Sample size

Dr Wan Nor Arifin

Outlines

- Estimation – One Mean/Percentage
- Hypothesis testing – Comparing Two Means (Independent/Paired)/Percentages

Software used

- Sample size calculation v1.7.1.xls
- Downloable from:

  `wnarifin.pancakeapps.com/cfcs`

Estimation

1. One mean

- Estimate mean of numerical variable in population e.g. blood pressure, BMI etc.

Estimation

- Conduct a study to estimate mean systolic blood pressure (SBP) among USM medical students. How many medical students should you measure? All of them?
- Sample! Need to calculate minimum number of students to measure → get precise estimation.

Estimation

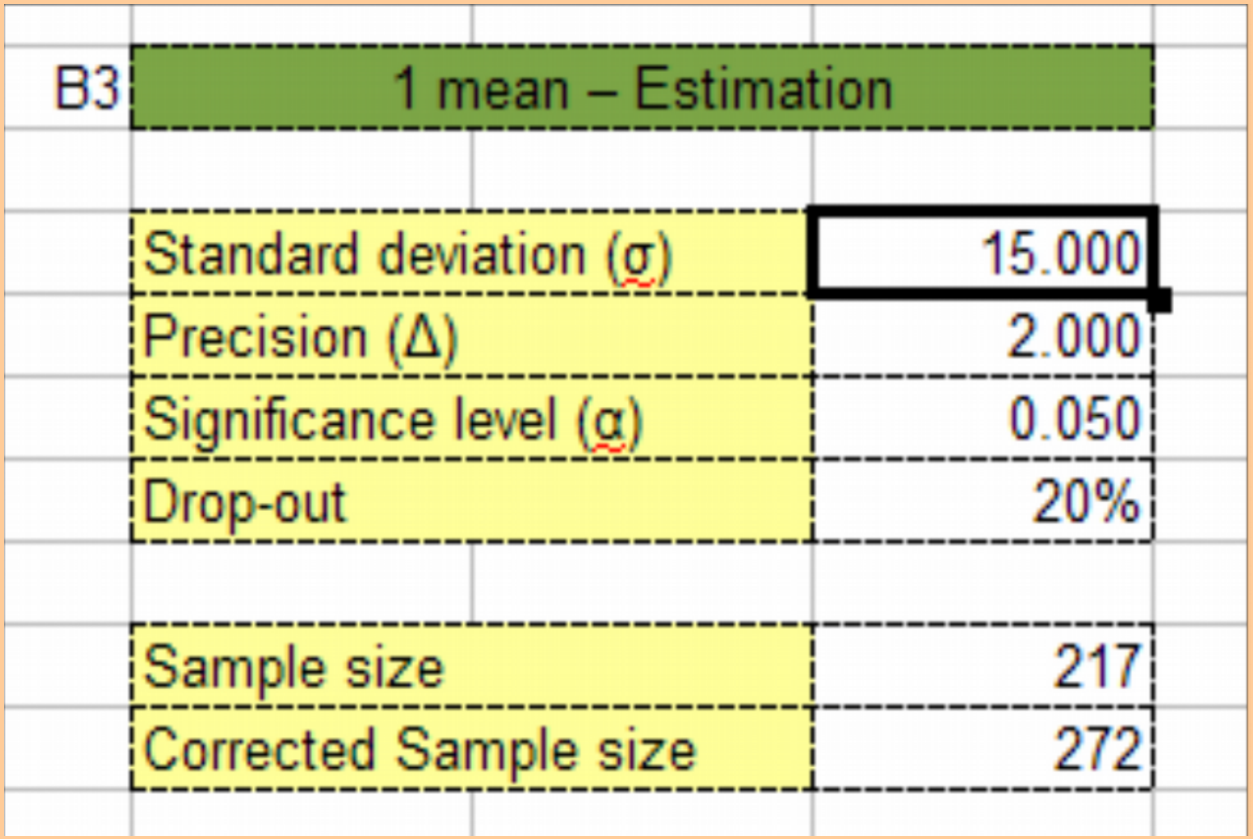
- What you need:
  - Standard deviation of SBP from other studies.
  - Set your precision (in unit of measurement, e.g. 1mmHg, 2mmHg, …).
  - Set Confidence level (90%, **95%**, 99%).
  - Determine Significance level = 100% – Confidence level.
  - Dropout % – % of your participants that run away from study.

Estimation

- Let say:
  - SD of SBP = 15mmHg.
  - Precision = 5mmHg
  - 95% Confidence level
  - Significance level = 5% (0.05)
  - 20% dropout

Estimation

- Use Means → Table B3 “1 mean – Estimation”:



B3	1 mean – Estimation	
	Standard deviation (σ)	15.000
	Precision (Δ)	2.000
	Significance level (α)	0.050
	Drop-out	20%
	Sample size	217
	Corrected Sample size	272

Estimation

- You have to sample 272 medical students to estimate mean SBP among USM medical students, with mean SBP ± 2mmHg.

Estimation

2. One percentage/proportion

- Estimate % of of categorical variable in population e.g. obesity status, HIV, diabetes etc.

Estimation

- Conduct a study to estimate % of obesity among USM medical students. How many medical students should you sample?

Estimation

- What you need:
  - % of obesity among medical students from other studies.
  - Set your precision (in percentage, 1%, 2%, 5%...)
  - Set Confidence level (90%, **95%**, 99%).
  - Determine Significance level = 100% – Confidence level.
  - Dropout % – % of your participants that run away from study.

