Webinar Series in Applied Quantitative Analysis - Updated

Date	Topic
	Session One
February 29	Potential Outcomes and Omitted Variable Bias I (Theory) ✓
March 7	Potential Outcomes and Omitted Variable Bias II (Application)
	Session Two
March 21	Difference-in-differences I (Theory) ✓
March 28	Difference-in-differences II (Application)
	Session Three
April 25	Power analysis, clustering and sample size calculations I (Theory)
May 2	Power analysis, clustering and sample size calculations II (Application)
	Session Four
May 23	Propensity score matching (Theory)
May 30	Propensity score matching (Application) **New**
	Session Five
June 20	Fixed-effects I (Theory)
June 27	Fixed-effects II (Application)
	Session Six
July 25	Instrumental variables I (Theory)
August 1	Instrumental Variables II (Application)
	Session Seven
August 22	Lagged dependent variables and the Arellano-Bond Estimator I (Theory)
August 29	Lagged dependent variables and the Arellano-Bond Estimator II (Application)

AIR®
Advancing Evidence.
Improving Lives.

https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series

AERC CREA

Power Analysis, Clustering and Sample Size Calculation I

Ashu Handa

Institute Fellow – AIR

Kenan Eminent Professor of Public Policy – UNC-CH

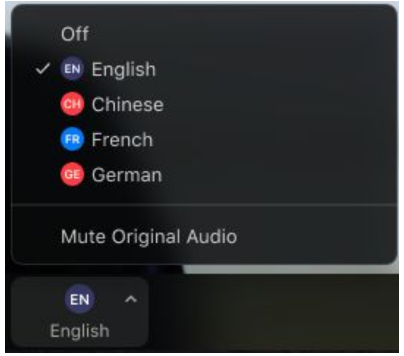
April 25, 2024

Advancing Evidence.
Improving Lives.

AERC
CREA

Écoute de l'interprétation d'une langue
Windows | macOS

1. Dans les contrôles de votre réunion/webinaire, cliquez sur Interprétation .
2. Cliquez sur la langue que vous souhaitez entendre. (Nous aurons le français) Pas besoin de choisir l'anglais, c'est la langue de la salle Zoom principale



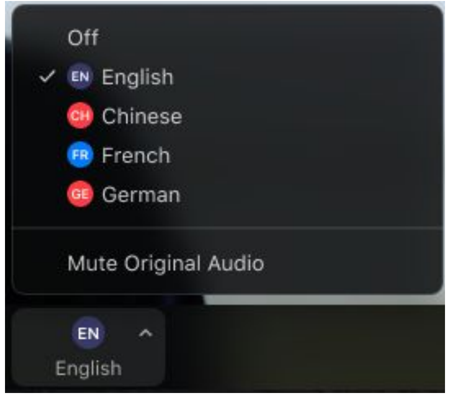
3. (Facultatif) Pour entendre uniquement la langue interprétée, cliquez sur Couper le son original.

Remarques:

- Vous devez rejoindre l’audio de la réunion via l’audio/VoIP de votre ordinateur. Vous ne pouvez pas écouter l’interprétation linguistique si vous utilisez les fonctions audio de connexion ou d’appel téléphonique.
- En tant que participant rejoignant une chaîne linguistique, vous pouvez retransmettre sur le canal audio principal canal si vous réactivez votre audio et parlez.

Listening to language interpretation
Windows | macOS

1. In your meeting/webinar controls, click Interpretation .
2. Click the language that you would like to hear. (We will have French) No need to choose English, that is the language in the main Zoom room



3. (Optional) To hear the interpreted language only, click Mute Original Audio.

Notes:

- You must join the meeting audio through your computer audio/VoIP. You cannot listen to language interpretation if you use the dial-in or call me phone audio features.
- As a participant joining a language channel, you can broadcast back into the main audio channel if you unmute your audio and speak.

Why do you need to understand power analysis?

- You are using a secondary data set (KIHBS, GLSS, NIDS) to answer a research question, but the data set was not designed for that question
  - Do you have enough statistical power to answer the question? **Calculate the minimum detectable effect size (MDES)**

- You are designing your own study with a specific research question
  - What sample size do you need to have enough statistical power? **Calculate sample size**

- In either of the above cases, if the sampling is cluster-based, that affects your statistical power
  - Clustering reduces statistical power, and thus requires a larger sample size
  - By how much? **Calculate sample size adjusting for clustering**

What is statistical power?

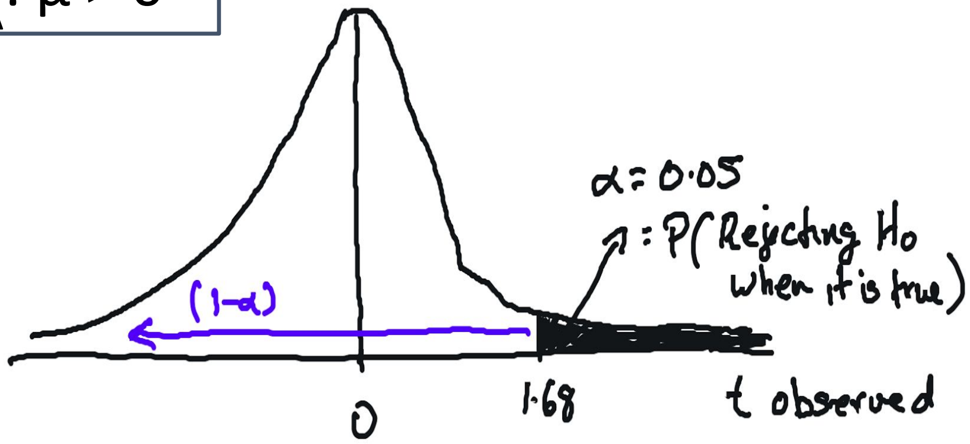
- ***The probability of finding a statistically significant effect when the alternative hypothesis H_A is true***

- In lay person's terms—is my sample big enough to detect an effect?

