ENGINEERING STATISTICS
(Lectures)

University of Technology,
Building and Construction Engineering Department
(Undergraduate study)

ESTIMATION THEORY

Dr. Maan S. Hassan
Lecturer: Azhar H. Mahdi

2010 – 2011
1.0 CONFIDENCE INTERVAL ON THE MEAN OF A NORMAL DISTRIBUTION, VARIANCE KNOWN

Suppose that $X_1, X_2, \ldots \ldots, X_n$ is a random sample from a normal distribution with unknown mean $\mu$ and known variance $\sigma^2$. We may standardize $\bar{X}$ as follow:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

A confidence interval estimate for $\mu$ is an interval of the form

$$l \leq \mu \leq u$$

where the end-points $l$ and $u$ are computed from the sample data. Because different samples will produce different values of $l$ and $u$, these end-points are values of random variables $L$ and $U$, respectively.

Suppose that we can determine values of $L$ and $U$ such that the following probability statement is true:

$$P\{L \leq \mu \leq U\} = 1 - \alpha$$

And

$$P\left\{-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right\} = 1 - \alpha$$

Or

$$P\left\{\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right\} = 1 - \alpha$$
Definition

If \( \bar{x} \) is the sample mean of a random sample of size \( n \) from a normal population with known variance \( \sigma^2 \), a \( 100(1 - \alpha)\% \) CI on \( \mu \) is given by

\[
\bar{x} - z_{\alpha/2} \sigma / \sqrt{n} \leq \mu \leq \bar{x} + z_{\alpha/2} \sigma / \sqrt{n}
\]

where \( z_{\alpha/2} \) is the upper \( 100\alpha/2 \) percentage point of the standard normal distribution.

EXAMPLE 1:

ASTM Standard E23 defines standard test methods for notched bar impact testing of metallic materials. The (CVN) technique measures impact energy and is often used to determine whether or not a material experiences a ductile-to-brittle transition with decreasing temperature. Ten measurements of impact energy (\( J \)) on specimens of A238 steel cut at 60\(^\circ\)C are as follows:

64.1, 64.7, 64.5, 64.6, 64.5, 64.3, 64.6, 64.8, 64.2, and 64.3

Assume that impact energy is normally distributed with \( \sigma = 1J \).

We want to find a 95\% CI for \( \mu \), the mean impact energy. The required quantities are
\[ z_{\alpha/2} = z_{0.025} = 1.96 \]
\[ n=10, \sigma=1, \text{ and } \bar{x}=64.46. \]

The resulting 95% CI is found from the Equation above as follows:

\[
\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}
\]

\[
64.46 - 1.96 \frac{1}{\sqrt{10}} \leq \mu \leq 64.46 + 1.96 \frac{1}{\sqrt{10}}
\]

\[
63.84 \leq \mu \leq 65.08
\]

2.0 Confidence Level and Precision of Estimation

The length of the 95% confidence interval is:

\[ 2(1.96\sigma/\sqrt{n}) = 3.92\sigma/\sqrt{n} \]

Whereas the length of the 99% CI is:

\[ 2(2.58\sigma/\sqrt{n}) = 5.16\sigma/\sqrt{n} \]

Thus, the 99% CI is longer than the 95% CI. This is why we have a higher level of confidence in the 99% confidence interval.

Generally, for a fixed sample size \( n \) and standard deviation \( \sigma \), the higher the confidence level, the longer the resulting CI.

The length of a confidence interval is a measure of the precision of estimation.
3.0 Choice of Sample Size

Definition

\[
n = \left( \frac{\frac{z_{\alpha/2} \sigma}{E}}{E} \right)^2
\]

If the right-hand side of the Equation above is not an integer, it must be rounded up. This will ensure that the level of confidence does not fall below 100(1-\(\alpha\)) %. Notice that 2\(E\) is the length of the resulting confidence interval.

EXAMPLE 2:

To illustrate the use of this procedure, consider the CVN test described in Example-1, and suppose that we wanted to determine how many specimens must be tested to ensure that the 95% CI on \(A_{238}\) steel cut at 60°C has a length of at most 1.0\(J\). Since the bound on error in estimation \(E\) is one-half of the length of the CI, to determine \(n\) with

\[\begin{align*}
E &= 0.5, \\
\sigma &= 1, \text{ and} \\
z_{\alpha/2} &= 0.025.
\end{align*}\]

The required sample size is 16

\[n = \left( \frac{\frac{z_{\alpha/2} \sigma}{E}}{E} \right)^2 = \left[ \frac{1.96 \cdot 1}{0.5} \right]^2 = 15.37\]

and because \(n\) must be an integer, the required sample size is \(n = 16\).
• As the desired length of the interval $2E$ decreases, the required sample size $n$ increases for a fixed value of $\sigma$ and specified confidence.

• As $\sigma$ increases, the required sample size $n$ increases for a fixed desired length $2E$ and specified confidence.

• As the level of confidence increases, the required sample size $n$ increases for fixed desired length $2E$ and standard deviation $\sigma$.

EXERCISES:

8.1. For a normal population with known variance $\sigma^2$, answer the following questions:
(a) What is the confidence level for the interval $\bar{x} - 2.14\sigma/\sqrt{n} \leq \mu \leq \bar{x} + 2.14\sigma/\sqrt{n}$?
(b) What is the confidence level for the interval $\bar{x} - 2.49\sigma/\sqrt{n} \leq \mu \leq \bar{x} + 2.49\sigma/\sqrt{n}$?
(c) What is the confidence level for the interval $\bar{x} - 1.85\sigma/\sqrt{n} \leq \mu \leq \bar{x} + 1.85\sigma/\sqrt{n}$?

