BIOS 662 Fall 2018

Power and Sample Size, Part I

David Couper, Ph.D.

david_couper@unc.edu

or

couper@bios.unc.edu

https://sakai.unc.edu/portal

Outline

- Introduction
- One sample:
  - continuous outcome, Z test
  - continuous outcome, t test
  - binary outcome, Z test
  - binary outcome, exact test

Introduction

- Choosing an appropriate sample size is not just a study design issue, it is an ethical issue
- For a (new) study to be ethical, it must be designed to have sufficient power to detect clinically meaningful differences
- There are ethical issues even if using already-collected data — wasting resources if the sample size is too small
- Power and sample size are mathematically related
- In some situations we can calculate sample sizes explicitly
- In complicated situations, may need to use simulation to determine the sample size

Introduction

- To estimate sample size we need to specify:
  - The study design
  - The significance level α
  - The null hypothesis
  - The test statistic and its distribution
  - The value θ_A that we want to be able to detect
  - The desired power to detect this θ_A
  - More complex models may require specifying other parameters, such as covariances (for measures taken at multiple time-points or if adjusting for confounders)
- These need to be specified when designing the study

Introduction

- How do we decide which values to use?
  - Some are reasonably standard, e.g. α
  - Obtain estimates from pilot studies or studies done elsewhere, e.g. θ_A, variances, covariances
  - θ_A may be what is regarded as the smallest clinically meaningful effect
  - We often calculate sample size for a few representative values of what the underlying parameters might be
  - We often calculate the sample size for two or more choices of the study power — typically 0.8 and 0.9
  - We often choose a few sample sizes and calculate the associated power

Introduction

- These choices need to be made regardless of how the power / sample sizes will be calculated
- These all assume subjects comply with the treatment group to which they are assigned and that we are able to obtain end-point information on all subjects

Introduction

- Read sections 5.8 and 6.3.3 of the text
- In a test of a hypothesis, we are testing whether some population parameter has a particular value

$$H_0 : \theta = \theta_0,$$

where θ_0 is a known constant

- Usually,

$$H_A : \theta \neq \theta_0$$

- Once the data are collected, we compute a statistic related to θ, say $S(\hat{\theta})$
- $S(\hat{\theta})$ is a random variable, because it is computed from a sample (and hence it has a probability distribution)

Power

$$\Pr[\text{Type I error}] = \alpha = \Pr[S(\hat{\theta}) \in C_\alpha \mid H_0]$$

$$\Pr[\text{Type II error}] = \beta = \Pr[S(\hat{\theta}) \notin C_\alpha \mid H_A]$$

$$\text{Power} = 1 - \beta = \Pr[S(\hat{\theta}) \in C_\alpha \mid H_A]$$

One Sample Z Test

- Example: One sample test
- Study: Collect data for a continuous outcome Y on N individuals
- $E(Y) = \mu, \;\; \mathrm{Var}(Y) = \sigma^2$

$$H_0 : \mu = \mu_0 \;\; \text{vs.} \;\; H_A : \mu > \mu_0$$

$$S(\hat{\theta}) = Z = \frac{\bar{Y} - \mu_0}{\sigma/\sqrt{N}}$$

One Sample Z Test

$$\Pr[S(\hat{\theta}) \in C_\alpha \mid H_0] = \Pr[Z > z_{1-\alpha} \mid H_0] = \alpha$$

$$\Pr[S(\hat{\theta}) \in C_\alpha \mid H_A] = \Pr[Z > z_{1-\alpha} \mid H_A] = 1 - \beta$$

- Choose a value $\mu_A \in H_A$
- The question of interest is: What sample size do we need in order to have power $1 - \beta$ to detect this alternative?
- The sample size depends on the particular choice of μ_A

