

Statistical Power

(morning)

Objectives

- Review the hypothesis testing and type I and type II errors; which error is more serious?
- Compute power; understand relationship and trade-offs against other factors (significance level, effect size, variance, and sample size) that influence power
- Morning assignment: Compute power and sample size

Hypothesis Testing: Possible Outcomes

	H_0 is true	H_0 is false
Fail to reject H_0	Correct Decision ($1 - \alpha$)	Type II Error (β)
Reject H_0	Type I Error (α)	Correct Decision ($1 - \beta = \pi$)

Type I and Type II Errors

Type I Error

$$= P(\text{Reject } H_0 \mid H_0 \text{ is true})$$

$$= \alpha \text{ (alpha, the significance level)}$$

Type II Error

$$= P(\text{Fail to reject } H_0 \mid H_0 \text{ is false})$$

$$= \beta \text{ (beta)} = 1 - \pi \text{ (one minus power)}$$

Which Type of Errors is More Serious?

Example 1: You have been put on trial for murder

H_0 : You are innocent (presumption of innocence)

H_A : You are guilty

The question for the jury is whether it finds enough evidence that you are guilty

Example 1: You have been put on trial for murder (cont.)

	H_0 is true You are truly innocent	H_0 is false You are truly guilty
Fail to reject H_0 Did not found enough evidence that you are guilty: You are declared innocent	Correct	Type II Error You are guilty but set free
Reject H_0 Found enough evidence that you are guilty: You are declared guilty	Type I Error You are found guilty of a murder that you did not commit (and the murderer is still free...) The American justice system puts a lot of emphasis on avoiding type I errors	Correct

Example 2: You are being screened for a disease

H_0 : You don't have the disease

H_A : You have the disease

The question for the test is whether it finds enough evidence (e.g., markers) that you have the disease

Example 2: You are being screened for a disease (cont.)

	H_0 is true You truly are disease-free	H_0 is false You truly have the disease
Fail to reject H_0 Did not found enough evidence that you are sick: You are declared to be disease-free	Correct	Type II Error You'll have the incorrect assurance that you don't have the disease As a result, you won't be treated for your disease
Reject H_0 Found enough evidence that you are sick: You are declared to have the disease	Type I Error You'll be anxious for a while but this will lead to other testing procedures. Eventually, you'll discover that the initial test was incorrect	Correct

Example 3: You are working on a new drug

H_0 : The new drug is not more effective than current drugs

H_A : The new drug more effective than current drugs

Underpowered study:

With a type II error, you would fail to reject that the new drug is not more effective than current drugs while said drug being really more effective

Minimizing the type II error (β)

Is equivalent to maximize π

$$= 1 - \beta$$

$$= 1 - P(\text{Fail to reject } H_0 \mid H_0 \text{ is false})$$

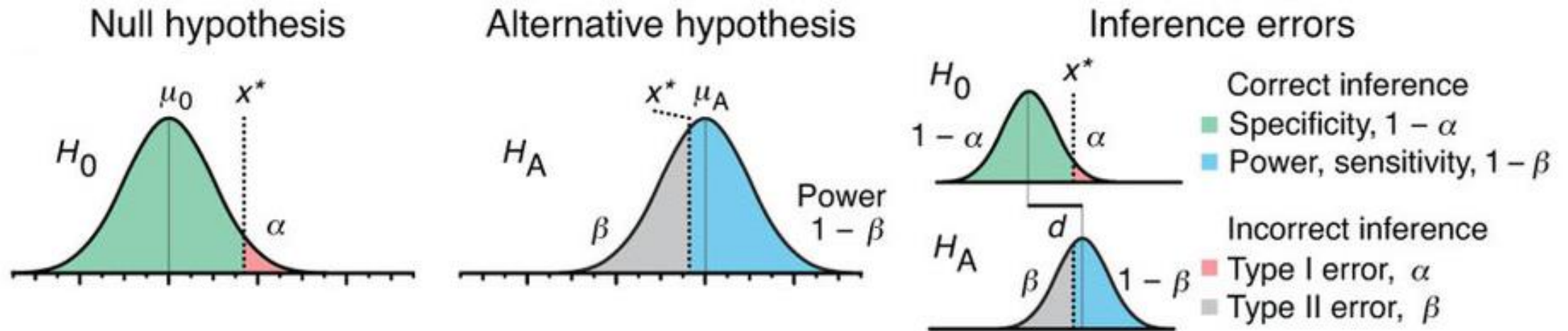
$$= P(\text{Reject } H_0 \mid H_0 \text{ is false})$$

