Confidence Interval for a Mean

Math 130: Introductory Statistics

1

1.1 Goals

Goals of this class meeting

- Learn how to estimate a population mean, and report a margin of error.
- Learn what a confidence interval is, and how to interpret it.

2 Terminology and interpretation

2.1 Margin of error

Margin of error

- **Point estimate**: a single value used to approximate a population parameter.
  - Example: a sample mean is a point estimate for the population mean.
- **Margin of error**: the maximum likely distance of a point estimate from the actual population parameter.
- The margin of error depends on the desired **level of confidence**.
  - The larger the level of confidence, the larger the margin of error.
  - Example: estimate today's high temperature.

2.2 Confidence interval

Confidence interval

- **Confidence interval**: a range of values (instead of a single point estimate) used to estimate a population parameter.
- The width of the confidence interval depends on the **level of confidence**.

- **Level of confidence**: a probability given by, $1 - \alpha$. It is equal to the proportion of times that similarly constructed confidence intervals will contain the true population parameter.
- Let E denote the margin of error for a sample mean, $\bar{x}$. The confidence interval is simply,

$$(\bar{x} - E, \bar{x} + E)$$

- Example, say $\bar{x} = 40$ and the margin of error for a 95% is 15. Then the *95% confidence interval is* $(25, 55)$.

2.3 Interpretation

Interpretation of CI

- Previous example: The 95% confidence interval for $\bar{x}$ is $(25, 55)$.
- Correct interpretation: 95% of similarly constructed confidence intervals will contain the true population mean.
- That is, if you were to repeat the experiment many, many times, and compute each confidence interval, 95% of these confidence intervals will contain the true mean.
- Wrong interpretation: there is a 95% chance the true mean is between 25 and 55.

3 Computing confidence intervals

3.1 Confidence interval formula

Computing confidence intervals

- For now, let's make the stupid assumption we know the population parameter, σ.
- The margin of error is given by:

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

- $z_{\alpha/2}$ is called a **critical value**. The larger the level of confidence, the larger is $z_{\alpha/2}$.
- Therefore the confidence interval is:

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

- This method of computing confidence intervals depends on the central limit theorem. So it is only applicable if the assumptions of the central limit theorem are met.

3.2 Critical values

Critical values

- We know from the Central Limit Theorem that the sampling distribution for $\bar{x}$ is normal and has a standard deviation of $\frac{\sigma}{\sqrt{n}}$.
- Therefore, if we want a 95% confidence interval, find values for $\bar{x}$ such that 95% of the sampling distribution will be within that range.
- What values for z include 95% of the normal distribution?
  - These are the *critical values* for a 95% confidence interval.

3.3 Desired sample sizes

Desired samples sizes

- Suppose before you conducted a study or gathered any data, you have a goal that the margin of error not be larger than some size.
- You need to know how much data to collect to make that happen.

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

- E is the largest margin of error you will tolerate, $1 - \alpha$ is the confidence level of that margin of error.
- You will need to have *at least* this many observations or else your margin of error will be larger than E.

4 Examples

Examples

- Suppose you take a sample of 65 Viterbo graduates and find the mean income is $37,600. It is known the population standard deviation is $3,500. Compute a 95% confidence interval for mean income of Viterbo graduates.
  - First: find critical z-values.
  - Plug and chug into the formula:

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

$$E = 1.96\frac{3500}{\sqrt{65}} = 850.88$$

$$(37600 - 850.88, 37600 + 850.88)$$

$$(36749.12, 38450.88)$$

Examples

- Use same example. Compute a 90% confidence interval. Is the range larger or smaller?
- Compute a 95% confidence interval, but suppose the sample size is 30. Is the range larger or smaller?
- Suppose you desire a 95% margin of error not to exceed $700. What sample size must be collected to make this happen?
- Suppose you desire a 99% margin of error not to exceed $700. What sample size must be collected to make this happen?
- Suppose you desire a 95% margin of error not to exceed $1000. What sample size must be collected to make this happen?

5 T-distribution

5.1 Assumptions

When σ is unknown.

- Assume we have a sample size $n > 30$ *or..*
- Assume the population has a normal distribution.
- The population standard deviation, σ is not known. Our best estimate is the sample standard deviation, s.

5.2 T-statistic

More uncertainty

- Since we need to estimate σ with s, this adds more uncertainty.
- More uncertainty means we need to widen the confidence interval.
- Use the T-statistic (instead of using Z).
- The degree of uncertainty $\rightarrow$ t-statistic depends on the degrees of freedom.

5.3 Confidence Intervals

Confidence Interval

- The margin of error is given by:

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}}$$

- Confidence interval is:

$$(\bar{x} - E, \bar{x} + E)$$