

Chapter Outlines for:

Frey, L., Botan, C., & Kreps, G. (1999). *Investigating communication: An introduction to research methods*. (2nd ed.) Boston: Allyn & Bacon.

Chapter 12: Inferring From Data – Estimation And Significance Testing

I. Introduction

- A. Often one needs to go beyond the usual descriptions of data acquired from particular people to infer conclusions about a much larger group of people.
- B. **Inferential statistics/inductive statistics**: The set of statistical procedures that allows a researcher to go beyond the group that has been measured, and make statements about the characteristics of a much larger group.
 - 1. Two purposes
 - a. **Estimation**: Used to generalize the results obtained from a sample to its parent population.
 - b. **Significance testing**: Examines how likely differences between groups and relationships between variables occur by chance.
- C. In this chapter, we first explore estimation and then focus on principles of significance testing that apply to both difference and relationship analysis.

II. Estimation

- A. The size of most populations usually makes it too difficult or impossible to conduct a census, so researchers must study a sample(s) drawn from a targeted population and use the results from that sample to describe the population.
- B. The ultimate purpose for many researchers is to generalize the results found for a sample back to its parent population.
- C. Researchers, therefore, often estimate population characteristics (called **parameters**) on the basis of characteristics found in a sample (called **statistics**).
 - 1. Estimation procedures are often referred to as parametric statistics. The statistics computed are called estimates and the formulas used to compute estimates are called estimators.
 - 2. Estimates of population parameters from sample statistics are possible as long as two assumptions are met:
 - a. The variable(s) of interest is assumed to be distributed “normally” in the population being studied.
 - b. A random sample has been selected from that population.
- D. Estimation procedures start with the assumption that scores on the variable(s) of interest in the population being studied are distributed in the shape of a symmetrical bell (see Figure 12.1); this type of distribution is called a **normal distribution, normal curve**, or, because of its symmetrical geometric shape, a **bell-shaped curve**.
 - 1. In a normal distribution, the center point is in the exact middle of the distribution and is the highest point of the curve.
 - a. The center point is the mean, median, and mode of the distribution.
 - 2. Because of the symmetrical shape of a normal distribution, researchers know the proportion of scores that falls within any particular area of the curve.
 - 3. In a normal distribution, 34.13% of scores fall in the area between the center point and +1SD (standard deviation), and 34.13% of scores fall between the center point and -1SD. Together, these two areas account for 68.26% of all the scores in a normal distribution.
 - 4. The curve continues to run on theoretically forever, a curve that runs on, getting ever closer to the line but never touching it is called asymptotic.
 - 5. The normal distribution, thus, tells researchers the relative probability of a score falling in

any given area of the curve.

6. Distributions not in the shape of a normal curve can be described using two characteristics.
 - a. **Kurtosis** refers to how pointed or flat is the shape of a distribution of scores.
 - i. **Mesokurtic**: The shape of the normal curve, with the center point being the highest peak and the two tails leveling off in a similar way.
 - ii. **Leptokurtic/peaked**: A curve that is tall and sharply pointed, with scores clustered around the middle.
 - iii. **Platykurtic/flat**: A curve that is flat and more spread out because scores do not cluster around the middle but are dispersed rather evenly across the distribution.
 - b. **Skewness**: The majority of scores are toward one end of a distribution and the curve tails off (becomes pointy, like a skewer) across the rest of the distribution, which creates an asymmetrical distribution
 - i. **Positively skewed distribution** occurs when the tail runs to the right side of the curve (Figure 12.2).
 - ii. **Negatively skewed distribution** occurs when the tail runs to the left side of the curve.

E. Use of random sampling: The second assumption that informs estimation procedures is the selection of a random sample from the population being studied.

1. The best assurance of a representative sample is random selection of population members, such that each population member has an equal chance of being selected for the sample.
2. The reason a random sample is representative of a population is that it is probable that if a population is normally distributed on a variable, then a random sample drawn from that population will also be normally distributed on that variable.
3. **Central Limit Theorem**: “For samples of a sufficiently large size, the real distribution of means is almost always approximately normal. The original variable can have any distribution. It doesn’t have to be bell-shaped in the least.”
 - a. Larger samples give more accurate results than do smaller samples.
4. Random samples are where it is at as far as inference is concerned; if researchers ever have to choose between a random and a nonrandom sample, they go with the random sample, even if they have to use a somewhat smaller sample size.

