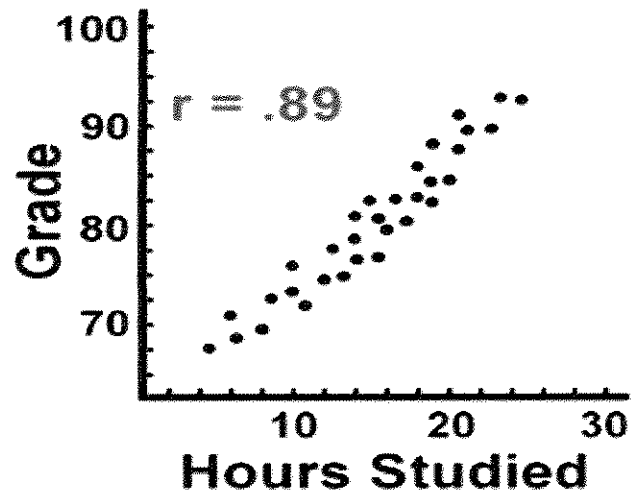


Notice that the dots are all in a straight line going up to the left. That indicates a positive relationship, as the number of hours studied increases, the exam grade is better.

Each dot represents one person, how many hours studied and grade.

Rarely, if ever, is the relationship between two variables perfect. Suppose the relationship was positive, but other factors influenced the grade, intelligence, amount of sleep before the exam, interest, etc. Maybe the relationship would look like this:

## Pearson Correlation Example

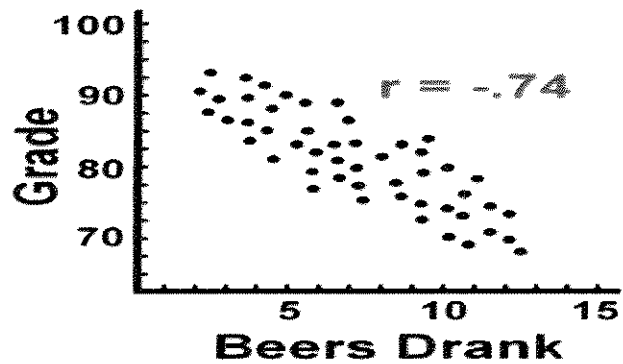


The grade of the student still increases, but some of the scores are above and some below where the straight line runs. But notice the straight line direction is still positive, that the line goes up as it goes from left to right.

But, suppose we look at another student characteristic, drinking alcohol. We know that is a negative characteristic, one that affects the individual's thinking process, so we expect thinking ability to decline. The result is this:



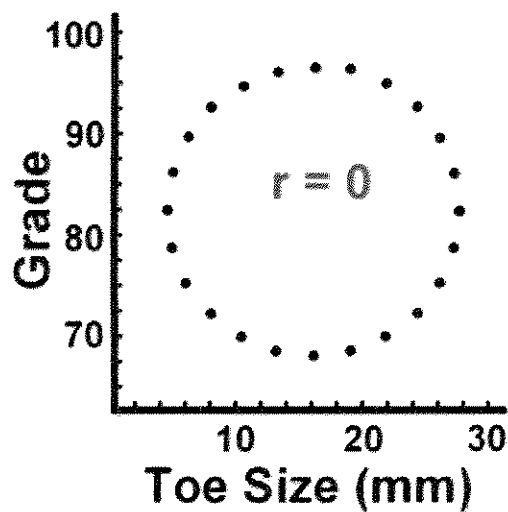
### Pearson Negative Correlation



The line goes down, from left to right, indicating a negative relationship, the more beers the student drank, the lower the student's grade. The minus sign in front of .74 says that the relationship is negative.

A relationship of 0.000, indicating absolutely no relationship, neither negative or positive would look like this:

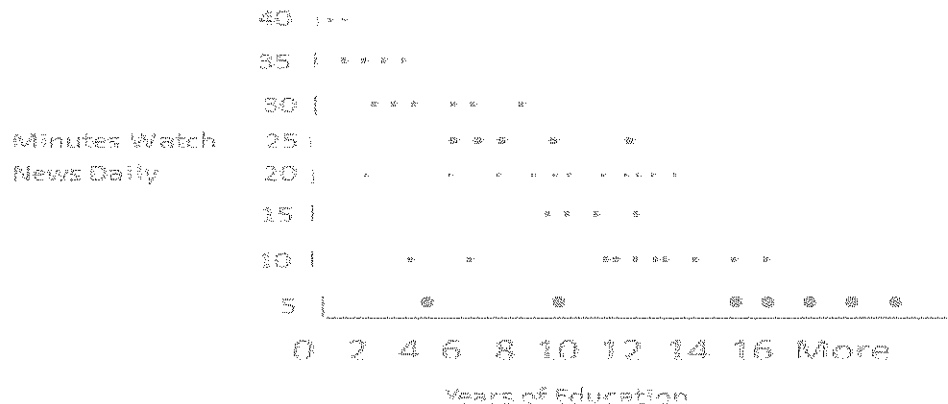
### No Correlation



Both visually and by statistical result, this chart states that there is no relationship between the size of the student's toe and the student's grade.

Before examining the data analysis of a correlation, let us consider what some journalistic correlations might look like. Notice that we have trouble trying to correlate political party identification with gender because the numbers do not represent any real reality. Suppose we want to know whether we should focus our attention on news to more educated, or less, educated individuals, and whether it should be the same for our newspaper as for our television station. We have two variables for which we need ratio level data, which is physical world data, education and viewing time.

Since it is obvious that the educational level precedes the amount of viewing, education will be on the horizontal axis, the straight line across our table and time with the media will be the vertical axis, the one rising up, up and up. It might look something like this on a scatterplot:



That table tells you that people with more education watch less television than people with less education. Although it is not exact, this table does somewhat describe the news viewing habits of Americans. However, each TV station must determine the viewing patterns in their own country, as well as in the community within their country.

The pattern for news reading should be just the opposite, a positive correlation stating that as educational level increased, so did news reading time. However, as Americans became more educated and more involved in business and industry, they worked more hours, so had less time, and needed more specific information than the newspaper was providing. The result on the chart was that about halfway up the straight started curving back down to zero again.

The interesting part of correlation is it sounds logical because it is based on a number scale ranging from +1.00 through 0 to -1.00. A rating of 1.00 would mean that every time variable A changed, variable B would also change by the same proportion. If the sign is +, it means both variables are increasing or both are decreasing. But if the sign is -, it means that while one variable is increasing, the other is decreasing.

A visual of a +1.00 would show a straight line either increasing from left to right while a -1.00 shows a straight line going down from left to right.

Correlation also has the power to determine how much one variable is related to the other. If the correlation of wearing warmer clothes in winter correlated at .700 it means that as it gets colder, people put on more clothes. The square of .700 is .49 or 49 percent. That means half of why you wear warmer clothes in winter is because of the drop in temperature.

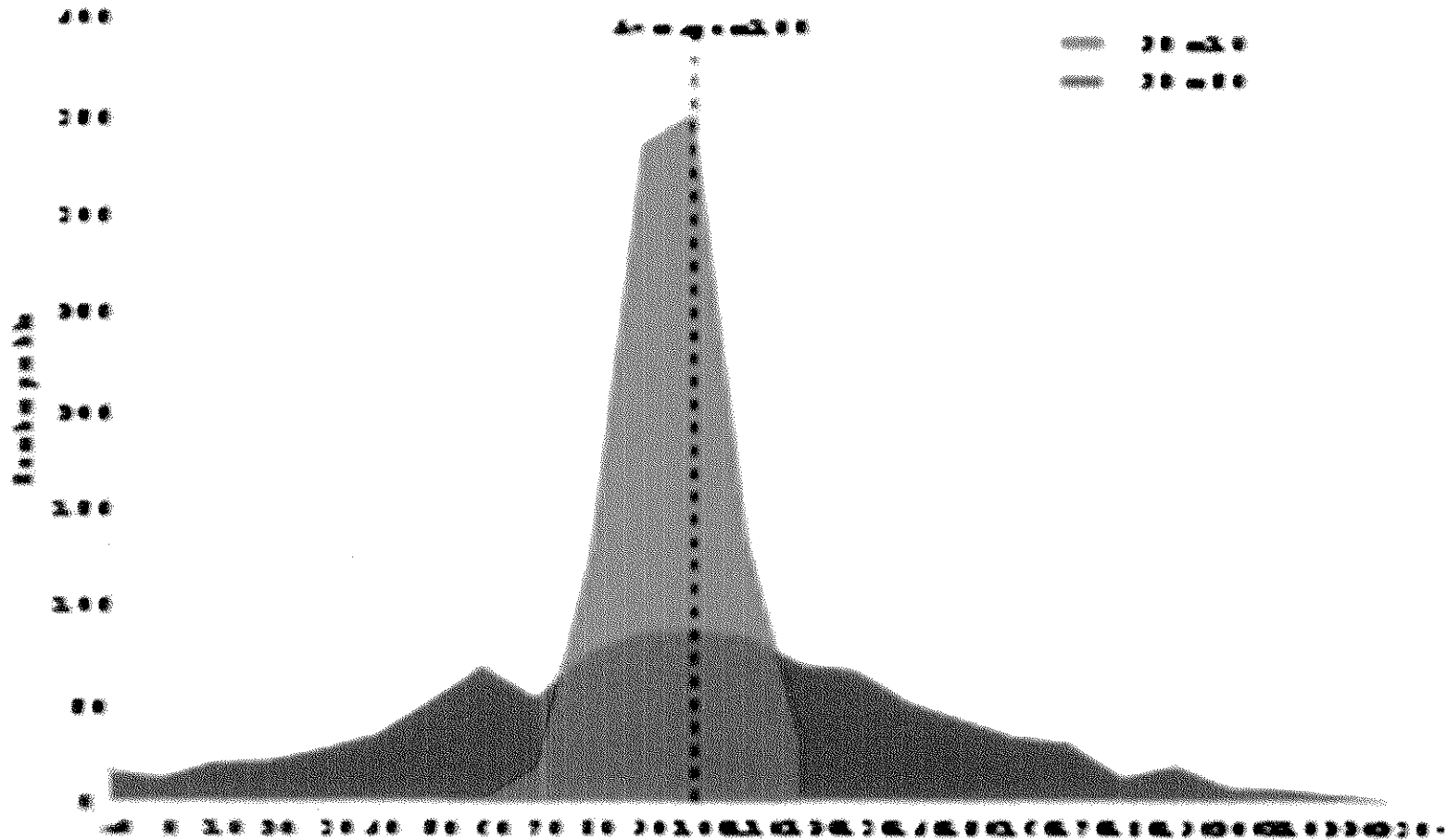
## Questions

1. Discuss the difference between things that you can content analyze and the things that you cannot content analyze. How, and when, can the second group be structured to be like the first group?
2. What are the advantages to station managers and newspaper editors in having their mediums content analyzed? For what categories?
3. How might you compare a content analysis of newspaper coverage with a content analysis of a television station in the same community?
4. What are the advantages and the disadvantages of comparing content analyzes of the same news organizations at two or more different time periods?
5. Why doesn't a high correlation between two variables always indicate a cause-and-effect relationship? How might you find an intervening variable?
6. What are some of the things that need to have a high correlation in order to be disseminated as news by your news medium? Such as distance from your area and number of people involved? Or others news considerations, such as significance, recognition, etc.?
7. How do you think some of the news categories would correlate with audience characteristics, such as education or income?