- [Note: In this presentation we will focus the example on a one-sided test rather than a two-sided test as the explanation and graphs are easier]

Refresh the basics of hypothesis testing
[we will use one-sided test for simplicity]

$H_0: \mu = 0$
$H_A: \mu > 0$



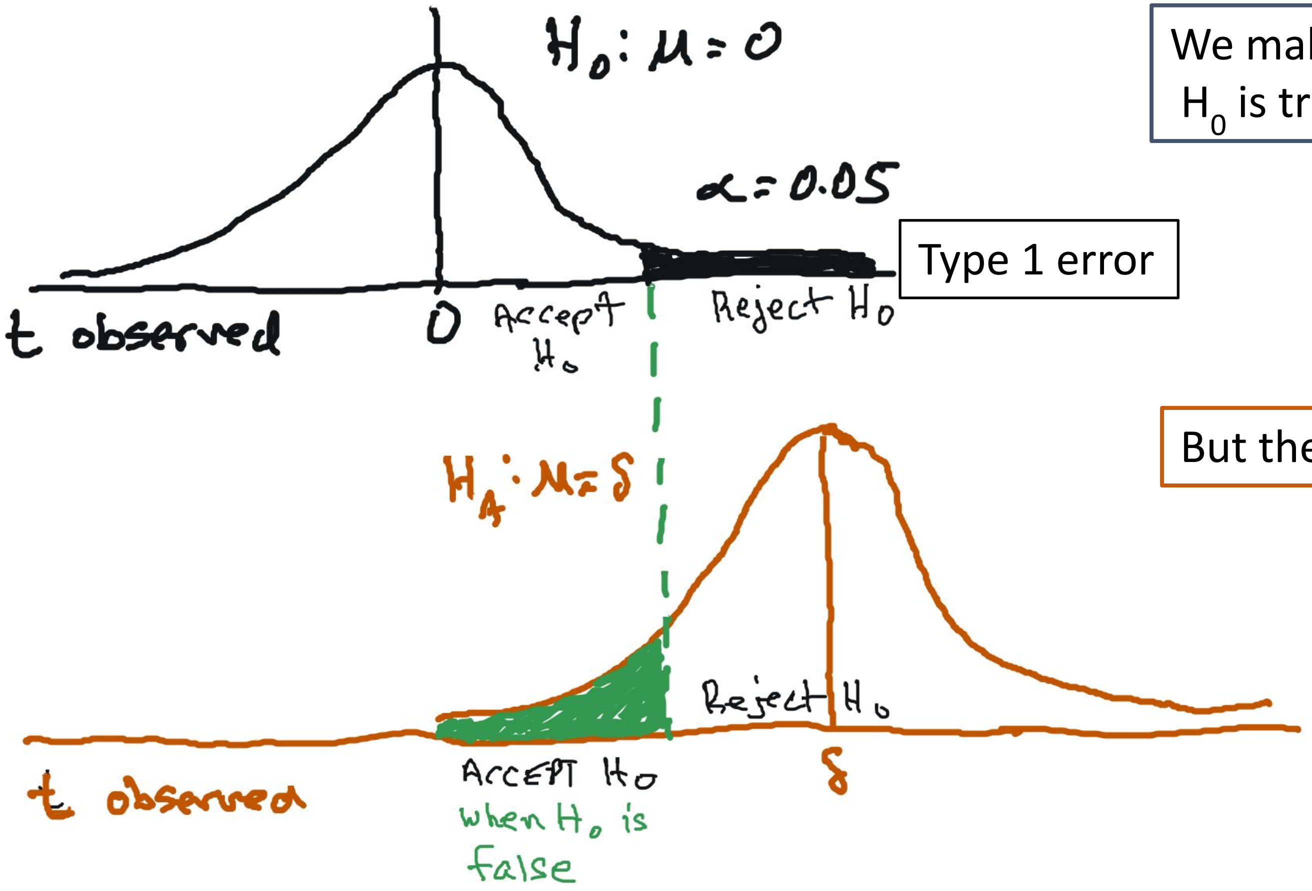
[We typically want to reject H_0 in favor of H_A]

alpha-level (α) is the significance level of the test
Usually set to 5% to minimize Type 1 error

"The probability of rejecting the null when it is true"

Type 1 error

Now consider if H_A is true ($\mu=\delta$)

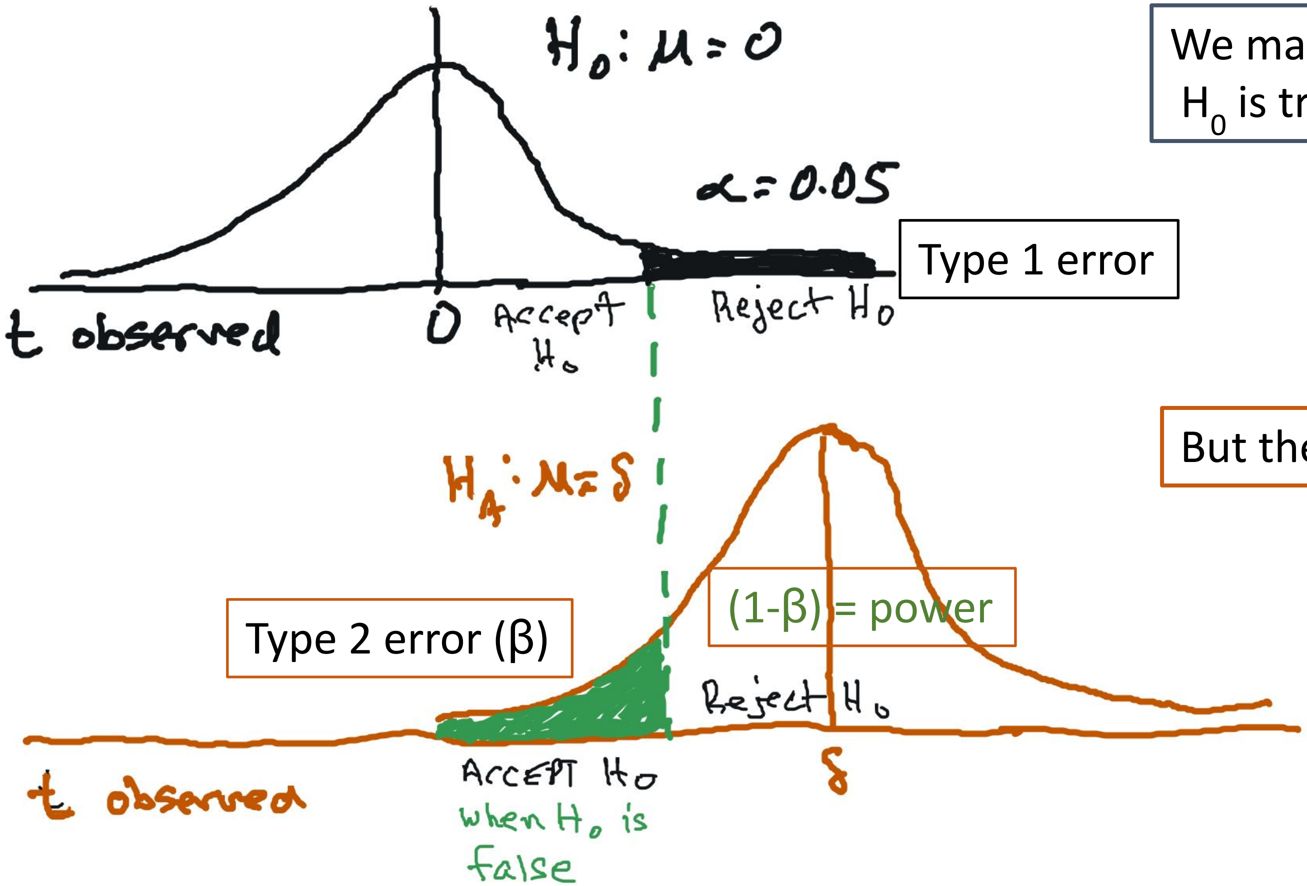


Type 1 error

We make our decision assuming H_0 is true ($\mu=0$)

