CHAPTER 8
POWER & EFFECT SIZE

For EDUC/PSY 6600

“ Cohen (1994): “Next, I have learned and taught that the primary product of research inquiry is one or more measures of effect size, not *p* values.” (p. 1310).

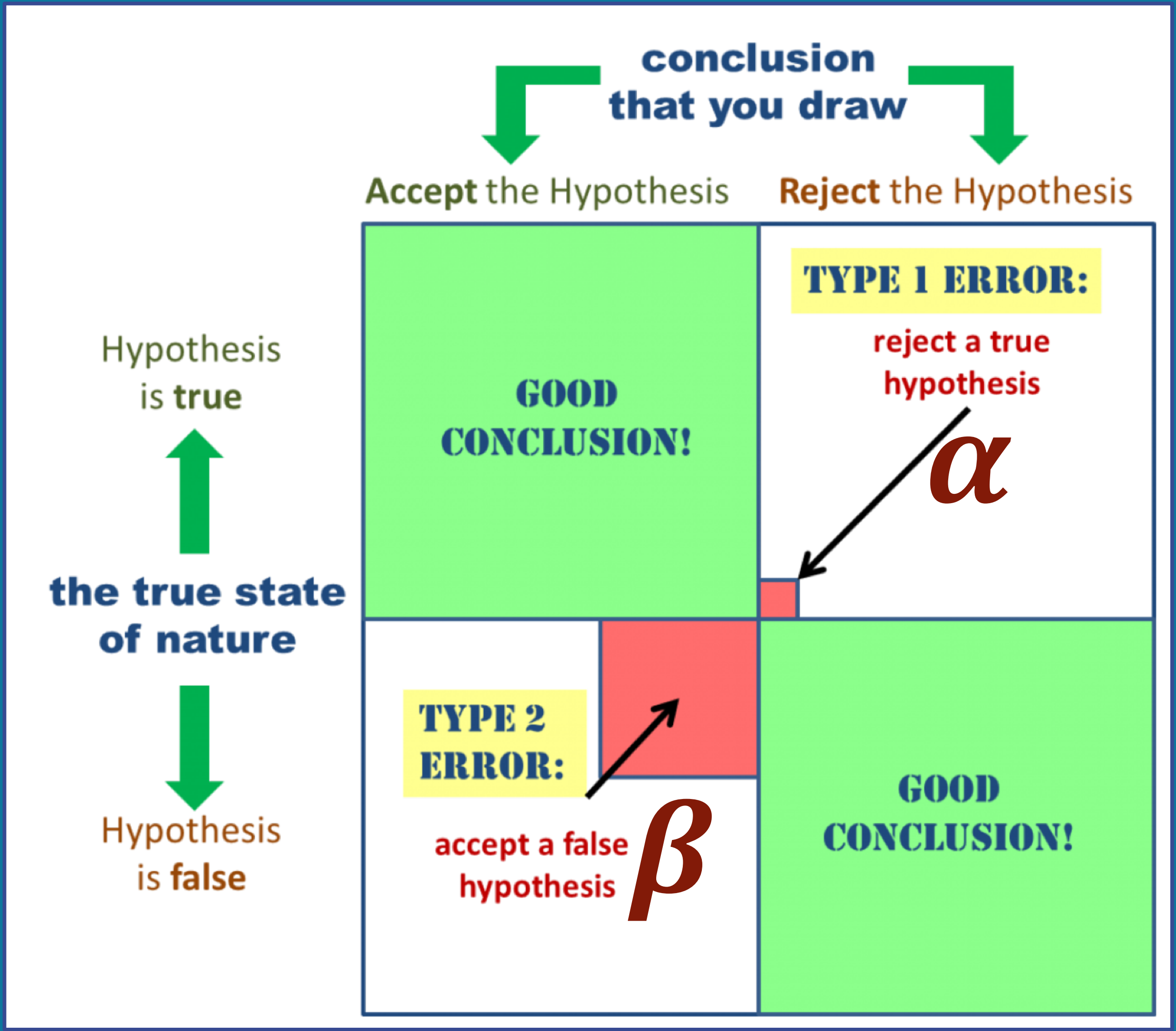
Abelson (1995): “However, as social scientists move gradually away from reliance on single studies and obsession with null hypothesis testing, effect size measures will become more and more popular” (p. 47). ”

Types of Errors

When we conduct a hypothesis test,

we wither reject or fail to reject the Null Hypothesis.

