Estimation

Population parameter or sample estimates of population parameters

- The average income of all adult Americans.
- The percentage who favored increasing affirmative action programs in the Gallup survey.
- The percentage of all Americans who favor increasing affirmative action programs.

Estimation

- When the GSS or Gallup or NBC takes a poll, they use the results they get to estimate the percentage of American adults who would agree with their results.
- Estimation is a process whereby we select a random sample from a population and use a sample statistic to estimate a population parameter

Why do we Estimate?

- It is much too expensive and time consuming and in some cases impossible to find information on a population.
- We can learn a lot about a population by randomly selecting a sample from that population and obtaining an estimate of the population parameter.
- With this knowledge we learn about ourselves.

Point and Interval Estimation

- There are two types of estimation.
  - Point estimation
  - Interval estimation
- Point estimates are sample statistics used to estimate the exact value of a population parameter.
- In Interval estimates (confidence intervals), we use a range of values within which the population parameter may fall.

Estimation

- Confidence Interval
  - A range of values defined by the confidence level within which the population parameter is estimated to fall
- Confidence Level
  - The likelihood, expressed as a percentage or a probability, that a specified interval will contain the population parameter

Confidence level

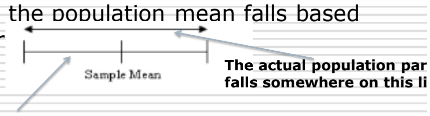
- When we use confidence intervals to estimate population parameters we can also evaluate the accuracy of this estimate by assessing the likelihood that any given interval will contain the mean.
- We express it as a percentage or a probability and we call it the confidence level

Confidence Intervals

- If we wanted to be 95% confident, all random sample means would fall within +\- 1.96 standard error (Z score).
- If we wanted to be 99% confident, all random sample means would fall within +\- 2.58 standard errors (Z scores).

Confidence Interval

- We want to construct an estimate of where the population mean falls based on our



Confidence Interval Formula

- The general formula for constructing a confidence interval (CI) for any level is:
- To calculate a confidence interval , we take the sample mean and add to or subtract from it the product of a Z value and the standard error.

$$CI = \bar{Y} \pm Z(\sigma_{\bar{y}})$$

Confidence Intervals

- The Z score we choose depends on the desired confidence level.
- For example, to obtain a 95% CI we would choose a Z of 1.96 because if we go to Appendix B it shows that 95% of the area under the curve is included between ± 1.96.
- You as the researcher get to choose how confident you would like to be.

Determining the Confidence Interval

- To determine the confidence interval for means, follow these steps:
  - Calculate the standard error of the mean.
  - Decide on the level of confidence, and find the corresponding Z value.
  - Calculate the confidence interval.
  - Interpret the results.

The Standard Error

- Standard Error of the Mean
  - The standard deviation of a sampling distribution

Standard Error $= \sigma_{\bar{y}} = \dfrac{\sigma_y}{\sqrt{N}}$

Estimating Standard Errors

- Because the standard error is generally not known, we work with the estimated standard error

$$s_{\bar{Y}} = \frac{s_Y}{\sqrt{N}}$$

Now let us try!

- Calculating the standard error of the mean.
- We are looking at the average age of 500 COD students.
- Lets say that the standard deviation for our population of COD students is 2.5
- We can now calculate the standard error of the mean.

Decide on a Level of Confidence

- We have decided to use a 80% confidence level.
- Find the corresponding Z value.

Calculate the Confidence Interval

- The confidence interval is calculated by adding the subtracting from the observed sample mean the product of the standard error and Z.

$$CI = \bar{Y} \pm Z(\sigma_{\bar{y}})$$

Interpreting the Results

- We can be 80% confident that the actual mean age of all COD students is not less than 20.66 years old and not greater than 20.94 years old.
- Remember, we can never be sure whether the population mean is actually contained within the confidence interval.

Reducing Risk

- We can reduce the risk of being incorrect by increasing the level of confidence.
- When we use the highest level of confidence, 99% confidence, there is only a 1% risk that we are wrong that the specified interval does not contain the true population mean.

Sample Size and confidence intervals

- By increasing the sample size, researchers increases the precision of their estimate. Meaning their confidence interval is smaller.
- Larger samples result in smaller standard errors and therefore, in sampling distributions that are more clustered around the population mean.
- A Closer Look 8.2 on page 249.

Confidence Intervals for Proportions and Percentages

- The procedures for estimating proportions and percentages are identical.
- The sampling distribution of proportions underlies the estimation of population proportions from the sample proportions.
- This is the same theory that we used for the sampling distribution of the mean.

Procedures for Estimating Proportions

- The general formula for constructing confidence intervals for proportion for any level of confidence is:
- $CI = p \pm Z (S_p)$
  - CI = confidence interval
  - P = the observed sample proportion
  - Z = the Z corresponding to the CI
  - S_p = the estimated standard error of proportions.

Steps for Estimating Proportions

- Calculate the standard error of the proportion.
- Decide on the level of confidence, and find the corresponding Z value.
- Calculate the confidence interval.
- Interpret the results.

Let's give it a try!

- We are going to use the data from the textbook.
- The textbook looked at the result of the CBS News Poll on satisfaction with the election outcome.

Calculating the Estimated Standard Error of the Proportion

- For our study, the observed sample proportion (p) is .50 (50%) with a sample (N) of 1,048.
- Find the Estimated standard error of the proportion based on the information given.

Formula for estimated standard error of the proportion is

$$S_p = \sqrt{\frac{(p)(1-p)}{n}}$$

Decide on the Level of Confidence

- We have decided to use a 90% confidence level.
- Find the corresponding Z value.

Calculate the Confidence Interval

- The confidence interval is calculated by adding the subtracting from the observed sample proportion the product of the standard error and Z.

$$CI = p \pm Z(s_p)$$

Interpreting the Results

- We are 90% confident that the true population proportion is somewhere between .475 and .525.
- We could also express the result in percentages and say that we are 90% confident that the true population percentage of satisfaction with the election outcome in between 47.5% and 52.5%.