Chapter 8 - Lecture 1
Basic Properties of Confidence Intervals

Yuan Huang

February 6, 2013

Overview of Chapter 8

It mainly discusses confidence intervals for a population mean for the following situations:

1. $N(\mu, \sigma)$ with σ known (8.1)
2. General sample with σ known/unknown under large sample size (8.2)
3. $N(\mu, \sigma)$ with σ unknown under small sample size (8.3)

Let n denote the sample size.

	σ known	σ unknown
Normal	8.1 n does not matter	8.3 under n is small/large
General	8.2 under n is large	8.2 under n is large

8.1 Case 1: $N(\mu, \sigma)$ with σ known

- From chapter 7, we know $\bar{X}$ is an estimator for μ.
- To gain the precision, we can report the stand error.
- What's the confidence level? – Use **confidence intervals (CI)** for the parameters.
- CI gives an estimated range of values which is likely to include an unknown population parameter

Definitions

A confidence interval is always calculated by first selecting a *confidence level*, which is a measure of the degree of reliability of the interval.

- The probability we allow ourselves to be wrong when we are estimating a parameter with a confidence interval, is called **significance/critical level** and is denoted with α.
- So if $\alpha = 0.05$ then $1 - 0.05 = 0.95$, so we call our confidence interval a **95% confidence interval**

General strategy to construct CI for θ

- Parameter θ
- Random sample $X_1, \ldots, X_n$
- Need a random variable Y such that
  1. Its functional form depends on $X_1, \ldots, X_n$ and θ. Hence we can denote $Y = h(X_1, \ldots, X_n, \theta)$.
  2. Its distribution does not depend on θ.

General strategy to construct CI for θ

If you can find such Y, then based on the distribution of Y, you should be able to write down

$$P\left(a < h(X_1, \ldots, X_n, \theta) < b\right) = 1 - \alpha$$

Both a and b do not depend on θ. By transformation, you can derive :

$$P\left(l(X_1, \ldots, X_n) < \theta < u(X_1, \ldots, X_n)\right) = 1 - \alpha$$

Here, $l(X_1, \ldots, X_n) < \theta < u(X_1, \ldots, X_n)$ is a random interval. We can plug in the observations to get the $1 - \alpha$ confidence interval

$$l(x_1, \ldots, x_n) < \theta < u(x_1, \ldots, x_n)$$

.

Now let's look at how to apply this general strategy

Assume that you have a random sample $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

- μ unknown .
- σ known.

Goal: To find confidence intervals for the mean μ.

Construct 95% confidence interval

- From Chapter 6, we know that $\bar{X} \sim N(\mu, (\sigma/\sqrt{n})^2)$

$$\Rightarrow Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

- We also know that the area under the standard normal curve between -1.96 and 1.96 is 0.95

$$\Rightarrow \quad P(-1.96 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 1.96) = 0.95$$

- Manipulate to get the equivalent form $l < \mu < u$.

$$P\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

- Until now, we have a random interval $\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right)$
- After observing $X_1 = x_1, \ldots, X_n = x_n$, we compute the observed sample mean $\bar{x}$ and then substitute into the above random interval in place of $\bar{X}$, the resulting fixed interval is called **95% confidence interval for** μ.

Definition

For $X_1, \ldots, X_n$ random sample from $N(\mu, \sigma)$, with σ known, the 95% confidence interval for μ can be expressed as

$$\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right) \text{ is a 95\% CI for } \mu$$

or as $\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}$ with 95% confidence.

Example 8.2 Assume the preferred height is normally distributed as $N(\mu, \sigma = 2)$. A random sample of size $n = 31$ is collected. The sample mean is $\bar{x} = 80$. What is the resulting 95% confidence interval ?

$$\begin{aligned}\bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}} &= 80 \pm 1.96\frac{2}{\sqrt{31}} \\ &= 80 \pm .7 \\ &= (79.3, 80.7)\end{aligned}$$

Interpretation

How do we interpret a 95% Confidence Interval?