One Sample Z Test

- Under $H_A : \mu = \mu_A$

$$Z' = \frac{\overline{Y} - \mu_A}{\sigma/\sqrt{N}} \sim N(0,1)$$

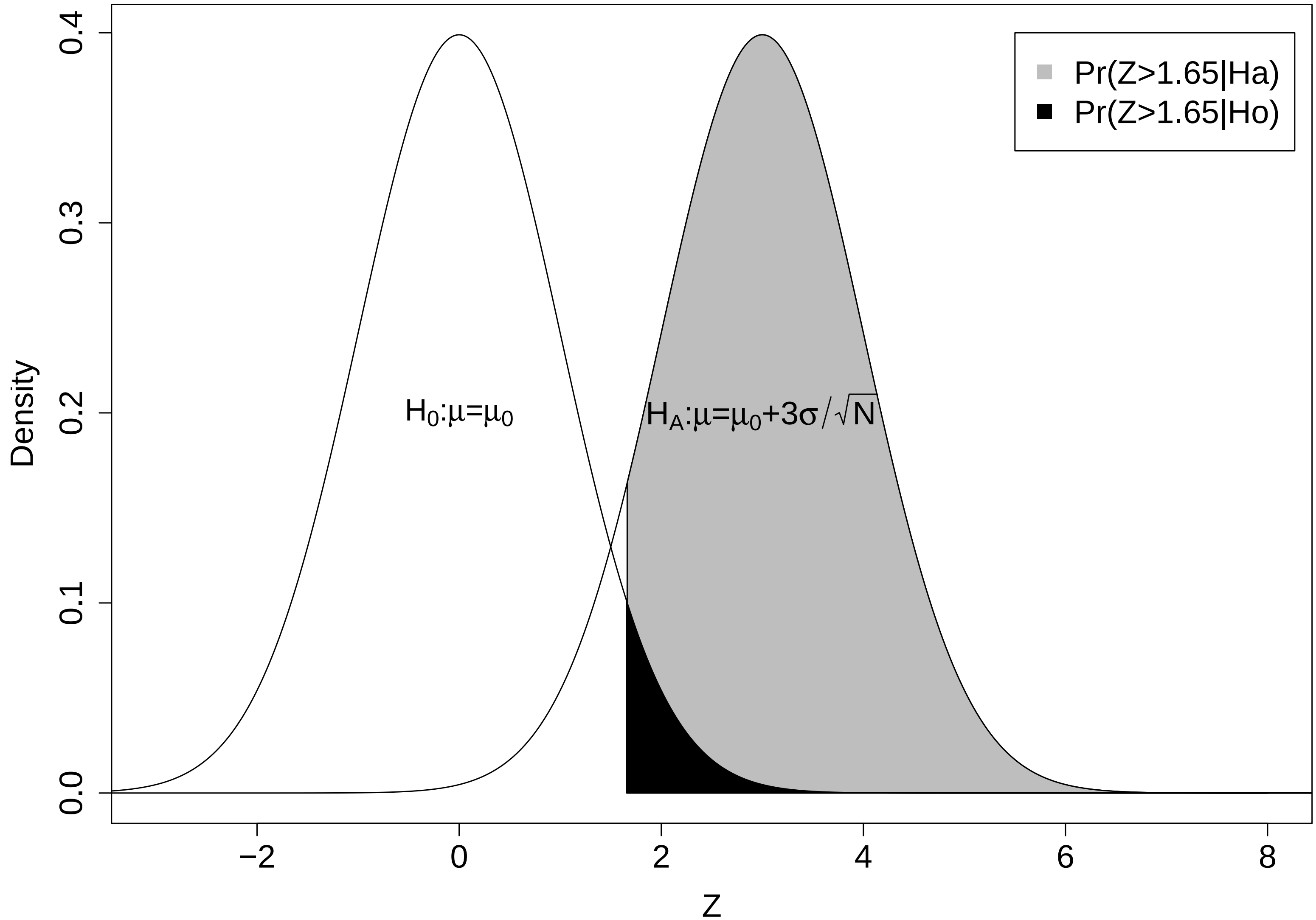
- Note

$$Z = \frac{\overline{Y} - \mu_0}{\sigma/\sqrt{N}} = Z' + \frac{\mu_A - \mu_0}{\sigma/\sqrt{N}}$$

- Thus

$$Z \sim N\left(\frac{\mu_A - \mu_0}{\sigma/\sqrt{N}}, 1\right)$$

Distribution of Z-statistic under H_0 and H_A



One Sample Z Test

- Power is given by

$$1-\beta = \Pr[Z > z_{1-\alpha} \mid \mu = \mu_A]$$

$$= \Pr[Z' > z_{1-\alpha} + \frac{\sqrt{N}(\mu_0 - \mu_A)}{\sigma} \mid \mu = \mu_A]$$

So

$$\beta = \Pr[Z' \leq z_{1-\alpha} + \frac{\sqrt{N}(\mu_0 - \mu_A)}{\sigma} \mid \mu = \mu_A]$$

- Therefore

$$z_{1-\alpha} + \frac{\sqrt{N}(\mu_0 - \mu_A)}{\sigma} = z_\beta = -z_{1-\beta}$$