Our decision usually causes four outcomes:



Types of Errors

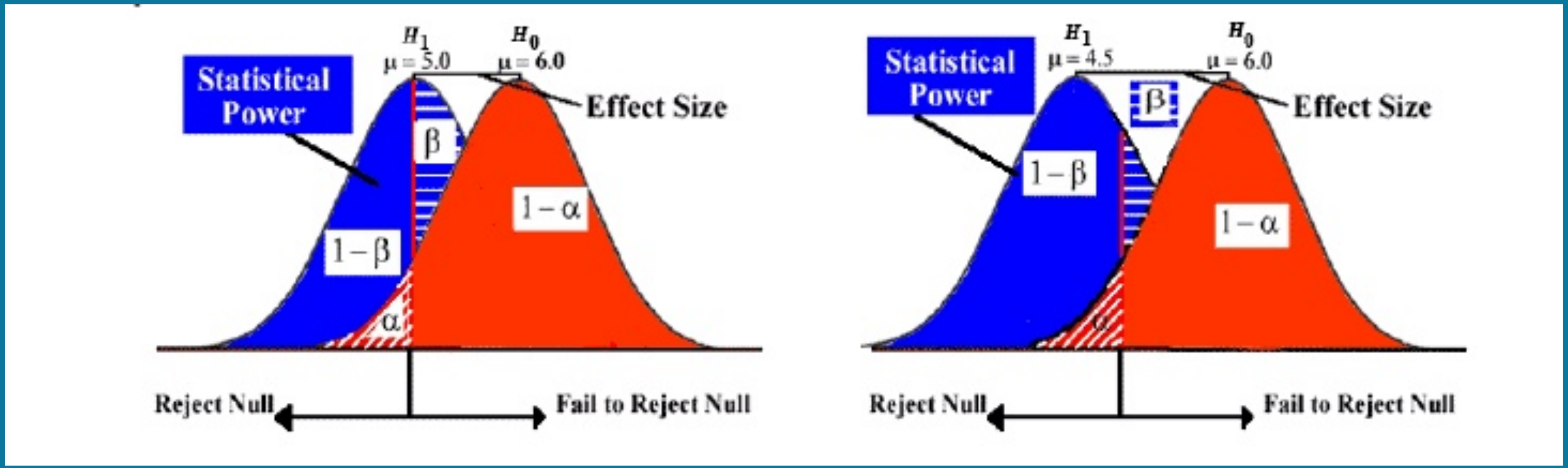
$$\underline{\text{Power} = 1 - \beta}$$

"the probability of
correctly rejecting
a false null hypothesis."

Effect Sizes

$$\text{Cohen's } d = \frac{\bar{X}_1 - \bar{X}_2}{s_p} \text{ or } t\sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

$$\eta^2 = r_{pb}^2 = \frac{t^2}{t^2 + (n_1 + n_2 - 2)}$$



Effect Sizes

$$\text{Cohen's } d = \frac{\bar{X}_1 - \bar{X}_2}{s_p} \text{ or } t\sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