But the reality is that H_A is true ($\mu=\delta$)

Now consider if H_A is true ($\mu=\delta$)



We make our decision assuming H_0 is true ($\mu=0$)

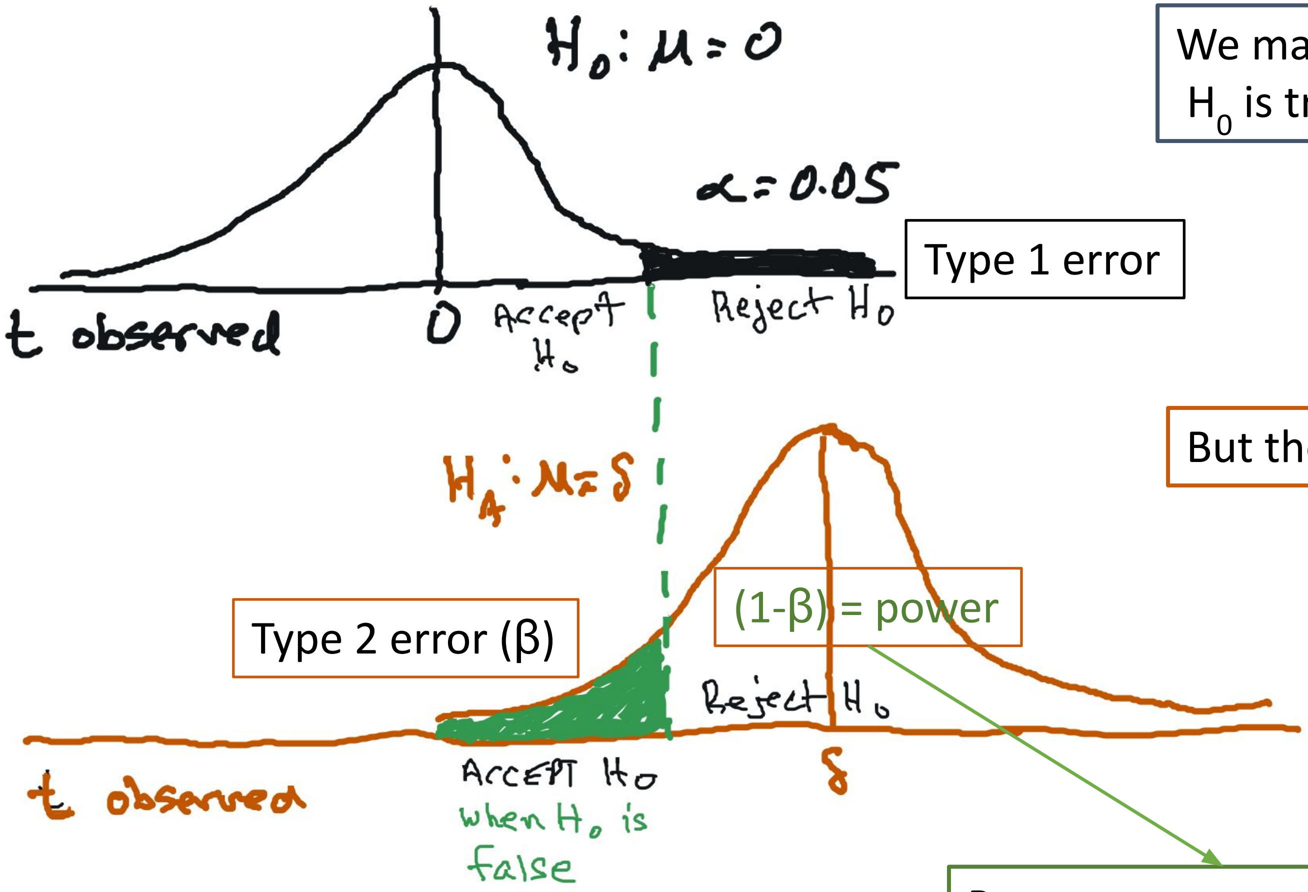
Type 1 error

But the reality is that H_A is true ($\mu=\delta$)

Type 2 error (β)

$(1-\beta)$ = power

Now consider if H_A is true ($\mu=\delta$)



We make our decision assuming H_0 is true ($\mu=0$)

Type 1 error

But the reality is that H_A is true ($\mu=\delta$)

Type 2 error (β)

$(1-\beta)$ = power

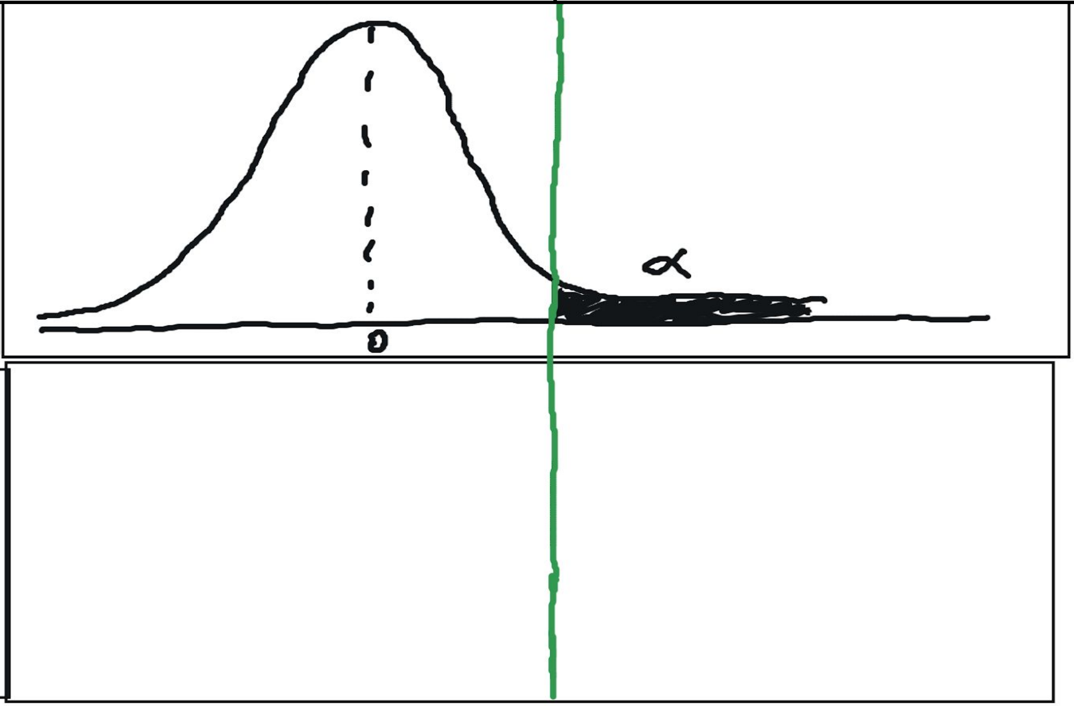
<u>Power</u>
Probability of accepting H_A when it is true