F. Inferring from a random sample to a population

1. **Sampling Distribution**: Occurs when treating the random sample means as a distribution of data; can also be formed of other values, such as percentages or median scores.
2. **Standard Error of the Mean (SEm)**: Tells researchers how much these random sample means are likely to differ from the grand mean of the sampling distribution.
3. Researchers assume that any characteristic normally distributed in a population will also be normally distributed in a random sample selected from that population.
4. **Confidence Level/Coefficient**: The range of scores needed to cover a given percentage of random sample means; the degree of assurance that an actual mean in a sampling distribution accurately represents the true population mean.
 - a. **Confidence Interval (CI)**: The range of scores of the random sample means (or other statistics) associated with a confidence level.
5. How confident a researcher is when estimating population parameters from a sampling distribution depends on the purpose of any particular estimation and the potential consequences of being wrong.
 - a. The 95% confidence level is an arbitrary level accepted by those in the social sciences.
6. Most researchers use a single random sample to stand for the large number of random

samples that should have been selected. That is, they select a random sample, calculate the mean and the standard error of the mean (or other appropriate statistics), and use the mean from that one sample to represent the sampling distribution mean and the standard error of the mean for that one sample to represent the sampling distribution's standard error of the mean.

a. Researchers select a random sample, calculate the mean and SD (or other statistics), move out $-1.96SD$ (standard deviation) and $+1.96SD$ away from the mean so that they are 95% confident, and use that range of scores as the confidence interval or "margin of error."

i. **Confidence Limits:** The actual values of the upper and lower limits of a confidence interval.

7. The size of the confidence interval is influenced by three features:

a. The variability that exists on the topic in the population being studied

b. The confidence level employed

c. The size of the random sample.

8. There isn't much that can be done about the variability that exists in a population; reducing the size of the confidence interval can be accomplished by decreasing the confidence level.

a. The best way to reduce the size of the confidence interval is to increase the size of the random sample selected.

i. **Law of Large Numbers:** The smaller the size of a random sample, the less accurately it approximates a true population parameter.

b. Most researchers rely on much smaller samples and accept larger confidence intervals (or errors) in their population estimates (Figure 12.3).

III. *Significance Testing:* "Significant" meaning "important; momentous."

A. Significance testing seeks to answer the question, "How is communication variable 'X' related to other variables?" by inferring relationships between variables within a population on the basis of relationships found in a sample selected from that population.

B. Unlike estimation, not all significance tests assume that the variable(s) being studied is normally distributed in the population of interest.

1. Significance tests that do assume this are called parametric statistics; significance tests that do not assume this are called **nonparametric statistics** (or **distribution-free statistics**).

C. Researchers often use a small, nonrandom sample, proceed to test for significant differences and relationships in that sample, and, then, in the discussion section of a journal article, generalize the results to the population

D. The Logic of Significance Testing

1. A one-tailed hypothesis about a relationship between variables predicts the direction of the suspected relationship as being either positive (as scores increase on one variable, so does the other variable) or negative (as scores increase on one variable, the other variable's score decreases).

2. Testing a null hypothesis: Researchers cannot prove a prediction by testing all possible instances in which it might apply—there might be millions; however, if researchers find a situation in which the idea does not work, they know the hypothesis is incorrect.

a. **Null hypothesis** (H_0): An implicit statement that underlies every hypothesis and predicts that there is no difference between the groups or no relationship between the variables being studied; represents chance occurrence—that the difference or relationship occurred simply by chance.

b. **Significance/hypothesis testing:** The process of analyzing quantitative data for the purpose of testing whether a null hypothesis is probably either correct or false.

i. On rare occasions, researchers may actually predict no significant difference or relationship, in which case the null hypothesis serves as a research hypothesis.

3. Rejecting a null hypothesis: The point at which the evidence suggests that a researcher can reject a null hypothesis of no difference or no relationship as being false and accept the most likely alternative explanation, the research hypothesis that predicts a difference or relationship.
 - a. The null hypothesis might be false because of one of the threats posed by researchers, the way the study was conducted, or because of the research participants studied.
 - b. If a researcher conducts a study in such a way as to confidently rule out or sufficiently minimize these validity threats, then the most likely explanation for the null hypothesis being false is the suspected difference between the groups or the relationship between the variables predicted by the research hypothesis.
 - c. Conditional probability: Where a preceding occurrence influences a subsequent occurrence.
 - d. Gambler's fallacy: Mistakenly thinking that independent events are conditional.

