Confidence Intervals (CIs)

Dr Trevor Bryant

Learning Outcomes

Following this session you should be able to:

• Understand the concepts and interpretation of confidence intervals;
• Explain how they are derived
• Understand how they can be used to assess precision
• Demonstrate how they are should be presented
• Use software to calculate them

Estimation Methods

We rarely measure the whole population

- Estimation
  - Point Estimation
  - Interval Estimation
Point Estimation

- Provides Single Value
  - Based on Observations from 1 Sample
- Gives No Information about how close our value is to the unknown Population Parameter
- Example: Sample Mean ($\bar{X}$) = 50
  Point Estimate of unknown Population Mean

Estimation Process

Population

- Mean, $\mu$, is unknown
- Sample

Random Sample

- Mean $ar{X}$ = 50

Estimation from a population

- The population is defined as the group about whom statements will be made
- If a representative sample is taken conclusions from the sample can be generalized to the wider group
Understanding Statistical notation

Population | Sample
---|---
Mean | \( \mu \) | \( \bar{x} \)
Standard Deviation | \( \sigma \) | \( s \) (SD Std Dev)

Estimation Process

Repeated sampling from the population gives samples means whose frequency distribution (sampling distribution) properties are:

- The mean of this distribution would be the population mean \( \mu \)
- The standard deviation of this distribution of sample means is called the Standard Error (SE)
Standard Error (SE)

- The Standard Error measures how precisely the population mean is estimated by the sample mean.
- SE is estimated by the sample SD divided by the square root of the number of observations.

\[ \text{SE} = \frac{\text{SD}}{\sqrt{n}} \]

Estimating the mean of a continuous variable

- Using the properties of the normal distribution we can estimate the range in which the unknown population mean lies.

Possible Range

Population Mean \[ \pm \] Sample Mean = Estimate of Population Mean

Estimating the mean of a continuous variable

- This range is called the 95\% confidence interval about the mean.
  - It is calculated as:
    - Sample mean \( \pm 1.96 \times \text{Standard Error} \)
  - All values within the confidence interval are reasonable values for the population mean that generated the observed sample.
  - It gives an idea of the precision of the estimate from the sample size available.
Different Degrees of Confidence

\[ X = \mu \pm z^* \text{SE}_x \]

Different forms of Confidence Intervals

- Continuous outcome variables: Means, Medians
  - One sample
  - Two sample (difference)
- Categorical outcome variables: Proportions
  - One sample proportion
  - Two sample proportion (difference)
- Correlation
- Odds ratio (OR) & Relative risk (RR)
- Standardised Mortality ratios

Alternative Distributions

Different Confidence Interval calculations require different theoretical distributions

- Means (small numbers): \( t \) distribution
- Standardised Mortality ratios: Poisson distribution
- Medians: Binomial Distribution

They all need a sample estimate and a standard error
Association between CI and P values

**Differences in Continuous measures or proportions**
- If 95% CI includes 0 then p value will be greater than 0.05
- If 95% CI does not include 0 then p value will be less than 0.05

**Ratios and Risks**
- If 95% CI includes 1 then p value will be greater than 0.05
- If 95% CI does not include 1 then p value will be less than 0.05

99% = 0.01

---

**CI & Hypothesis Testing**

If CIs do not cross at a significance level (say 5%), then hypothesis testing is significant but the opposite is not always true?

- Confidence intervals NO overlap: Can conclude that there is a real difference between the two groups
- Confidence intervals overlap: Cannot draw any conclusions about difference without further information

---

**Example 1: Interpreting a rate**

- Sample of 1106 pregnancies, estimated rate of congenital abnormality was 4.2% (95% CI 3.0% to 5.3%)
- The ‘true’ population rate could be as low as 3.0%
- The ‘true’ population could be as high as 5.3%
- There is a 1 in 20 chance that our estimate is wrong and that the true population value is outside this range
- Our best estimate of congenital abnormality is 4.2%
Example 2: Interpreting a difference between two means

- Mean birthweight was measured in a sample of 15 non-smokers (3.59Kg) and 14 heavy smokers (3.20Kg)
- The difference in the mean weight was 390g (95%CI 60g to 721g)
- The 95% CI excludes 0, therefore the difference is statistically significant (P will be less than 0.05)
- Although the difference is significant, our estimate of the Percentage difference is 390 / 3400 = 11.5%
- Is this clinically important?