Put everything together

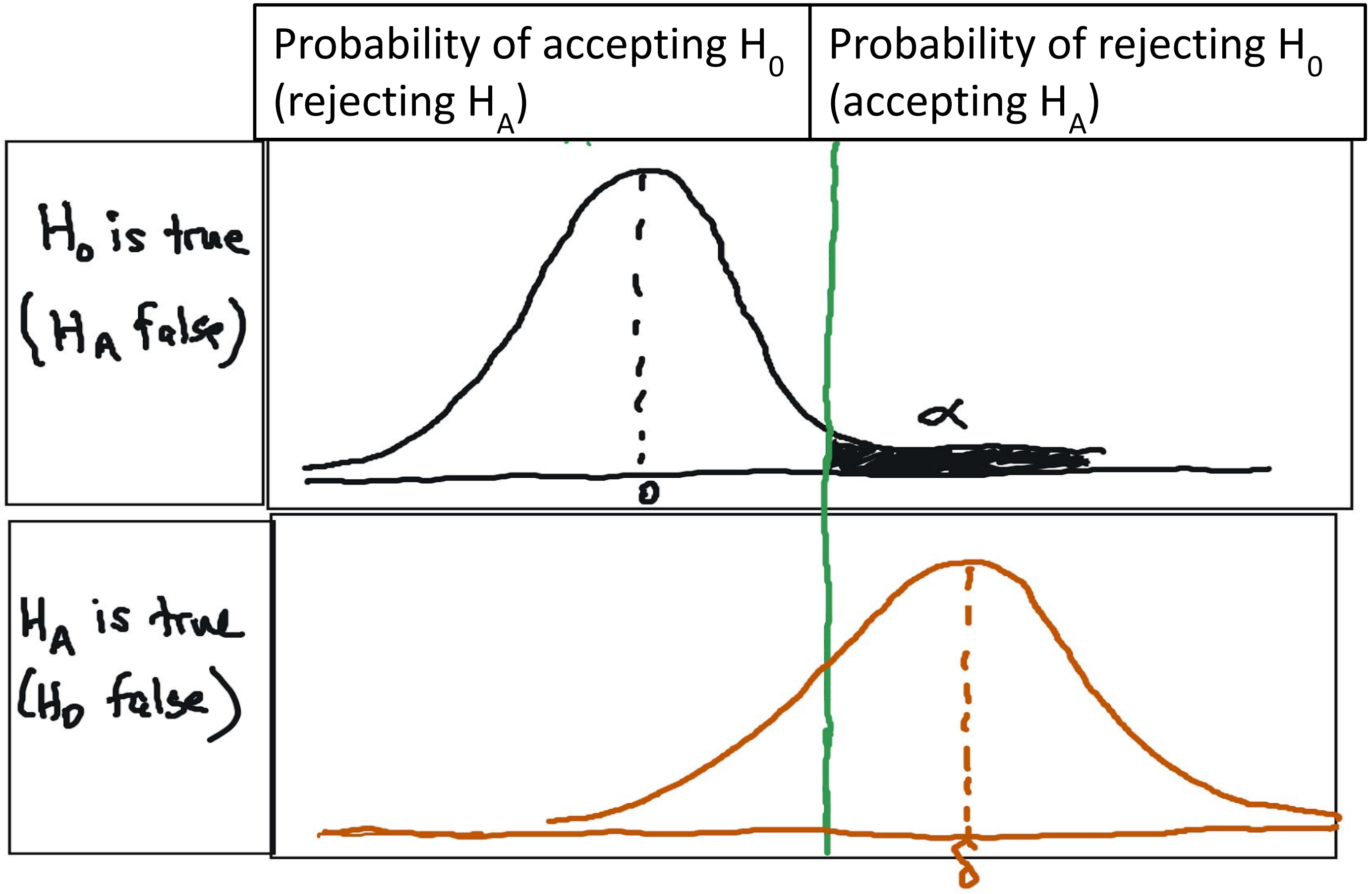
Column: The decision made

Row: The truth

	Probability of accepting H_0 (rejecting H_A)	Probability of rejecting H_0 (accepting H_A)
H_0 is true (H_A false)		
H_A is true (H_0 false)		

