Statistical inference may be divided into two major areas:

a) The estimator

1- Characteristics of a good estimator: (unbiased, with the least variance)

2- How to construct an estimator: (maximum likelihood)

b) Making decision or drawing conclusions

3- Constructing Confidence intervals.

4- Conducting the testing of hypothesis.

A **confidence interval** estimate for μ 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:

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

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

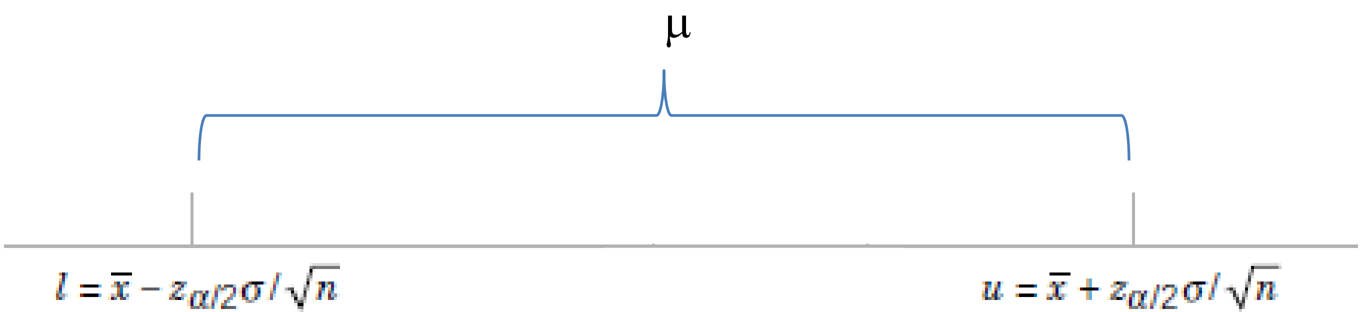
$$P\left\{\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right\} = 1 - \alpha$$

Population with known variance

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

$$\overline{x} - z_{\alpha/2}\sigma/\sqrt{n} \leq \mu \leq \overline{x} + z_{\alpha/2}\sigma/\sqrt{n}$$

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



If $\overline{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error $|\overline{x} - \mu|$ will not exceed a specified amount E when the sample size is

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

$$\hat{\theta} - z_{\alpha/2}\sigma_{\hat{\Theta}} \leq \theta \leq \hat{\theta} + z_{\alpha/2}\sigma_{\hat{\Theta}}$$

Example1

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).

Solution

a) 95% two sided CI on the mean compressive strength

$z_{\alpha/2} = z_{0.025} = 1.96$, and $\bar{x} = 3250$, $\sigma^2 = 1000$, n=12

$$\bar{x} - z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$3250 - 1.96\left(\frac{31.62}{\sqrt{12}}\right) \le \mu \le 3250 + 1.96\left(\frac{31.62}{\sqrt{12}}\right)$$

$$3232.11 \le \mu \le 3267.89$$

b.) 99% Two-sided CI on the true mean compressive strength

$z_{\alpha/2} = z_{0.005} = 2.58$

$$\bar{x} - z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$3250 - 2.58\left(\frac{31.62}{\sqrt{12}}\right) \le \mu \le 3250 + 2.58\left(\frac{31.62}{\sqrt{12}}\right)$$

$$3226.5 \le \mu \le 3273.5$$

Population with unknown variance

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal distribution with unknown variance σ^2, a $100(1 - \alpha)$ **percent confidence interval** on μ is given by

$$\bar{x} - t_{\alpha/2,n-1} s/\sqrt{n} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1} s/\sqrt{n}$$

where $t_{\alpha/2,n-1}$ is the upper $100\alpha/2$ percentage point of the t distribution with $n - 1$ degrees of freedom.

If the population variance is unknown we use the T-distribution (if n<30), but if n become large, we can use the Z-distribution

When n is large, the quantity

$$\bar{x} - z_{\alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{s}{\sqrt{n}}$$

is a **large sample confidence interval** for μ, with confidence level of approximately $100(1 - \alpha)\%$.

Example 2:

An article in the journal *Materials Engineering* (1989, Vol. II, No. 4, pp. 275–281) describes the results of tensile adhesion tests on 22 U-700 alloy specimens. The load at specimen failure is as follows (in megapascals):

19.8	10.1	14.9	7.5	15.4	15.4
15.4	18.5	7.9	12.7	11.9	11.4
11.4	14.1	17.6	16.7	15.8	
19.5	8.8	13.6	11.9	11.4	

The sample mean is $\bar{x} = 13.71$, and the sample standard deviation is $s = 3.55$.

Solution We want to find a 95% CI on μ. Since $n = 22$, we have $n - 1 = 21$ degrees of freedom for t, so $t_{0.025,21} = 2.080$. The resulting CI is

$$\bar{x} - t_{\alpha/2,n-1}s/\sqrt{n} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1}s/\sqrt{n}$$

$$13.71 - 2.080(3.55)/\sqrt{22} \leq \mu \leq 13.71 + 2.080(3.55)/\sqrt{22}$$

$$13.71 - 1.57 \leq \mu \leq 13.71 + 1.57$$

$$12.14 \leq \mu \leq 15.28$$

The CI is fairly wide because there is a lot of variability in the tensile adhesion test measurements.

Example 3: A postmix beverage machine is adjusted to release a certain amount of syrup into a chamber where it is mixed with carbonated water. A random sample of 25 beverages was found to have a mean syrup content of $\bar{x} = 1.10$ fluid ounces and a standard deviation of $s = 0.015$ fluid ounces. Find a 95% CI on the mean volume of syrup dispensed.

Solution

$$\bar{x} = 1.10 \quad s = 0.015 \quad n = 25$$

95% CI on the mean volume of syrup dispensed

For $\alpha = 0.05$ and $n = 25$, $t_{\alpha/2,n-1} = t_{0.025,24} = 2.064$

$$\bar{x} - t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right)$$

$$1.10 - 2.064\left(\frac{0.015}{\sqrt{25}}\right) \le \mu \le 1.10 + 2.064\left(\frac{0.015}{\sqrt{25}}\right)$$

$$1.093 \le \mu \le 1.106$$

Ex: An article in the 1993 volume of the *Transactions of the American Fisheries Society reports* the results of a study to investigate the mercury contamination in largemouth bass. A sample of fish was selected from 53 Florida lakes and mercury concentration in the muscle tissue was measured (ppm). The mercury concentration values are

1.230	0.490	0.490	1.080	0.590	0.280	0.180	0.100	0.940
1.330	0.190	1.160	0.980	0.340	0.340	0.190	0.210	0.400
0.040	0.830	0.050	0.630	0.340	0.750	0.040	0.860	0.430
0.044	0.810	0.150	0.560	0.840	0.870	0.490	0.520	0.250
1.200	0.710	0.190	0.410	0.500	0.560	1.100	0.650	0.270
0.270	0.500	0.770	0.730	0.340	0.170	0.160	0.270	

We want to find an approximate 95% CI on μ

Solution:

We want to find an approximate 95% CI on μ. Because $n > 30$, the assumption of normality is not necessary .

The required quantities are $n = 53, \bar{x} = 0.5250, s = 0.3486$, and $z_{0.025} = 1.96$. The approximate 95% CI on μ is