Example 3: Interpreting differences

- Length of stay in hospital
  - Group 1 (n= 392) Mean stay 37 days
  - Group 2 (n= 368) Mean stay 41 days
- Difference = 4 days (95%CI = -2 to 9) days
- 95%CI includes 0, not statistically significant (P>0.05)
- The study has been unable to rule out that the true difference could be 9 days
- Lack of evidence of a difference is NOT EVIDENCE of no difference!

Example 4: Interpreting proportions

- RCT of flu vaccine
- Infection rate of placebos 80/220 (36%)
- Infection rate for subjects 20/240 (8%)
- Difference in rates 28% (95%CI 21% to 35%)
- 95% CI excludes 0, difference was significant (P<0.001)
- The true difference is at least 21% best estimate is 28%
- Vaccine clearly demonstrates protective effect
- But…… side effects, consider costs, generalisability
Confidence Intervals (CIs) or P values?

- Leading medical journals recommended both when reporting the main study results
- Use of CIs recommend by the ICMJE
- Over emphasis on the P values detracts from more useful approaches when interpreting study results

The problem with P values

- Wrong type of thinking through use of arbitrary cut off at a predefined level (5%)
- Low quality information with $P<0.05$, $P>0.05$, $P=NS$
- $P = 0.049$ is declared as significant and $P = 0.051$ as not significant
- Cut off leads to statistical significance being equated with clinical significance

The problem with P values continued

- A very small improvement, 1% of one treatment compared to another may be statistically significant ($P < 0.001$)
- Only quoting P values may lead uncritical reader into thinking that treatment A was more effective than treatment B
- A clinically important effect may be non-significant because of a small sample size
Confidence intervals from SPSS

Descriptives

<table>
<thead>
<tr>
<th>Statistic</th>
<th>95% Confidence Interval for Mean</th>
<th>95% Confidence Interval for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>144.7078</td>
<td>150.1122</td>
</tr>
<tr>
<td>Lower Bound</td>
<td>142.7078</td>
<td>146.2123</td>
</tr>
<tr>
<td>Upper Bound</td>
<td>150.1122</td>
<td>154.6177</td>
</tr>
<tr>
<td>Median</td>
<td>146.3333</td>
<td>148.4833</td>
</tr>
<tr>
<td>Variance</td>
<td>340.654</td>
<td>360.654</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>18.4568</td>
<td>19.4568</td>
</tr>
<tr>
<td>Minimum</td>
<td>77</td>
<td>77</td>
</tr>
<tr>
<td>Maximum</td>
<td>113</td>
<td>113</td>
</tr>
<tr>
<td>Range</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>Interquartile Range</td>
<td>23.7580</td>
<td>29.3860</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.422</td>
<td>0.241</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>1.683</td>
<td>1.683</td>
</tr>
</tbody>
</table>

Chi-Square Tests

<table>
<thead>
<tr>
<th>Value</th>
<th>df</th>
<th>Asymp. Sig. (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
<th>Exact Sig. (1-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.040b</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>11.996</td>
<td>1</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>13.201</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>12.970</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Extracorporeal membrane oxygenation * 1yr survival Crosstabulation

<table>
<thead>
<tr>
<th>Extracorporeal membrane oxygenation</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>101</td>
<td>84</td>
</tr>
<tr>
<td>1yr survival</td>
<td>63</td>
<td>38</td>
</tr>
<tr>
<td>Count</td>
<td>93</td>
<td>92</td>
</tr>
<tr>
<td>% within Extracorporeal membrane oxygenation</td>
<td>67.7%</td>
<td>41.3%</td>
</tr>
<tr>
<td>% within 1yr survival</td>
<td>54.6%</td>
<td>54.6%</td>
</tr>
</tbody>
</table>