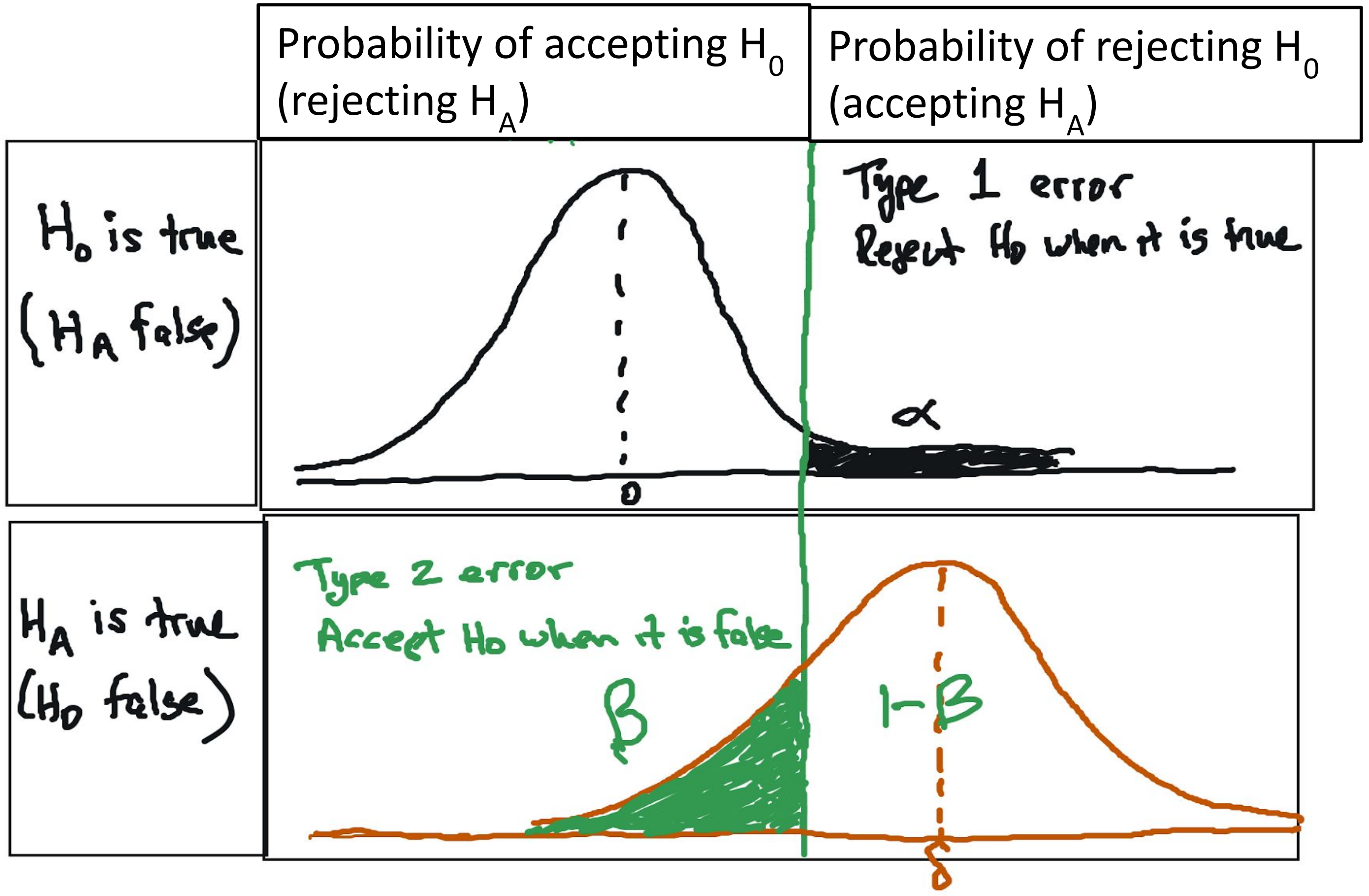


Column: The decision made
Row: The truth



Distribution if H_A is true ($\mu=\delta$)

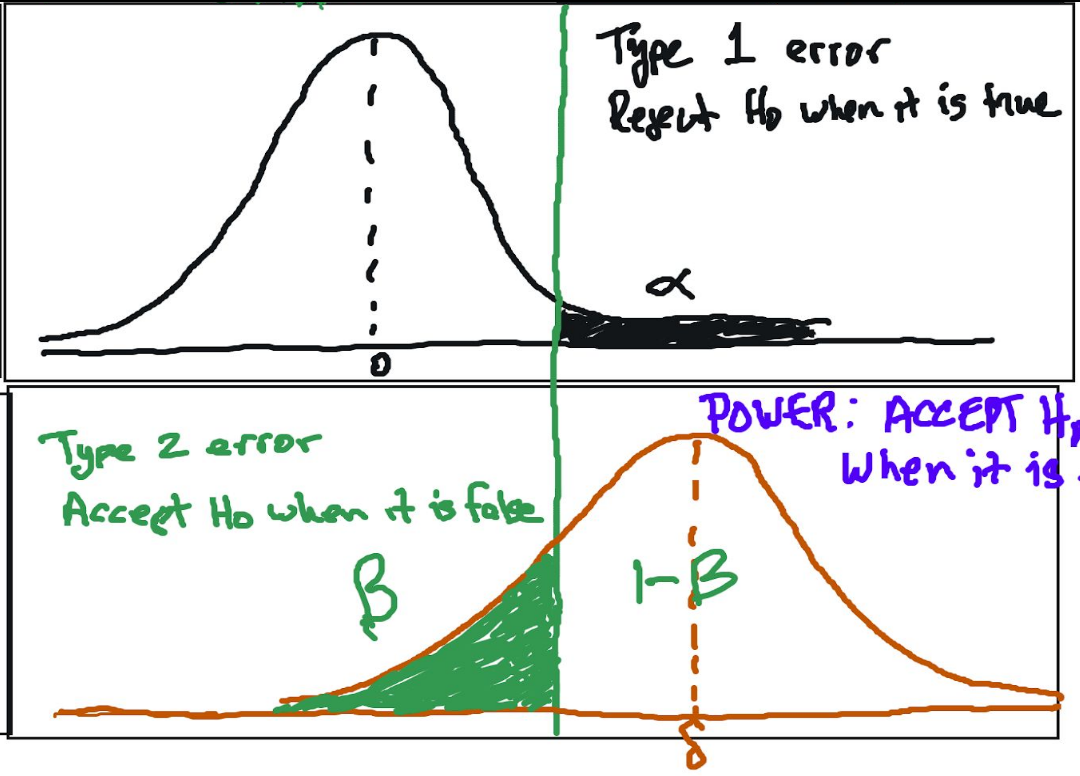
Column: The decision made
Row: The truth



Distribution if H_A is true ($\mu=\delta$)

Column: The decision made
Row: The truth

	Probability of accepting H_0 (rejecting H_A)	Probability of rejecting H_0 (accepting H_A)
H_0 is true (H_A false)		Type 1 error Reject H_0 when it is true
H_A is true (H_0 false)	Type 2 error Accept H_0 when it is false	POWER: ACCEPT H_A when it is true!



Power
Probability of accepting H_A when it is true ($1-\beta$)

Distribution if H_A is true ($\mu=\delta$)

