Graphics for Power and Sample Size

Course developed by
Deborah H. Glueck and Keith E. Muller

Slides developed by Jessica R. Shaw, Keith E. Muller, Albert D. Ritzhaupt and Deborah H. Glueck

 1

Learning Objectives

Describe power curves.

Decide how to choose a graphic that best tells the desired story.

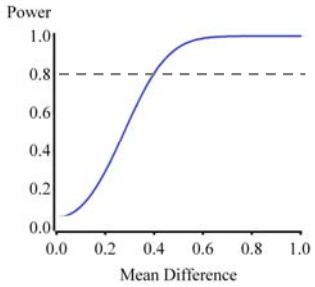
Describe how design inputs affect power.

2

POWER CURVES

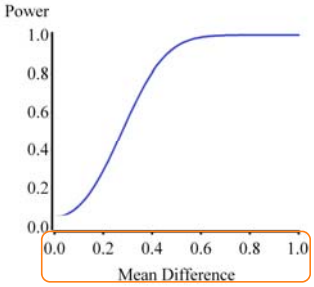
3

Recall a power curve across mean differences



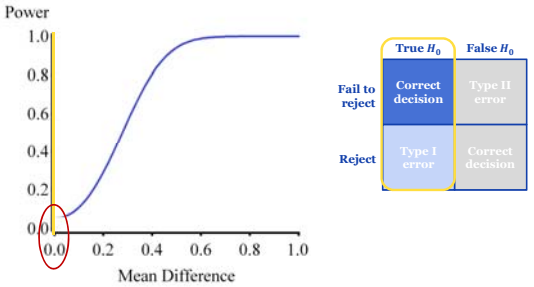
4

Null and alternative hypotheses are described in terms of mean difference



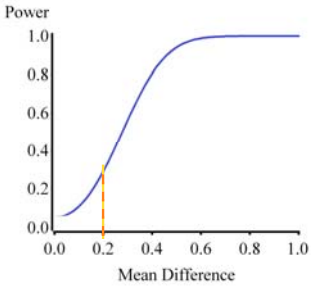
5

A mean difference of zero represents the null hypothesis being true



6

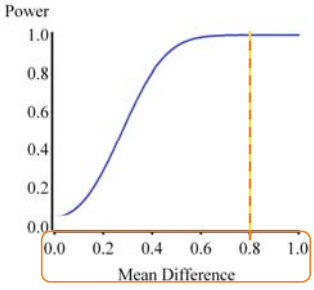
A mean difference not equal to zero represents an alternative hypothesis



	True H_0	False H_0
Fail to reject	Correct decision	Type II error
Reject	Type I error	Correct decision

7

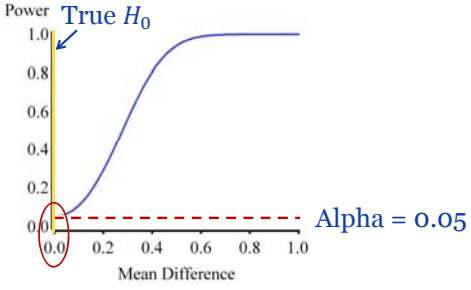
The x axis of a power curve represents many possible alternative hypotheses



	True H_0	False H_0
Fail to reject	Correct decision	Type II error
Reject	Type I error	Correct decision

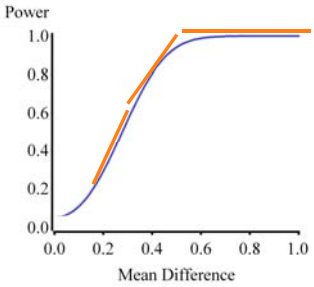
8

Power is equal to the Type I error rate when the null hypothesis is true (for an exact test)



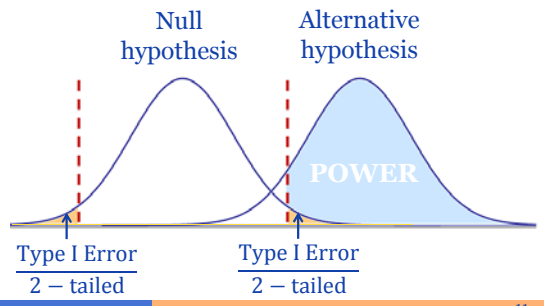
9

The power curve eventually flattens out as the mean difference gets larger



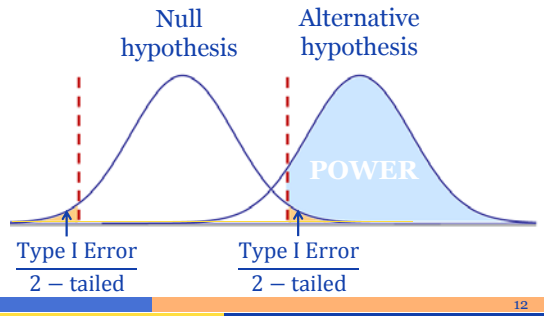
10

As the Type I error rate decreases, power also decreases (less orange -> less blue)



11

A lower Type I error requires a larger sample size to achieve the same power



12

GRAPHICS WHICH TELL STORIES

13

Graphical representations of power calculations tell stories

Each graph should convey one idea.

