15.482 Healthcare Finance
Spring 2017

Andrew W. Lo, MIT

Unit 6, Part 2: Statistical Framework for Clinical Trials

Unit Outline

- Overview of the Drug Development Process
- Randomized Clinical Trial Design
- Size, Power, and Cost
- Formal Statistical Analysis

Size, Power, and Cost

Example: Balanced Two-Arm RCT

Consider Drug That Has Quantifiable Range of Impact

$$T_i \overset{\text{IID}}{\sim} \mathcal{N}(\mu_t, \sigma^2) \quad , \quad X_i \overset{\text{IID}}{\sim} \mathcal{N}(\mu_x, \sigma^2) \quad , \quad i = 1, \ldots, n$$

- Null hypothesis H_0: $\delta \equiv \mu_t - \mu_x = 0$ (no effect)
- Alternative hypothesis H_1: $\delta > 0$ (positive effect)
- How to decide between H_0 and H_1?

$$\hat{\delta} \;\equiv\; \hat{\mu}_t - \hat{\mu}_x \;=\; \frac{1}{n}\sum_{i=1}^{n} T_i - \frac{1}{n}\sum_{i=1}^{n} X_i$$

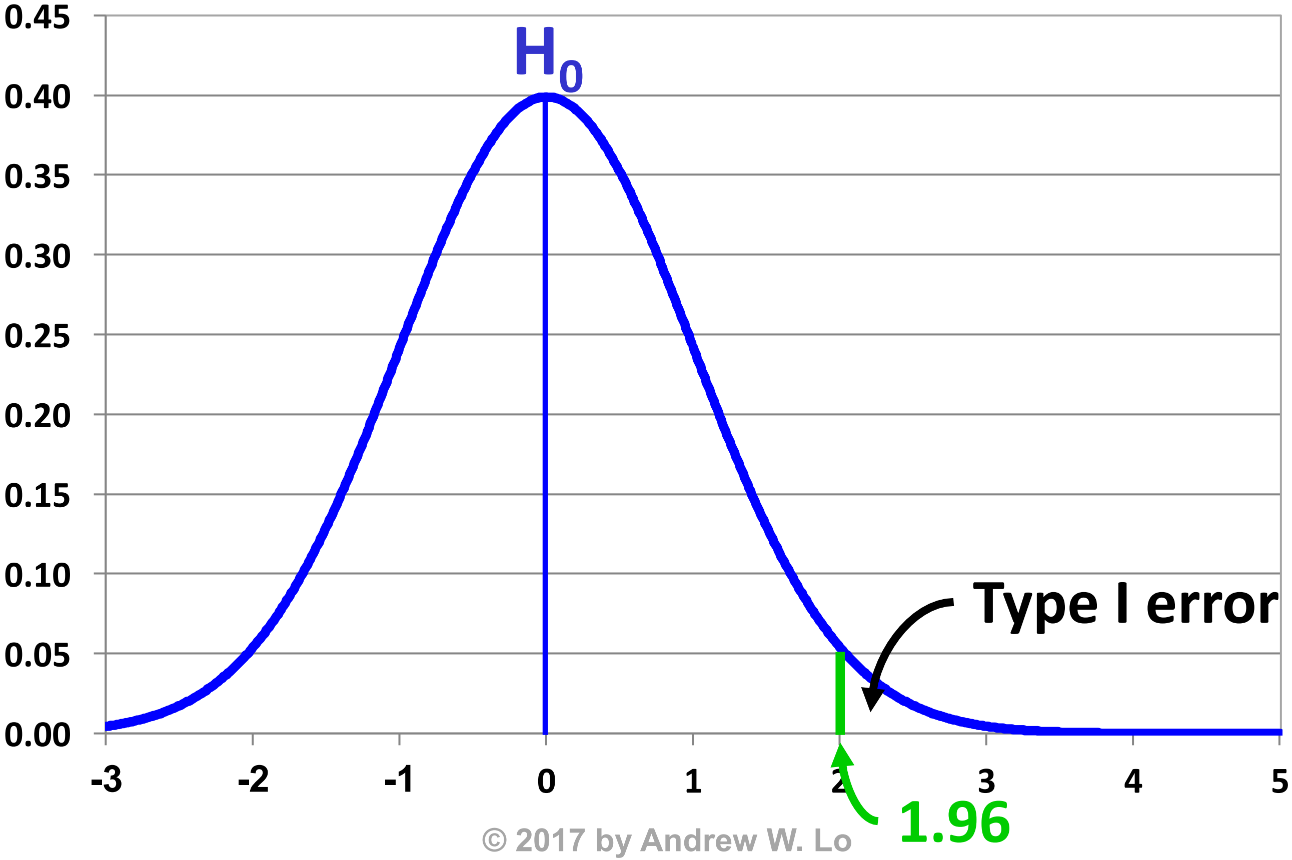
Is this difference statistically significant?

Example: Balanced Two-Arm RCT

Consider Drug That Has Quantifiable Range of Impact

- $\hat{\delta}$ is a random variable; we need its distribution to determine whether data is consistent with null or alternative
- $Z_n \equiv (\hat{\delta} - \mathrm{E}[\hat{\delta}]) \,/\, \mathrm{SD}[\hat{\delta}] = \hat{\delta} \,/\, \mathrm{SD}[\hat{\delta}] \sim \mathcal{N}(0,1)$ under H_0
- Decision rule: if $Z_n \leq c$, accept H_0 (reject drug)
  if $Z_n > c$, accept H_1 (approve drug)
- How to pick c?? Usual choice is 1.96, but **why**??