Estimation

- Let say:
  - % of obesity = 25%.
  - Precision = ±5%.
  - 95% Confidence level
  - Significance level = 5% (0.05)
  - 10% dropout (i.e. those who won't let you know their BMI).

Estimation

- Use Proportion → Table B3 "1 proportion – Estimation":

C3	1 proportion – Estimation	
	Proportion (p)	25.00%
	Precision (Δ)	5.00%
	Significance level (α)	0.050
	Drop-out	10%
	Sample size	289
	Corrected Sample size	322

Estimation

- You have to sample 322 medical students to estimate % of obesity among USM medical students, with % obesity ± 5%.

Hypothesis Testing

3. Comparing two means of two populations

- Testing hypothesis that means of a continuous variable for two different populations are actually different → Using independent t-test.

Hypothesis Testing

- Conduct a study to compare mean BMI of Year 5 with Year 2 medical students. How many medical students should you sample from each population?

Hypothesis Testing

- What you need:
  - SD of BMI of medical students from other studies (preferably Year 5/Year 2, take the largest you could find).
  - Set Expected difference in BMI between the two → Known as Effect Size.
  - Determine Significance level = **0.05**, 0.01, 0.001
  - Power/sensitivity of the test – usually 80%.
  - Dropout %.

Hypothesis Testing

- Let say:
  - Largest SD you could find = 1.5
  - Expected Difference = 1.
  - Significance level = 5% (0.05)
  - Leave Power = 80% (0.8) – default value.
  - 30% dropout (i.e. as some weight themselves while only one foot was on the scale...).

Hypothesis Testing

- Use Means → Table B1 "2 means – Hypothesis Testing":

B1	2 means – Hypothesis Testing	
	Standard deviation (σ)	1.500
	Effect size (Δ)	1.000
	Significance level (α)	0.050
	Power (1-β)	0.800
	Drop-out	30%
	Sample size	36
	Corrected Sample size	52

Hypothesis Testing

- You have to sample 52 Year 5 students and 52 Year 2 students to make the comparison, expecting a difference of 1 unit BMI between the two.

Hypothesis Testing

4. Comparing two means of same population (before-after)

- Testing hypothesis that means of a continuous variable for a population measured before and after an intervention are actually different → Using paired t-test.

Hypothesis Testing

- Conduct a study to compare mean BMI of Year 2 medical students before a lose weight campaign and 2 months after the campaign. How many medical students should you sample from the population?

Hypothesis Testing

- What you need:
  - SD of BMI of medical students from other studies (preferably Year 2, take the largest you could find).
  - Set Expected difference in BMI (after – before) → Effect Size.
  - Determine Significance level = **0.05**, 0.01, 0.001
  - Power/sensitivity of the test – usually 80%.
  - Dropout %.

Hypothesis Testing

- Let say:
  - Largest SD you could find = 2
  - Expected Difference = 1.2.
  - Significance level = 5% (0.05)
  - Power = 80% (0.8)
  - 30% dropout (i.e. as you are unable to find the same participants again for after campaign measurement, it's quite common though).

Hypothesis Testing

- Use Means → Table B2 "2 means (paired/cross-over) – Hypothesis Testing":

B2	2 means (paired/cross-over) – Hypothesis Testing		
	Standard deviation of difference(σ)	2.000	*
	Effect size (Δ)	1.200	
	Significance level (α)	0.050	
	Power (1-β)	0.800	
	Drop-out	30%	
	Sample size	22	
	Corrected Sample size	32	

Hypothesis Testing

- You have to sample 32 Year 2 students to make the before-after comparison, expecting a difference of 1.2 unit BMI between the two.

Hypothesis Testing

5. Comparing percentages/proportions two populations

- Testing hypothesis that percentages of a categorical variable for two different populations are actually different → Using Chi-square test.

Hypothesis Testing

- Conduct a study to % of obesity among Year 2 with Year 5 medical students. How many medical students should you sample from each population?

Hypothesis Testing

- What you need:
  - % of obesity of Year 2 medical students (as control, p0) from other studies (or Year 5 as control, you decide).
  - Set expected % of obesity of Year 5 students (as case, p1).
  - Determine Significance level = **0.05**, 0.01, 0.001
  - Power/sensitivity of the test – usually 80%.
  - Dropout %.

Hypothesis Testing

- Let say:
  - P0 = 35% → Year 2/medical students in general as control.
  - P1 = 50% → If you think this could be the % for Year 5 students.
  - Significance level = 5% (0.05)
  - Power = 80% (0.8)
  - 10% dropout.

Hypothesis Testing

- Use Means → Table C1 "2 proportions – Hypothesis Testing":

	2 proportions – Hypothesis Testing	
	Proportion in control (p0)	35.00%
	Proportion in case (p1)	50.00%
	Significance level (α)	0.050
	Power (1-β)	0.800
	Drop-out	10%
C1	Sample size	167
	Corrected Sample size	186

Hypothesis Testing

- You have to sample 186 Year 2 students and 186 Year 5 students to make the comparison, expecting a difference of 15% for % of obesity between the two.

Reference

Naing, N. N. (2011). A practical guide on determination of sample size in health sciences research. Kelantan: Pustaka Aman Press.