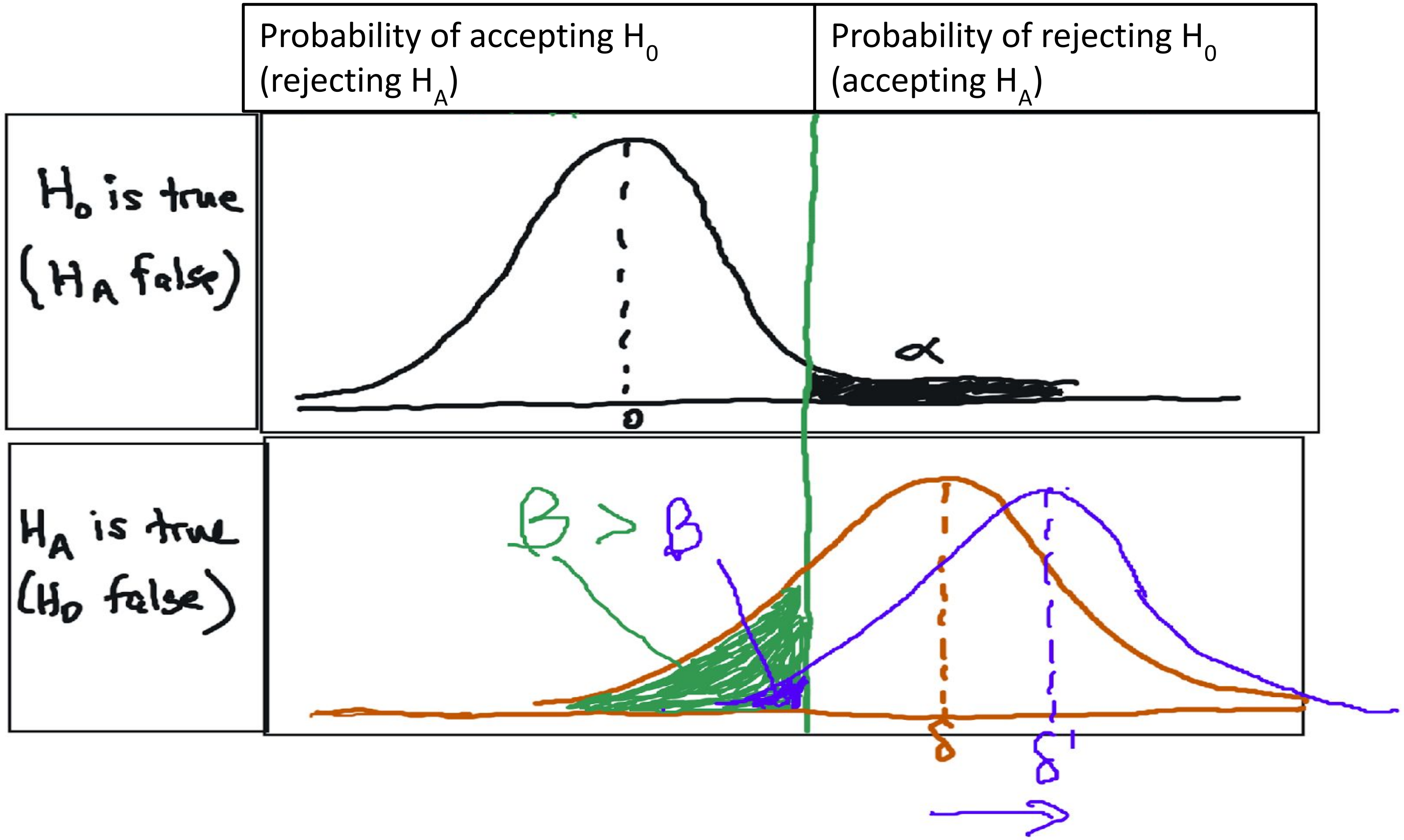
POLL

Two important facts about statistical power

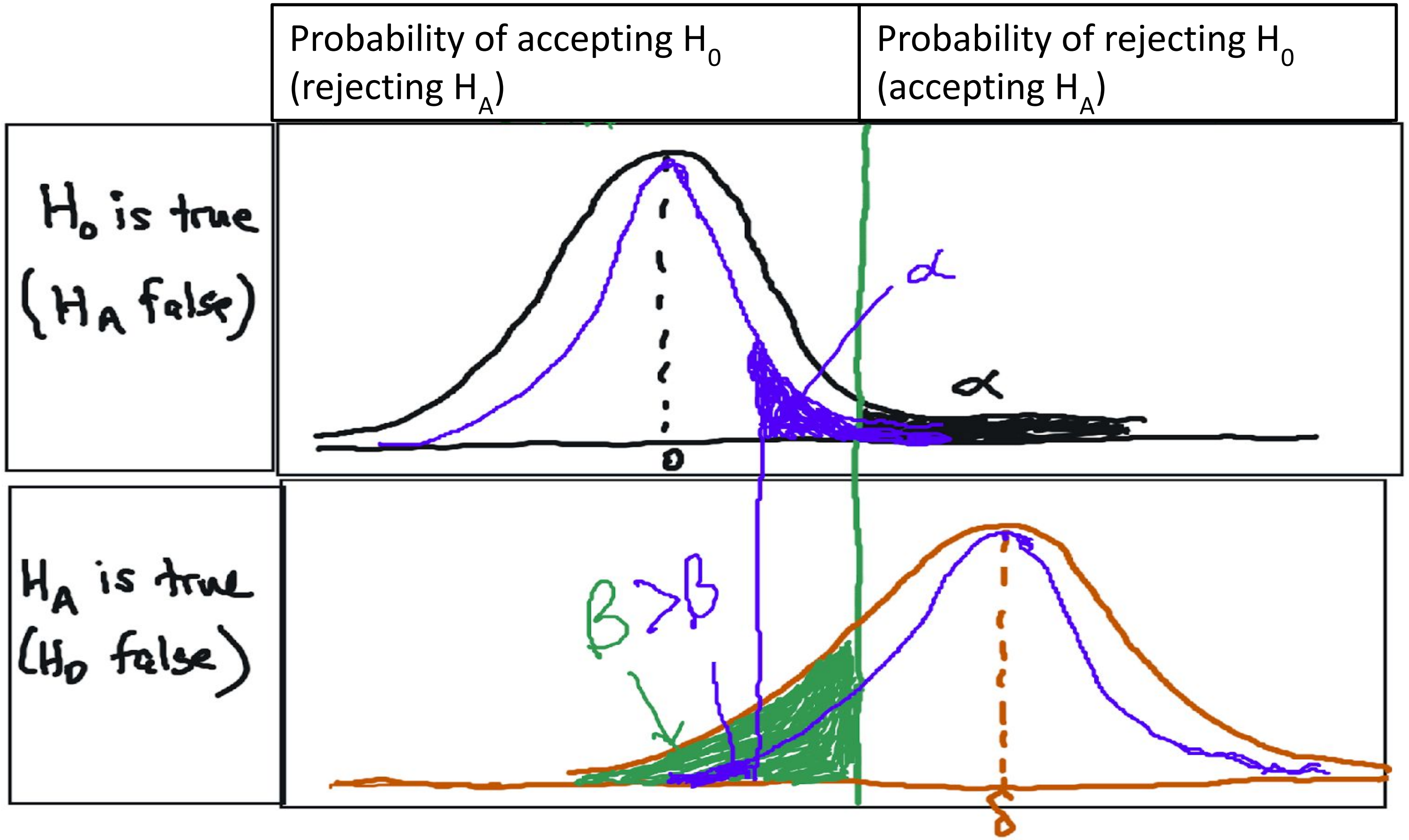
- Power increases as your hypothesized H_A (δ) increases
  - The larger the true effect, the larger your power to detect it
  - It is always easier to detect a large effect than a small effect!