Words or concepts to know and remember

Manifest content  
Positive correlation  
Inverse correlation  
Scatterplot  
 $r^2$

# Same Mean, Different Standard Deviations



# Content Analysis looks at . . .

Content analysis is an examination of one theme of a selected text; or an analysis of foreign language media using software translation. Content is:

- Print—newspapers, magazines, books, etc.
- Other writing – the web, ads, posters, etc.
- Broadcast – radio, television, movies, etc.
- Other recordings – music, photos, drawings
- Live situations -- speeches, plays, interviews
- Observations – gestures, products, rooms
-

# Why Do Content Analysis?

:

In order to:

- Find links between causes (ad content) and effects (consumer behavior)
- Evaluate and improve programming
-



# Before Conducting Content Analysis . . .

1. Formulate your research questions or hypotheses
2. Define the universe
3. Select a sample from the population
4. Select a unit of analysis
5. Construct the categories to be analyzed
6. Establish a quantification system
7. Train coders and conduct a pilot study
8. Code the content according to definitions
9. Analyze the data
10. Interpret the results

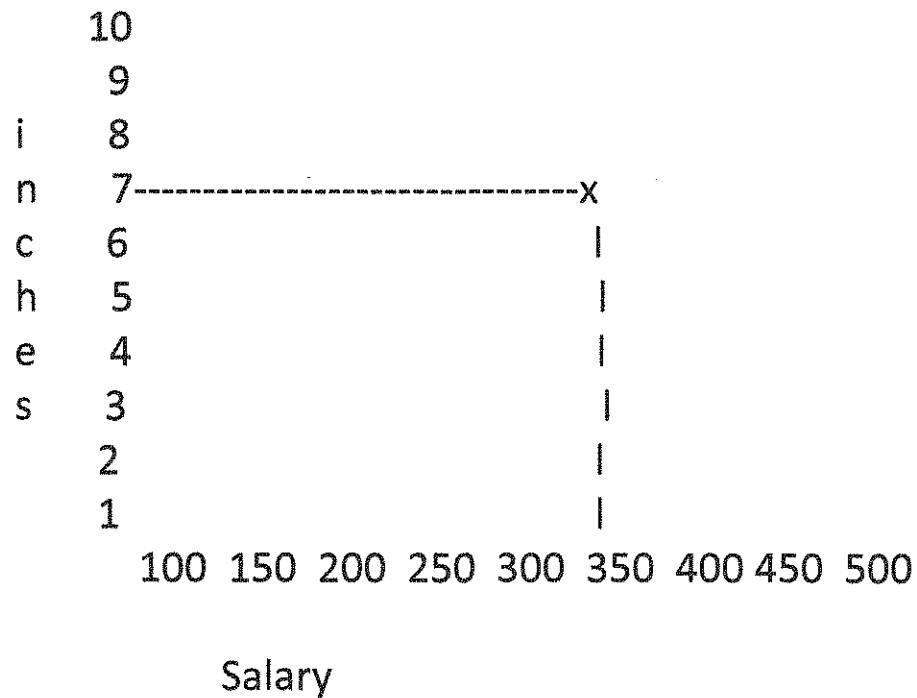
# Content Analysis . . .

- Can find central aspects of social interaction
- Can utilize both quantitative and qualitative
- Can provide historical, over-time trends
- Is unobtrusive so avoids subject structuring responses

## But, Content Analysis . . .

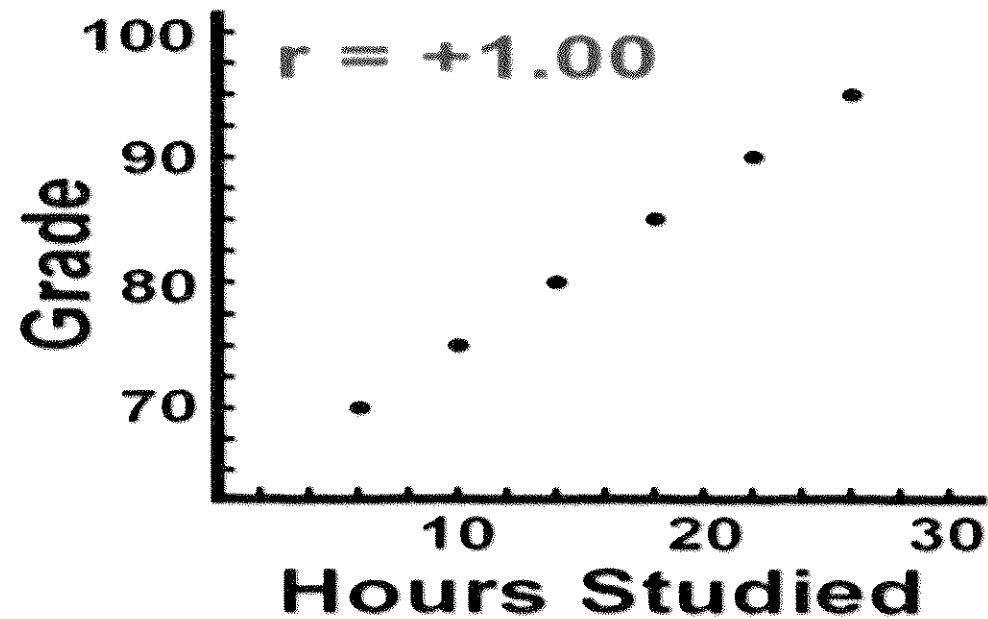
- Can be extremely time consuming
- Often devoid of theoretical base
- Is simplistic, often just word count
- Often disregards content that produced text
- Can be difficult to computerize

If we construct a graph where the horizon line is salary and the vertical line is inches, it would like a bit like this, but larger:



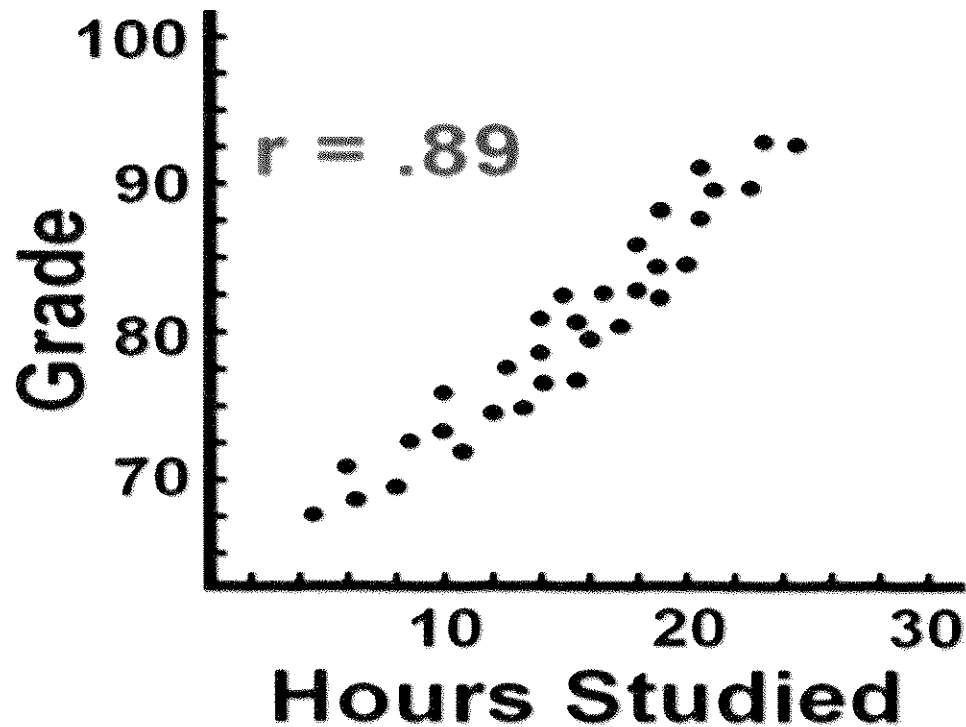
# A Perfect Correlation

Perfect positive (all points fall on a straight line)



