Confidence Intervals (CIs)

Dr Trevor Bryant

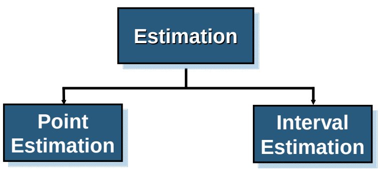
Learning Outcomes

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

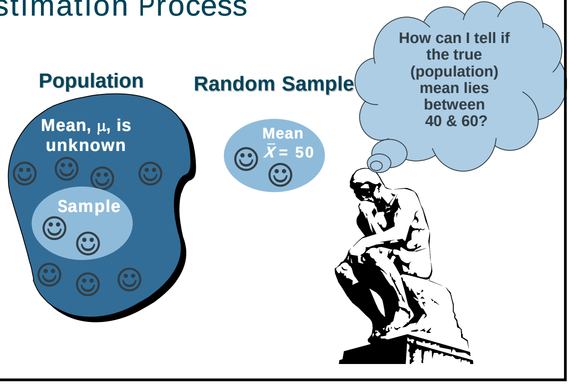
We rarely measure the whole population



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



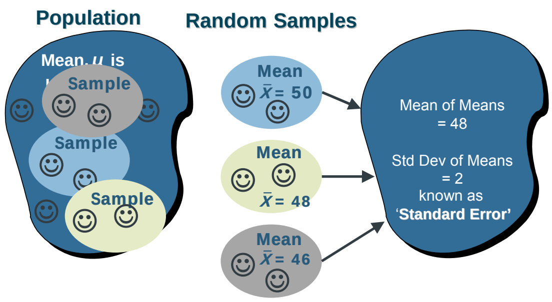
Estimation from a population

- 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	μ	$\bar{x}$
Standard Deviation	σ	s (SD Std Dev)

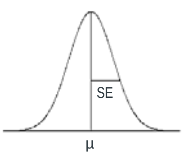
Estimation Process



Estimating the mean of a continuous variable

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 μ
- The standard deviation of this distribution of sample means is called the Standard Error (SE)



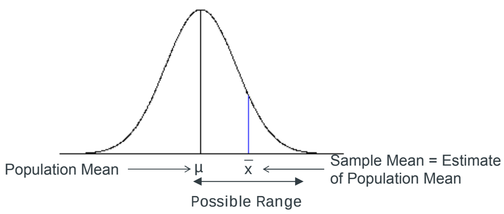
Standard Error (SE)

- 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

$$SE = \frac{SD}{\sqrt{n}}$$

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

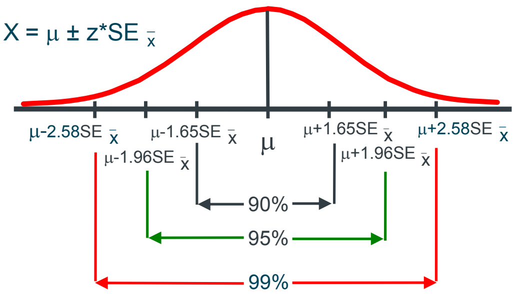


Estimating the mean of a continuous variable

- This range is called the 95% confidence interval about the mean
  - It is calculated as:

    Sample mean ± 1.96 * 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



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 distrubution
Medians	Binomial Distribution

They all need a sample estimate and a standard error

Associaton 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

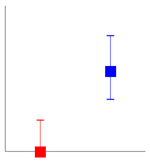
- If 95% CI includes **1** then p value will be greater than 0.05
- If 95% CI does not include **1** then p value will be less than 0.05

99% = 0.01

CI & Hypothesis Testing

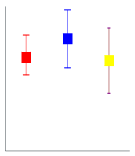
If CIs do not cross at a significance level (say 5%), then hypothesis testing is significant but the opposite is not always true?

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

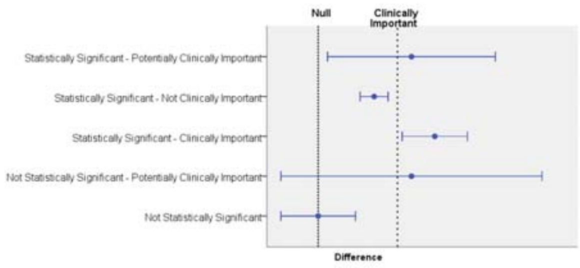
The problem with P values

- 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

Statistical and Clinical Importance



Redrawn from Machin, Campbell and Walters – Medical Statistics 2007

Confidence intervals from SPSS

Descriptives

			Statistic	Std. Error
DIABETIC	Mean		146.4500	1.8457
	95% Confidence Interval for Mean	Lower Bound	142.7878	
		Upper Bound	150.1122	
	5% Trimmed Mean		146.7333	
	Median		146.0000	
	Variance		340.654	
	Std. Deviation		18.4568	
	Minimum		77	
	Maximum		190	
	Range		113	
	Interquartile Range		23.7500	
	Skewness		-.422	.241
	Kurtosis		1.683	.478