We will look at several example graphs.

The examples will serve as templates for presenting power analysis in grant proposals and manuscripts.

14

Recall the group randomized trial example

Vignette

A single level study examined the efficacy of a workplace training program designed to reduce alcohol consumption. Researchers randomized workplaces to one of two treatment groups.

Adapted from Reynolds, G. Shawn and Joel B. Bennett, 2015

15

Workplaces were randomized to two different treatments

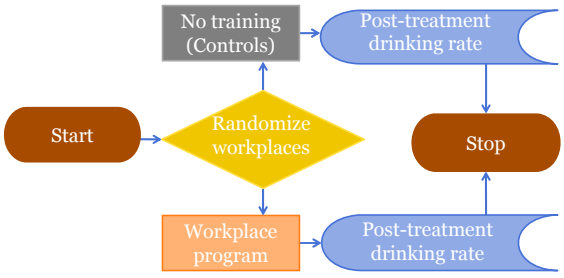
Vignette, continued

The first treatment included a workplace training program and the second treatment included no training. Post-treatment drinking rate for each worker was measured as the outcome of interest.

Adapted from Reynolds, G. Shawn and Joel B. Bennett, 2015

16

Employees were surveyed after training or no training



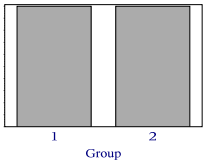
Adapted from Reynolds, G. Shawn and Joel B. Bennett, 2015

17

Researchers hoped to observe an association between treatment and reduction in drinking

Null hypothesis:

There is no difference in post-treatment drinking frequency between workers who receive no training and workers who receive the workplace program.



18

This was a single level design with one outcome measurement in time

Independent sampling unit: Workplace

Unit of observation: Drinking rate for each employee after treatment

Adapted from Reynolds, G. Shawn and Joel B. Bennett, 2015

19

This study evaluated the impact of treatment on employee behaviors

Between-independent sampling unit factor:
Intervention (standard of care or workplace drinking program)

Within-independent sampling unit factor:
Cluster member, person

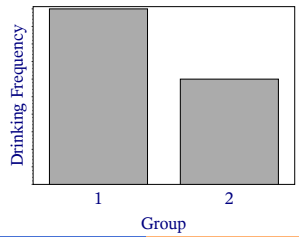
Interest in factor depends on setting

Adapted from Reynolds, G. Shawn and Joel B. Bennett, 2015

20

Analysis compared post-treatment drinking frequency between the two treatment groups

Scientific goal: **REJECT** the null hypothesis



21

Measurements of drinking frequency of workers within a workplace are correlated

Some workers drink together.

Some workers attend sobriety programs together.

Workers discuss drinking habits with each other.

22

Randomization allowed researchers to make two important assumptions

1. The average correlation between workers in the intervention groups **is equal to** the average correlation between workers in the control groups.

2. Pre-existing employee factors did not bias study outcomes.

23

What is the best way to present a power analysis?

The choice of graphic depends on the question one is trying to answer.

We present multiple questions and associated graphs.

24

What is the effect on power of correlation within a cluster (ICC)?

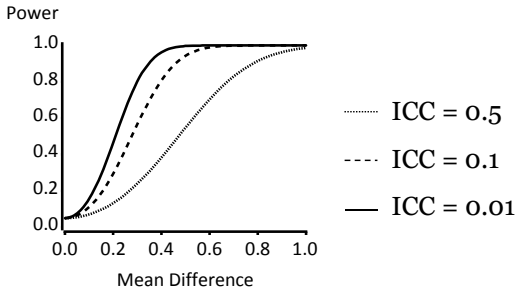
Suppose we have 20 workplaces assigned to no intervention, and 20 assigned to a workplace treatment program.

Suppose we have 10 workers in each workplace.

Assume we have the Type I error rate set at 0.05, and that the standard deviation of the outcome is 1.

25

Power decreases as correlation within a cluster (ICC) increases



26

For a given cluster size, what is the effect of the intracluster correlation (ICC) on power?

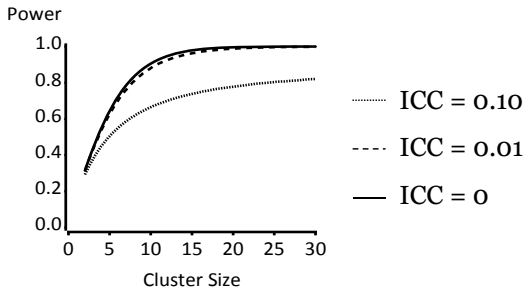
Suppose we have 5 workplaces assigned to no intervention, and 5 assigned to a workplace treatment program.

Assume that the standard deviation is 1, the mean difference is 0.75, the workplace size is 10, and the Type I error rate is 0.05

We vary the ICC from 0 to 0.1.

27

Power decreases as correlation within the cluster (ICC) increases (size 0-30)



28

What is the effect on power of cluster size?

