



Section C

Sample Size Determination for Comparing Two Proportions

Power for Comparing Two Proportions

- Same ideas as with comparing means

2

Pilot Study

- We can find the necessary sample size(s) of this study if we specify the following:
 - α -level of test
 - Specific values for p_1 and p_2 (specific H_0) and hence $d = p_1 - p_2$: this usually represents the minimum scientific difference of interest
 - The desired power

3

Pilot Study

- Given this information, we can use Stata to do the sample size calculation
- “sampszi” command
 - Command syntax (items in italics are numbers to be supplied by researcher)
 - sampszi p_1 p_2 , alpha(α) power(*power*)

4

Another Example

- Two drugs for treatment of peptic ulcer compared (Familiari, et al. [1981])
 - The percentage of ulcers healed by pirenzepine (drug A) and trithiozine (drug B) was 77% and 58% based on 30 and 31 patients respectively (p-value = .17)
 - 95% CI for difference in proportions healed ($P_{DRUG A} - P_{DRUG B}$) was (-.04, .42)
 - The power to detect a difference as large as the sample results with samples of size 30 and 31 respectively is only 25%

	Healed	Not healed	Total
Drug A	23	7	30
Drug B	18	13	31

Source: Campbell, J., Machin (1993), Familiari, et al. (1981), Clinical Trial.

5

Example

- As a clinician, you find the sample results intriguing
- You want to do a larger study to better quantify the difference in proportions healed
- Redesign a new trial, using the afore mentioned study results to “gustimate” population characteristics
 - Use $p_{DRUG A} = .77$ and $p_{DRUG B} = .58$
 - 80% power
 - $\alpha = .05$
- Command in Stata
 - sampszi .77 .58, alpha (.05) power (.8)

6

Example

- Command in Stata
 - sampsi .77 .58, alpha (.05) power (.8)
 - . sampsi .77 .58, alpha (.05) power (.8)

Estimated sample size for two-sample comparison of proportions

Test Ho: $p_1 = p_2$, where p_1 is the proportion in population 1 and p_2 is the proportion in population 2

Assumptions:

```
alpha = 0.0500 (two-sided)
power = 0.8000
p1 = 0.7700
p2 = 0.5800
n2/n1 = 1.00
```

Estimated required sample sizes:

n1 =	105
n2 =	105

7

Example

- What would happen if you change power to 90%?
 - sampsi .77 .58, alpha (.05) power (.9)
 - . sampsi .77 .58, alpha (.05) power (.9)

Estimated sample size for two-sample comparison of proportions

Test Ho: $p_1 = p_2$, where p_1 is the proportion in population 1 and p_2 is the proportion in population 2

Assumptions:

```
alpha = 0.0500 (two-sided)
power = 0.9000
p1 = 0.7700
p2 = 0.5800
n2/n1 = 1.00
```

Estimated required sample sizes:

n1 =	136
n2 =	136

8

Example

- Suppose you wanted two times as many people on trithiozone ("Drug B") as compared to pirenzepine ("Drug A")
 - Here, the ratio of sample size for Group 2 to Group 1 is 2.0
- Can use "ratio" option in "sampsi" command

9

Example

- Changing the ratio
 - sampsi .77 .58, alpha (.05) power (.9) ratio(2)
 - . sampsi .77 .58, alpha (.05) power (.9) ratio(2)

Estimated sample size for two-sample comparison of proportions

Test Ho: $p_1 = p_2$, where p_1 is the proportion in population 1 and p_2 is the proportion in population 2

Assumptions:

```
alpha = 0.0500 (two-sided)
power = 0.9000
p1 = 0.7700
p2 = 0.5800
n2/n1 = 2.00
```

Estimated required sample sizes:

n1 =	103
n2 =	206

10

Sample Size for Comparing Two Proportions

- A randomized trial is being designed to determine if vitamin A supplementation can reduce the risk of breast cancer
 - The study will follow women between the ages of 45-65 for one year
 - Women were randomized between vitamin A and placebo
- What sample sizes are recommended?