Confidence intervals from SPSS

:xtracorporeal membrane oxygenation * 1yr survival Crosstabulation

			1yr survival		
			Yes	No	Total
Extracorpore membrane oxygenation	Yes	Count	63	30	93
		% within Extracorpor membrane oxygenati	67.7%	32.3%	100.0%
	No	Count	38	54	92
		% within Extracorpor membrane oxygenati	41.3%	58.7%	100.0%
Total		Count	101	84	185
		% within Extracorpor membrane oxygenati	54.6%	45.4%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	13.040[b]	1	.000		
Continuity Correction[a]	11.996	1	.001		
Likelihood Ratio	13.201	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	12.970	1	.000		
N of Valid Cases	185				

?

a. Computed only for a 2x2 table

b. 0 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

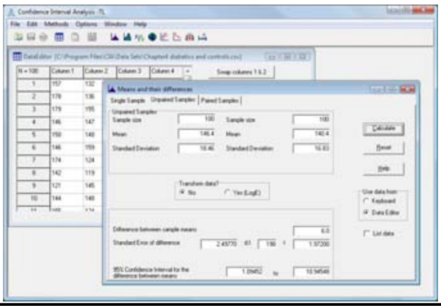
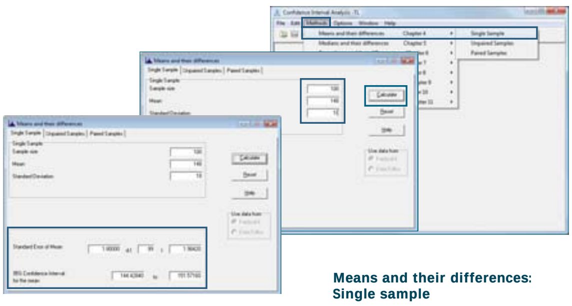


Table of CIA features

Chapter in Statistics with Confidence	Confidence Intervals calculated
4 Mean and their differences	Single & Two samples
5 Median and their differences	Single & Two samples
6 Proportions and their differences	Single & Two samples
7 Epidemiological studies	Incidence study Case-control study Ratio of two standardised ratios Standardised rates
8 Regression and correlation	Single & Two samples
9 Time to Event studies	Single & Two sample Hazard Ratio
10 Diagnostic studies	Sensitivity & Specificity Positive & Negative Predictive values Likelihood ratio Area under ROC curves Kappa
11 Clinical trials and Meta-Analysis	Numbers needed to treat Parallel group & Crossover trials

Continuous variable: single sample

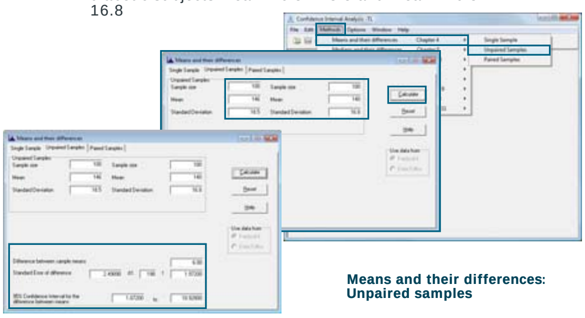
Example 100 diabetics aged between 30-59 with mean BP=146(SD=18)



Means and their differences: Single sample

Continuous variable: paired/unpaired

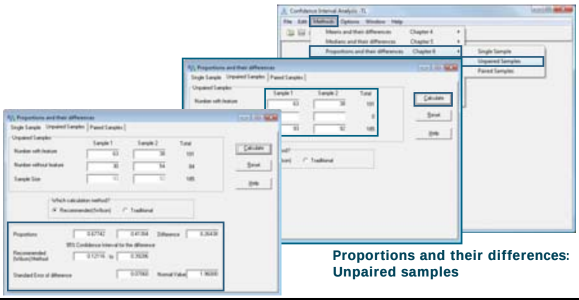
Example : Difference in systolic BP in 100 diabetic and 100 non diabetic subjects mean146 SD 18.5 and mean 140 SD 16.8



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

- Altman D.G., Machin D., Bryant, T.N. & Gardner M.J. *Statistics with Confidence.* 2nd Edition. BMJ Books 2000
- Altman, D.G. *Practical Statistics for Medical Research.* Chapman and Hall 1999. Chapter 8.
- Bland M. *An Introduction to Medical Statistics.* 3rd Edition. Oxford: Oxford Medical Publications 2000.
- Machin D., Campbell M.J., & Walters S.J. *Medical Statistics: A textbook for the Health Sciences.* 4th Edition, 2007. Chapters 7 & 8.
- Kirkwood B.R. & Sterne J.A.C. *Essential Medical Statistics.* 2nd Edition. Oxford: Blackwell Science Ltd 2003. http://www.blackwellpublishing.com/essentialmedstats/default.htm Chapters 6 to 8 (and others).
- Altman DG, Bland JM. (1995) Absence of evidence is not evidence of absence. *BMJ* 311 485.

Questions?