Example: Balanced Two-Arm RCT



Size and Power

	Decision	
	Approve	Reject
Ineffective Drug		
Effective Drug		

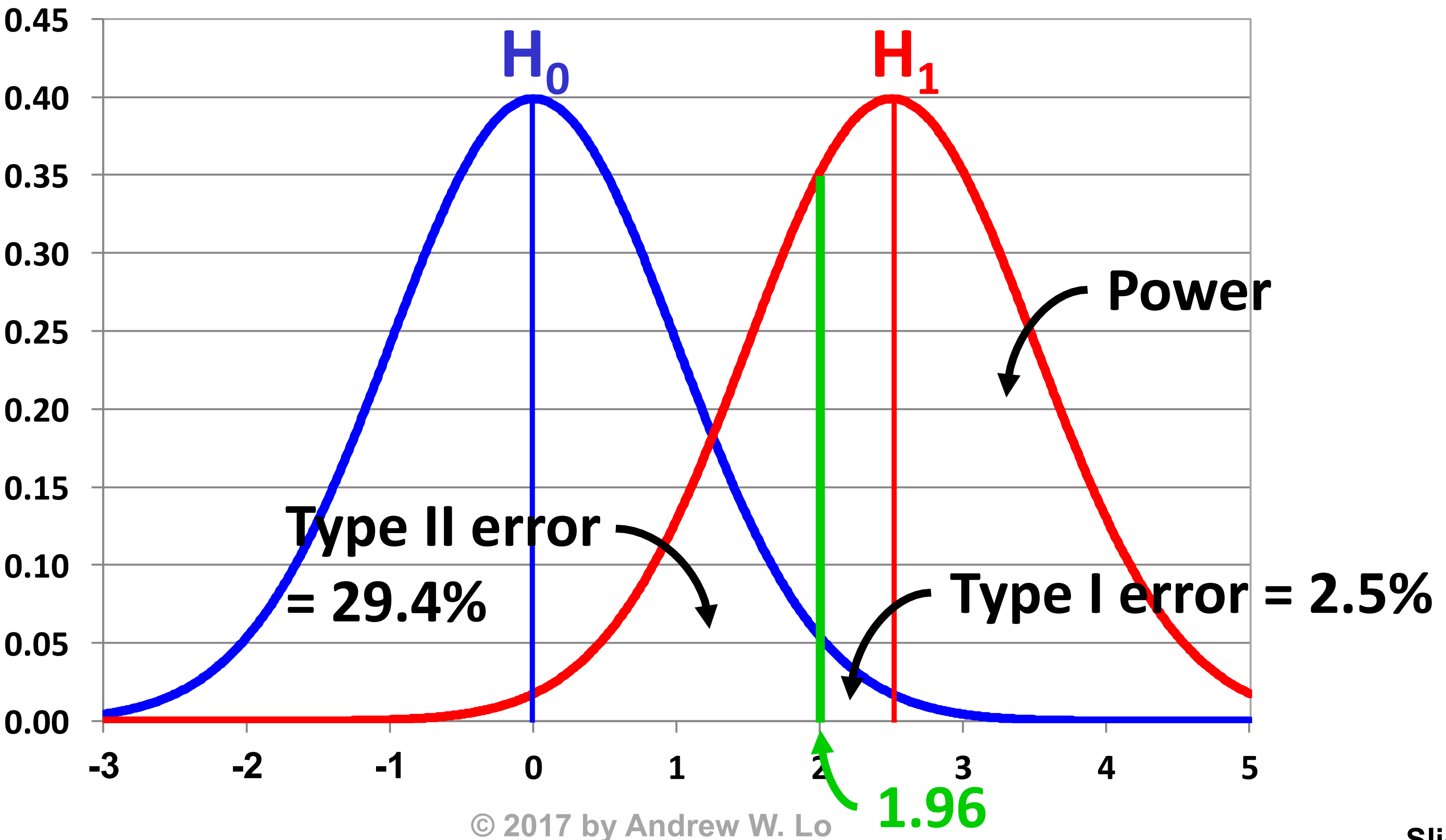
- Goal: minimize Type I and II errors by setting the threshold for approval (critical size of treatment effect)
- But there's a trade-off between these errors

