

Running Randomized Evaluations: A Practical Guide

Chapter 6: Statistical Power (檢定力)

LIN, Shuan-Wen 林宣文

LIN, Chien-Hsun 林建勳

FONG, Meng-Jhang 馮孟章

Why should we care about this?

- We can use power analysis to determine...
 1. What **sample size (樣本數)** we need,
 2. The **level at which to randomize (隨機層級)**,
 3. How many different **treatment groups (實驗組數)** we can test,
- And other design issues.

Statistical Background

For further “statistics” help, go find STATISTICS 101 class!

LIN, Chien-Hsun 林建勳



Two things we care about.....

1. Is the effect FALSE? Significance Test

- Do students make difference?

2. Is the effect TRUE? Power Test

- Do students become better?

Hypothesis Testing

- $\begin{cases} H_0: \text{Effect size} = 0 \\ H_1: \text{Effect size} > 0 \end{cases}$

How can we decide which is true?

TRUE STATE

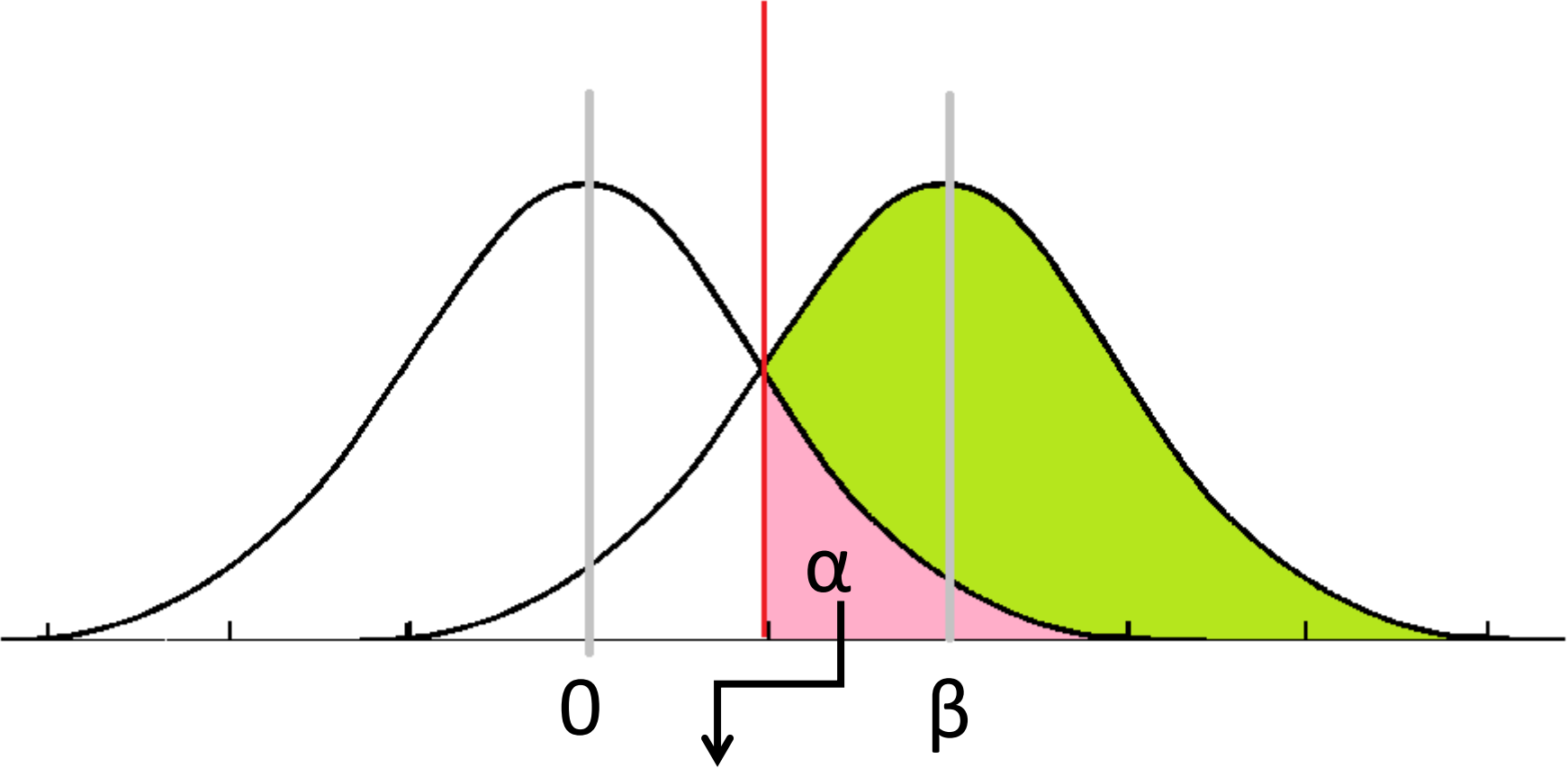
DATA

	H0 is true	H1 is true
Not rejects H0	NOT FALSE	Data not support H1
Rejects H0	FALSE Positive	TRUE ENOUGH