$$= P(\text{Reject } H_0 \mid H_A \text{ is true})$$

Hypothesis Testing: Possible Outcomes

	H_0 is true	H_0 is false
Fail to reject H_0	Correct Decision ($1 - \alpha$)	Type II Error (β)
Reject H_0	Type I Error (α)	Correct Decision ($1 - \beta = \pi$)

Hypothesis Testing: Outcomes and Regions



Compute Statistical Power of a Test

Example: One-sample Test of Mean

$H_0: \mu = \mu_0$ (e.g., the test finds no evidence of the better effectiveness of the new drug)

$H_A: \mu = \mu_1 (> \mu_0)$ (e.g., the test finds evidence of the better effectiveness of the new drug)

Compute Statistical Power of a Test

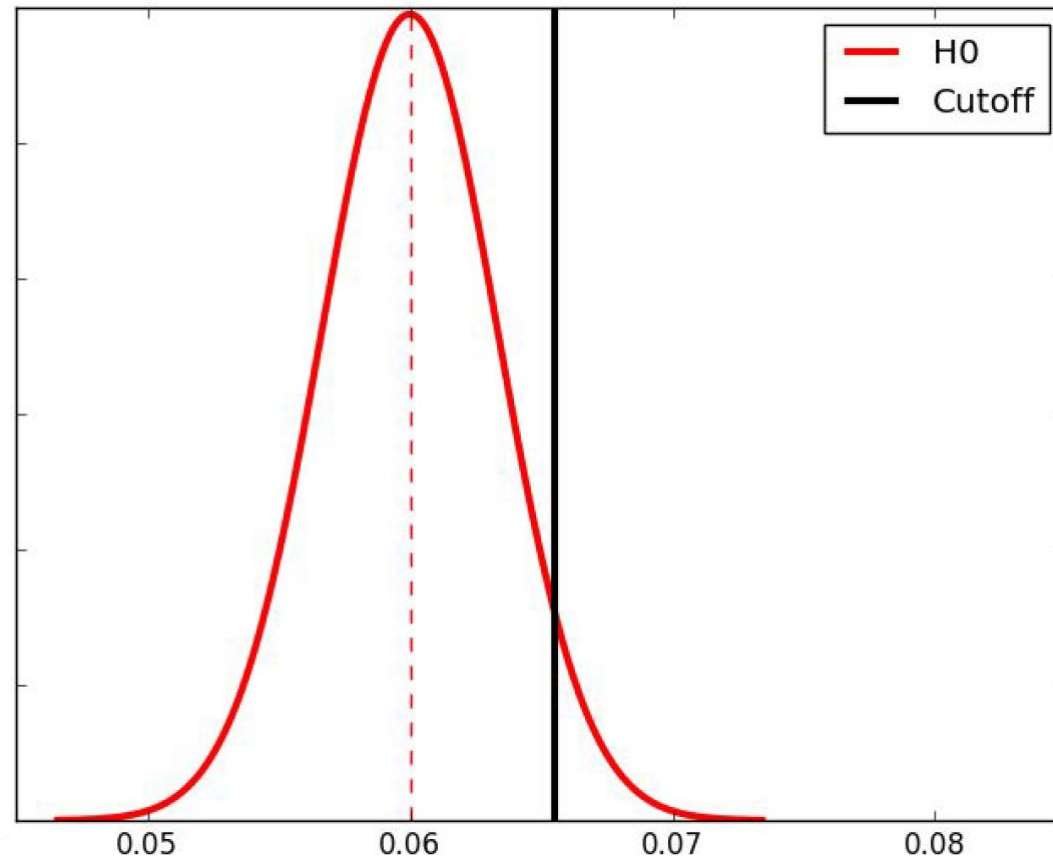
- First, we assume that the null hypothesis (H_0) is ***false***