11

Breast Cancer/Vitamin A Example

- Design a study to have 80% power to detect a 50% relative reduction in risk of breast cancer w/vitamin A
 - (i.e., $RR = \frac{P_{VITA}}{P_{PLACEBO}} = .50$)
- Using a (two-sided) test with significance level α -level = .05
- To get estimates of proportions of interest:
 - Using other studies, the breast cancer rate in the controls can be assumed to be 150/100,000 per year

12

Breast Cancer/Vitamin A Example

- A 50% relative reduction: if $RR = \frac{P_{VITA}}{P_{PLACEBO}} = .50$ then, $P_{VITA} = .50 \times P_{PLACEBO}$
- So, for this desired difference in the relative scale:

$$P_{VITA} = \frac{150}{100,000} \times 0.5 = \frac{75}{100,000}$$
- Notice, that this difference on the absolute scale, $P_{VITA} - P_{PLACEBO}$, is much smaller in magnitude:

$$-\frac{75}{100,000} = -0.00075 = -0.075\%$$

13

Breast Cancer Sample Size Calculation in Stata

- “sampsi” command


```
sampsi .00075 .0015, alpha(.05) power(.8)
. sampsi .00075 .0015, alpha(.05) power(.8)
```

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1 and p2 is the proportion in population 2

Assumptions:

```
alpha = 0.0500 (two-sided)
power = 0.8000
p1 = 0.0008
p2 = 0.0015
n2/n1 = 1.00
```

Estimated required sample sizes:

n1	=	33974
n2	=	33974

14

Breast Cancer Sample Size Calculation in Stata

- You would need about 34,000 individuals per group
- Why so many?
 - The difference between two hypothesized proportions is very small (= .00075)
 - We would expect about 50 cancer cases among the controls and 25 cancer cases among the vitamin A group

$$\text{placebo: } \frac{150}{100,000} \times 34,000 = 51$$

$$\text{vitamin A: } \frac{75}{100,000} \times 34,000 = 25$$

15

Breast Cancer/Vitamin A Example

- Suppose you want 80% power to detect only a 20% (relative) reduction in risk associated with vitamin A
- 20% relative reduction: if $RR = \frac{P_{VITA}}{P_{PLACEBO}} = .80$ then $P_{VITA} = .80 \times P_{PLACEBO}$
- So, for this desired difference in the relative scale:

$$P_{VITA} = \frac{150}{100,000} \times 0.8 = \frac{120}{100,000}$$
- Notice, that this difference on the absolute scale, $P_{VITA} - P_{PLACEBO}$, is much smaller in magnitude:

$$-\frac{30}{100,000} = -0.0003 = -0.03\%$$

16

Breast Cancer Sample Size Calculation in Stata

- “sampsi” command


```
sampsi .0012 .0015, alpha(.05) power(.8)
. sampsi .0012 .0015, alpha(.05) power(.8)
```

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1 and p2 is the proportion in population 2

Assumptions:

```
alpha = 0.0500 (two-sided)
power = 0.8000
p1 = 0.0012
p2 = 0.0015
n2/n1 = 1.00
```

Estimated required sample sizes:

n1	=	241769
n2	=	241769

17

Breast Cancer–Vitamin A Example Revisited

- You would need about 242,000 per group!
- We would expect 360 cancer cases among the placebo group and 290 among vitamin A group

18

An Alternative Approach—Design a Longer Study

- Proposal
 - Five-year follow-up instead of one year
 - Here:
 - ▶ $P_{VITA} = 5 \times .0012 = .006$
 - ▶ $P_{PLACEBO} = 5 \times .0015 = .0075$
- Need about 48,000 per group
 - Yields about 290 cases among vitamin A and 360 cases among placebo
- Issue
 - Loss to follow-up