Interval Estimation of the Population Mean

[redacted]

Confidence Interval for the Population Mean: σ is Known

In many real-life situations, errors of estimation result since most of the time, the estimates will not be exactly equal to the true value of the parameter. The size of this error is the absolute value of the difference between the estimate and the true parameter value. You can be $(1 - \alpha)100\%$ certain that this difference will not exceed $e = Z_{\sigma/2} \frac{\sigma}{\sqrt{n}}$, which is referred to as the *maximum error* or *margin of error*. [redacted]

[redacted]

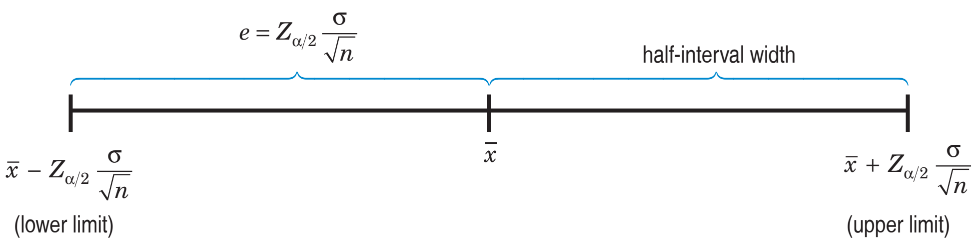

Figure 4.5 A $(1 - \alpha)100\%$ confidence interval for the population mean μ

The $(1 - \alpha)100\%$ confidence interval estimator illustrated in figure 4.5 (Case 1) for the population mean μ is applied when the population is normally distributed and σ is known, whatever the sample size n is.

Oftentimes, the population standard deviation σ and the distribution of the population are unknown. The Central Limit Theorem assures you that as long as the sample size is large, that is, $n \geq 30$, the sample standard deviation s can be used to estimate the population standard deviation σ, and hence, the large sample confidence interval estimator can be used.

[redacted]

Example 4.2

The scores of a random sample of 100 high school students on a standardized mathematics test in school A gave a mean of 78 and a standard deviation of 20.

1. What is the best point estimate of the true average score in this standardized mathematics test?
2. What is the standard error of this point estimate?
3. What is the margin of error or maximum allowable error?
4. Construct a 95% confidence interval estimate for the true average score in mathematics in this standardized test.
5. If the average score in mathematics in this standardized test is 73 in school B, can you conclude that there is a significant difference between the average scores in the standardized mathematics test for the two schools?

Solution.

It is given that for school A, $\overline{x} = 78$, $s = 20$, and $u = 100$.

1. The value of the sample mean $\overline{x}$, which is 78, is the best point estimate of the true average score of school A in this standardized mathematics test.

2. The standard error given by this point estimate, with the sample standard deviation s estimating the population standard deviation, is

$$\text{stderr}(\overline{x}) = \frac{s}{\sqrt{n}} = \frac{20}{\sqrt{100}} = \frac{20}{10} = 2.$$

3. The margin of error or maximum error e at 95% level of confidence ($\alpha = 0.05$) is given by

$$e = Z_{\alpha/2}\text{stderr}(\overline{x}) = Z_{0.05/2}\text{stderr}(\overline{x}) = z_{0.025}(2) = 1.96(2) = 3.92.$$

   Hence, the size of the error should not exceed 3.92.

4. A 95% confidence interval estimate of the true average score in this standardized mathematics test of school A is given by

$$\overline{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

   or simply

$$\overline{x} - e < \mu < \overline{x} + e.$$

This results to

$$\bar{x} - e < \mu < \bar{x} + e$$

$$78 - 3.92 < \mu < 78 + 3.92$$

$$74.08 < \mu < 81.92.$$

or simply 95% confidence interval [74.08, 81.92].

This means that you are 95% confident that the true average score in this standardized mathematics test of school A is between 74.08 and 81.92.

5. There is a significant difference between the true average scores in the standardized mathematics test of schools A and B since 73 is not contained in the 95 confidence interval [74.08, 81.92]. Moreover, you can conclude that school A has a significantly higher average score in this mathematics test than school B.

[redacted]

[redacted]

Confidence Interval for the Population Mean: σ is Unknown

In case the population deviation σ is unknown and sample size is small ($n < 30$), confidence intervals for the population mean μ can be constructed based on the t-distribution provided that the population from which the random sample was obtained is normally distributed. You learned from chapter 3 that

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

follows a t-distribution with $n - 1$ degrees of freedom. Analogously, the confidence interval for μ can be constructed by

$$P\left(-t_{\alpha/2} < z < t_{\alpha/2}\right) = 1 - \alpha,$$

where $t_{\alpha/2}$ is the t-value where you find an area of $\frac{\alpha}{2}$ in the right tail of a t-distribution with $n - 1$ degrees of freedom. Then

$$P\left(-t_{\alpha/2} < \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} < t_{\alpha/2}\right) = 1 - \alpha$$

and

$$P\left(\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right) = 1 - \alpha.$$

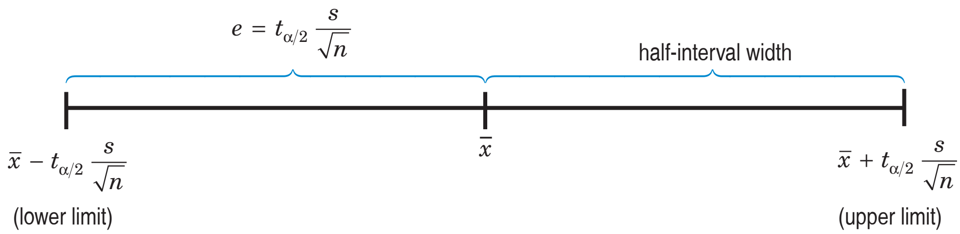

Figure 4.6 A $(1 - \alpha)100\%$ confidence interval for population mean μ (small sample with $n < 30$ and σ unknown)

The confidence interval estimator, which you can classify as case 2, for the population mean μ is applied when the sample size n is small ($n < 30$) and the population standard deviation σ is unknown. In this case, the population is assumed to be (approximately) normally distributed.

[redacted]