Chapter 8: Notes

Evan Hunt

November 28, 2019

Section 8.1

Point Estimate

A point estimate of a parameter is the value of a statistic used to estimate the parameter. Roughly speaking, a point estimate of a parameter is our best guess for the value of the unknown parameter based on sample data. For instance, a sample mean is a point estimate of a population mean, and a sample standard deviation is a point estimate of a population standard deviation.

Confidence-Interval Estimate

A confidence-interval estimate for a parameter provides a range of numbers along with a percentage confidence that the parameter lies in that range.

Confidence interval (CI): An interval of numbers obtained from a point estimate of a parameter.

Confidence level: The confidence we have that the parameter lies in the confidence interval (i.e., that the confidence interval contains the parameter).

Confidence-interval estimate: The confidence level and confidence interval.

Section 8.2

One-Mean z-Interval Procedure

Purpose To find a confidence interval for a population mean, μ.

Assumptions

1. Simple random sample
2. Normal population or large sample size
3. σ is known

Step 1 For a confidence level of $1 - \alpha$, use a z-table to find $z_{\alpha/2}$.

Step 2 The confidence interval for μ is from

$$\bar{x} - z_{\alpha/2} * \frac{\sigma}{\sqrt{n}} to \bar{x} + z_{\alpha/2} * \frac{\sigma}{\sqrt{n}}$$

where $z_{\alpha/2}$ is found in Step 1, n is the sample size, and $\bar{x}$ is computed from the sample data.

Step 3 Interpret the confidence interval.

Note: The confidence interval is exact for normal populations and is approximately correct for large samples from nonnormal populations.

When to Use the One-Mean z-Interval Procedure

- For small samples, less than 15 roughly, the z-interval procedure should be used only when the variable under consideration is normally distributed or very close to being so.
- For samples of moderate size, between 15 and 30 roughly, the z-interval procedure can be used unless the data contain outliers or the variable under consideration is far from being normally distributed.
- For large samples, around 30 or more, the z-interval procedure can be used without restriction, generally. However, if outliers are present and their removal is not justified, you should compare the confidence intervals obtained with and without the outliers to see what effect the outliers have. If the effect is substantial, use a different procedure or take another sample, if possible.
- If outliers are present but their removal is justified and results in a data set for which the z-interval procedure is approximate (as previously stated), the procedure can be used.

A Fundamental Principle of Data Analysis

Before performing a statistical-inference procedure, examine the sample data. If any of the conditions required for using the procedure appear to be violated, do not apply the procedure. Instead use a different, more appropriate procedure, if one exists.

Margin of Error for the Estimate of μ

The margin of error for the estimate of μ is $z_{\alpha/2} * \sigma/\sqrt{n}$, which is denoted by the letter E. Thus,

$$E = z_{\alpha/2} * \sigma/\sqrt{n}.$$

The margin of error for the estimate of population mean indicates the accuracy with which a sample mean estimates the unknown population mean.

Confidence and Accuracy

For a fixed sample size, decreasing the confidence level decreases the margin of error and, hence, improves the accuracy of a confidence-interval estimate.

Sample Size and Accuracy

For a fixed confidence level, increasing the sample size decreases the margin of error and, hence, improves the accuracy of a confidence-interval estimate.

Sample Size for Estimating μ

The sample size required for a $(1-\alpha)$-level confidence interval for μ with a specified margin of error, E, is given by the formula

$$n = (\frac{z_{\alpha/2} * \sigma}{E})^2,$$

rounded up to the nearest whole number.

Section 8.3

Studentized Version of the Sample Mean

Suppose that a variable x of a population is normally distributed with mean μ. Then, for samples of size n, the variable

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

has the t-distribution with $n-1$ degrees of freedom (df). For a normally distributed variable, the studentized version of the sample mean has the t-distribution with degrees of freedom 1 less that the sample size.

Basic Properties of t-Curves

Property 1: The total area under a t-curve equals 1.

Property 2: A t-curve extends indefinitely in both directions, approaching, but never touching, the horizontal axis as it does so.

Property 3: A t-curve is symmetric about 0.

Property 4: As the number of degrees of freedom becomes larger, t-curves look increasingly like the standard normal curve.

One-Mean t-Interval Procedure

Purpose To find a confidence interval for population mean, μ

Assumptions
1. Simple random sample
2. Normal population or large sample
3. σ unknown

Step 1 For a confidence level of $(1 - \alpha)$, use a t-table to find $t_{\alpha/2}$ with $df = n - 1$, where n is the sample size.

Step 2 The confidence interval for μ is from

$$\bar{x} - t_{\alpha/2} * \frac{s}{\sqrt{n}} to \bar{x} + t_{\alpha/2} * \frac{s}{\sqrt{n}},$$

where $t_{\alpha/2}$ is found in Step 1 and $\bar{x}$ and s are computed from sample data.

Step 3 Interpret the confidence interval.

Note: The confidence interval is exact for normal populations and is approximately correct for large samples from nonnormal populations.