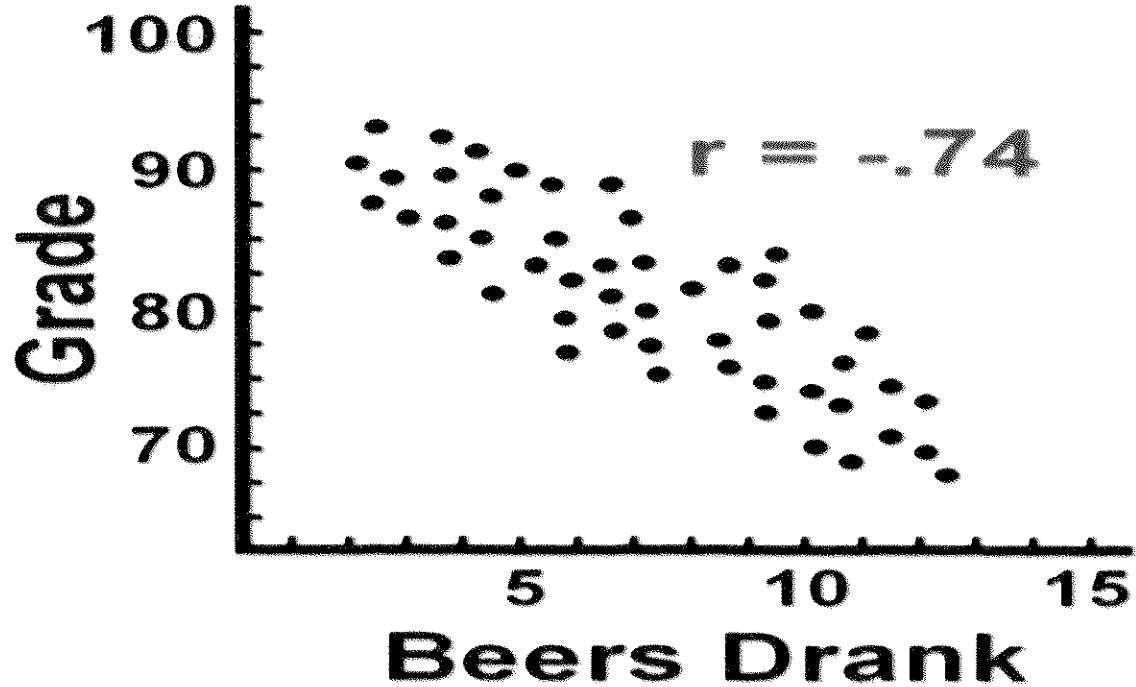
# A Strong Positive Correlation

## Pearson Correlation Example



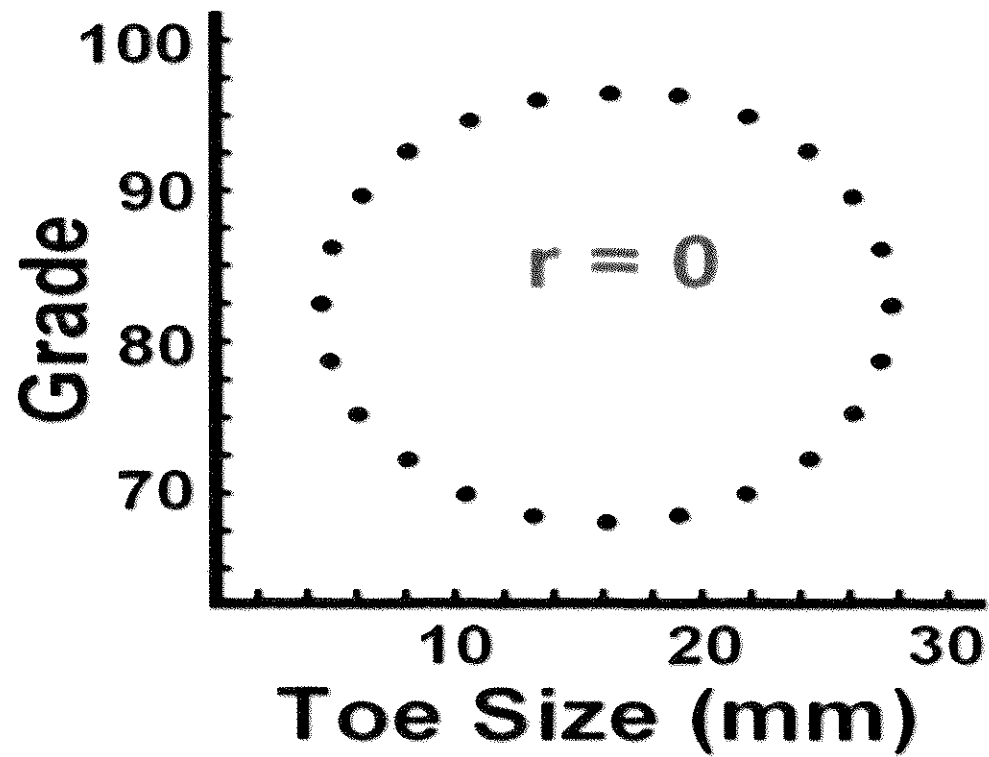
# A Strong Negative Correlation

Pearson Negative Correlation



# No Correlation At All

**No Correlation**





NAME \_\_\_\_\_

Score \_\_\_\_\_

Quiz #1

- |      |       |   |
|------|-------|---|
| TRUE | FALSE | 1. The null hypothesis states that there is a significant difference between two attributes of a variable, caused by a second variable. |
| TRUE | FALSE | 2. The .01 level of probability means study results may be based on random error one out of ten times.                                  |

There are at least eight types of probability and non-probability samples. They are, in alphabetical order, (A) availability (also called convenience); (B) cluster; (C) Purposive; (D) Quota; (E) simple random; (F) stratified random; (G) systematic random; and (H) volunteer. Match each to the one example that best describes it.

- 3. \_\_\_\_\_ draw names from a hat containing names of class members.
- 4. \_\_\_\_\_ advertise for students to participate in a study on advertising methods
- 5. \_\_\_\_\_ draw student sample from SJSU computer proportionate to ethnic percentages.
- 6. \_\_\_\_\_ Professor gives his class a questionnaire on election issue awareness.
- 7. \_\_\_\_\_ Take every 75<sup>th</sup> name from a city computer list of registered voters.

TRUE	FALSE	8. Content Analysis is a method for comparing the means of two groups of students in order to evaluate the effective of an ad.
TRUE	FALSE	9. Content Analysis counts the number of times something is seen, observed or otherwise present.
TRUE	FALSE	10. Content Analysis examines the manifest content, that is whatever is actually present, and not just implied.
TRUE	FALSE	11. Content Analysis can be done at midnight if the researcher so desires and has access to the materials to be counted and coded.
TRUE	FALSE	12. You can't do content analysis of television programming Because there is nothing written that is available to be counted.

13. Which one of the following items or descriptions cannot be measured by Content Analysis?

- A. A preacher's sermon for religious words
- B. Types of trees on the San Jose State University campus
- C. Use of vulgar words in rap music
- D. Patriotic thoughts of the Taliban
- E. Graffiti on San Jose downtown

14. Which one of the following items or descriptions can be measured by Content Analysis?

- A. A political leader's thoughts
- B. Types of trees on the San Jose State University campus
- C. Taliban's moral conduct philosophy
- D. Patriotic thoughts of the Taliban
- E. Results of an experiment

- |      |       |   |
|------|-------|---|
| TRUE | FALSE | 15. The Chi Square can test the significance of group scores on a zero-to-one ratio                           |
| TRUE | FALSE | 16. You can conduct a Chi Square test with as few as ten subjects   |
| TRUE | FALSE | 17. The accuracy of the Chi Square is based on the size of the sample and not the size of the population base |
| TRUE | FALSE | 18. You cannot do a Content Analysis of broadcast news because it is oral.                                    |
| TRUE | FALSE | 19. Content Analysis can analyze graffiti as well as books and records, DVDs, etc.                            |
| TRUE | FALSE | 20. In Content Analysis, the subjects cannot lie to the researcher.   |

ANSWERS:

- |          |           |
|----------|-----------|
| 1. FALSE | 11. TRUE  |
| 2. FALSE | 12. FALSE |
| 3. E     | 13. D     |
| 4. H     | 14. B     |
| 5. F     | 15. FALSE |
| 6. A     | 16. FALSE |
| 7. G     | 17. TRUE  |
| 8. FALSE | 18. FALSE |
| 9. TRUE  | 19. TRUE  |
| 10. TRUE | 20. TRUE  |

*"There are three kinds of lies: lies, damned lies, and statistics."*

Benjamin Disraeli, British politician, 1804-1881

*"Statistics: The only science that enables different experts using the same figures to draw different conclusions."*

Evan Esar, American humorist, 1899-1999

STUDENTS: The three main things you should learn from this course are:

1. What are the types of samples and the characteristics of each?
2. Impact on accuracy on increased number of subjects.
3. Criteria to use in selecting a type of sample.

Some numbers are not precise. They are estimates. The numbers come from surveys or polls, which gather information from, and about, particular population groups. Researchers do gather more precise numbers from experiments and content analysis. These three research forms are the backbone of the quantitative method, which focuses much more on numbers than on words.

Back to surveys, which normally focus on trends while polls zero in on one brief time period. But to most researchers, the terms are interchangeable. It is virtually impossible to ask every single person in large, often scattered groups about their lifestyles and opinions. Not everyone can be reached. It would be too time-consuming. It would be far too expensive.

Also, it is unnecessary.

Instead, social scientists, governments, pollsters and political candidates take samples, a small part of that population. How the sample was selected, as well as the

numbers, and what they represent, determine the value and the accuracy of the numbers generated.

The first question for the reporter incorporating these numbers into a news story is whether the sample is representative of the particular population. To ensure a qualified representation, individuals conducting surveys use the EPSEM structure, equal probability of selection. Its foundation is that if every member of the population has the same chance of being randomly chosen, the sample results can be generalized back to the parent population.

By contrast, the ABC-TV news magazine *Nightline* as well as many local television stations and newspapers reported information, as news, from a volunteer sample, people who called in to give their views. An example of this is having viewers or readers call one phone number if they favored a particular issue or candidate, and call another number if they opposed, such as gun control or abortion. The results, say 8,428 in favor and 9,091 against would then be broadcast.

The numbers are worthless. People who call in or otherwise volunteer are a biased or non-representative sample and the results cannot be generalized to the over-all population.

In other words, sampling selects a random group of Americans whose answers would be the same obtained if every member of the population were interviewed.

How big a sample do you need? Is a sample of 100 enough? 200? 400? Or more? Do you need bigger samples if you have bigger populations, a college, the city of San Francisco, a country?

Sample size depends on a number of things, money, time, purpose . . . and level of accuracy needed, the margin of error. The more people that are sampled, the more confident pollsters can be that the "true" percentage is close to the observed percentage. The margin of error is a measure of how close the results are likely to be. Since the numbers obtained in a survey are an estimate, we need to know how far off our answer might be, that is how much more or less might the population answer be. Or what is the margin of error?

Suppose that 60 per cent of a sample of 100 put ketchup on their eggs at breakfast and a ketchup company wanted to know the margin of error. Divide the number one by the square root of the sample size and both add and subtract it from 60 percent. One divided by 10 is .10 or ten per cent. Therefore the answer to how many people in the population put ketchup on their eggs ranges 50 per cent to 70 per cent. For a sample size of 400, you divide one by 20 and the answer is .05 or five percent cutting the margin of for the population to between 55 and 65 per cent.

But some things, like elections, the difference of whether issues or candidates are favored or opposed can be much closer, say 53 percent for and 47 per cent against. You can see that if the sample size is 400, adding and subtracting results in saying that between 48 per cent and 58 percent favor or that between 42 per cent and 52 per cent oppose.

That is like saying Candidate A will win or lose? Or that the voters favor or oppose the issue. But you knew that before the survey.

You need a bigger survey. Because of the square root factor in determining the margin of error, or size of the spread, the sample size has to be quadrupled in order to cut the margin of error in half. If the sample size is 1,600, its square is 40 and one divided by 40 is .025 which is 2.5 percent. Now the answer to whether Candidate A with 51 per cent will win is yes, because 2.5 per cent added and subtracted to 53 per cent is within 50.5 per cent and 55.5 per cent.

If you guessed that you need larger samples for larger populations, you guessed wrong. One divided by the square of the size of the sample provides the margin of error, regardless of whether the sample was drawn randomly from Kabul or from from all of Afghanistan.

The major advantage of surveys is that a lot of information can be obtained from a variety of people at a reasonable cost. However, they are getting harder to conduct on a much-surveyed United States, sometimes the wrong respondents are included and bad questions can produce biased results. Surveys can cost between \$30 and \$50 per respondent.

There are about five types of surveys that you should know the strengths and weaknesses of: (1) personal (face-to-face); (2) telephone; (3) mail; (4) Internet; and (5) mail.

The advantages of the personal interview include: flexibility; Interviewer observations and rapport of an identifiable respondent; and higher response rate. The drawbacks include that it is the most costly, takes more time than other survey forms,

can have interviewer biasing responses and requires interviewer training. It is also the least safe.

The advantages of telephone surveys are that they are the quickest to complete, cheaper than personal interviews and the interviewer can clarify any misunderstanding. The response rate is higher than all but mail surveys. The disadvantages of telephone surveys include the suspicion you are a telemarketer, the inability to show photos, charts, and graphs plus the difficulty of reaching cell phone users who may differ significantly from users of home or business telephones.