So, we reject H_0 when:

$$Z = \frac{X - \mu_0}{s / \sqrt{n}} \geq Z_{1-\alpha}$$

$$X \geq X^* = \mu_0 + Z_{1-\alpha} \frac{s}{\sqrt{n}} = \mu_0 - Z_{\alpha} \frac{s}{\sqrt{n}}$$

Compute Statistical Power of a Test, Step 1



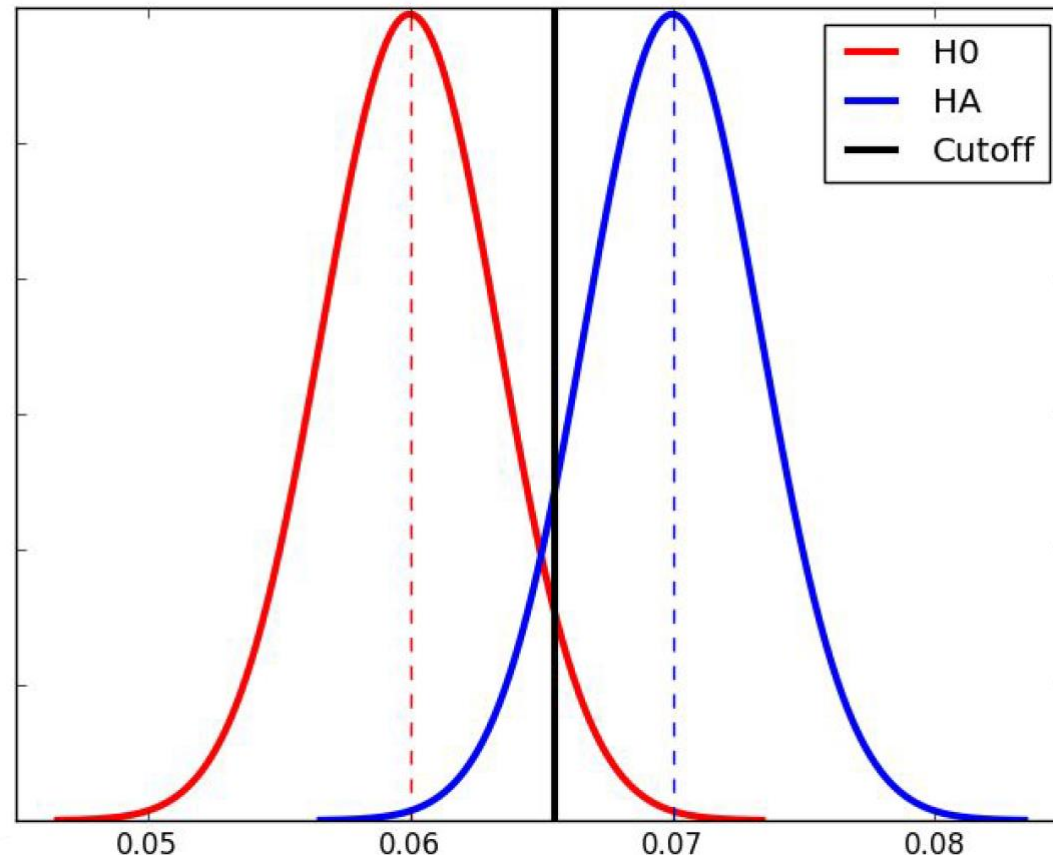
Compute Statistical Power of a Test (cont.)