Interpretation

How do we interpret a 95% Confidence Interval?

Attention!!

Interpretation

How do we interpret a 95% Confidence Interval?

Attention!!

For random interval:
With probability 0.95, the random interval $(\bar{X}_n - 1.96\sigma/\sqrt{n}, \bar{X}_n + 1.96\sigma/\sqrt{n})$ will cover the true value μ.

Interpretation

How do we interpret a 95% Confidence Interval?

Attention!!

For random interval:
With probability 0.95, the random interval $(\bar{X}_n - 1.96\sigma/\sqrt{n}, \bar{X}_n + 1.96\sigma/\sqrt{n})$ will cover the true value μ.

For confidence interval:
If the experiment is taken (the random sample is drawn) independently over and over again, about 95% of the intervals derived from this formula will cover the true value μ.

Interpretation

How do we interpret a 95% Confidence Interval?

Attention!!

For random interval:
With probability 0.95, the random interval $(\bar{X}_n - 1.96\sigma/\sqrt{n}, \bar{X}_n + 1.96\sigma/\sqrt{n})$ will cover the true value μ.

For confidence interval:
If the experiment is taken (the random sample is drawn) independently over and over again, about 95% of the intervals derived from this formula will cover the true value μ.

Misspecified interpretation: with probability 0.95, μ will take value in the confidence interval.

Confidence interval of other confidence levels

A $\mathbf{100(1-\alpha)}$% **confidence interval** for the mean μ of a normal population when the value of σ is known is given by

$$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

where $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha$

Confidence interval of other confidence levels

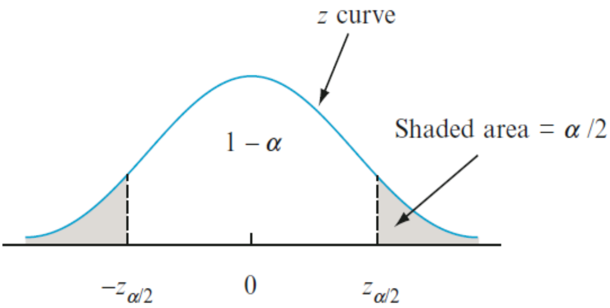

Figure 8.4 $P(-z_{a/2} \leq Z \leq z_{a/2}) = 1-\alpha$

Figure: Modern Mathematical Statistics with Applications, 2rd, P387

Example 8.2 (revised) Assume the preferred height is normally distributed as $N(\mu, \sigma = 2)$. A random sample of size $n = 31$ is collected. The sample mean is $\bar{x} = 80$. What is the resulting 90% confidence interval ?

For 90% confidence interval, we need to find $z_{0.05}$. By checking table or use R, we can get $z_{0.05} = 1.645$.

$$\begin{aligned} \bar{x} \pm 1.645 \frac{\sigma}{\sqrt{n}} &= 80 \pm 1.645 \frac{2}{\sqrt{31}} \\ &= 80 \pm .59 \\ &= (79.41, 80.59) \end{aligned}$$

- The width of confidence interval

$$w = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- Larger confidence levels means more reliable. However, for a given sample size n, the width w is increasing with $z_{\alpha/2}$.
- Therefore, the gain in reliability entails a loss in precision.
- Solution: calculate the sample size.
- Given the width w_0, the sample size that ensures w_0 is

$$n = \left(2z_{\alpha/2}\frac{\sigma}{w_0}\right)^2$$

Example: We want to get a 95% confidence interval for the mean μ of a random sample coming from $N(\mu, \mathrm{var} = 4)$. Find the sample size that we need in order to get interval with width 4.

$$\begin{aligned} n &= \left(2z_{\alpha/2}\frac{\sigma}{w}\right)^2 \\ &= \left(2z_{0.025}\frac{2}{4}\right)^2 \\ &= z_{0.025}^2 \\ &= 1.96^2 \\ &= 3.84 \end{aligned}$$