Mail surveys have the advantage of being the cheapest method, eliminating any interviewer bias, are safe, provides the respondent with anonymity, and can cover a large geographic area. However, mail surveys have the lowest response rate, take a longer time, must be completely self-explanatory, the researcher never knows for sure who answered the questions, and if only interested individuals respond, it may be biased concerning the content.

The newest type of survey is the Internet survey. It is generally inexpensive, easy to conduct even with visuals and it can be done quickly. However, Internet surveys are more difficult to generalize to, or be like, the general population, provide no control over the data-gathering process and there is no assurance that the requested respondent is the one who completes the interview.

The mall survey is, as its name implies, conducted generally in shopping malls where researchers roam the corridors or sometimes stake out positions by particular

stores in their search for selected individuals. This survey is quick, relatively cheap, but cannot be used to reflect the overall population.

But then, the good mall survey is not interested in reflecting the people in general. In a good mall survey, the researcher is looking for people with the same interest, such as wearing running or jogging shoes and therefore goes where he can find them, going in and out of shoe and sporting goods stores in shopping malls.



Questions:

1. What does margin of error tell you?
2. Which type of survey is the cheapest? Why?
3. What are the comparative advantages of telephone surveys?
4. What factors would influence you to use particular survey methods?
5. When might you use a mail survey or a volunteer sample?

Words or concepts to know and remember:

EPSEM

Generalizability

Margin of error

Volunteer samples

Mail surveys

STUDENTS: The three main things you should learn from this lecture are:

1. Demonstrate when the Chi Square test is applicable in examining data;
2. Demonstrate how to do a Chi Square;
3. Distinguish whether sample subset differences are statistically significant

Once you have completed your data gathering, it is time to input the information into a computer statistical package and run significance tests to see if any differences are merely by chance or if the subjects are significantly different. With really large differences you can tell a glance the difference must be significant, but you don't know how significant. And if there are differences, but they are small, you don't know for sure one way or the other.

We are going to start our search for statistical significance by examining the chi square test for the goodness of fit. This test is used when you have large groups that are measured using the nominal level where numbers merely classify, such as asking if there a difference in male and female views on candidates for office or preference for sweets, such as ice cream.

Chi (rhymes with sky) square compares two or more groups of subsets of a given population. If 50 percent of a sample prefers chocolate and the other 50 percent prefers vanilla, the next question might be whether different groups, such as men and women each have half and half preferences.

Suppose we take a survey of voters and find that 864 of 1,600 voters intend to vote for Candidate A and the rest, 736, plan on voting for Candidate B. The voters obviously prefer Candidate A.

But what about different groups, men and women, or professional workers and laborers, or college-educated people and less-educated people. Intuitively, we probably think there are differences, but we need science to show us.

The chi square test will examine the numbers in the computer on the assumption that if there is no difference between men and women in their political views, they will vote the same, and thus since 54 percent of the total sample voted for Candidate B, then 54 percent of the men and 54 percent of the women will vote for Candidate B.

In other words, the chi square test expects the numbers to look like this:

	Candidate A	Candidate B	Total
Men	378	486	864
Women	322	414	736
Total	700	900	1,600

Just about 54 percent of both men and women prefer Candidate A. The conclusion, there is no difference in support for the candidates by gender. But what if more than 54 percent of men prefer Candidate A? Suppose 60 percent do. If one set of numbers and its percentage goes up, the other must go down in order to maintain the overall numbers and percentages. That means the table would look like this:

	Candidate A	Candidate B	Total
Men	346	518	864
Women	354	382	736
Total	700	900	1,600

Notice that it is awkward looking at the numbers. That is because the two group totals is not equal. It is better to show the numbers in percentage form, like this:

	Candidate A (n=700)	Candidate B (n=900)	Total
Men	40%	60%	54%
Women	48	52	46

There is something wrong with this table. The numbers add across, because the independent variable, men and women are placed in the row position. It is easier to read and is usual for the independent variable to be presented in the column position, because the columns of percentages should then total 100%, our results should look like this:

	Men	Women
Candidate A	40%	48%
Candidate B	60	52

Before we look at how these statistics are handled in a compute, we have to discuss another aspect first. Computers are stupid. They are fast. They are accurate. But the numbers are meaningless to them so when they indicate that there is a significant difference, a human mind has to look it at.

For example, before we had television, most of the important discussions in American political conventions were in the smoke-filled side rooms where party leaders negotiated for support and patronage. The reporters were out in the main room with all the state delegates. After a while one reporter noticed that the delegates to the Democratic convention had more hair than the delegates to the Republican convention.

He mentioned this to other reporters who felt likewise, so at the next conventions, they counted whether the men were bald (complete or partial) or had hair. They found that of all of the 1,950 delegates, the number with a full head of hair totaled 1,014 while 936 were bald. In percentages, 52 percent had hair and 49 percent were bald. When the computer separated the numbers by party identification and hair, this is what it revealed:

	Democrats	Republicans	Total
Had Hair	539	237	776
Bald	281	593	874
Total	820	830	1,950

Since it is much to compare something by using percentages because they eliminate the problem of different sizes, we should revise our table to look like this:

	Democrats (n=820)	Republicans (n=830)	Total (n=1,950)
Had Hair	65.7%	28.6%	52%
Bald	34.3	71.4	48

It sure looks like Democrats have more hair than Republicans. To find out, we look at what the Chi Square test did with the numbers.

	Hair Observed	-	Hair Expected	=	Difference	Difference Squared	Dif <sup>2</sup> /Expected	=	Chi Square
Dem (hair)	539	-	426	=	113	12,769	12,769/426	=	28.002
Dem (bald)	281	-	394	=	-113	12,769	12,769/394	=	32.409
Rep (hair)	237	-	432	=	-113	12,769	12,769/432	=	29.558
Rep (bald)	593	-	398	=	113	12,769	12,769/398	=	<u>32.083</u>
TOTAL CHI SQUARE VALUE =									122.052

In order to interpret what 122.052 means, you need to consult a table of Chi Square values. Here is an excerpt:

Probability of Occurrence by Chance						
DF	.20	.10	.05	.02	.01	.001
1	1.642	2.706	3.841	5.412	6.635	10.827
2.	3.219	4.605	5.991	7.824	9.210	13.815

First, you need to know what DF stands for. It means degrees of freedom, which is how much freedom is there for changing the number in a cell. Here is an example of one degree of freedom.

	Men	Women	Total
Liberal			100
Conservative			200
TOTAL	120	180	300

Can you tell how many men are liberal? Or how many women are conservative? You can't. All you know is that the totals. But, suppose you know that the table looks like this:

	Men	Women	Total
Liberal		70	100
Conservative			200
TOTAL	120	180	300

By inserting the one number, you know the other three because  $100 - 70$  means there are 30 men who are conservative. And subtracting 30 from 120 gives you 90 male conservatives and subtracting 70 from 180 means 110 women are conservative.

There are more degrees of freedom as you increase the number of columns and rows, such as liberal, moderate and conservative, by 20-35 years old, 36-50, and over 50 looks like this:

	20-35	36-50	Over 50	Total
Conservative	_____	_____	_____	60
Moderate	_____	_____	_____	210
Liberal	_____	_____	_____	90
TOTAL	120	150	90	360

There are nine numbers there. How many do we need to know until we know all the rest. HINT: It is the number of rows minus one times the number of columns minus one. Since we have three columns and three rows, it will be  $(3 - 1)(3 - 1)$  or 2 times 2 equals 4. We need to know four numbers. For example, if we know the first two in the conservative row, then we know the third one. If we know the first one in the moderate row, then we know the first in the liberal row, and if we know the last one in the liberal row, we know the second, and then all of them. What has just been said looks like this visually:

	20-35	36-50	Over 50	Total
Conservative	10	40	_____	60
Moderate	80	_____	_____	210
Liberal	_____	_____	35	90
TOTAL	120	150	90	360

The rest of the numbers can be determined by subtractin these individual row and column totals. Because overall chi square values get larger with more groups, it requires more levels, or degrees of freedom.



With one degree of freedom, any number less than .05 means the results would have more than five percent chance of occurring randomly. Less than the number under .20, 1.642 means the results have more than a 20 percent chance of occurring randomly. Since science considers the .05 level to be where statistical significance begins, it means if our chi square number exceeds 3.841, the results are significant, meaning that changes in one variable will result in changes in the other variable. Our number, 122.052, exceeds 3.841 considerably. In other words there is a relationship between hair loss and political party identification.

But that cannot be. It would be silly to say that if you become a Republican, you are more likely to lose your hair. But it would be equally silly to say that if you lose your hair you are more likely to become a Republican.

So how can the test show a relationship? Remember, the computer is stupid. It crunches numbers and spits them out. A computer can add the temperature at the North Pole to the national income of Greece and get a number. That number will be useless.

What it means is that there is an intervening variable, something that cause both age and hair loss in men. What is it? Think about it one at a time. What causes men to lose their hair? When does a man become a Republican, or in European and other countries, when does a man become a conservative? What could cause both?

The answer is: Age!

But let us consider a more realistic example, whether people become more conservative as they age. Our hypothesis is that age and political ideology are related.

This is called  $H_1$ . Chi Square tests the null hypothesis, the opposite which says there is no connection. It is listed as  $H_0$ . The two are, respectively,  $H_1$  and  $H_0$ .

Consider this table of results. What does it tell you?

Political	Age Group			Over 50 Totals
	20-35	36-50		
Conservative	8.3%	26.7%	11.1%	16.7%
Moderate	66.7	56.7	50.0	58.3
Liberal	25.0	16.6	38.9	25.0

The first thing this table tells you is that, by far, most of the people in this sample are moderate and, secondly, that there are more liberals than conservatives. When you look at the percentages inside the table, it suggests that young people are mostly moderate (66.7%), become slightly more conservative than liberal as they become somewhat older, and the switch back to becoming liberal when they get to be over 50. But is this just an accident of this sample? Chi Square can tell us the answer.

The actual numbers for each group can be seen in the observed cells below:

$$\text{Observed} - \text{expected} = \text{Differences Squared} / \text{Expected} = \text{Chi Square}$$

Conservative (20-35)	10	-	20	-10	100	/20	=	5.000
Conservative (36-50)	40	-	25	15	225	/25	=	9.000
Conservative (51+)	10	-	15	-5	25	/15	=	1.667
Moderate (20-35)	80	-	70	10	100	/70	=	1.428
Moderate (36-50)	85	-	87.5	-2.5	6.25	/87.5	=	0.071
Moderate (51+)	45	-	52.5	-7.5	56.25	/52.5	=	1.071
Liberal (20-35)	30	-	30	---	---	/30	=	0.000
Liberal (36-50)	25	-	37.5	-12.5	156.25	/37.5	=	4.167
Liberal (51+)	35	-	22.5	12.5	156.25	/22.5	=	6.944
TOTAL	360		360					29.348

As before, we need to turn to the table of Chi Square values. However, this time we need to look at the fourth row, the line of 4 under DF. What number is 29.348 higher than in the fourth row?