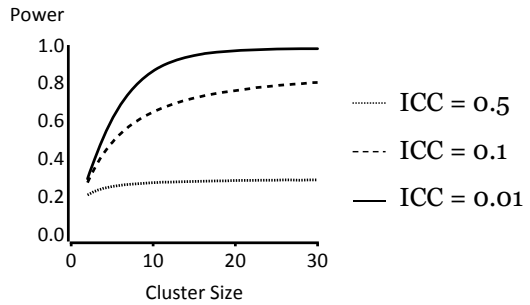
Again suppose we have 20 workplaces assigned to no intervention, and 20 assigned to a workplace treatment program.

Assume we have the Type I error rate set at 0.05.

Vary the cluster size from 10 to 30.

29

Power increases as cluster size increases, with diminishing return



30

What is the effect on power of standard deviation (σ)?

Again suppose we have 20 workplaces assigned to no intervention, and 20 assigned to a workplace treatment program.

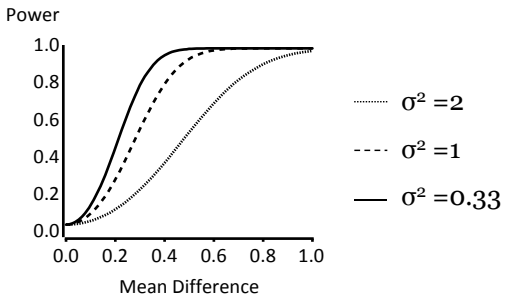
Assume that the cluster size is 10.

Assume the Type I error rate is 0.05.

Change the variance (σ^2) from 0.33 to 2.

31

Power decreases as variance increases



32

For a given mean difference, what cluster size should I choose?

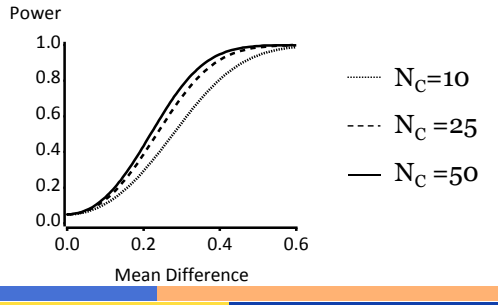
Again suppose we have 20 workplaces assigned to no intervention, and 20 assigned to a workplace treatment program.

Assume that the standard deviation is 1, the Type I error rate is 0.05 and the intracluster correlation (ICC) is 0.1.

Change the cluster size from 10 to 50.

33

For a fixed number of clusters, power increases as cluster size (N_C) increases, with diminishing return



34

For a given mean difference, how does Type I error rate affect power?

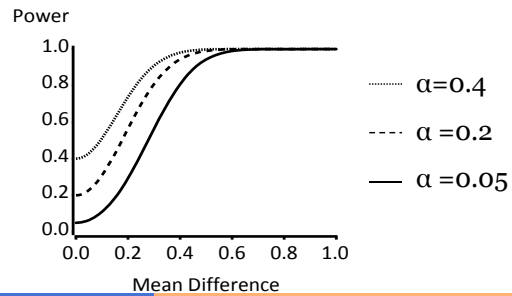
Again suppose we have 20 workplaces assigned to no intervention, and 20 assigned to a workplace treatment program.

Assume that the standard deviation is 1, the workplace size is 10, and the intracluster correlation (ICC) is 0.1.

Change the Type I error rate from 0.05 to 0.4.

35

Power increases as Type I error rate(α) increases



36

Is there a big difference in power between two designs with different Type I error rates?

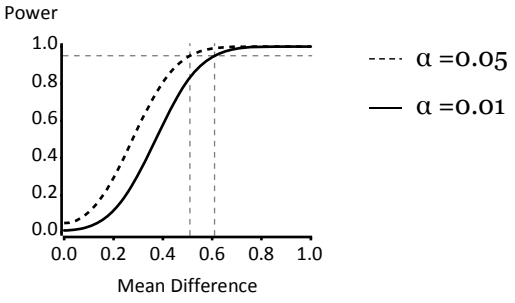
Again suppose we have 20 workplaces assigned to no intervention, and 20 assigned to a workplace treatment program.

Assume that the standard deviation is 1, the workplace size is 10, and the intracluster correlation (ICC) is 0.1.

Vary the Type I error rate from 0.01 to 0.05.

37

There is little change in power for different Type I error rates



38

What is the relationship between the following design inputs and power?

Intracluster correlation ↑ Power ↓

39

What is the relationship between the following design inputs and power?

Cluster size Power

Increase follows a law of diminishing return.

40

What is the relationship between the following design inputs and power?

Standard deviation Power

41

What is the relationship between the following design inputs and power?

Type I error rate Power

42

Review Summary

- Power curves are graphical representations of power calculations that "tell stories"
- Design inputs affect power, and this can be displayed through power curves:
  - Direct relationship with power:
    - Cluster size
    - Type I error rate
  - Inverse relationship with power:
    - Intracluster correlation
    - Standard deviation

43