Size and Power

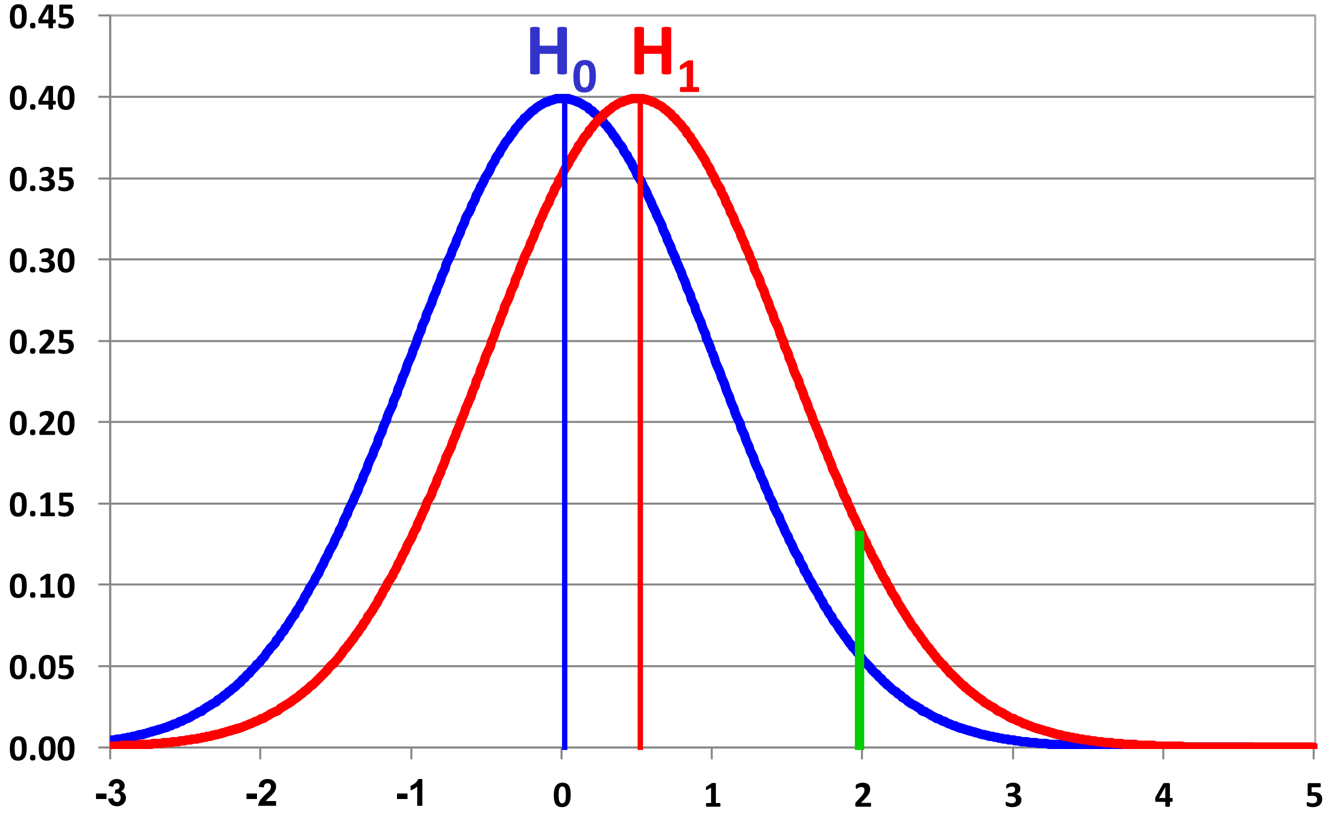
Trade-Off Also Exists In Criminal Justice

- "It is better that ten guilty persons escape than that one innocent suffer," W. Blackstone (1765) ("Blackstone ratio," 10:1)
- "It is better 100 guilty Persons should escape than that one innocent Person should suffer," B. Franklin (1785)
- But this has implications for the amount of crime that may occur (and criminals on the loose)