4. Deciding on the probability level

- a. **Significance level:** The probability level (p value) researchers set for rejecting a null hypothesis; establishes a point at which a researcher is confident enough of a statistical difference or relationship to reject a null hypothesis.
- b. **Decision rule:** The specific significance level set by a researcher for rejecting a null hypothesis prior to conducting a study and analyzing the data.
- c. In significance testing, the 95% confidence level is referred to as the .05 significance level, which references the probability of the difference or relationship being significant. Researchers set the significance level quite high to minimize reporting a significant difference or relationship that occurs due to chance.
- d. Medical researchers typically make it much more difficult to reject the null hypothesis by setting the significance level much higher, such as .0001 (see additive error).
- e. Occasionally, social science researchers use a less stringent significance level, such as .10, which means they are 90% confident of rejecting the null hypothesis; (between .10 and .05 is sometimes considered a trend, or, if very close to .05, marginally significant).

E. The Practice of Significance Testing: Three steps:

1. **Posing a research question or hypothesis and a null hypothesis:** A hypothesis is posed when a researcher has enough evidence from theory, relevant literature, or sometimes logic and/or observations from everyday life to make a prediction.
2. **Conducting the study:** To investigate the suspected difference between groups or relationship between variables. the population is identified and a random sample size is selected from it.
3. Testing the null hypothesis: (Entails three items:
 - a. Setting the significance level: Most often set at .05
 - b. **Computing the calculated value:** There are many types of significance-testing procedures, each resulting in a final number, a numerical value called the calculated value. This calculated value is then assessed in terms of the probability of its occurrence.
 - c. **Comparing the calculated value to the critical value** needed to reject the null hypothesis: if the calculated value has a 5% or less probability of occurrence by chance, the researcher can reject the null hypothesis; if more that 5%, must accept.
 - i. Researchers can assess the probability of obtaining the calculated value they found for the data analyzed from any particular study by constructing a table for each statistical significance test.
 - ii. This is done by first locating the critical value that corresponds to the .05 significance level needed to reject the hypothesis.

- iii. Running down the side of the table are what are called **degrees of freedom**, the number of scores that are “free to vary.”
- iv. If the calculated value does not reach the critical value it falls into the **region of acceptance** and the null hypothesis of no difference or no relationship is accepted; if the calculated value reaches or exceeds the critical value, it falls into the **region of rejection/critical region**.
- v. When degrees of freedom are tied to sample size, the larger the degrees of freedom, the smaller the critical value needed for rejecting the null hypothesis.
- vi. The critical value needed for accepting/rejecting a null hypothesis also depends on whether a *one-tailed hypothesis*, *two-tailed hypothesis*, or *research question* is posed (Figure 12.4).
- vii. **NS** is often used to signify a result that is not [statistically] significant.

F. Type I and Type II Error

- 1. Two kinds of decision errors researchers might make when employing significance testing:
 - a. **Type I error (alpha error)** occurs when researchers reject a null hypothesis and accept a research hypothesis when, in fact, the null hypothesis is probably true and should have been accepted.
 - b. **Type II error (beta error)** is exactly the opposite, and occurs when researchers accept a null hypothesis when, in fact, it is probably false and thereby, reject a sound research hypothesis.
 - c. As Figure 12.5 reveals, there are two correct and two incorrect decisions that can be made concerning significance testing.
 - d. The chance of committing a Type I error is easy to calculate; it is equal to the significance level employed.
 - i. If multiple tests are made, sum the significance levels (e.g., $.05 + .05 + .05 = .15$); and these are called **familywise error** or **experimenter error**.
 - e. Type II error is far too complex to be explained here as there are numerous factors that affect its calculation.

G. Statistical Power

- 1. This factor affects the strength of inferences about significant differences and relationships.
 - a. **Statistical power** estimates the probability of rejecting a null hypothesis that is, in fact, probably false and should be rejected.
 - b. Statistical power calculates, before conducting a study, the number of participants needed to have a reasonable chance of rejecting the null hypothesis.
 - c. Statistical power is equal to 1 minus the probability of committing a Type II error (failing to reject a false null hypothesis).
 - i. The range of statistical power, then, is from a minimum of 0 to a maximum of 1.0
 - ii. Authorities generally agree that statistical power should be set at .80.

IV. Conclusion

- A. We explored how researchers infer from a relatively small sample to its parent population. Inferring in this way certainly is not limited to research; it's a common feature of everyday life.