DF	.20	.10	.05	.02	.01	.001
1	1.642	2.706	3.841	5.412	6.635	10.827
2.	3.219	4.605	5.991	7.824	9.210	13.815
3.	4.642	6.251	7.815	9.837	11.245	16.266
4.	5.989	7.779	9.488	11.668	13.277	18.469

The answer is that 29.348 is higher than all of them. That means that the results in our sample of conservative-liberal ideology by age would occur by chance, less than one in .001, which is 1,000 times. That means that age definitely is a factor in a person's conservative-liberal ideology.

Social scientists do not do these statistics by hand, as was done for this exercise. But they can. By looking at these examples, you now know what the computer does when it examines data using the chi square test.

The Chi Square test is one of the principal tools used by mass media researchers in examining the relationship of readership, credibility and many, many other questions about the news media. It is suggested that you go online, google in Chi Square test. It will show you many examples which will help you get a feel for its use. Also, look at studies in journalism periodicals and you will understand the tables.

## Questions

1. When the chi square test separates a variable, such as gender, into two or more attributes, such as men and women, to look at the attributes of another variable, such as preference for soccer or dramas on television, what is the test doing? What is it comparing? Comparing to what?
2. Give an example to show how the law of large numbers improves the likelihood of statistical significance in chi square tests?
3. Demonstrate how degrees of freedom are determined. What does it mean to say that there is one degree of freedom in a two-by-two table?
4. What is meant by the .05 level? Compare it to the .20 level.
5. What is the difference between the numbers in a correlational analysis and the numbers in a chi square analysis?



Words or concepts to know and remember

Degrees of freedom

Chi square value

Level of significance

Margin of error

Attribute

Variable



# What Chi Square Expects

	Candidate A	Candidate B	Total
Men	378	486	864
Women	322	414	736
Total	700	900	1,600

# But what if more than 54 percent of men prefer Candidate A?

Suppose 60 percent do. If one set of numbers and its percentage goes up, the other must go down in order to maintain the overall numbers and percentages. That means the table would look like this:

	Candidate A	Candidate B	
Total			
Men	346	518	864
Women	354	382	736
Total	700	900	1,600

It is better to show the numbers in percentage form, like this:

	Candidate A (n=700)	Candidate B (n=900)	Total
Men	40%	60%	54%
Women	48	52	46



## IV At Top of Columns

It is usual for the independent variable to be in the column position. The column percentages then total 100%, like this:

	Men	Women
Candidate A	40%	48%
Candidate B	60	52

# Democrats and Republicans

	Democrats	Republicans	Total
Had Hair	539	237	776
Bald	281	593	874
Total	820	830	1,950

	Democrats (n=820)	Republicans (n=830)	Total (n=1,950)
Had Hair	65.7%	28.6%	52%
Bald	34.3	71.4	48

# Chi Square calculation . . .

	Hair Observed	-	Hair Expected	=	Difference	Difference Squared	Dif <sup>2</sup> /Expected i	=	Chi Square
Dem (hair)	539	-	426	=	113	12,769	12,769/426	=	28.002
Dem (bald)	281	-	394	=	-113	12,769	12,769/394	=	32.409
Rep (hair)	237	-	432	=	-113	12,769	12,769/432	=	29.558
Rep (bald)	593	-	398	=	113	12,769	12,769/398	=	<u>32.083</u>
TOTAL CHI SQUARE VALUE =									122.052

# Chi Square Table

Probability of Occurrence by Chance

DF	.20	.10	.05	.02	.01	.001
1	1.642	2.706	3.841	5.412	6.635	10.827
2	3.219	4.605	5.991	7.824	9.210	13.815



	20-35	36-50	Over 50	Total
Conservative	10	40	<u>          </u>	60
Moderate	80	<u>          </u>	<u>          </u>	210
Liberal	<u>          </u>	<u>          </u>	35	90
TOTAL	120	150	90	360



Political	Age Group			Totals
	20-35	36-50	Over 50	
Conservative	8.3%	26.7%	11.1%	16.7%
Moderate	66.7	56.7	50.0	58.3
Liberal	25.0	16.6	38.9	25.0
•				

# Chi Square calculation

Observed – expected = Differences Squared /Expected = Chi Square							
Conservative	(20-35)	10	-	20	-10	100 /20	= 5.000
Conservative	(36-50)	40	-	25	15	225 /25	= 9.000
Conservative	(51+)	10	-	15	-5	25 /15	= 1.667
Moderate	(20-35)	80	-	70	10	100 /70	= 1.428
Moderate	(36-50)	85	-	87.5	-2.5	6.25 /87.5	= 0.071
Moderate	(51+)	45	-	52.5	-7.5	56.25 /52.5	= 1.071
Liberal	(20-35)	30	-	30	---	--- /30	= 0.000
Liberal	(36-50)	25	-	37.5	-12.5	156.25 /37.5	= 4.167
Liberal	(51+)	35	-	22.5	12.5	156.25 /22.5	= 6.944
TOTAL		360		360			29.348



DF	.20	.10	.05	.02	.01	.001
1	1.642	2.706	3.841	5.412	6.635	10.827
2.	3.219	4.605	5.991	7.824	9.210	13.815
3.	4.642	6.251	7.815	9.837	11.245	16.266
4.	5.989	7.779	9.488	11.668	13.277	18.469

NAME \_\_\_\_\_

SCORE \_\_\_\_\_

Quiz #2

Match each of these three tests: (A) chi square; (B) t-tests, and (C) correlation with their use, or description below that best describes them.

1. \_\_\_\_\_ measures changes in one variable relative to changes in another
2. \_\_\_\_\_ assumes the percentage of a subset is like the percentage of the total
3. \_\_\_\_\_ uses a +1.00 to -1.00 scale to measure the strength of association of two variables
4. \_\_\_\_\_ measures nominal data, such as men and women, Republicans and Democrats, etc.
5. \_\_\_\_\_ compares the scores of one group with the scores of another group from the same test
6. \_\_\_\_\_ can determine the percentage that one variable change produces in the other variable
7. \_\_\_\_\_ Which one of the above three tests can best determine whether membership in a political party varies significantly by religion?
8. \_\_\_\_\_ Which one of the above three tests is able to use nominal level data?
9. \_\_\_\_\_ Which one of the above three tests compares the means of two groups?
10. \_\_\_\_\_ can analyze subsets, that is comparing not just religion with political Party Identification, but comparing individual religions within a particular political party or across two or more parties.

Indicate whether each of the following is more appropriate for (A) chi square testing or for (B) t-tests.

11. \_\_\_\_\_ Comparing views of left-wing, moderates, and right-wing on questions of social justice.
12. \_\_\_\_\_ A group of 10 girls and a group of 10 boys on the ability to lose pounds
13. \_\_\_\_\_ Comparing two groups of have average scores of 76 and 81 to determine statistical significance

14. \_\_\_\_\_ Assessing the ability of Afghanistan's military readiness before and after training by U.S. military forces, comparing Army, Navy and Special Forces
15. \_\_\_\_\_ Determining if the prices of vegetables in the farmers' markets increased significantly in a six-month period.
16. \_\_\_\_\_ Testing whether men or women were more likely to remember copy points in a particular advertisement.
- TRUE      FALSE      17. You don't need to know averages in order to conduct a chi square test.
- TRUE      FALSE      18. Chi Square and t-tests can both analyze three or four groups at the same time.
- TRUE      FALSE      19. The t-test can compare how the individual members in two groups change scores over time, but Chi Square cannot do that.
- TRUE      FALSE      20. The t-test is more powerful than the Chi Square test because it uses a higher level of measurement, averages, than does Chi Square.

#### Answers

- |     |   |     |       |
|-----|---|-----|-------|
| 1.  | C | 11. | A     |
| 2.  | A | 12. | B     |
| 3.  | C | 13. | B     |
| 4.  | A | 14. | A     |
| 5.  | B | 15. | B     |
| 6.  | C | 16. | A     |
| 7.  | A | 17. | TRUE  |
| 8.  | A | 18. | FALSE |
| 9.  | B | 19. | TRUE  |
| 10. | A | 20. | TRUE  |

## Advanced Mass Media Research – 5 Experiments and t-tests

STUDENTS: The main things you should learn from this lecture are:

1. What are the elements required in order to have a good experiment.
2. The advantages and disadvantages of experiments
3. The reason for preferences for pre-test-post-test or for post-test only
4. What kinds of things the t-test can be used to test
5. The concept of variability or why groups with the same means are very different from each other
6. How to conduct a t-test

Advertisers and quite a few public researchers and probably a fair number of newspapers choose the third form of quantitative research, the experiment because it allows you to find out whether one variable, a treatment, **influences (causes)** an effect in another variable by **randomly assigning** some participants to get the **treatment**, while other participants do not.

Experiments that uses random assignment can avoid problems of measurement and manipulation. Others cannot. Some may not need to.

There are three steps in conducting an experiment:

1. Randomly assign subjects to two (or more) groups, the control group and the test group (the dependent variable).
2. Second is the manipulation or treatment. You do something to the test group that you don't do to the test group.
3. Measure Control and Test groups for any differences.

### **Canons of Causality:**

- 1. Time Order:** Causation is present if and only if the cause precedes the effect.
- 2. Association:** Causation occurs only if some tendency for change in A results in change in B.
- 3. Non-spuriousness:** All other alternatives must be ruled out (rival causes).

There are several experimental structures. But good experiments, those that can determine if there are significant effects, usually contain these three elements:

- R = randomization of subjects into test or control groups
- O = observation or measurement of subject on some topic, etc.
- X = treatment, manipulation, or something done to test group subjects

The following three groups occasionally generally are unacceptable because they lack the randomization or measurement:

1. One-Shot Case Study

X O

We do not know the structure of the test group (no R), no control group for comparison, and since there has been no measurement (O) taken before the treatment, we do not know what difference, if any, that the treatment (X) made.

2. One-Group pre-test/post-test

O X O

3. Static-Group Comparison

X O  
O

Again, Group 2 has neither randomization nor any control group for comparison. Group 3 also fails to randomize although it does compare groups after a treatment.

Below are the three experimental formats that lead to accurate testing.

1. Pre-test/Post-test Control Group

R O X O  
R O O

2. Post-Test Only Control Group

R X O  
R O

### 3. Solomon-Four Group

R	O	X	O
R	O		O
R		X	O
R			O

Group 1 randomizes for both groups, tests both, then manipulates the test group and measures again. Group 2 has no pre-tests. The argument is that pre-testing sensitizes the subjects. In addition, you don't need a test group pre-test to compare with its post-test, because its score can be compared to the control group's score. The logic is that since the control group did not receive any manipulation, its post-test score should be about the same as its pre-test score if it had been given a pre-test. Both groups are the same due to randomization; therefore the test-group's pre-test score would have been the same as the control group's pre-test score.

The Solomon-Four is simply the other two combined which means you can examine the pre-test scores if you wish, or you can examine the post-test scores only and also compare both sets.