8.2. For a normal population with known variance $\sigma^2$:
(a) What value of $z_{0.9}$ in Equation 8-7 gives 98% confidence?
(b) What value of $z_{0.95}$ in Equation 8-7 gives 80% confidence?
(c) What value of $z_{0.75}$ in Equation 8-7 gives 75% confidence?

8.3. Consider the one-sided confidence interval expressions, Equations 8-9 and 8-10.
(a) What value of $z_{0.01}$ would result in a 90% CI?
(b) What value of $z_{0.05}$ would result in a 95% CI?
(c) What value of $z_{0.025}$ would result in a 99% CI?

8.4. A confidence interval estimate is desired for the gain in a circuit on a semiconductor device. Assume that gain is normally distributed with standard deviation $\sigma = 20$.
(a) Find a 95% CI for $\mu$ when $n = 10$ and $\bar{x} = 1000$.
(b) Find a 95% CI for $\mu$ when $n = 25$ and $\bar{x} = 1000$.
(c) Find a 99% CI for $\mu$ when $n = 10$ and $\bar{x} = 1000$.
(d) Find a 99% CI for $\mu$ when $n = 25$ and $\bar{x} = 1000$.

8.5. Consider the gain estimation problem in Exercise 8.4. How large must $n$ be if the length of the 95% CI is to be 40?

8.6. Following are two confidence interval estimates of the mean $\mu$ of the cycles to failure of an automotive door latch mechanism (the test was conducted at an elevated stress level to accelerate the failure).

$$3124.9 \leq \mu \leq 3215.7 \quad 3110.5 \leq \mu \leq 3230.1$$

(a) What is the value of the sample mean cycles to failure?
(b) The confidence level for one of these CIs is 95% and the confidence level for the other is 99%. Both CIs are calculated from the same sample data. Which is the 95% CI? Explain why.

8.7. $n = 100$ random samples of water from a fresh water lake were taken and the calcium concentration (milligrams per liter) measured. A 95% CI on the mean calcium concentration is $0.49 \leq \mu \leq 0.82$.
(a) Would a 99% CI calculated from the same sample data be longer or shorter?
(b) Consider the following statement: There is a 95% chance that $\mu$ is between 0.49 and 0.82. Is this statement correct? Explain your answer.
(c) Consider the following statement: If $n = 100$ random samples of water from the lake were taken and the 95% CI on $\mu$ computed, and this process was repeated 1000 times, 950 of the CIs will contain the true value of $\mu$. Is this statement correct? Explain your answer.

8.8. The breaking strength of yarn used in manufacturing drapery material is required to be at least 100 psi. Past experience has indicated that breaking strength is normally distributed and that $\sigma = 2$ psi. A random sample of nine specimens is tested, and the average breaking strength is found to be 98 psi. Find a 95% two-sided confidence interval on the true mean breaking strength.

8.9. The yield of a chemical process is being studied. From previous experience yield is known to be normally distributed and $\sigma = 3$. The past five days of plant operation have resulted in the following percent yields: 91.6, 88.75, 90.8, 89.95, and 91.3. Find a 95% two-sided confidence interval on the true mean yield.
8-10. The diameter of holes for cable harness is known to have a normal distribution with \( \sigma = 0.01 \) inch. A random sample of size 10 yields an average diameter of 1.5045 inch. Find a 99% two-sided confidence interval on the mean hole diameter.

8-11. A manufacturer produces piston rings for an automobile engine. It is known that ring diameter is normally distributed with \( \sigma = 0.001 \) millimeters. A random sample of 15 rings has a mean diameter of \( \bar{x} = 74.036 \) millimeters.
(a) Construct a 99% two-sided confidence interval on the mean piston ring diameter.
(b) Construct a 95% lower-confidence bound on the mean piston ring diameter.

8-12. The life in hours of a 75-watt light bulb is known to be normally distributed with \( \sigma = 25 \) hours. A random sample of 20 bulbs has a mean life of \( \bar{x} = 1014 \) hours.
(a) Construct a 95% two-sided confidence interval on the mean life.
(b) Construct a 95% lower-confidence bound on the mean life.

8-13. A civil engineer is analyzing the compressive strength of concrete. Compressive strength is normally distributed with \( \sigma^2 = 1000(\text{psi})^2 \). A random sample of 12 specimens has a mean compressive strength of \( \bar{x} = 3250 \) psi.

(a) Construct a 95% two-sided confidence interval on mean compressive strength.
(b) Construct a 99% two-sided confidence interval on mean compressive strength. Compare the width of this confidence interval with the width of the one found in part (a).

8-14. Suppose that in Exercise 8-12 we wanted to be 95% confident that the error in estimating the mean life is less than five hours. What sample size should be used?

8-15. Suppose that in Exercise 8-12 we wanted the total width of the two-sided confidence interval on mean life to be six hours at 95% confidence. What sample size should be used?

8-16. Suppose that in Exercise 8-13 it is desired to estimate the compressive strength with an error that is less than 15 psi at 99% confidence. What sample size is required?

8-17. By how much must the sample size \( n \) be increased if the length of the CI on \( \mu \) in Equation 8-7 is to be halved?

8-18. If the sample size \( n \) is doubled, by how much is the length of the CI on \( \mu \) in Equation 8-7 reduced? What happens to the length of the interval if the sample size is increased by a factor of four?