Cohen's d	Interpretation
.2	Small
.5	Moderate
.8	Large

Effect Sizes

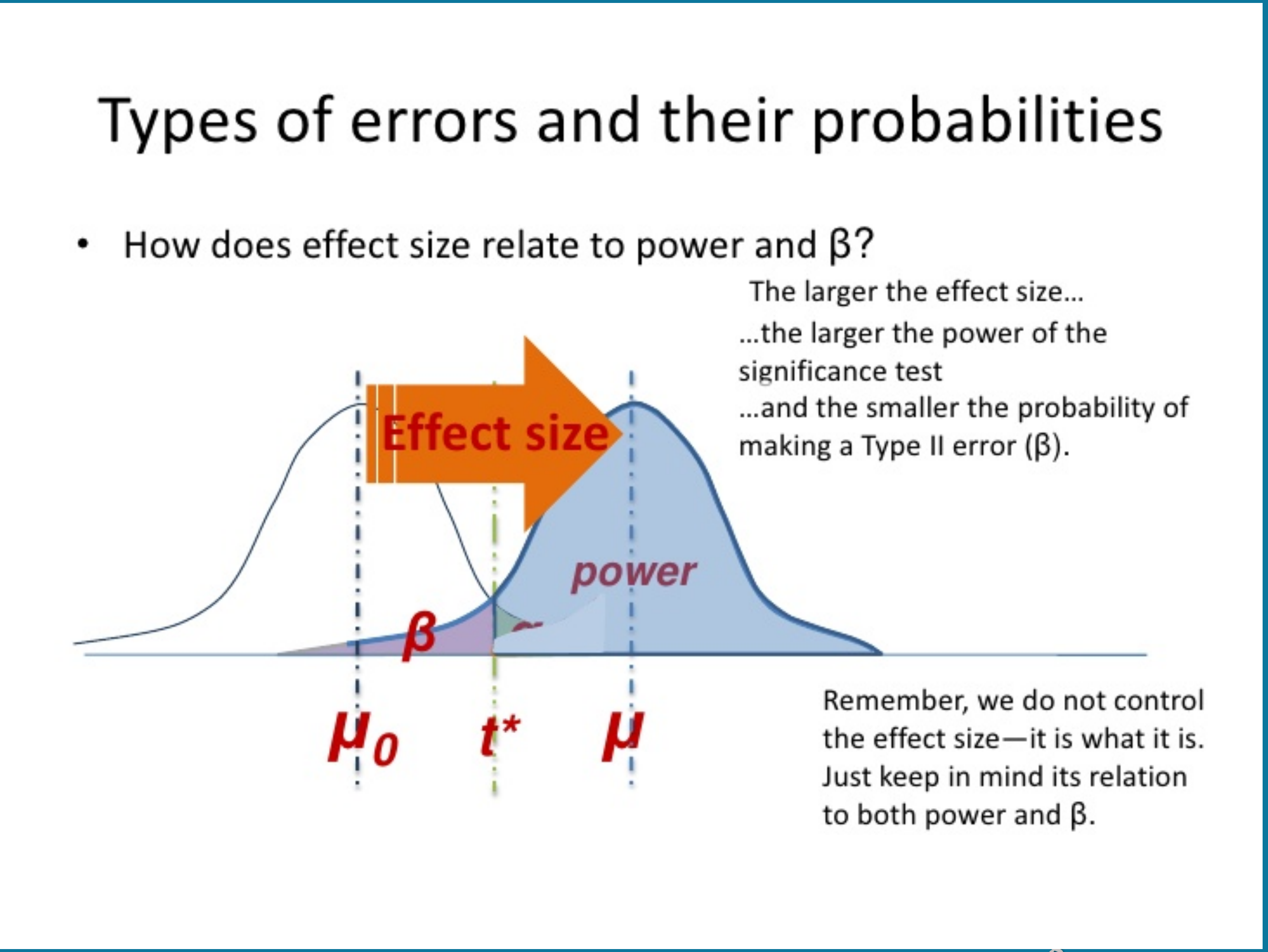
$$\eta^2 = r_{pb}^2 = \frac{t^2}{t^2 + (n_1 + n_2 - 2)}$$

η^2 (eta squared) and r_{pb}^2

- association between grouping variable (IV) and continuous DV
- Ranges from 0 to 1
- With only 2 groups, results are same

What affects power?