185

Pearson Chi-Square

<table>
<thead>
<tr>
<th>Value</th>
<th>df</th>
<th>Asymp. Sig. (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
<th>Exact Sig. (1-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.040b</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>11.996</td>
<td>1</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>13.201</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>12.970</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

No cells (.0%) have expected count less than 5. The minimum expected count is 41.77.
Software: Confidence Interval Analysis - CIA

- On public workstations
- Downloadable from module website

Table of CIA features

<table>
<thead>
<tr>
<th>Chapter/Statistics with Confidence</th>
<th>Confidence intervals calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Mean and their differences</td>
<td>Single &amp; Two samples</td>
</tr>
<tr>
<td>5 Median and their differences</td>
<td>Single &amp; Two samples</td>
</tr>
<tr>
<td>6 Proportions and their differences</td>
<td>Single &amp; Two samples</td>
</tr>
<tr>
<td>7 Epidemiological studies</td>
<td>Incidence study</td>
</tr>
<tr>
<td></td>
<td>Case-control study</td>
</tr>
<tr>
<td></td>
<td>Ratio of two standardised ratios</td>
</tr>
<tr>
<td>8 Regression and correlation</td>
<td>Single &amp; Two samples</td>
</tr>
<tr>
<td>9 Time to Event studies</td>
<td>Single &amp; Two sample Hazard Ratios</td>
</tr>
<tr>
<td>10 Diagnostic studies</td>
<td>Sensitivity &amp; Specificity</td>
</tr>
<tr>
<td></td>
<td>Positive &amp; Negative Predictive values</td>
</tr>
<tr>
<td></td>
<td>Likelihood rates</td>
</tr>
<tr>
<td></td>
<td>Area under ROC curves</td>
</tr>
<tr>
<td>11 Clinical trials and Meta-Analysis</td>
<td>Numbers needed to treat</td>
</tr>
<tr>
<td></td>
<td>Parallel group &amp; Crossover trials</td>
</tr>
</tbody>
</table>

Continuous variable: single sample

Example 100 diabetics aged between 30-59 with mean BP=146(SD=18)
Continuous variable: paired/unpaired
Example: Difference in systolic BP in 100 diabetic and 100 non-diabetic subjects mean 146 SD 18.5 and mean 140 SD 16.8

Means and their differences: Unpaired samples

Proportions: two unpaired sample
Example: Trial of asthma treatment which measured respiratory failure. Control group it was 63/93 and in treated it was 38/92

Proportions and their differences: Unpaired samples

A Common Question
What is the difference between Reference Range and Confidence Interval?
Reference Ranges refer to Individual values and Confidence Intervals to Estimates
Reference Range uses Standard Deviation
Mean ± 1.96 * Std Deviation
Confidence Interval uses Standard Error
Mean ± 1.96 * Std Error
Quoting Confidence Intervals

• They are not required for all results
• Not required for the mean response of subjects to treatments A and B, if major outcome was the difference between treatments A and B
• Generally restricted to the main outcome of the study which is usually a contrast (difference) between means or proportions

Quoting Confidence Intervals

The difference between the sample mean systolic blood pressure in diabetics and non-diabetics was 6.0 mmHg, with a 95% confidence interval from 1.1 to 10.9 mmHg, the t-test statistic was 2.4 with 198 degrees of freedom and an associated P value of 0.02

Mean difference was 6.0 mmHg (95%CI 1.1 to 10.9 mmHg)

Summary

• Indicate the (im)precision of sample estimates as population values
• They give a range of values for the estimated population parameter (difference)
• They depend on
  – Sample size (larger sizes give narrower CIs)
  – Variability of parameter being estimated
  – Degree of confidence required (90% 95% 99%)
References


Questions?