Significance
Level

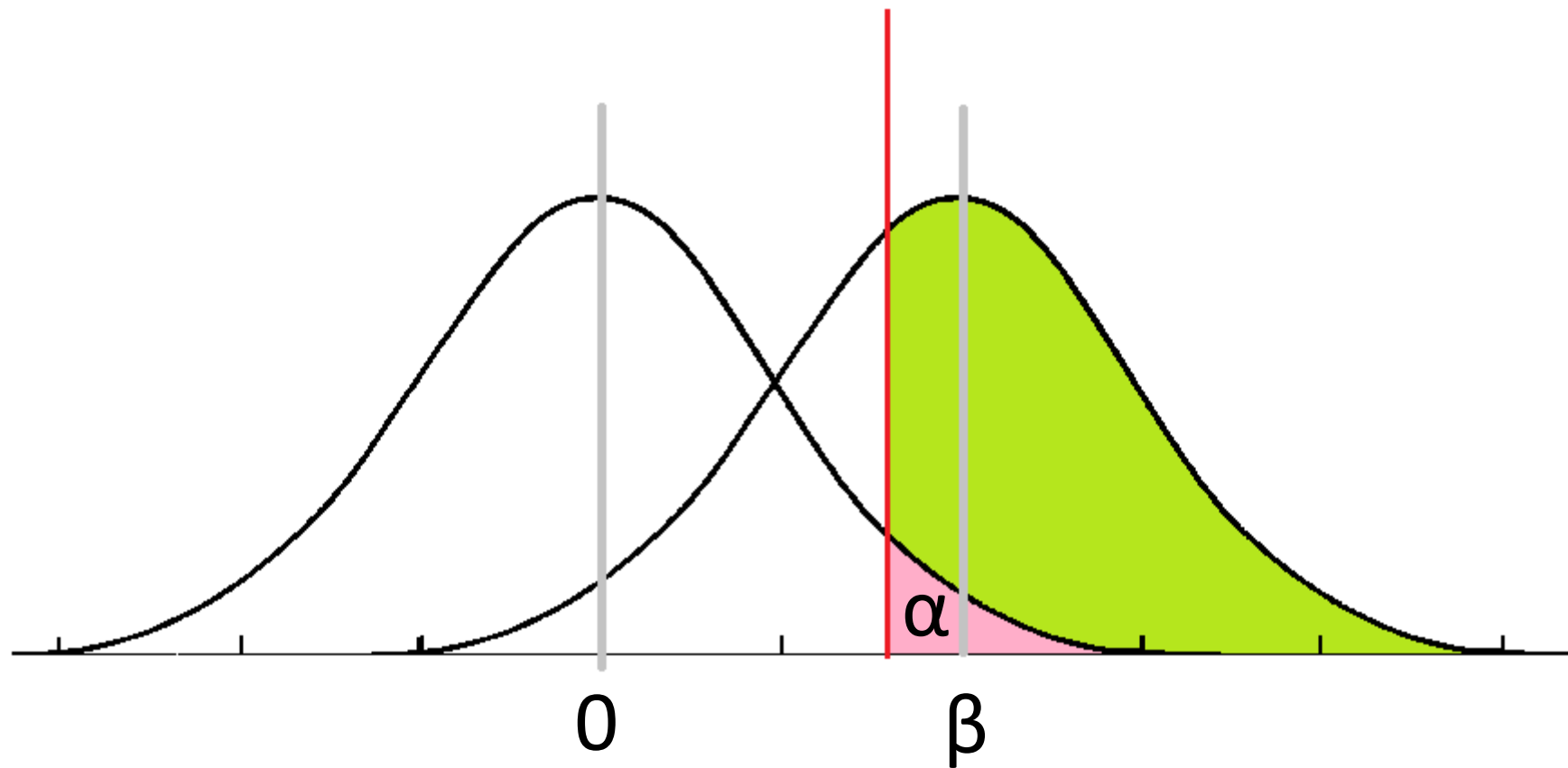
POWER
of the test

Critical Value (Determined by significance level)