1. Sample Size
   - Larger sample = more power
2. Effect Size
   - Larger Effect size = more power
3. Alpha Level
   - Higher Alphas = more power
4. Directionality
   - One tail = more power



Power Analysis

- Non-centrality parameter is calculated by:

$$\delta = \frac{d}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- Since it's assumed that the…
  - Variances are same in 2 groups
  - *N*'s are same in 2 groups
- ...and since σ is often assumed to be 1…
- …the equation is simplified…

	One-Tailed Test (α)			
	.05	.025	.01	.005
	Two-Tailed Test (α)			
δ	.10	.05	.02	.01
0.5	.14	.08	.03	.02
0.6	.16	.09	.04	.02
0.7	.18	.11	.05	.03
0.8	.21	.13	.06	.04
0.9	.23	.15	.08	.05
1.0	.26	.17	.09	.06
1.1	.29	.20	.11	.07
1.2	.33	.22	.13	.08
1.3	.37	.26	.15	.10
1.4	.40	.29	.18	.12
1.5	.44	.32	.20	.14
1.6	.48	.36	.23	.16
1.7	.52	.40	.27	.19
1.8	.56	.44	.30	.22
1.9	.60	.48	.33	.25
2.0	.64	.52	.37	.28
2.1	.68	.56	.41	.32
2.2	.71	.60	.45	.35
2.3	.74	.63	.49	.39
2.4	.77	.67	.53	.43
2.5	.80	.71	.57	.47
2.6	.83	.74	.61	.51
2.7	.85	.77	.65	.55
2.8	.88	.80	.68	.59
2.9	.90	.83	.72	.63

When $n_1 = n_2$

$$\delta = \mathbf{d}\sqrt{\frac{n_k}{2}}$$

$$n_k = 2\left(\frac{\delta}{\mathbf{d}}\right)^2$$

When $n_1 \neq n_2$

$$\bar{n}_h = \frac{2}{\frac{1}{n_1} + \frac{1}{n_2}} = \frac{2n_1 n_2}{n_1 + n_2}$$

$$\delta = \mathbf{d}\sqrt{\frac{\bar{n}_h}{2}}$$

FORMULA SHEET

$$\delta = \text{"EXPECTED T OR Z"}\ (population\ parameters)$$

$$1\ group\!:\ \ \delta = \frac{\mu}{\sigma}\sqrt{n} \xrightarrow{d=\frac{\mu}{\sigma}} \delta = d\sqrt{n}$$

$$2\ groups\!:\ \ \delta \xRightarrow{equal\ n's} \frac{\mu_1 - \mu_2}{\sigma}\sqrt{\frac{n}{2}} \xrightarrow{d=\frac{\mu_1-\mu_2}{\sigma}} \delta = d\sqrt{\frac{n}{2}}$$