- Then, we find the corresponding cut-off of this value under H_A :

$$X^* = \mu_1 - Z_{1-\beta} \frac{s}{\sqrt{n}} = \mu_1 + Z_{\beta} \frac{s}{\sqrt{n}}$$

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s/\sqrt{n}} = 0$$

Compute Statistical Power of a Test, Step 2



Statistical Power and its relationship with other Factors

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s / \sqrt{n}} = 0$$

The equation above links the following variables:

- α (type I error; significance level)
- β (type II error; $\pi = 1 - \beta$, the statistical power)
- $\mu_1 - \mu_0$ (effect size)
- s (standard deviation)
- n (sample size)

Define 4 of these variables and solve for the remaining one

Power Calculation

1. Assume that the null hypothesis (H_0) is *false*
2. Given:
 - a. Significance level, α
 - b. Effect size, $\mu_1 - \mu_0$
 - c. Sample size, n
 - d. Standard deviation, s

Calculate Power: $P[\text{reject } H_0 \mid H_0 \text{ is false}] = 1 - \beta$

Power Calculation Example

Given (for A/B testing of website)

- a. Significance level, $\alpha = .05$
- b. Effect size, $\mu_0 = .06$; $\mu_1 = .07$
- c. Sample size, $n = 5,000$
- d. Standard deviation, $s = .237$ (*known since $s^2 = p(1-p)$ for proportions*)

Calculate Power: $P[\text{reject } H_0 \mid H_0 \text{ is false}] = 1 - \beta$

Compute Statistical Power of a Test (cont.)

- $\alpha = .05$
- $Z_\alpha = -1.645$
- $Z_\beta = -\left(\frac{\mu_1 - \mu_0}{s/\sqrt{n}} + Z_\alpha\right) = -\left(\frac{.07 - .06}{.237/\sqrt{5000}} + 1.645\right) = -1.351$
- $\beta = .088$
- $\pi = 1 - \beta = .912$

The power of this test is 91%

A. Type I Error and Type II Error Trade-Off

Assuming all other variables are fixed

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s / \sqrt{n}} = 0$$

If $\alpha \downarrow$ (Type I Error increases):
?

A. Type I Error and Type II Error Trade-Off (cont.)

Assuming all other variables are fixed

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s / \sqrt{n}} = 0$$

If $\alpha \downarrow$ (Type I Error increases):

- $Z_{\alpha} \downarrow$
- $Z_{\beta} \uparrow$
- $\beta \uparrow$ (Type II error increases)
- $1 - \beta \downarrow$ (Power decreases)

B. Power and Effect Size Trade-Off

Assuming all other variables are fixed

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s / \sqrt{n}} = 0$$

If $1 - \beta \uparrow$ (Power increases):
?

B. Power and Effect Size Trade-Off (cont.)

Assuming all other variables are fixed

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s / \sqrt{n}} = 0$$

If $1 - \beta \uparrow$ (Power increases):

- $\beta \downarrow$
- $Z_{\beta} \downarrow$
- $\mu_1 - \mu_0 \uparrow$ (Effect size increases)

C. Power and Sample Size Trade-Off

Assuming all other variables are fixed

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s / \sqrt{n}} = 0$$

If $1 - \beta \uparrow$ (Power increases):
?

C. Power and Sample Size Trade-Off (cont.)

Assuming all other variables are fixed

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s/\sqrt{n}} = 0$$

If $1 - \beta \uparrow$ (Power increases):

- $\beta \downarrow$
- $Z_{\beta} \downarrow$
- $\frac{\mu_1 - \mu_0}{s/\sqrt{n}} \uparrow$
- $s/\sqrt{n} \downarrow$
- $n \uparrow$ (Sample size increases)

(as your sample size increases, the power of your test increases)

D. Sample Size and Effect Size Trade-Off

Assuming all other variables are fixed

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s / \sqrt{n}} = 0$$

If $n \uparrow$ (Sample size increases):
?

D. Sample Size and Effect Size Trade-Off (cont.)

Assuming all other variables are fixed

$$Z_{\alpha} + Z_{\beta} + \frac{\mu_1 - \mu_0}{s / \sqrt{n}} = 0$$

If $n \uparrow$ (Sample size increases):

- $s / \sqrt{n} \downarrow$
- $\mu_1 - \mu_0 \downarrow$ (Effect size decreases)

Power analysis allows you to determine the sample size needed to detect a particular effect (with more samples, you can detect smaller size effects)

Morning assignment