**Evidence of Causality:** An experiment is a valuable source of information because it is the only method that allows us to make causal statements about what makes something happen.

There two major advantages to experiments:

1. Control: Researchers have control over the environment, the variables and the subjects.

2. Replication: The conditions of the study are clearly spelled out in the description of the experiment.

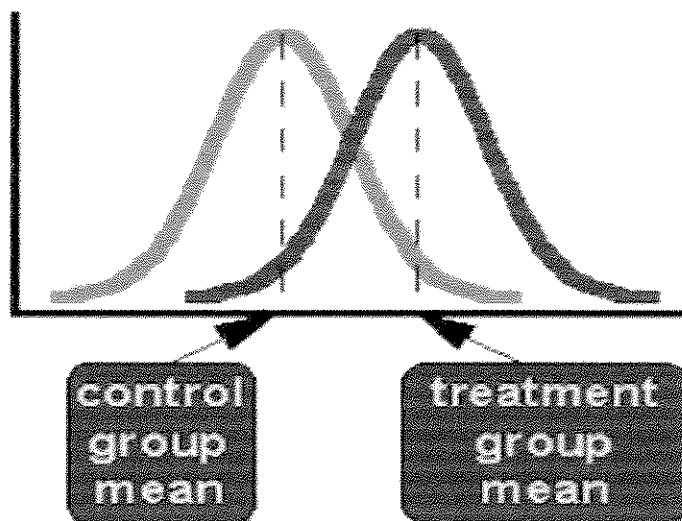
Disadvantages of experiments include:

1. Artificiality
2. Experimenter bias
3. Limited scope
4. Lack of external validity
5. Subject's awareness of the testing situation

What does a t-test do?

It compares the difference between two groups that are measured on the same variable. It must be a ratio level variable, although the two groups can both be nominal. What the t-test does is examine the mathematical distance between the two average scores. The examples below demonstrate visually what it looks

## Test of Means



Most mass media research would be one of two types:

1. Comparing two groups on the same thing
2. Comparing one group before and after

The important thing to remember is that the two groups could be male and female, or reader and viewer as the independent variable, and the dependent variable would be ratio level number, such as number of minutes read a newspaper daily.

Let us look at an example. Suppose our TV station, or the local newspaper, wanted to see if there was a difference in the ability of men and women to recall the copy points in an advertisement. Both men and women are shown the commercial, probably as part of a longer broadcast, or read it on a news page. They can all be in the same room at the same time. Then they fill out a paper test that asks them what copy points they can remember.

Suppose we have 10 people in each group, although one group can be larger than the other. Below are their scores on how many copy points they remember.

### T-test example

Female Recall Scores			Male Recall Scores		
X	x	$x^2$	X	x	$x^2$
4	-4	16	2	-4	16
4	-4	16	3	-3	9
5	-3	9	4	-2	4
7	-1	1	4	-2	4
7	-1	1	4	-2	4
8	0	0	6	0	0
9	1	1	6	0	0
9	1	1	8	2	4
12	4	16	10	4	16
15	7	49	13	7	49
80		110	60		106



The score is in the first column for Females and the first column under Males. The scores are added, to 80 and 60, respectively. Next we need the mean. We divide the 80 and 60 by 10 members of each group, and the mean for women is 8 and for men it is 6. That seems like that should be a significant difference, but we need to test it to be find out.

We subtract the mean from each score which is listed in column two for both females and males. Since the first woman remembered four copy points, we subtract 8, the mean, from 4 and the answer is -4. The t-test cannot deal with negative numbers, so we square each difference. The -4 now becomes 16. It is in the third column. We add all the differences squared, the third columns and the totals are 110 for women and 106 for men.

We are testing the null hypothesis of no difference that the mean of Group 1 will not differ significantly from the mean of Group 2.

To test it, we take the square root of the following: we divide the sums of the square scores by the sum of the group subjects minus 2 and multiply the result by one divided by the first sample size plus one divided by the second sample size. If we look at the parts we have:

--sums of squares are 110 and 106

--sums of group subjects minus two are  $10 + 10 - 2 = 18$

--first group sample is 10 and so is second group sample size.

Therefore we add 110 and 106 divided by 18 and multiply that one divided by 10 plus one divided by 10. That equals  $216/18$  times  $1/10 + 1/10$ .

The result is 12 times  $0.2 = 2.4$ .

The square root of 2.4 is 1.549.

Now we have to divide the difference of the means, 8 and 6, by 1.549. That is  $2/1.549 = 1.291$ .

But so what does this number tell us? By itself, nothing. We have to go to a mathematical table which illustrates the values of the normal curve.

## Table of T-test Values

_degrees Of freedom	.10	.05	.01	.001
17	1.740	2.110	2.898	3.965
18	1.734	2.101	2.978	3.992
19	1.729	2.093	2.861	3.883

Since we have two groups, our degrees of freedom are the 20 people minus 2 or 18.

We look at row 18 to see how many of the numbers that is larger. We see that it is not larger than even the first number. This tells us that there is so statistically significant difference in the recall of the advertisement by men and women.

# Experiments

An experiment is a research that allows scientists to find out whether one variable, a treatment, **influences (causes)** an effect in another variable by **randomly assigning** some participants to get the **treatment**, while other participants do not.