- Power increases as your sample size increases
  - Larger samples lead to more accuracy and less noise/error
  - This reduces β and therefore increases $(1- \beta)$

Larger effect size (δ) shifts H_A distribution to the right, β gets smaller, (1-β) increases



Larger sample size reduces variance and narrows the distribution, β gets smaller, (1-β) increases

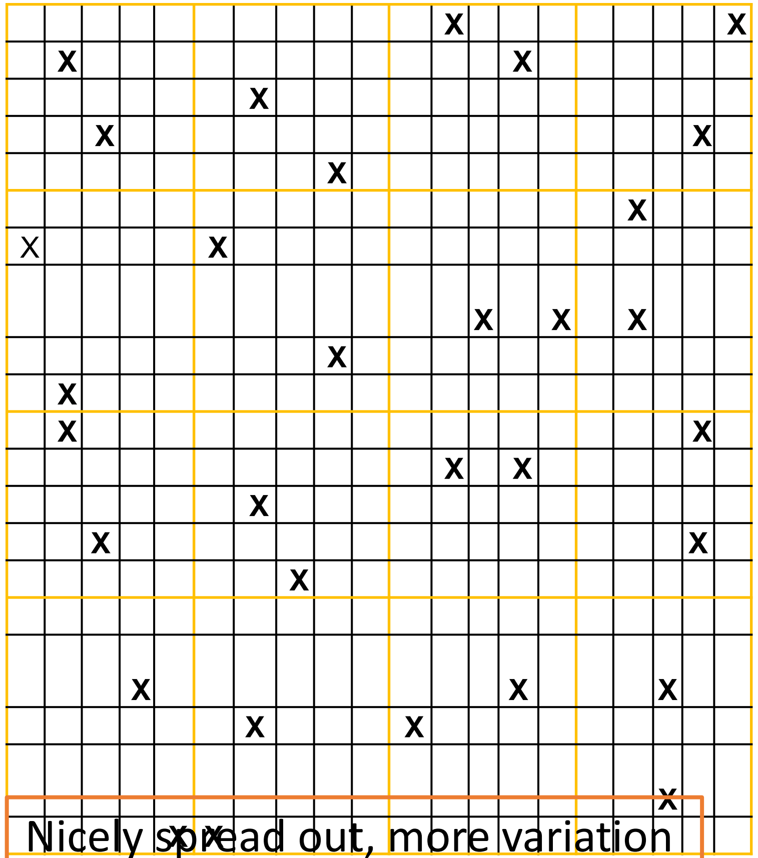


Poll

How does cluster-based sampling affect power?

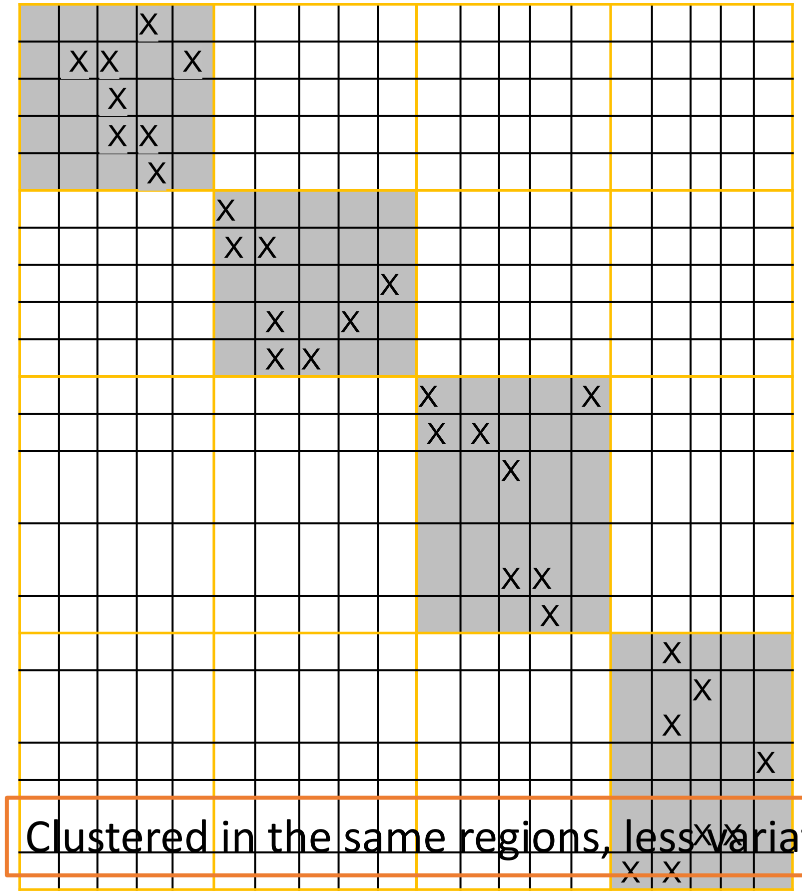
Reduces power due to homogeneity within clusters

Simple random sample (N=32)



Nicely spread out, more variation

Cluster-based sample (N=32 in 4 clusters)



Clustered in the same regions, less variation

Clustering requires that we increase sample size to account for correlation within clusters

- **Design effect tells by how to increase sample size = [1 + rho*(k-1)]**
  - rho is intra-cluster correlation (degree of homogeneity within the cluster)
  - k is the number of units in each cluster on average

- **Say rho=0.12 and k=15 on average**
  - Design effect (DEFF) = [1 + 0.12*(15-1)]
  - DEFF = [1 + 0.12*(14)] = [1 + 1.68] = 2.68
  - Sample size must be 2.68 times larger due to clustering relative to a simple random sample

- **We will do some examples!**

POLL

The recipe for calculating sample size for a study

1. What is the expected change or effect size (δ)?
   - 0.20SD is a small effect size, 0.30SD-0.50SD is medium, >0.50SD is large
2. What is the initial (or baseline) mean and SD of the indicator?
   - [Why do you need the SD?] [Affects shape of the distribution-skinny or fat]
3. Minimum power (1-β) is 0.80, α=0.05
   - 0.90 is better
4. Calculate DEFF if cluster-based sampling
   - Need estimate of rho and average number of units per cluster (k)

Example: Calculate sample size for a new study

- Study design features
  - Cash transfer to poor/vulnerable households to support food consumption in rural Kenya
  - Treatment group and comparison group
  - Primary indicator: food consumption per capita

- What does the recipe say? We need….
  - **Anticipated effect size or change (δ)**
  - **Baseline mean and SD**
  - **Clustering? Assume none (for now)**

Retrieve mean and SD from Kenya Integrated Household Budget Survey (KIHBS 2015/16)

Use rural households in bottom half of consumption distribution (DO file is on website)

```
. sum pcfood if resid==1 & pcexp<3218

Variable      Obs      Mean   Std. dev.   Min       Max

pcfood      6,546  1423.073   534.298       0  3068.786
```

