Unit 3: Foundations for inference

2. Confidence intervals

GOVT 3990 - Spring 2020

Cornell University

Outline

1. Main ideas

    1. Statistical inference methods based on the CLT depend on the same conditions as the CLT
    2. Use confidence intervals to estimate population parameters
    3. Critical value depends on the confidence level
    4. Calculate the sample size a priori to achieve desired margin of error

2. Summary

Outline

1. Main ideas

    1. Statistical inference methods based on the CLT depend on the same conditions as the CLT
    2. Use confidence intervals to estimate population parameters
    3. Critical value depends on the confidence level
    4. Calculate the sample size a priori to achieve desired margin of error

2. Summary

1. Statistical inference methods based on the CLT depend on the same conditions as the CLT

Always check these in context of the data and the research question!

1. *Independence:* Sampled observations must be independent.
   * This is difficult to verify, but is more likely if
     - random sampling/assignment is used, and,
     - if sampling without replacement, $n <$ 10% of the population.
2. *Sample size/skew:* Either the population distribution is normal or $n > 30$ and the population distribution is not extremely skewed (the more skewed the distribution, the higher n necessary for the CLT to apply).
   * This is also difficult to verify for the population, but we can check it using the sample data, and assume that the sample mirrors the population.

Outline

1. Main ideas

    1. Statistical inference methods based on the CLT depend on the same conditions as the CLT

    2. Use confidence intervals to estimate population parameters

    3. Critical value depends on the confidence level

    4. Calculate the sample size a priori to achieve desired margin of error

2. Summary

2. Use confidence intervals to estimate population parameters

$$CI : point\ estimate \pm margin\ of\ error$$

2. Use confidence intervals to estimate population parameters

$$CI : point\ estimate \pm margin\ of\ error$$

If the parameter of interest is the population mean, and the point estimate is the sample mean,

$$\bar{x} \pm Z^{\star}\frac{s}{\sqrt{n}}$$

Outline

1. Main ideas

    1. Statistical inference methods based on the CLT depend on the same conditions as the CLT

    2. Use confidence intervals to estimate population parameters

    3. Critical value depends on the confidence level

    4. Calculate the sample size a priori to achieve desired margin of error

2. Summary

3. Critical value depends on the confidence level

Your turn

What is the critical value ($Z^\star$) for a confidence interval at the 91% confidence level?

(a) $Z^\star = 1.34$

(b) $Z^\star = 1.65$

(c) $Z^\star = 1.70$

(d) $Z^\star = 1.96$

(e) $Z^\star = 2.33$

Your turn

What is the critical value ($Z^\star$) for a confidence interval at the 91% confidence level?

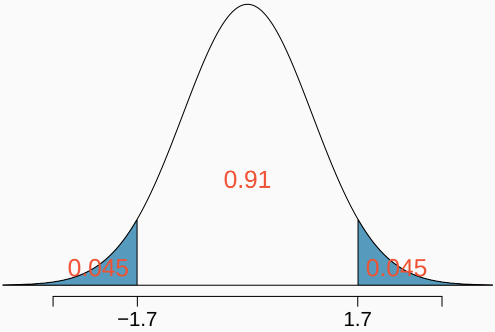
(a) $Z^\star = 1.34$

(b) $Z^\star = 1.65$

(c) $Z^\star = 1.70$

(d) $Z^\star = 1.96$

(e) $Z^\star = 2.33$



Common misconceptions about confidence intervals

1. The confidence level of a confidence interval is the probability that the true population parameter is in the confidence interval you construct for a single sample.

Common misconceptions about confidence intervals

1. The confidence level of a confidence interval is the probability that the true population parameter is in the confidence interval you construct for a single sample.
   The confidence level is equal to the proportion of random samples that result in confidence intervals that contain the true pop. parameter.

Common misconceptions about confidence intervals

1. The confidence level of a confidence interval is the probability that the true population parameter is in the confidence interval you construct for a single sample.
   The confidence level is equal to the proportion of random samples that result in confidence intervals that contain the true pop. parameter.
2. A narrower confidence interval is always better.

Common misconceptions about confidence intervals

1. The confidence level of a confidence interval is the probability that the true population parameter is in the confidence interval you construct for a single sample.
   The confidence level is equal to the proportion of random samples that result in confidence intervals that contain the true pop. parameter.
2. A narrower confidence interval is always better.
   This is incorrect since the width is a function of both the confidence level and the standard error.

Common misconceptions about confidence intervals

1. The confidence level of a confidence interval is the probability that the true population parameter is in the confidence interval you construct for a single sample.
   The confidence level is equal to the proportion of random samples that result in confidence intervals that contain the true pop. parameter.
2. A narrower confidence interval is always better.
   This is incorrect since the width is a function of both the confidence level and the standard error.
3. A wider interval means less confidence.

Common misconceptions about confidence intervals

1. The confidence level of a confidence interval is the probability that the true population parameter is in the confidence interval you construct for a single sample.
   The confidence level is equal to the proportion of random samples that result in confidence intervals that contain the true pop. parameter.
2. A narrower confidence interval is always better.
   This is incorrect since the width is a function of both the confidence level and the standard error.
3. A wider interval means less confidence.
   This is incorrect since it is possible to make very precise statements with very little confidence.

Outline

1. Main ideas

    1. Statistical inference methods based on the CLT depend on the same conditions as the CLT

    2. Use confidence intervals to estimate population parameters

    3. Critical value depends on the confidence level

    4. Calculate the sample size a priori to achieve desired margin of error

2. Summary

4. Calculate the sample size a priori to achieve desired margin of error

$$ME = z^{\star} \frac{s}{\sqrt{n}}$$

So if we know the desired ME, and confidence level (and hence $z^{\star}$), and the sample standard deviation, we can solve for n.

4. Calculate the sample size a priori to achieve desired margin of error

$$ME = z^\star \frac{s}{\sqrt{n}}$$

So if we know the desired ME, and confidence level (and hence $z^\star$), and the sample standard deviation, we can solve for n.

$$n = (\frac{z \star s}{ME})^2$$

Application exercise: 3.1 Confidence interval for a single mean

Outline

1. Main ideas

    1. Statistical inference methods based on the CLT depend on the same conditions as the CLT

    2. Use confidence intervals to estimate population parameters

    3. Critical value depends on the confidence level

    4. Calculate the sample size a priori to achieve desired margin of error

2. Summary

Summary of main ideas

1. Statistical inference methods based on the CLT depend on the same conditions as the CLT
2. Use confidence intervals to estimate population parameters
3. Critical value depends on the confidence level
4. Calculate the sample size a priori to achieve desired margin of error