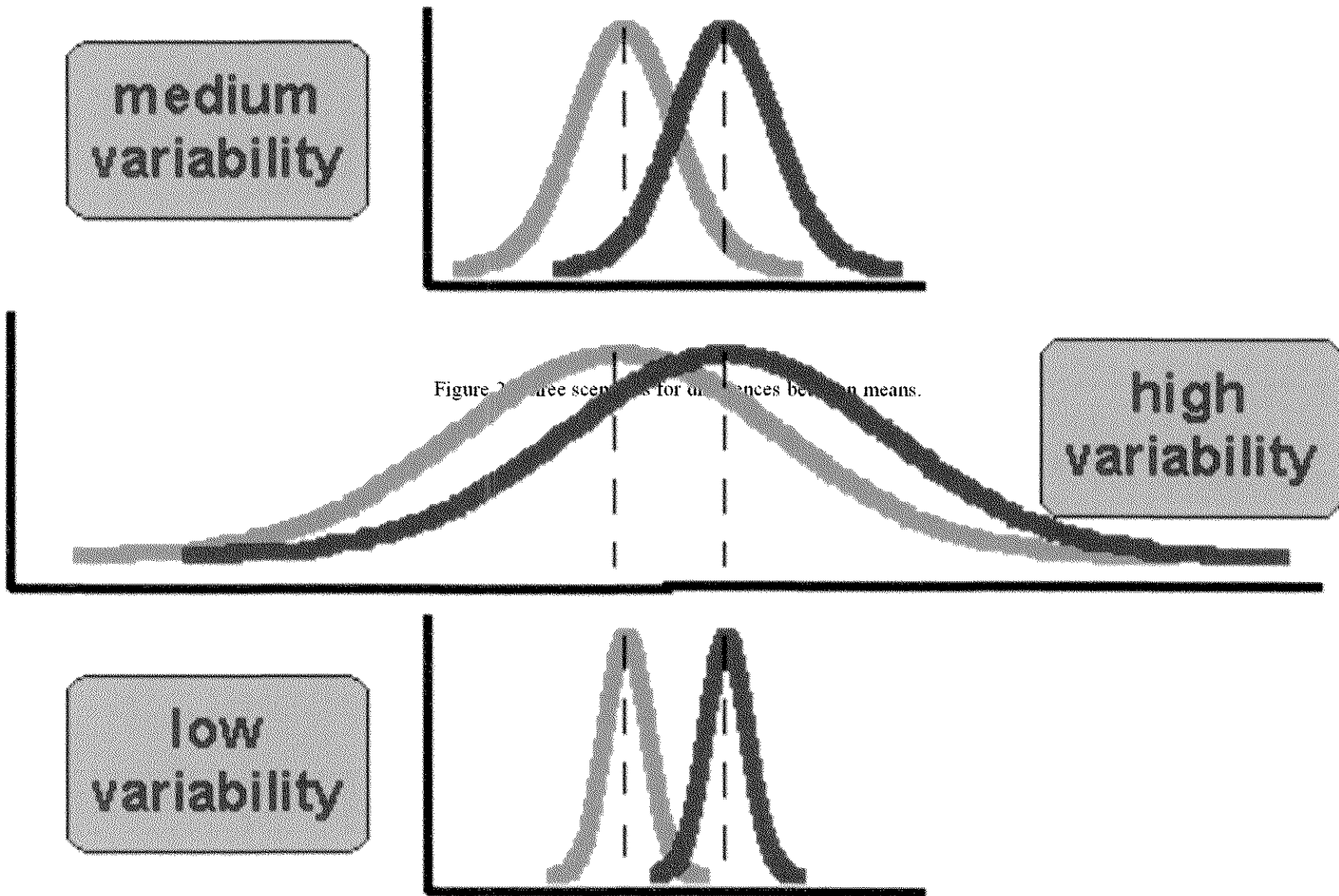
# Experiments: Three Stages

1. Random assignment (sampling)
2. Manipulation
3. Measurement

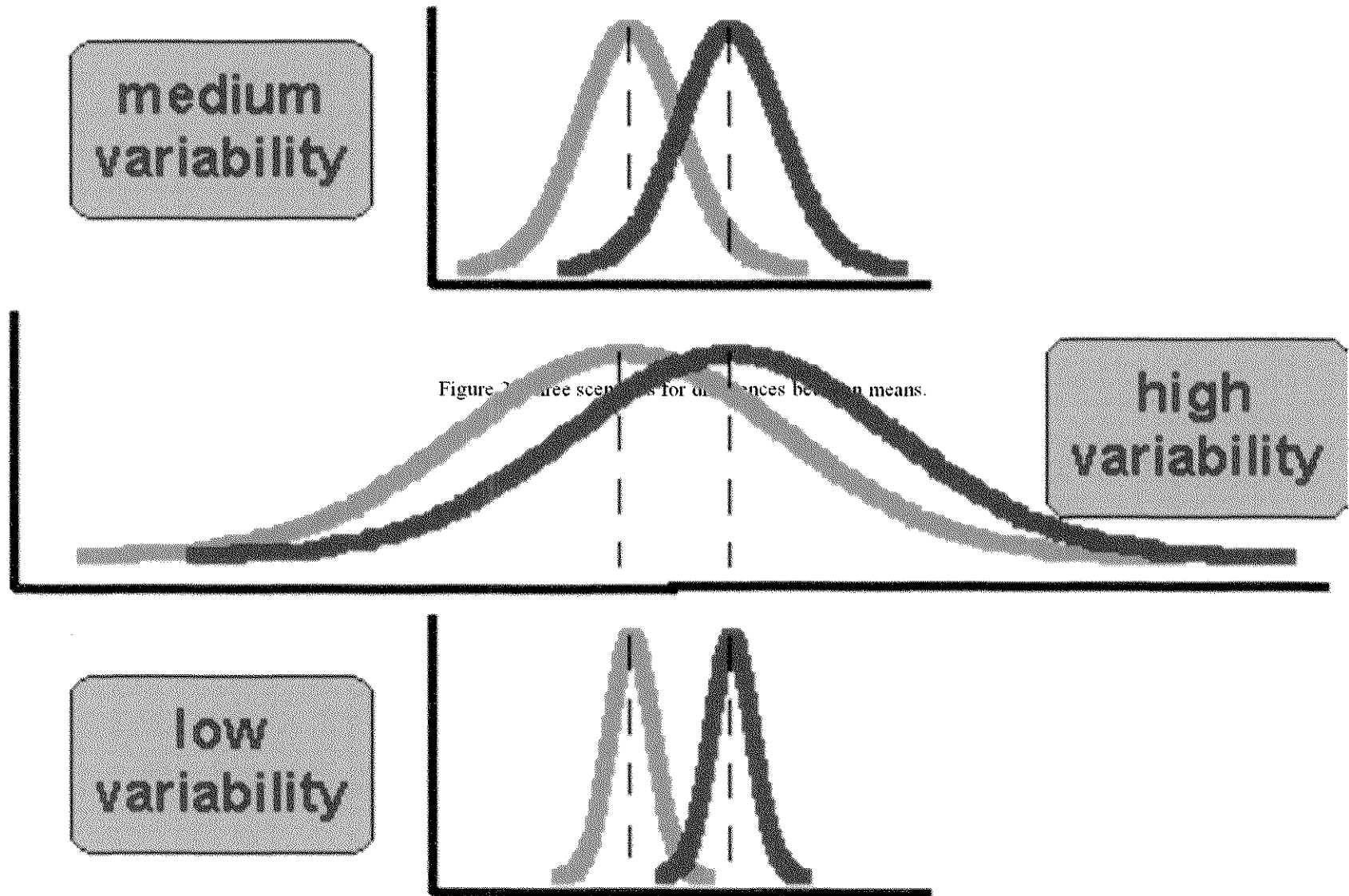
# Experiments: Three Stages

1. Random assignment to two (or more) groups, the dependent variable:
  1. Control
  2. Test
2. Take, give, or otherwise do something, the treatment, (manipulation) to the Test group.
3. Measure Control and Test groups for any differences.

# Test of Means



# Test of Means



# Experiment Types

1. Quasi-experimental – No randomization and no pre-test
2. Experimental – Randomization, may, or may not, pre-test



# Experiment Types

## Canons of Causality:

- 1. Time Order:** Causation is present if and only if the cause precedes the effect.
- 2. Association:** Causation occurs only if some tendency for change in A results in change in B.
- 3. Non-spuriousness:** All other alternatives must be ruled out (rival causes).

# Three Quasi-Experimental Types

1. One-Shot Case Study

X O

2. One-Group pre-test/post-test

O X O

3. Static-Group Comparison

X O

O

R = Randomization

O = Observation

(measurement)

X = treatment

# Three Experimental Types

:

## 1. Pre-test/Post-test Control Group

R	O	X	O
---	---	---	---

R	O		O
---	---	--	---

# Three Experimental Types

## 2. Post-test Only Control Group

R	X	O
R		O

# Experiment Advantages

**Evidence of Causality:** An experiment valuable sources of information because it is the only method that allows us to make causal statements about what makes something happen.

# Experiment Advantages

**Control:** Researchers have control over the environment, the variables, and the subjects.

**Replication:** The conditions of the study are clearly spelled out in the description of the experiment.

# Experiment Disadvantages

1. Artificiality
2. Experimenter bias
3. Limited scope
4. Lack of external validity
5. Subject's awareness of the testing situation

# Experiment Design Elements

R = random assignment of subjects

O = observation (measurement)

X = treatment (the manipulation)



# Quasi (Pre)-Experiment Designs

1. One-shot case study:

X O

2. One-group pre-test/post-test:

O X O

3. Static-group comparison:

X O

O

# Three Experimental Types

## 1. Pre-test/post-test control group

R	O	X	O
R	O		O

# Three Experimental Types

## 2. Post-test Only Control Group

R	X	O
R		O

# Three Experimental Types

## 3. Solomon-Four Group

R    O    X    O

R    O            O

R            X    O

R                    O

# T-tests

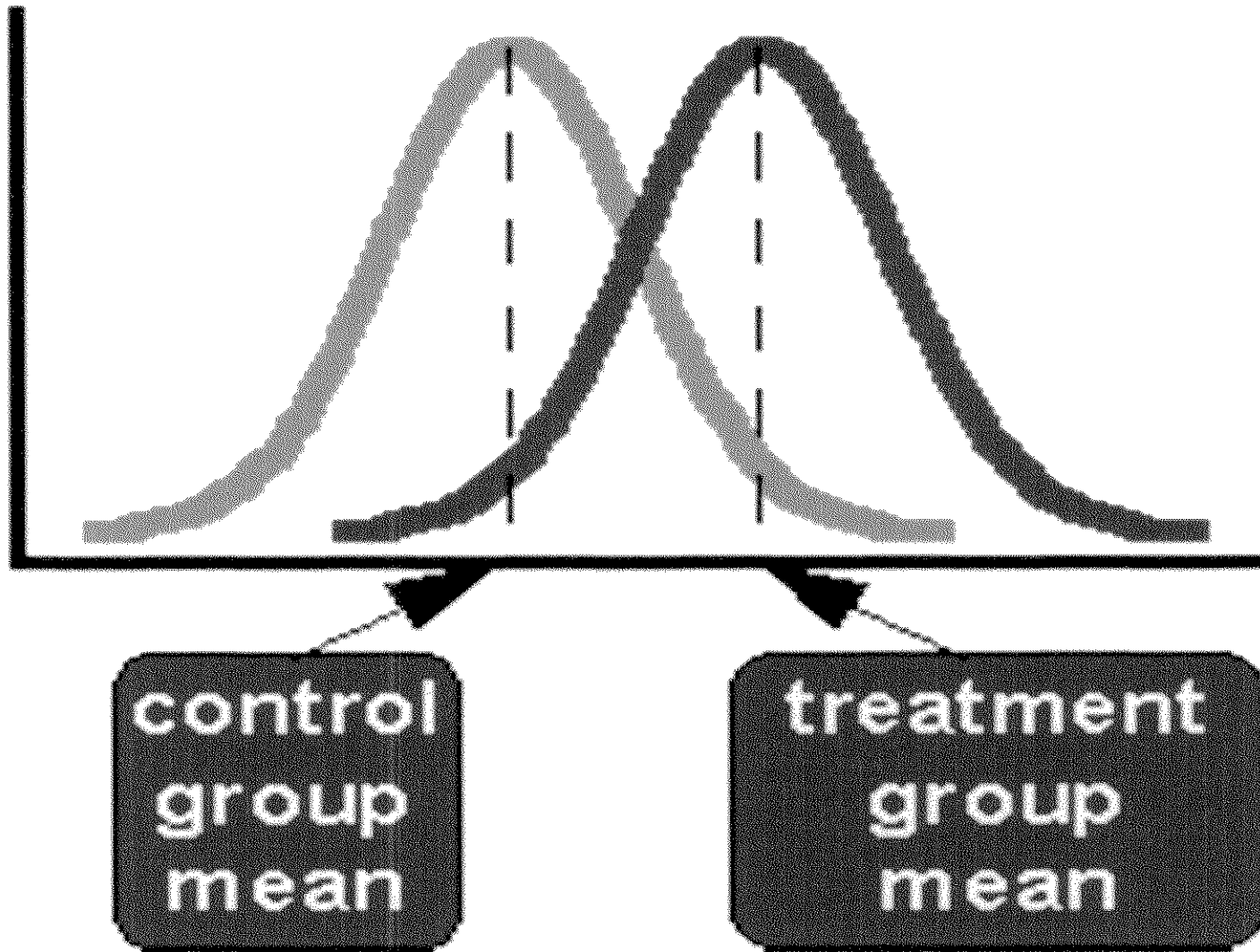
QUESTION: When do we use the t-test?

ANSWER: When we compare the difference between two groups measured on a ratio level dependent variable.

Independent Variable: Categorical

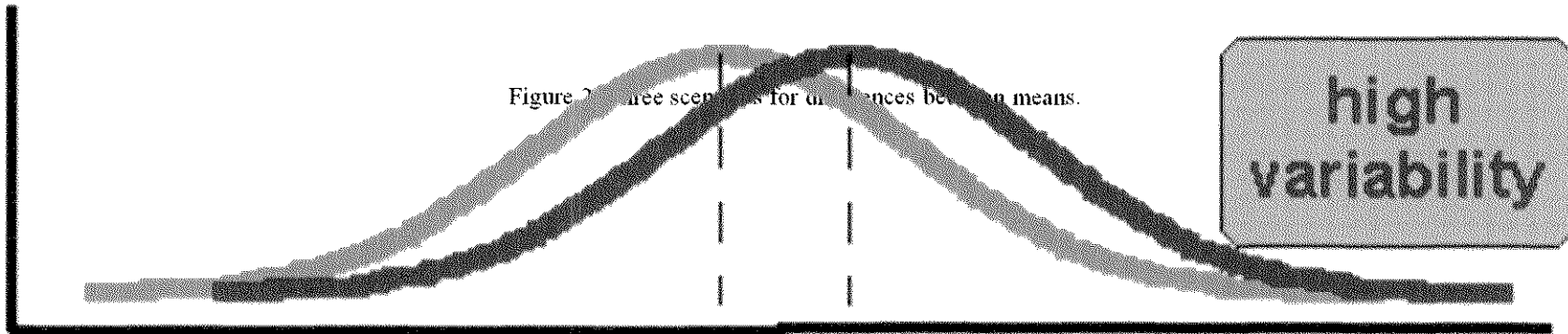
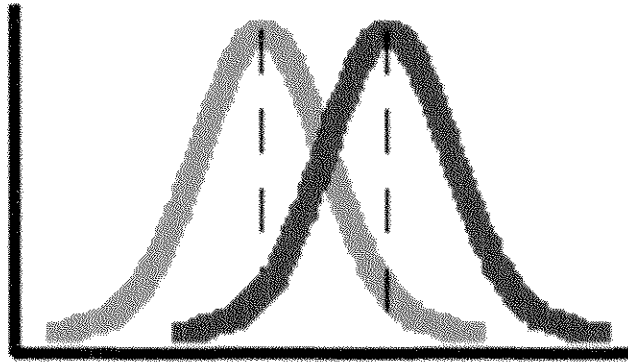
Dependent Variable: Numerical

# Test of Means



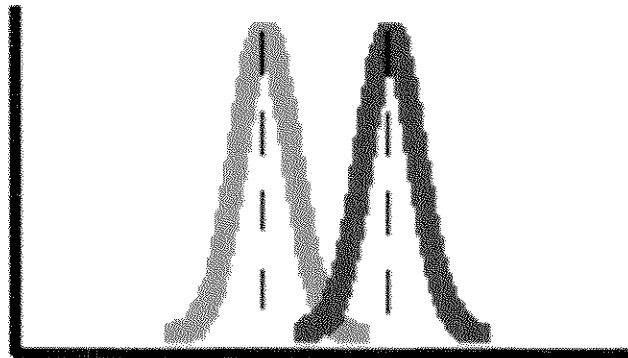
# Test of Means

medium  
variability



high  
variability

low  
variability



# T-test example

## Female Recall Scores

X	x	$z^2$
4	-4	16
4	-4	16
5	-3	9
7	-1	1
7	-1	1
8	0	0
9	1	1
9	1	1
12	4	16
15	7	49
80		110

## Male Recall Scores

X	x	$x^2$
2	-4	16
3	-3	9
4	-2	4
4	-2	4
4	-2	4
6	0	0
6	0	0
8	2	4
10	4	16
13	7	49
60		106



# T-test formula

The null hypothesis is that Group 1 is like Group 2, or more specifically that the mean of Group 1 will not differ significantly from the mean of Group 2.