One Sample Z Test

- Equivalently

$$N = \frac{\sigma^2(z_{1-\alpha} + z_{1-\beta})^2}{(\mu_0 - \mu_A)^2}$$

- For a two sided test

$$N = \frac{\sigma^2(z_{1-\alpha/2} + z_{1-\beta})^2}{(\mu_0 - \mu_A)^2}$$

One Sample Z Test

- Equivalent form

$$N = \left(\frac{z_{1-\alpha/2} + z_{1-\beta}}{\Delta} \right)^2$$

where

$$\Delta = \frac{|\mu_0 - \mu_A|}{\sigma}$$

- Δ is called the *standardized distance* or *difference*

One Sample Z Test

- To apply this formula, we need to know α, β, $|\mu_0 - \mu_A|$, and σ^2
- Example: A study is planned to determine the effect of a drug on blood pressure.
- BP will be measured, a drug administered, and BP measured again 2 hours later
- $Y = (\mathrm{BP}_{\mathrm{after}} - \mathrm{BP}_{\mathrm{before}})$

$$H_0 : \mu = 0 \quad \text{vs.} \quad H_A : \mu \neq 0$$

One Sample Z Test

- Choose $\alpha = 0.05, \;\; 1 - \beta = 0.9$
- Estimate of σ^2: Need data from the literature or a pilot study; note that we need the variance of the difference (after $-$ before)
- The choice of μ_A is a subject-matter decision. In this example, a drug that changes BP by just 1 mmHg would not be of practical importance, but a drug that changes BP by 5 mmHg might be of interest

One Sample Z Test

- Suppose $\alpha = 0.05$, $1 - \beta = 0.9$, $\sigma^2 = 225$, and $\mu_A = 5$

$$N = \frac{225(1.96 + 1.28)^2}{5^2} = 94.5 \approx 95$$

- Often compute N for various different values of α, β, σ^2, and μ_A
- Note that $(1.96 + 1.28)^2 \approx 10.5$, so that

$$N \approx \frac{10.5}{\Delta^2}$$

One Sample Z Test

α	$1-\beta$	σ^2	μ_A	N	α	$1-\beta$	σ^2	μ_A	N
0.05	0.90	225	5	95	0.01	0.90	225	5	133
		225	6	66			225	6	93
		256	5	108			256	5	151
		256	6	75			256	6	105
0.05	0.80	225	5	71	0.01	0.80	225	5	105
		225	6	49			225	6	73
		256	5	81			256	5	119
		256	6	56			256	6	83

One Sample Z Test

$$\alpha \downarrow \Rightarrow \; N \uparrow$$

$$1 - \beta \uparrow \Rightarrow \; N \uparrow$$

$$\sigma^2 \uparrow \Rightarrow \; N \uparrow$$

$$|\mu_0 - \mu_A| \downarrow \Rightarrow \; N \uparrow$$

One Sample Z Test

- Sometimes N is fixed and we want to estimate the power of the test

$$N = \frac{\sigma^2 (z_{1-\alpha/2} + z_{1-\beta})^2}{(\mu_0 - \mu_A)^2}$$

$$z_{1-\beta} = \frac{|\mu_0 - \mu_A|\sqrt{N}}{\sigma} - z_{1-\alpha/2}$$

$$1 - \beta = \Phi\left(\frac{|\mu_0 - \mu_A|\sqrt{N}}{\sigma} - z_{1-\alpha/2}\right)$$

One Sample Z Test: Example

- An investigator says that the budget can accommodate just 50 patients
- Suppose $\alpha = 0.05$, $\sigma^2 = 225$, $|\mu_0 - \mu_A| = 5$
- Then

$$z_{1-\beta} = \frac{5\sqrt{50}}{15} - 1.96 = 0.40$$

$$1 - \beta = 0.65$$

One Sample t Test

- In practice, σ is not known, so we use a t test instead of a Z test
- The previous results should be viewed as approximations in this case
- Now derive the power for a t test
- Need the following result: If $U \sim N(\lambda, 1)$ and $V \sim \chi^2_\nu$ with $U \perp V$, then

$$\frac{U}{\sqrt{V/\nu}} \sim t_{\nu,\lambda}$$

that is, a non-central t distribution with ν degrees of freedom and non-centrality parameter λ

One Sample t Test

- Consider $H_0 : \mu = 0$ versus $H_A : \mu = \mu_A$ for $\mu_A \neq 0$
- Under H_A,

$$\overline{Y} \sim N(\mu_A, \sigma^2/N)$$

- Thus

$$\frac{\overline{Y}\sqrt{N}}{\sigma} \sim N\left(\mu_A \frac{\sqrt{N}}{\sigma}, 1\right)$$