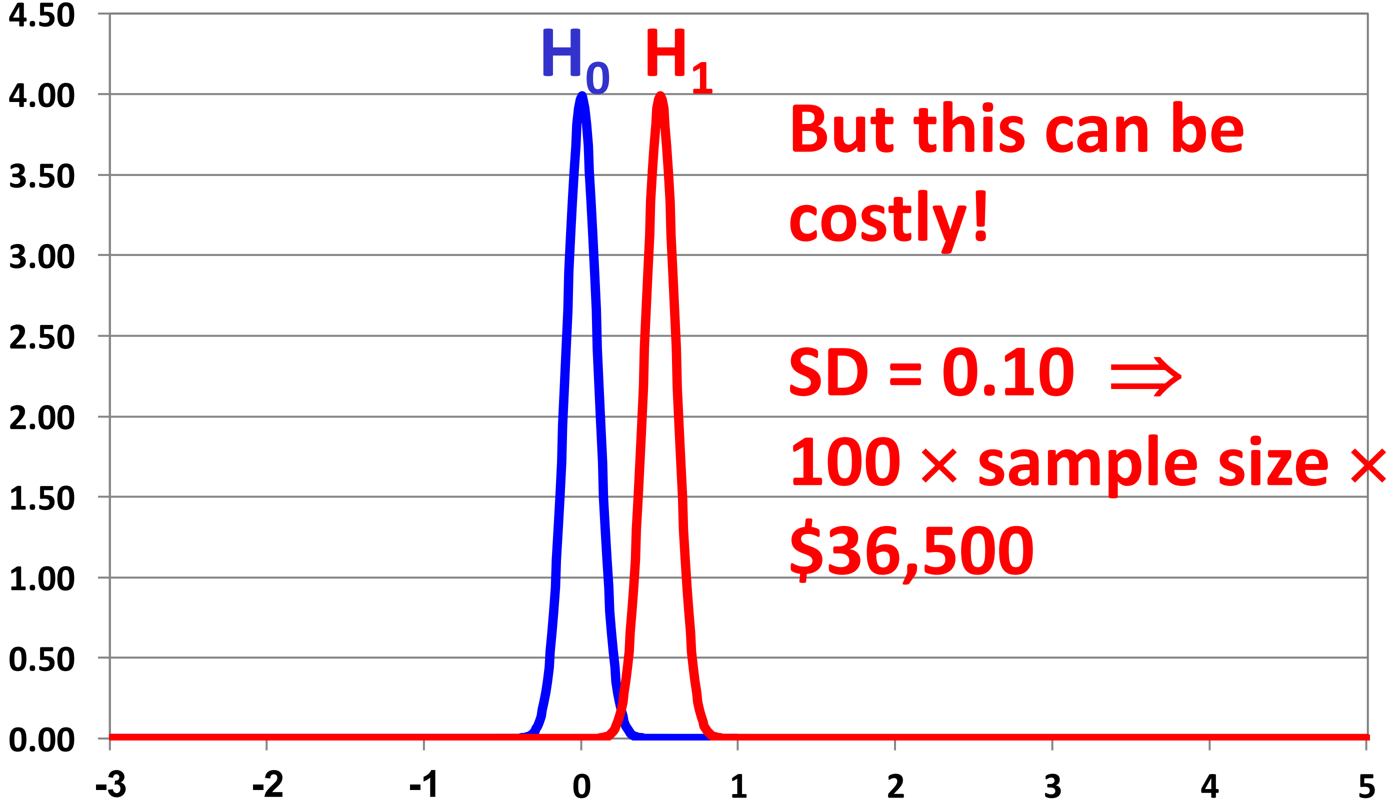
Trade-Off Between Size and Power



Weak Treatment Effect ⇒ Lower Power



More Data Increases Power, At A Cost



But this can be costly!

SD = 0.10 $\Rightarrow$

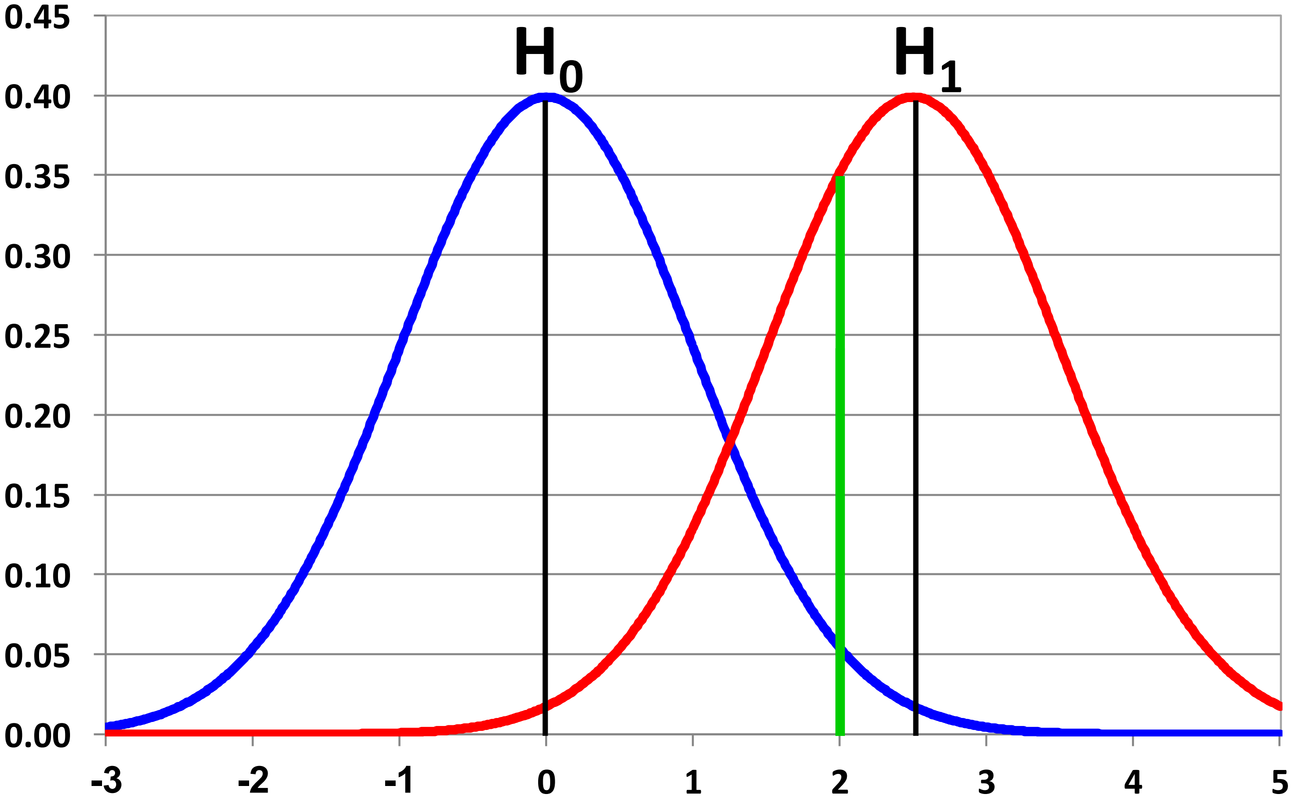
100 × sample size ×

$36,500

Changing Threshold *c*

- Should all drug candidates be held to the same Type I error?
- Pancreatic cancer vs. toenail fungus
- "Patient values vs. *p*-values" D. Berry
- Is there a "Blackstone ratio" for clinical trials?
- See Montazerhodjat, Chaudhuri, Sargent, Lo (*JAMA Oncology*, 2017)