To test it, we divide the sums of the squared scores by the sum of the group subjects minus 2 and multiply the result by one divided by the first sample size plus one divided by the second sample size.

# T-test formula

$$\frac{\underline{110 + 106}}{10 + 10 - 2} \times \frac{\underline{1} + \underline{1}}{10 \quad 10} = \underline{216} \times \underline{2} = 12 \times .2$$

12 X .2 = 2.4 Square root 2.4 is 1.549

$$\underline{8 - 6} = 1.29$$

1.55

# Table of T-test Values

\_degrees

Of freedom	.10	.05	.01	.001
17	1.740	2.110	2.898	3.965
18	1.734	2.101	2.978	3.992
19	1.729	2.093	2.861	3.883

# Quasi (Pre)-Experiment Designs

1. One-shot case study:

X O

2. One-group pre-test/post-test:

O X O

3. Static-group comparison:

X O

O

# Three Experimental Types

## 1. Pre-test/post-test control group

R	O	X	O
R	O		O

# Three Experimental Types

## 2. Post-test Only Control Group

R	X	O
R		O

# Three Experimental Types

## 3. Solomon-Four Group

R	O	X	O
R	O		O
R		X	O
R			O

# Design Validity Problems

1. Selection
2. Testing
3. History
4. Maturation/Mortality
5. Instrumentation
6. Statistical Regression
7. Hawthorne Effect (compensation, rivalry, demoralization)



# T-Tests

1. Selection
2. Testing
3. History
4. Maturation/Mortality
5. Instrumentation
6. Statistical Regression
7. Hawthorne Effect (compensation, rivalry, demoralization)

## Questions

1. How do you compensate for the disadvantages of the experimental method?
2. Weigh the advantages and disadvantages of the pre-test-post-test method with the advantages and disadvantages of the post-test only. When might one be better than the other? Why? Should a researcher use the combination method instead, the Solomon Four, and, if so, why?
3. Discuss the relative uses, merits, and drawbacks of the three information-gathering techniques, content analysis, surveys, and experiment, and when do you think each should be the preferred methods in particular types of mass media research
4. What are some media questions that might be answered differently and more accurately through content analysis, surveys and experiments?
5. Compare the t-test with chi square and correlation in terms of the number of subjects needed, the real-life aspect of the information-gathering, and the types of information acquired with each.
6. Explain, by example, how chi square analysis can be used on data from t-tests, but a t-test cannot be used on chi-square data.

Words and concepts to know and remember

Random assignment

Control group

Test group

Treatment

Causality

Deviations squared

Sums of square

Advanced Media Research – 6  
Data Input, Analysis and Report Writing

STUDENTS: The three main things you should learn from this lecture are:

1. The basics of inputting data and statistical analysis
2. The sections of a social science report and the general content of each section

If you are writing a paper on research of a qualitative nature, you can structure it in somewhat the same way you structure a report on quantitative research. Omit the sections that obviously do not apply to qualitative research – like this first on one data input.

Unlike qualitative research, most quantitative research involves quite a few subjects, 600 people or more. It would take too long a time to add up each answer, and impossible to then break it down by subgroups, such as how many answers were by men, and how many by women. Most researchers use a particular statistical program. Statistical Package for the Social Sciences, simply known as SPSS, is the main data input and data analysis tool used by mass media and mass communication researchers.

The data sheet in the computer looks like a spread sheet, lots of rows and columns. Each column represents a different variable in your study, usually a different question. Each row represents a different person, article, or whatever population you are examining. It might look something like this for a content analysis:

Article	Magazine	Year	Length	Main Source	Climate Focus	Supporter/Denier
1	1	1980	15	1	2	1
2	1	1980	27	2	2	2
3	1	1990	11	1	3	1
4	2	1980	35	1	1	1
5	3	1990	18	3	1	1

Magazines: 1 = Time; 2 = Economist; 3= Newsweek; 4 = Other

Length in inches

Main Source: 1 = government spokesperson; 2 = industry representative; 3 = citizens group

Climate focus: 1 = weather change; 2 = possible costs; 3 = social impact;

Supporters/deniers: 1 = people who believe climate change is coming; 2 = people who don't believe

You first physically input into the computer the language you want to describe each variable, and the attributes of that variable. The computer will then print them when it provides you with the data analysis. Next you type in the numbers, one row at a time until you have no more articles, persons or other cases to include.

The data analysis section focuses on two aspects: knowing what you are looking for, and how to perform the analysis in the computer.

If a newspaper wanted to know how people rated the various content categories that it printed every day, it would be interested in the reader characteristics: gender, age, income, occupation, religion, political ideology and others. It would have to ask people about international news, regional news, local news, sports, business, fashion, and what other sections it had. It might possibly want to ask about crime stories, stories on UFOs, health issues and so on.

You probably will do one or two tests, the chi square and the t-test we have already discussed. There are more advanced statistics but those can come later as you progress. Remember these two points:

1. If both variables in the analysis are nominal, such as gender and favorite news section, you use the chi square test.
2. If you are comparing just two groups, and the dependent variable is ratio, you can use a t-test. For example, you want to compare the reading of page 1 time in minutes (ratio and independent variable) of men and women (nominal and independent variable); you use the t-test. A test called ANOVA can do three or more groups.

The computer will provide you with a series of numbers. The most important one is the probability level. If the probability level is .05 or less, such as .02, or .001, the differences you see in the data are statistically significant. At the .05 level, those results would occur by chance less than five times in 100. At the .001 level, the results would occur by chance less than one in 1,000 times.

If you are going to use a computer for data analysis, it is crucial that you first learn how to use the statistical package as well as understand some of the other numbers that are associated with the analysis.

The purpose of conducting all the research that you have gathered, coded, put in the computer and then analyzed is at least two-fold:

1. Your desire to learn something new in your areas of interest
2. The desire to share your results with other interested people

The second reason leads to the conclusion of a particular research project which is the writing of the research report. It is not hard to do, especially if you do it one section at a time. Scientific reports for academic journals generally follow a pattern, a structure, or a frame-work of writing. The stages of the frame-work are:

1. The abstract
2. The title
3. Introduction

4. Literature review
5. Methods
6. Results
7. Discussion
8. Appendix
9. References

**Abstract.** This is a short (usually fewer than 200 words), concise description of your study. State the question. Link it to major research area. Discuss methods. State results. STOP. No interpretation. Nothing else.

**Introduction.** No more than three or four pages. Give the background. Cite the primary theoretical area. Indicate the subjects of your research, the method for gathering the data and your hypotheses or research questions.

**Literature Review.** You need to discuss at least two areas. First, the situation, person, event, or problem you are studying. Then discuss the research or theoretical area. Cite sources in all areas of other than general knowledge. End this section with a short, terse, summary of the research which can be viewed as the foundation for your research.

**Method.** This is another short section. Describe how the variables were defined and the information gathered and then what statistical tests were used and why. A major purpose of this section is to allow other researchers to examine, and if they wish, repeat (replicate) the research to see if they obtain the same results.

**Results.** This is usually the longest section that is more pages, of your report. Begin with a summation of your findings. You usually start by indicating how many of

your hypotheses were statistically confirmed. Each confirmed hypothesis gets its own table which looks something like this:

TABLE 1

Coverage of the Taliban in Three Afghanistan Newspapers

News Coverage	Afghanistan Times Daily (n = 248)	Daily Annis (n = 197)	The Outlook Afghanistan (n = 321)
Favorable	70%	45%	10%
Neutral	20	50	60
Unfavorable	10	5	30

This is a stand-alone table. It does not need text to explain it. The title informs the reader of the topic. The number of newspaper articles is listed under the newspaper's name. I can calculate how many articles in the Afghanistan Times Daily were favorable simply by multiplying the 248 articles by 70 percent, which is .7. The answer is 173.6 or probably rounded to 174 articles. After you present this table, you describe briefly the major findings which are: The Afghanistan Times Daily is the most favorable and the least likely to be either neutral or unfavorable while The Outlook Afghanistan is the most neutral as well as the most unfavorable and least favorable. Generally you include a line under the table listing the chi square value and the probability of occurrence by chance. You do this for every hypothesis. Then you discuss any other findings you consider to be of interest

**Conclusion.** Now you get to give your opinion. You cannot interpret the results in the results section, but you can state here how they all fit together and what the larger picture is. But first you must restate the original problem and research



question usually just a paragraph. Then you need to link your findings to the literature review. If you conducted a study replicating a previous study, say one on bias in newspapers, you need to discuss how your results confirmed or denied your hypotheses. Did your research confirm and thus strengthen an area of knowledge, or does your research cast doubt by finding contrary results. When you have finished connecting your research to the literature, you can expound, that is pontificate or brag a bit about your findings and how the study is a contribution to the literature.

After you have discussed the value of your research, you need to discuss its limitations. No other country but Afghanistan was involved. All three newspapers are in Kabul. Then, based on your research, suggest what future research might cover.

**References.** List every source that you mention in your research article. It is in alphabetical order by name of author, or if no author by publication. It might look like this:

Brown, Amos, (2010). News Bias in Middle East Newspapers,  
*International Gazetteer*, Vol. 32 Spring, pp. 343-356.

**Appendix.** First, check with the publication. It may not want an appendix. But if it does, you can put additional tables and the specific method sheet you used to analyze the articles.


When you are finished, and your article is published, you probably still won't make any more money, but you will know that you have contributed new knowledge and help advance the field of mass communications.

TABLE 1  
 Coverage of the Taliban in Three Afghanistan Newspapers

News Coverage	Afghanistan Times Daily (n = 248)	Daily Annis (n = 197)	The Outlook Afghanistan (n = 321)
Favorable	70%	45%	10%
Neutral	20	50	60
Unfavorable	10	5	30

## Questions

1. What is the value of an abstract?
2. How does the literature review link help link your research to prior applied and theoretical research?
3. The inability to confirm, find support, for the study's hypotheses is not necessarily a study failure. Why is that so?
4. What steps need to be taken during a research project to ensure that the written report will be of value?



Words and concepts to know and remember

Abstract

Literature review

Data sheet

Appendix

References