- Recall

$$\frac{(N-1)s^2}{\sigma^2} \sim \chi^2_{N-1}$$

One Sample t Test

- Because $\overline{Y} \perp s^2$,

$$T = \frac{\overline{Y}}{s/\sqrt{N}} \sim t_{N-1,\lambda}$$

where $\lambda = \mu_A \sqrt{N}/\sigma$

- So the power of a two-sided t test for $\mu_A > 0$ is

$$\Pr[T \geq t_{N-1;1-\alpha/2}]$$

where $T \sim t_{N-1,\mu_A \sqrt{N}/\sigma}$

One Sample t Test: R

```
# by hand
> 1-pt(qt(0.975,49), 49, 5*sqrt(50)/15)
[1] 0.6370846

> power.t.test(n=50, sd=15, delta=5, type="one.sample")

     One-sample t test power calculation

              n = 50
          delta = 5
             sd = 15
      sig.level = 0.05
          power = 0.6370846
    alternative = two.sided
```

One Sample t Test: SAS

```
proc power;
   onesamplemeans
   mean   = 5
   ntotal = 50
   stddev = 15
   power  = .;
```

```
One-sample t Test for Mean

     Fixed Scenario Elements
Distribution                Normal
Method                       Exact
Mean                             5
Standard Deviation              15
Total Sample Size               50
Number of Sides                  2
Null Mean                        0
Alpha                         0.05

Computed Power

0.637
```

One Sample Z Test: Binary Outcome

- Null hypothesis

$$H_0 : \pi = \pi_0$$

- Test statistic

$$Z = \frac{\hat{p} - \pi_0}{\sqrt{\pi_0(1-\pi_0)/N}}$$

- Sample size formula

$$N = \frac{\left(z_{1-\alpha/2}\sqrt{\pi_0(1-\pi_0)} + z_{1-\beta}\sqrt{\pi_A(1-\pi_A)}\right)^2}{(\pi_A - \pi_0)^2}$$

One Sample Z Test, Binary Outcome: Example

- Study of risk of breast cancer in siblings.
- Prevalence in population of women 50-54 years old: 2%
- Plan to sample sisters of women with breast cancer and test $H_0 : \pi = 0.02$ versus $H_A : \pi \neq 0.02$
- Suppose $\alpha = 0.05$, $1 - \beta = 0.9$, $\pi_A = 0.05$
- Then

$$N = \frac{\left(1.96\sqrt{0.02(0.98)} + 1.28\sqrt{0.05(0.95)}\right)^2}{(0.05 - 0.02)^2} = 340.24$$

- Round up to 341

One Sample Z Test With Binary Outcome: SAS

```
proc power;
   onesamplefreq test = z
   method = normal
   nullp  = 0.02
   p      = 0.05
   power  = 0.9
   ntotal = .;
```

```
The POWER Procedure
Z Test for Binomial Proportion

          Fixed Scenario Elements
Method                  Normal approximation
Null Proportion                         0.02
Binomial Proportion                     0.05
Nominal Power                            0.9
Number of Sides                            2
Alpha                                   0.05

Actual        N
 Power    Total
 0.900      341
```

One Sample Exact Test: Binary Outcome

- What is the power of the exact test?
- Let Y, the number of successes, be the test statistic.
- Under H_0, $Y \sim \text{Binomial}(N, \pi_0)$

  Under H_A, $Y \sim \text{Binomial}(N, \pi_A)$
- Power

$$\Pr[Y \geq y_{1-\alpha/2} \mid \pi = \pi_A] + \Pr[Y \leq y_{\alpha/2} \mid \pi = \pi_A]$$

  where $y_{\alpha/2}$ and $y_{1-\alpha/2}$ are determined as in the notes on "Count Data"

Exact Test, Binary Outcome: Example

- Suppose $N = 20$, $\pi_0 = 0.2$, $\pi_A = 0.5$, $\alpha = 0.05$
- What is exact power of a two-sided exact test?
- Based on the CDF of $\text{Binomial}(20, 0.2)$, choose $y_{\alpha/2} = 0$ and $y_{1-\alpha/2} = 9$
- Then the power is

$$\Pr[Y \geq 9 \mid \pi = 0.5] + \Pr[Y \leq 0 \mid \pi = 0.5] = 0.748$$

Exact Test With Binary Outcome: SAS

```
proc power;
   onesamplefreq test = exact
   method = exact
   nullp  = 0.2
   p      = 0.5
   power  = .
   ntotal = 20;

          Computed Power

 Lower    Upper
 Crit     Crit     Actual
  Val      Val      Alpha    Power

    0        9     0.0215    0.748
```