7 Confidence Intervals for One Population Parameter

Introduction

Chapter 6 discussed how to derive a point estimate. Since each sample can potentially lead to a different point estimate for a given population parameter, it is important to be able to describe the amount of randomness associated with a point estimate. Confidence Intervals are one method to quantify the uncertainty that comes with a point estimate.

7.1: Basic Properties of Confidence Intervals

In Chapter 6, we examined how to find a point estimate for an unknown population parameter. When reporting point estimates, it is important to note the amount of uncertainty associated with the point estimate. This is especially key when we want to do Statistical Inference.

Def: Statistical Inference is the process of inferring properties of a population by use of sample data.

We will examine two methods of Statistical Inference:
(i) Confidence Intervals
(ii) Hypothesis Tests

Note: Confidence Intervals (CI) can be one-sided or two-sided.

Note: The variance (σ^2) of the population may be known or unknown. We will begin by assuming that we know σ^2 (and hence also σ).

General Form of a Confidence Interval (Two-Sided)

Example:

Note: α is the probability that a confidence interval does not contain the true population parameter given that the null hypothesis is true. It is also the probability of a Type I error.

Example:

Example:

Deriving a Confidence Interval

Example: Suppose we want to create a two-sided 95% confidence interval for μ.

Writing and Interpreting a Confidence Interval

Here's how to write confidence intervals:

Here's how to interpret confidence intervals:

Sample Size Necessary for a CI

Suppose we want to know how big of sample we need for a desired width. We can use the following formula (which can be derived from previous CI calculations).

$$\boxed{n = \left(z_{\alpha/2} \cdot \frac{\sigma}{E}\right)^2}$$

In the above equation, E is the desired margin of error. The width of the interval, w, is twice the margin of error ($w = 2 \cdot E$).

<u>Note:</u> Since a sample size must be a whole number, round the result of the previous equation to the next highest integer. For example, if n is calculated to be 10.1, use n=11.

<u>Example:</u> Suppose we want to create a 90% confidence interval for μ and we know that the population standard deviation is $\sigma = 3$. Our study requires a margin of error of at most 2. Calculate the required sample size.

7.2: Large-Sample Confidence Intervals for a Population Mean and Proportion

The Central Limit Theorem states that $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ has an approximately standard normal distribution.

We can use this result to produce a confidence interval since $P(-z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}) \approx 1 - \alpha$.

Theorem: If our random sample comes from a normal distribution and σ is known, then a $100(1-\alpha)\%$ confidence interval for μ is given by:

$$\boxed{\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}}$$

Our textbook notes that if n is large, then we don't require normality. We will come to this case in the next section. This type of confidence interval can also be called a large-sample confidence interval.

Example: A random sample of 16 CPU chips has a total of 342 defects. Suppose we know the population standard deviation is 4.52 and also that the number of defects/chip is approximately normally distributed.

(a) Calculate a point estimate for μ, the true mean number of defects/chip.

(b) Calculate a 95% CI for μ.

Example: Suppose that GPAs are approximately normally distributed with population variance $\sigma^2 = 1.2$. If we take a random sample of 12 students and find a sample mean of 2.87. Construct a 90% confidence interval for μ.

We can also create confidence intervals for proportions.

Theorem: Provided that n is sufficiently large, a $100(1-\alpha)\%$ confidence interval for a proportion p is given by:

$$\boxed{\bar{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}}$$

Example: Suppose a random sample of 80 students finds 45 of them spend more than 20 hours/week on homework. Construct a 95% confidence interval for the true proportion of students that spend more than 20 hours/week on homework.

7.3: Intervals Based on a Normal Population Distribution

In this section, we will assume that both μ and σ are unknown.

Recall that the Central Limit Theorem stated that $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ has an approximately standard normal distribution. If σ is unknown, then $t = (\bar{X} - \mu)/(s/\sqrt{n})$ does <u>not</u> have a normal distribution.

t-distribution

<u>Notation:</u> Let t_ν be a t-distribution with ν degrees of freedom.

<u>Properties of the t-distribution</u>

(1) Each t_ν curve is bell-shaped and centered at zero.
(2) Each t_ν curve has more spread than the standard normal curve.
(3) As ν increases, the spread of the corresponding t_ν curve decreases.
(4) As $\nu \to \infty$, the sequence of t_ν curves approaches the standard normal curve.

<u>Example:</u> Find t^* such that the area under the curve to the right of t^* is 0.025 for $df = 10$.

<u>Example:</u> Find t^* such that the area under the curve to the left of t^* is 0.05 for $df = 14$.

<u>Example:</u> Find the two values, t_1 and t_2, such that the area under the curve is the middle 95% for $df = 24$.

<u>Example:</u> Suppose $t^* = 2.1$ and $df = 9$. Calculate $P(t(9) \geq t^*)$.

Theorem: If $\bar{X}$ is the mean of a random sample of size n from an (approximately) normal distribution with mean μ, then $t = (\bar{X} - \mu)/(s/\sqrt{n})$ follows a t-distribution with (n-1) degrees of freedom.

Theorem: Let $\bar{x}$ and s be the sample mean and sample standard deviation from a random sample from a (approximately) normal population with mean μ. Then a $100(1-\alpha)\%$ confidence interval for μ is given by:

$$\boxed{\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}, \text{ df=n-1}}$$

Example: Suppose a random sample of 30 Hummers have a mean of 9.8 miles/gallon (MPG) and a sample standard deviation of 2.2. Compute a 95% confidence interval for μ.

Example: Suppose a random sample of 17 Priuses have a mean of 52.3 MPG and variance of 14.9. Construct a 99% confidence interval for μ.

One-Sided Confidence Intervals

Example: Suppose a random sample of 30 Hummers have a mean of 9.8 miles/gallon (MPG) and a sample standard deviation of 2.2. Find the upper bound for a 95% lower confidence interval for μ.

Example: Suppose a random sample of 17 Priuses have a mean of 52.3 MPG and variance of 14.9. Find the lower bound for a 99% upper confidence interval for μ