Changing Threshold c



Formal
Statistical Analysis

Key Parameters

$$
\begin{aligned}
\alpha &= \text{Type I error} \\
\beta &= \text{Type II error} \quad , \quad 1-\beta \quad = \quad \text{Power} \\
n &= \text{Sample Size for Each Arm} \\
X_i &\sim \mathcal{N}(\mu_x, \sigma^2) \quad = \quad \text{Patient } i\text{'s Response, Control Arm} \\
T_i &\sim \mathcal{N}(\mu_t, \sigma^2) \quad = \quad \text{Patient } i\text{'s Response, Treatment Arm} \\
\delta &= \mu_t - \mu_x \quad = \quad \text{Expected Treatment Effect Size} \\
\sigma^2 &= \text{Variance of Patient Outcome} \\
\text{H}_0 &= \text{Null Hypothesis, } \delta = 0 \\
\text{H}_1 &= \text{Alternative Hypothesis, } \delta = \kappa\sigma > 0
\end{aligned}
$$

Null Hypothesis and Size

- Under null hypothesis H_0: $\delta \equiv \mu_t - \mu_x = 0$ (no effect):

$$
\begin{aligned}
\alpha &= \text{Prob}(-c < \hat{\delta} < c) \quad (\alpha\% \text{ Confidence Interval}) \\
\frac{\alpha}{2} &= \text{Prob}(\hat{\delta} > c) = \text{Prob}\left(\frac{\hat{\delta}}{\sqrt{\text{Var}[\hat{\delta}]}} > \frac{c}{\sqrt{\text{Var}[\hat{\delta}]}}\right) \\
&= \text{Prob}\left(Z > \frac{c}{\sigma\sqrt{2/n}}\right) = 1 - \text{Prob}\left(Z \le \frac{c}{\sigma\sqrt{2/n}}\right) \\
\frac{\alpha}{2} &= 1 - \Phi\left(\frac{c}{\sigma\sqrt{2/n}}\right) \Rightarrow \frac{c}{\sigma\sqrt{2/n}} = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \equiv z_{\alpha/2}
\end{aligned}
$$

Null Hypothesis and Size

- Under null hypothesis H_0: $\delta \equiv \mu_t - \mu_x = 0$ (no effect):

$$c \;=\; \sigma z_{\alpha/2}\sqrt{2/n}$$

- For $\alpha = 5\%$, $z_{\alpha/2} = 1.96$
- Threshold c decreases with sample size n, and increases with σ and $z_{\alpha/2}$ (does this make sense?)

Alternative Hypothesis and Power

- Under alternative hypothesis H_1: $\delta \equiv \mu_t - \mu_x = \kappa\sigma > 0$

$$\beta = \text{Prob}(\hat{\delta} \leq c) \quad \text{(Type II error)}$$

$$= \text{Prob}\left(\frac{\hat{\delta} - \kappa\sigma}{\sqrt{\text{Var}[\hat{\delta}]}} \leq \frac{c - \kappa\sigma}{\sqrt{\text{Var}[\hat{\delta}]}}\right) = \text{Prob}\left(Z \leq \frac{c - \kappa\sigma}{\sigma\sqrt{2/n}}\right)$$

$$\beta = \Phi\left(\frac{c - \kappa\sigma}{\sigma\sqrt{2/n}}\right) = \Phi\left(\frac{\sigma z_{\alpha/2}\sqrt{2/n} - \kappa\sigma}{\sigma\sqrt{2/n}}\right) = \Phi\left(z_{\alpha/2} - \frac{\kappa}{\sqrt{2/n}}\right)$$

$$\Phi^{-1}(\beta) = z_{\alpha/2} - \frac{\kappa}{\sqrt{2/n}} \Rightarrow z_\beta = z_{\alpha/2} - \frac{\kappa}{\sqrt{2/n}}$$

