

Lecture Outline 1 - 23 - 09

Estimated Margin of Error for \bar{x} (when sampling from a *normal* population). This method applies exactly for every $n > 1$ provided the population from which we are sampling has an x-score distribution which is *normal*. Consult pg. 606, pp. 586-596. Refer to 1-14-09. For *every* $n > 1$, all that is needed is to replace 1.96 by an entry from the t-Table A-97, Appendix D, for "degrees of freedom" $df = n-1$.

Question: It is thought that diameters x_i of a production run of steel pins are random and follow a **normal distribution** having (unknown) mean and standard deviation μ, σ . A random sample of 5 such pins is selected from production and their sample mean diameter \bar{x} is calculated. What is the margin of error for \bar{x} and what is the 95% confidence interval for μ ? The x-measurements are

1.0030, 1.0034, 1.0031, 1.00292, 1.0031

1a. As with 1-14-09, we require the sample mean and sample standard deviation s.

$$\bar{x} = 1.0031$$

$$s = 0.000181879$$

Care must be taken when calculating s . Rounding and other errors can easily affect the result. If you use a calculator's built-in sample standard deviation routine be sure that it is using the $n-1$ divisor. It is especially important for very small n since it *does* make a difference.

1b. The estimated margin of error for \bar{x} is calculated:

$$2.776 \frac{s}{\sqrt{n}} = 2.776 \frac{0.000181879}{\sqrt{5}} \sim 0.000225797.$$

Value $t = 2.776$ is from t-Table (A-97, Appendix D). You find 2.776 by locating confidence 95% at the bottom of the table and running up the column to row 4 (degrees of freedom (df) = $5-1 = 4$).

The t-table uses df to denote "degrees of freedom." In our application of the table $df = n-1$ because, in assessing the margin of error of \bar{x} , we have to estimate the parameter σ by s . Statistically, this has the effect of reducing the sample size by 1, thus $df = n-1$.

1c. Claim made for estimated margin of error:

Around 95% of random samples of $n = 5$ from a *normal population* have $\bar{x} \pm$ estimated margin of error cover the population mean μ .

Our random sample of $n = 5$ parts has produced the interval

$$\begin{aligned} & \bar{x} \pm \text{estimated margin of error} \\ & = 1.0031 \pm 0.000225797 \\ & = [1.00287, 1.00333] \end{aligned}$$

This interval is called a "95% confidence interval for μ ."

1d. If the population is normal and all calculations are made with perfect accuracy the EXACTLY 95% of samples of 5 are "good samples" whose 95% confidence interval $\bar{x} \pm \text{estimated margin of error}$ covers the true value of μ . Is our sample of 5 a "good one?" We don't know. Nonetheless, the method of reporting the sampling results does give a pretty good idea of the reliability of the findings.

2. Other confidence levels. Look at the next to bottom row of the t-table (it has an ∞ symbol at the left of that row, denoting large sample size). Notice the entry 2.576 of this row and that directly below this is confidence level 99%. We will however use the entry for $df = 5 - 1 = 4$. To obtain a 99% confidence interval μ in the $n = 5$ case you would substitute 4.604 for 2.576. So the 99% confidence interval for μ takes the form

$$\bar{x} \pm (4.604 \text{ not } 2.576) \frac{s}{\sqrt{n}}$$

Claim: Around 99% of samples of n from a normal population will produce a 99% confidence interval covering the population mean μ .

Exercises for 1-21-09.

1. #23 pg. 644
2. #31 pg. 646 (use z-based CI for $\mu_1 - \mu_2$).
3. #39 pg. 647.
4. #44 pg. 648 (just give 95% CI for $\mu_{\text{sparrow}} - \mu_{\text{robin}}$)

Exercises for 1-23-09.

1. #36 pg. 614. Assume population \sim normal (see pg. 592).
2. #33 pg. 613. ((b) should say sample mean and sample sd s).
3. #34. pg. 613. ((b) should use word "sample").
4. #39 pg. 614. Actually, the distribution of $x = \#$ chips in a bag is approximately normal distributed if chips are well mixed in the batter.
5. #42 pg. 614. Give a 95% CI for population mean stopping distance. If this CI falls entirely to the right of 125 feet let's reject the supposition that mean stopping distance is 125 or shorter. If mean stopping distance is actually 125, what is the chance we'll reject? Do we reject based on this data?