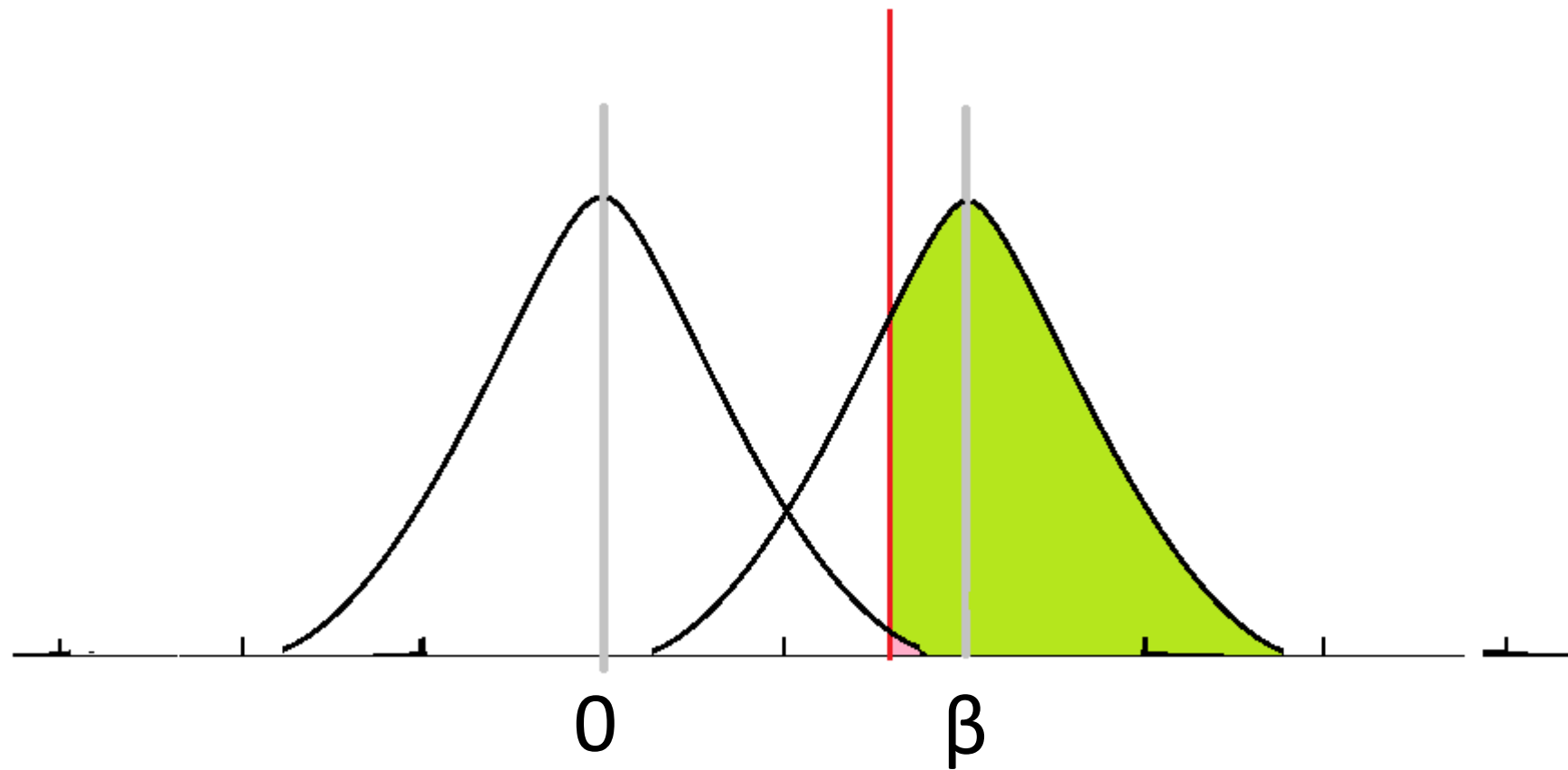


Significance level

What happens if we lower significance level...



Less variant?



When will we say :

“There is really a positive effect” ?

Minimum Detectable Effect (MDE) size:

Given a set of

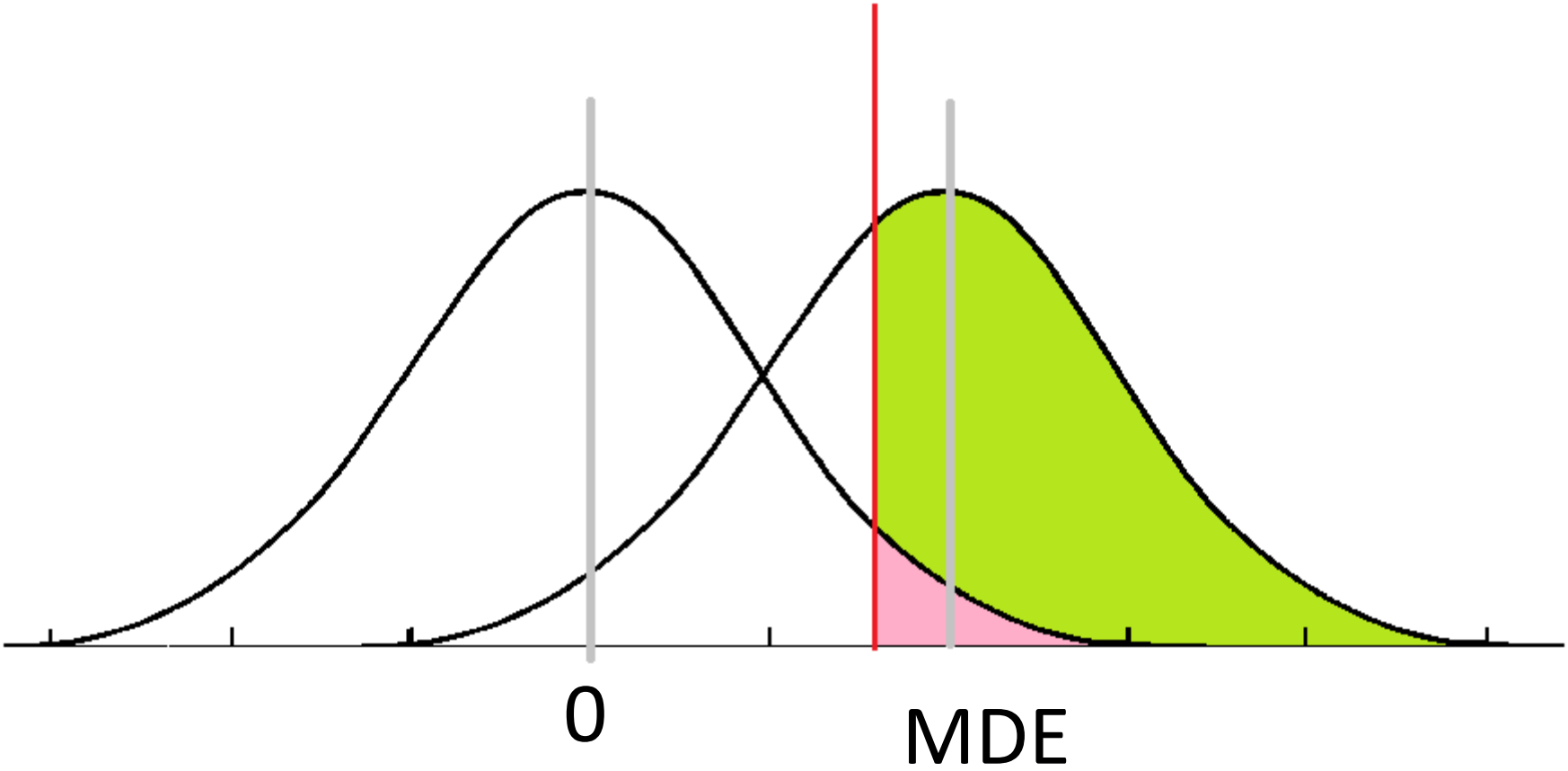
significance size and power size,

at which we can say

“The effect is really detectable!!”

Significant size: 10%

Power size: 70%



How to choose MDE size?

- What are MDE of similar treatments?



MDE= 10 points



MDE= 10 points

- “One penny for little.”
 - One banana for monkey RA.

Sample and Power

FONG, Meng-jhang 馮孟章

Budget Size and Power

- Money : the more, the better.
- We don't have enough budget
- How to allocate/randomize the sample, that is the question!

Allocation Ratio and Power







- More allocation, more accurate measure
 - diminishing returns
 - Principle: **equal allocation** among all groups maximizes the power
- Exceptions always exist.



Choosing allocation fractions

- Key point: **budget constraint**
1. When one budget pays for both **program** and **evaluation**
 2. When MDE size **varies** by treatment group
 3. When the comparison group (baseline) plays a particular **important** role

- When one budget pays for both **program** and **evaluation**

	Treatment	Control
Program costs	\$1,000 	\$0 
Evaluation costs	\$100 	\$100 
Total costs	\$1,100 	\$100 