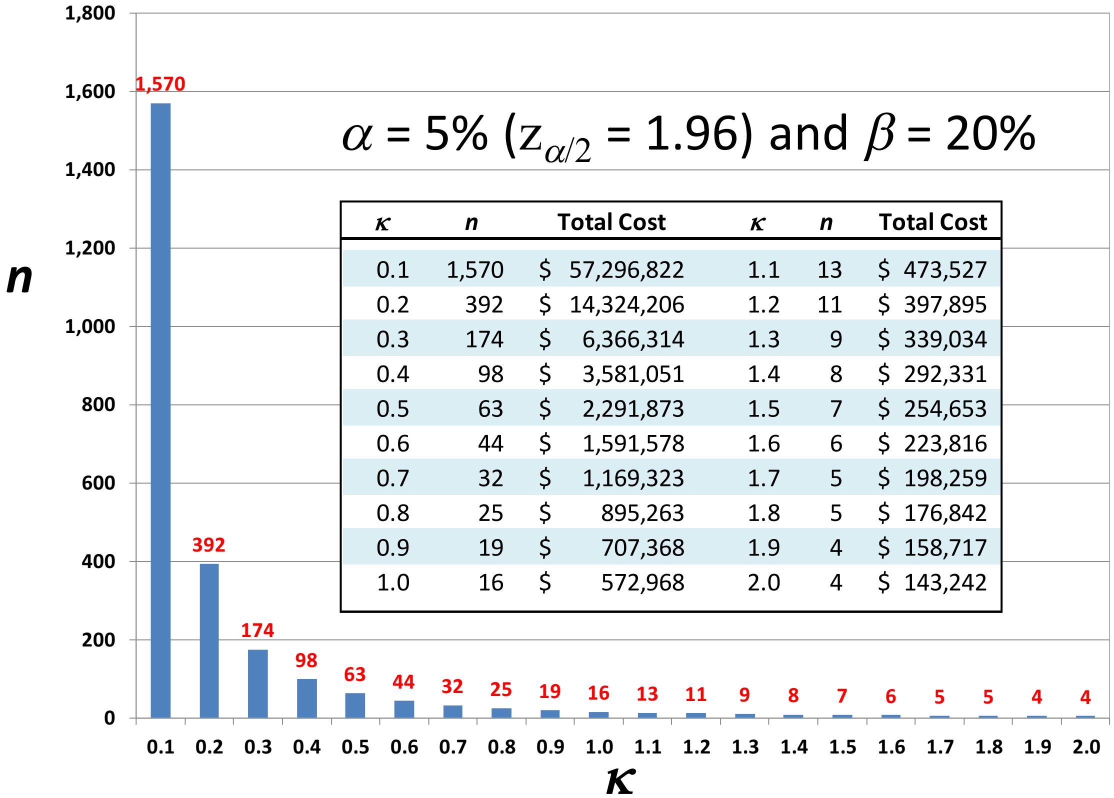
Alternative Hypothesis and Power

- Under alternative hypothesis H_1: $\delta \equiv \mu_t - \mu_x = \kappa\sigma > 0$

$$n = 2\left(\frac{z_{\alpha/2} - z_{\beta}}{\kappa}\right)^2$$

- Sample size *n* decreases with κ, increases with power $(1-\beta)$
- How should we choose size and power?
- Suppose α = 5% ($z_{\alpha/2}$ = 1.96) and β = 20% (8:1 ratio)

Alternative Hypothesis and Power



α = 5% ($z_{\alpha/2}$ = 1.96) and β = 20%

κ	*n*	Total Cost	κ	*n*	Total Cost
0.1	1,570	$ 57,296,822	1.1	13	$ 473,527
0.2	392	$ 14,324,206	1.2	11	$ 397,895
0.3	174	$ 6,366,314	1.3	9	$ 339,034
0.4	98	$ 3,581,051	1.4	8	$ 292,331
0.5	63	$ 2,291,873	1.5	7	$ 254,653
0.6	44	$ 1,591,578	1.6	6	$ 223,816
0.7	32	$ 1,169,323	1.7	5	$ 198,259
0.8	25	$ 895,263	1.8	5	$ 176,842
0.9	19	$ 707,368	1.9	4	$ 158,717
1.0	16	$ 572,968	2.0	4	$ 143,242

So How Big Is Your κ?

More Later:

- How does financing affect these decisions?
- Biotech vs. pharma?
- Should we consider different values of α?

Review of Clinical Trial Statistical Analysis

Example: Balanced Two-Arm RCT

Consider Drug That Has Quantifiable Range of Impact

$$T_i \overset{\text{IID}}{\sim} \mathcal{N}(\mu_t, \sigma^2) \quad , \quad X_i \overset{\text{IID}}{\sim} \mathcal{N}(\mu_x, \sigma^2) \quad , \quad i = 1, \ldots, n$$

- Null hypothesis H_0: $\delta \equiv \mu_t - \mu_x = 0$ (no effect)
- Alternative hypothesis H_1: $\delta > 0$ (positive effect)
- How to decide between H_0 and H_1?

$$\hat{\delta} \;\equiv\; \hat{\mu}_t \;-\; \hat{\mu}_x \;=\; \frac{1}{n}\sum_{i=1}^{n} T_i \;-\; \frac{1}{n}\sum_{i=1}^{n} X_i$$