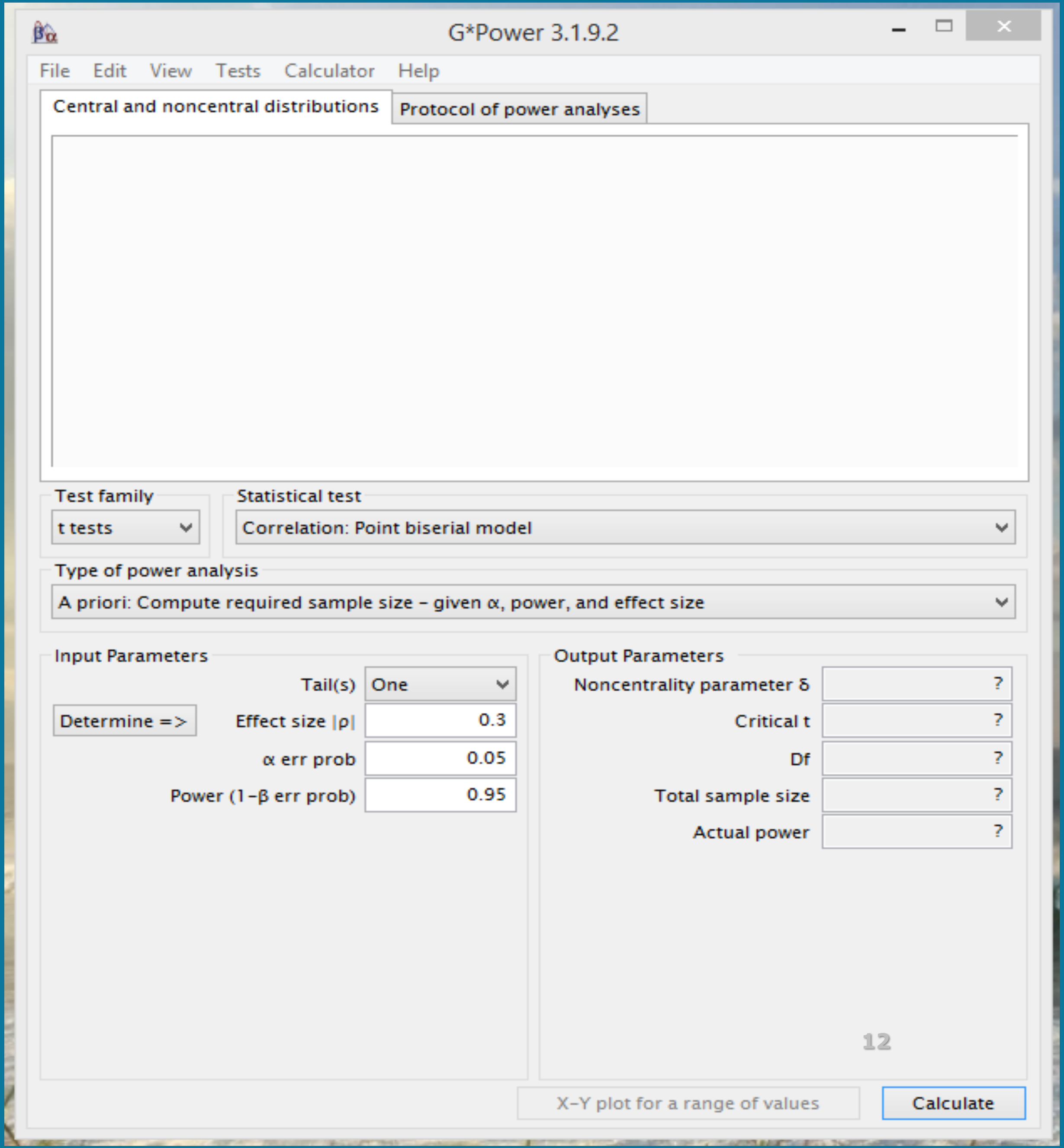
$$\xLeftrightarrow{est.\ \mathbf{d} = g\left(1 - \frac{3}{4\,df - 1}\right)}$$

$$g = \text{"}effect\ size\text{"}\ (sample\ statistics)$$

$$g = \frac{\bar{X}_1 - \bar{X}_2}{s_p} \begin{cases} = t\sqrt{\dfrac{2}{n}} \xleftarrow{n_1 = n_2} \\ = t\sqrt{\dfrac{n_1 + n_2}{n_1 n_2}} \xleftarrow{n_1 \neq n_2} \end{cases}$$

G-POWER

Download at: http://www.gpower.hhu.de/



CHAP 8: SECTION A

- **d** is just the number of standard deviations that separate two population means

- ***g*** is the number of standard deviations (based on pooling the sample variances and taking the square-root) separating the sample means.

- connection between a calculated *t* and delta;
  - large *t's* are *usually* associated with large deltas
  - small *t's* *usually* with small deltas.
  - Of course, the alternate hypothesis distribution shows that *t* can occasionally come out very differently from delta

An estimate of power
is only as good as
the estimate of effect size upon which it is based

…BUT determining the effect size is usually
the purpose (or should be) of the experiment.