- When MDE size **varies** by treatment group

- Treatment 1: Pagamo

- Treatment 2: free learning supplements

- less expensive

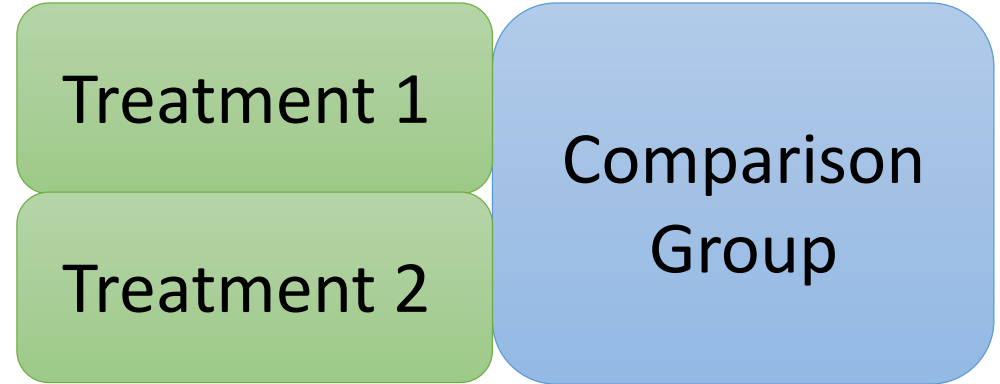


- When the comparison group (baseline) plays a particular **important** role

- Treatment 1 vs. the comparison

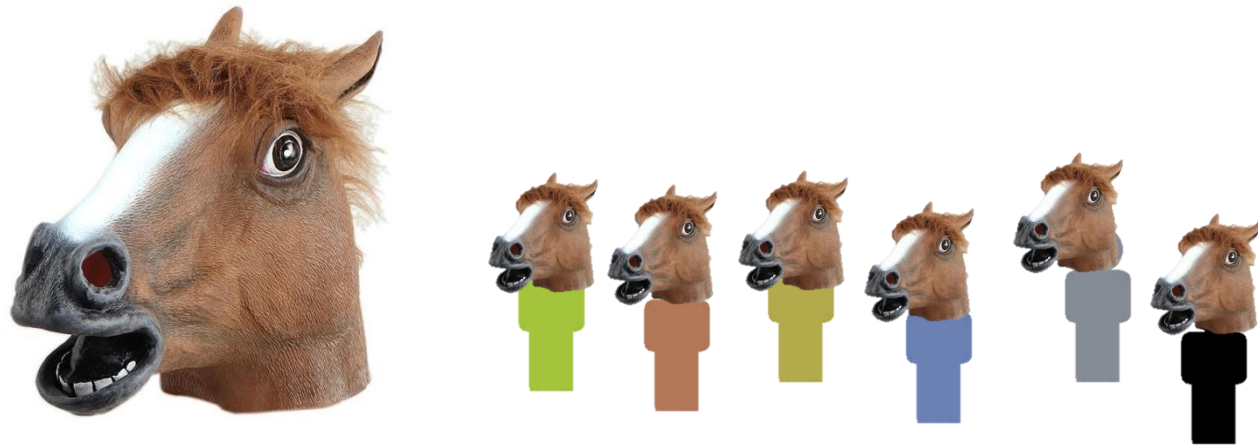
- Treatment 2 vs. the comparison

- Pooled data from treatments vs. the comparison



Power with Clustering

- Randomize at the level of, for example, the school
- Reduce power (for a given sample size)
 - intraclass correlation



Choosing the Number of Clusters

- Key point: **budget constraint**
 - Fixed cost

➤ Trade-off:

- **marginal power vs. marginal cost** of each additional person per cluster
 - clusters = N vs. clusters = 1

How to Design a High-Powered Study

LIN, Shuan-Wen 林宣文

When Designing the Evaluation... (實驗設計時...)

- Choose a decent sample size (適當的樣本數)
- Use fewer treatment group (減少實驗組)
- Randomize at the lowest level possible (降低隨機層級)
- Use a design that increases compliance (確實分隔實驗組與對照組的實驗設計)

When Designing the Evaluation (cont.) (實驗設計時...)

- Use a design that limits attrition
(避免樣本遺漏)
(Will be introduced in detail in chapter 7.)
- Use stratification (分層抽樣)
(Has been introduced in chapter 4.)
- Choose an allocation fraction (適當的分組方式)

When Planning the Data Collection...

(準備資料蒐集時...)

- Choose proximate outcome measure
(近似的結果測量)
- Collect data on control variables (使用控制變數)
- Collect multiple observations on one person
(對同一樣本重複試驗)
- Plan data collection to limit attrition (避免樣本遺漏)
- Limit procedural variation (減少操作誤差)

When Implementing the Evaluation

--Managing Threats (實驗進行時)

- Increasing compliance (保持實驗組與對照組的隔離)
- Limit attrition (避免樣本遺漏)

When Undertaking the Impact Analysis... (分析結果時...)

- Use control variables (使用控制變數)
- Choose a significance level (挑選適當的顯著水準)

Take-Home Knowledge

- After the presentation, you should know...
 1. What is power? What is MDE?
 2. How MDE work? How MDE affects your evaluation?
 3. What is important to design your study?