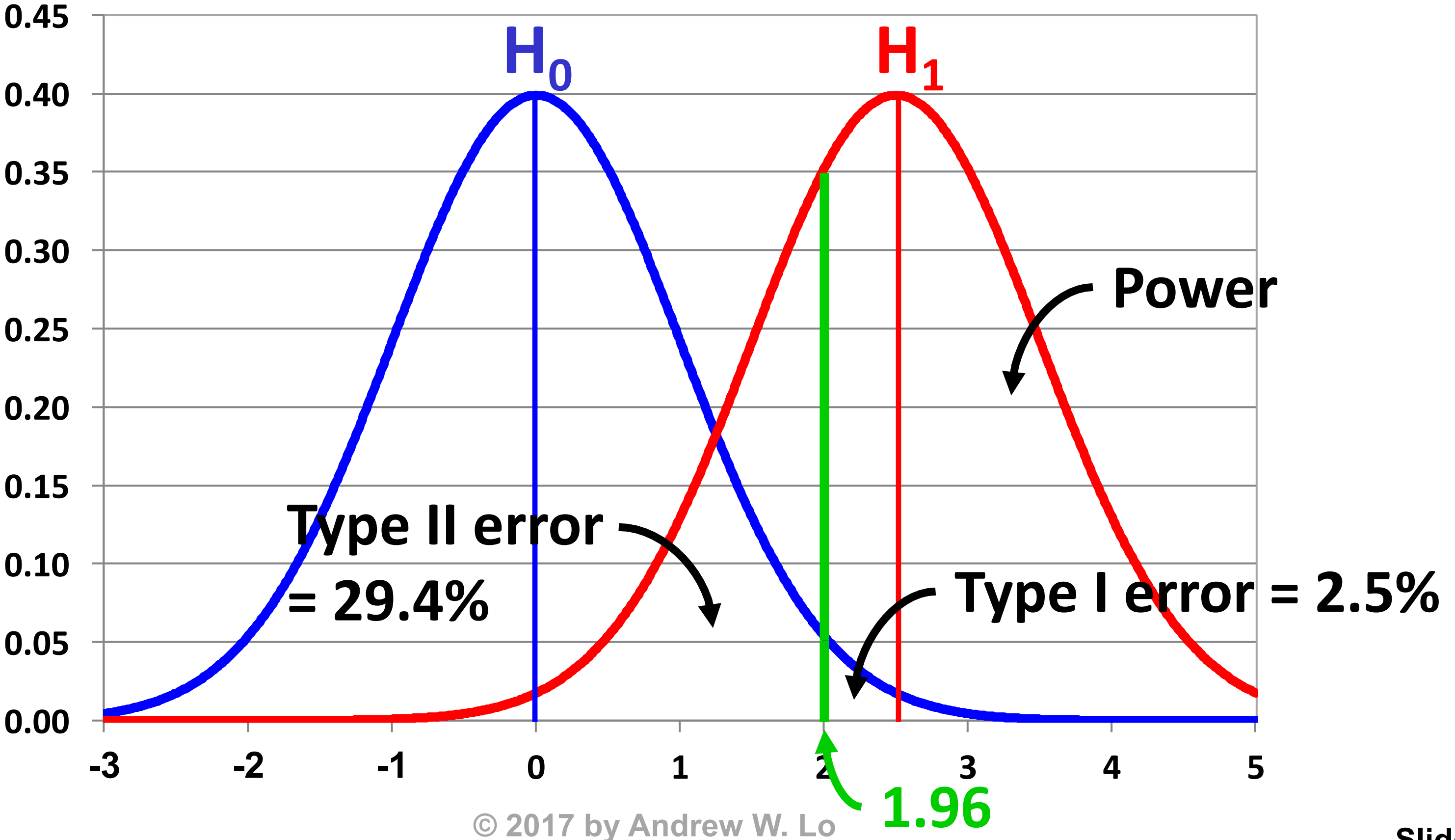
Is this difference statistically significant?

Size and Power

	Decision: Approve	Reject
Ineffective Drug	Type I Error (False Positive), α	Correct
Effective Drug	Correct	Type II Error (False Negative), β

- Goal: minimize Type I and II errors by setting the threshold for approval (critical size of treatment effect)
- But there's a trade-off between these errors

Trade-Off Between Size and Power



Key Parameters

$$\begin{aligned}
\alpha &= \text{Type I error} \\
\beta &= \text{Type II error} \quad , \quad 1-\beta \;=\; \text{Power} \\
n &= \text{Sample Size for Each Arm} \\
X_i &\sim \mathcal{N}(\mu_x, \sigma^2) \;=\; \text{Patient } i\text{'s Response, Control Arm} \\
T_i &\sim \mathcal{N}(\mu_t, \sigma^2) \;=\; \text{Patient } i\text{'s Response, Treatment Arm} \\
\delta &= \mu_t - \mu_x \;=\; \text{Expected Treatment Effect Size} \\
\sigma^2 &= \text{Variance of Patient Outcome} \\
\mathrm{H}_0 &= \text{Null Hypothesis, } \delta = 0 \\
\mathrm{H}_1 &= \text{Alternative Hypothesis, } \delta = \kappa\sigma > 0
\end{aligned}$$

Null Hypothesis and Size

- Under null hypothesis H_0: $\delta \equiv \mu_t - \mu_x = 0$ (no effect):

$$
\begin{aligned}
\alpha &= \text{Prob}(-c < \hat{\delta} < c) \quad (\alpha\% \text{ Confidence Interval}) \\
\frac{\alpha}{2} &= \text{Prob}(\hat{\delta} > c) = \text{Prob}\left(\frac{\hat{\delta}}{\sqrt{\text{Var}[\hat{\delta}]}} > \frac{c}{\sqrt{\text{Var}[\hat{\delta}]}}\right) \\
&= \text{Prob}\left(Z > \frac{c}{\sigma\sqrt{2/n}}\right) = 1 - \text{Prob}\left(Z \leq \frac{c}{\sigma\sqrt{2/n}}\right) \\
\frac{\alpha}{2} &= 1 - \Phi\left(\frac{c}{\sigma\sqrt{2/n}}\right) \Rightarrow \frac{c}{\sigma\sqrt{2/n}} = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \equiv z_{\alpha/2}
\end{aligned}
$$

Null Hypothesis and Size

- Under null hypothesis H_0: $\delta \equiv \mu_t - \mu_x = 0$ (no effect):

$$c = \sigma z_{\alpha/2}\sqrt{2/n}$$

- For $\alpha = 5\%$, $z_{\alpha/2} = 1.96$
- Threshold c decreases with sample size n, and increases with σ and $z_{\alpha/2}$ (does this make sense?)