Identify a range of effect sizes in SD units

Effect size in SD (1)	Initial mean (Sh) (2)	Change (3)	Final mean= (2) + (3) (4)
0.20SD	1423	0.20*534 = 107	1530
0.30SD	1423	0.30*534 = 160	1583
0.40SD	1423	0.40*534 = 214	1637
0.50SD	1423	0.50*534 = 267	1690
0.60SD	1423	0.60*534 = 320	1743
0.70SD	1423	0.70*534 = 374	1797

STATA command

Type of statistical analysis

Initial and final mean [0.20SD]

SD

```
. power twomeans 1423 1530, sd(534)

Performing iteration ...

Estimated sample sizes for a two-sample means test
t test assuming sd1 = sd2 = sd
H0: m2 = m1  versus  Ha: m2 != m1

Study parameters:

alpha =    0.0500
power =    0.8000
delta =  107.0000
m1 = 1423.0000
m2 = 1530.0000
sd =  534.0000

Estimated sample sizes:

          N =        784
N per group =        392
```

```
. power twomeans 1423 1530, sd(534)

Performing iteration ...

Estimated sample sizes for a two-sample means test
t test assuming sd1 = sd2 = sd
H0: m2 = m1  versus  Ha: m2 != m1

Study parameters:

alpha =     0.0500
power =     0.8000
delta =   107.0000
m1  = 1423.0000
m2  = 1530.0000
sd  =  534.0000

Estimated sample sizes:

          N =         784
N per group =         392
```

alpha and (1-β)

Estimated sample size!

```
. power twomeans 1423 1583, sd(534)

Performing iteration ...

Estimated sample sizes for a two-samplemeans
   test
t test assuming sd1 = sd2 = sd
H0: m2 = m1  versus  Ha: m2 != m1

Study parameters:

alpha =    0.0500
power =    0.8000
delta =  160.0000
m1 = 1423.0000
m2 = 1583.0000
sd =  534.0000

Estimated sample sizes:

         N =        352
N per group =        176
```

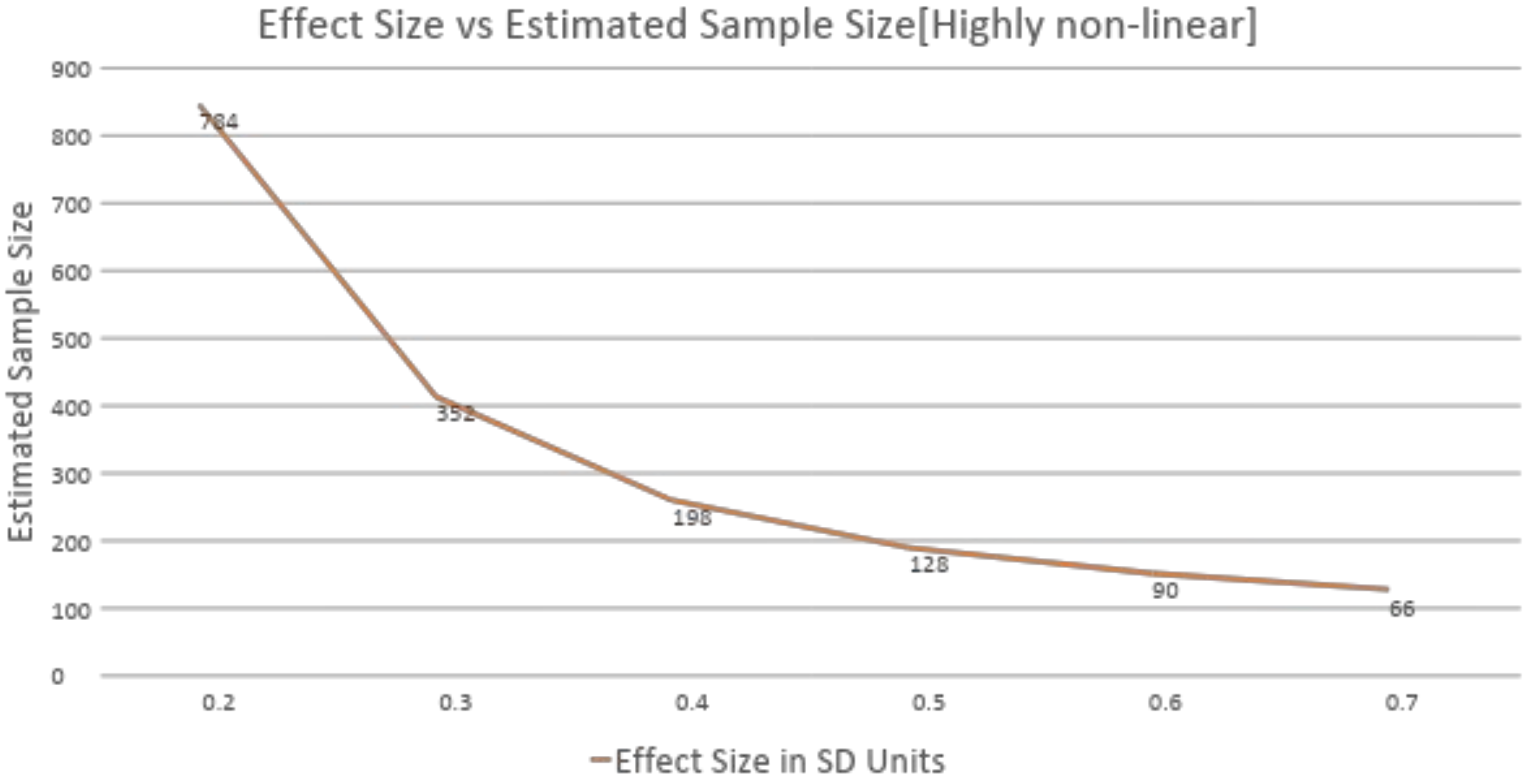
This new mean represents a 0.30SD change or effect size

No change here

New estimated sample size!

Estimated (required) sample size for a range of effect sizes in SD units

Effect size in SD units (1)	Initial mean (Sh) (2)	Change (3)	Final mean= (2) + (3) (4)	Sample Size (5)
0.20SD	1423	0.20*534 = 107	1530	**784**
0.30SD	1423	0.30*534 = 160	1583	**352**
0.40SD	1423	0.40*534 = 214	1637	**198**
0.50SD	1423	0.50*534 = 267	1690	**128**
0.60SD	1423	0.60*534 = 320	1743	**90**
0.70SD	1423	0.70*534 = 374	1797	**66**


Effect Size vs Estimated Sample Size[Highly non-linear]
Estimated Sample Size
900
800
700
600
500
400
300
200
100
0
784
352
198
128
90
66
0.2
0.3
0.4
0.5
0.6
0.7
Effect Size in SD Units