Alternative Hypothesis and Power

- Under alternative hypothesis H_1: $\delta \equiv \mu_t - \mu_x = \kappa\sigma > 0$

$$
\begin{aligned}
\beta &= \text{Prob}(\hat{\delta} \leq c) \quad \text{(Type II error)} \\
&= \text{Prob}\left(\frac{\hat{\delta} - \kappa\sigma}{\sqrt{\text{Var}[\hat{\delta}]}} \leq \frac{c - \kappa\sigma}{\sqrt{\text{Var}[\hat{\delta}]}} \right) \quad = \quad \text{Prob}\left(Z \leq \frac{c - \kappa\sigma}{\sigma\sqrt{2/n}} \right) \\
\beta &= \Phi\left(\frac{c - \kappa\sigma}{\sigma\sqrt{2/n}} \right) \quad = \quad \Phi\left(\frac{\sigma z_{\alpha/2}\sqrt{2/n} - \kappa\sigma}{\sigma\sqrt{2/n}} \right) \quad = \quad \Phi\left(z_{\alpha/2} - \frac{\kappa}{\sqrt{2/n}} \right) \\
\Phi^{-1}(\beta) &= z_{\alpha/2} - \frac{\kappa}{\sqrt{2/n}} \quad \Rightarrow \quad z_\beta \; = \; z_{\alpha/2} - \frac{\kappa}{\sqrt{2/n}}
\end{aligned}
$$

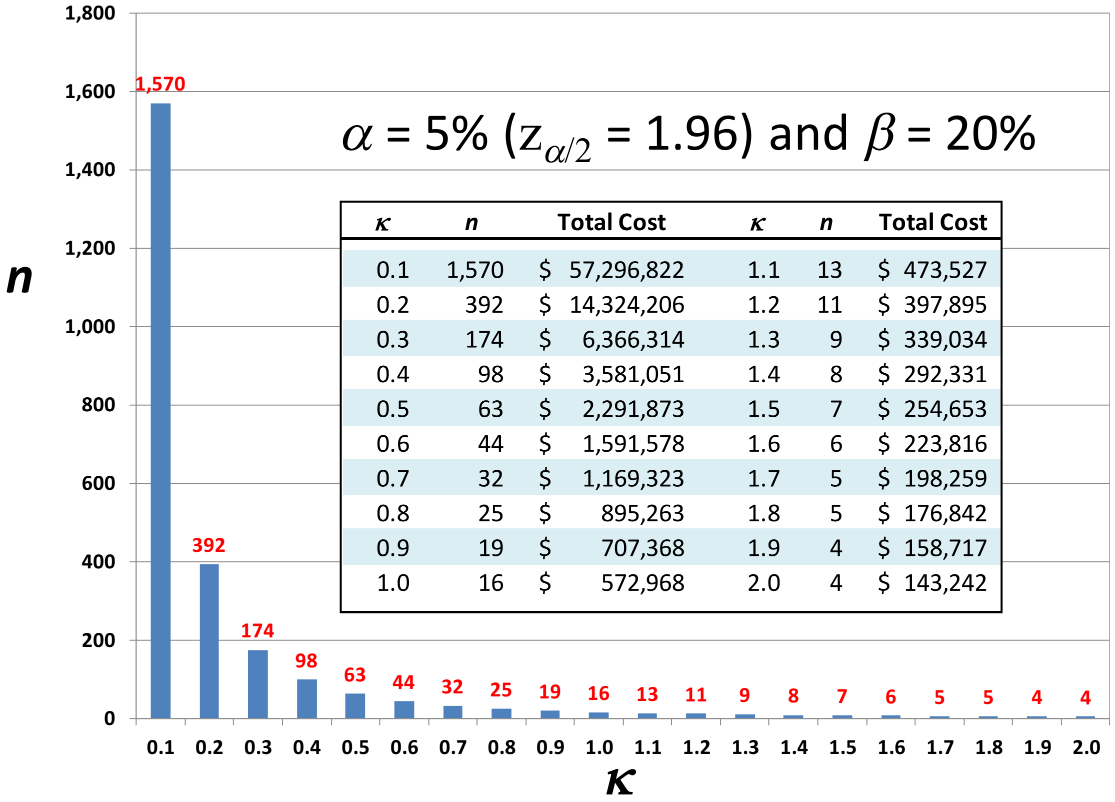
Alternative Hypothesis and Power

- Under alternative hypothesis H_1: $\delta \equiv \mu_t - \mu_x = \kappa\sigma > 0$

$$n = 2\left(\frac{z_{\alpha/2} - z_{\beta}}{\kappa}\right)^2$$

- Sample size n decreases with κ, increases with power $(1-\beta)$
- How should we choose size and power?
- Suppose $\alpha = 5\%$ ($z_{\alpha/2} = 1.96$) and $\beta = 20\%$ (8:1 ratio)

Alternative Hypothesis and Power



α = 5% ($z_{\alpha/2}$ = 1.96) and β = 20%

κ	n	Total Cost	κ	n	Total Cost
0.1	1,570	$ 57,296,822	1.1	13	$ 473,527
0.2	392	$ 14,324,206	1.2	11	$ 397,895
0.3	174	$ 6,366,314	1.3	9	$ 339,034
0.4	98	$ 3,581,051	1.4	8	$ 292,331
0.5	63	$ 2,291,873	1.5	7	$ 254,653
0.6	44	$ 1,591,578	1.6	6	$ 223,816
0.7	32	$ 1,169,323	1.7	5	$ 198,259
0.8	25	$ 895,263	1.8	5	$ 176,842
0.9	19	$ 707,368	1.9	4	$ 158,717
1.0	16	$ 572,968	2.0	4	$ 143,242

So How Big Is Your κ?

- How does financing affect these decisions?