

Power Analysis Basics

Dr. Mark
Williamson



BERDC
UND



1. What is Power Analysis?

- Power is the probability of detecting an effect, given that the effect is truly there [1]
 - Example, the effect of a treatment, measured by comparing the mean between a treatment group and a control group
- The most common use of power analysis is to determine sample sizes for experiments
 - Too few, you could miss a true effect
 - Too many, you could waste resources or unethically expose more patients to risk
- Requirements for some studies

2. Components needed.

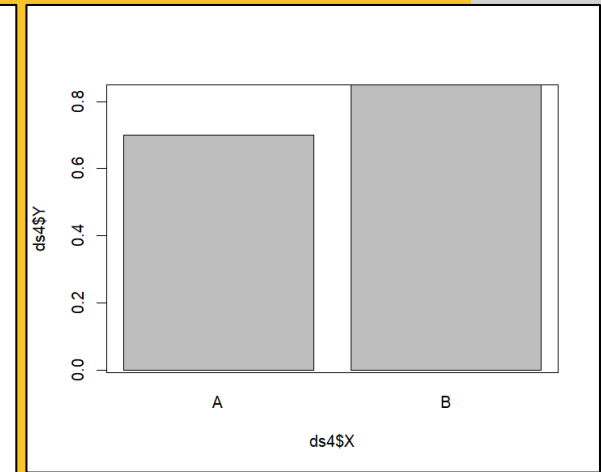
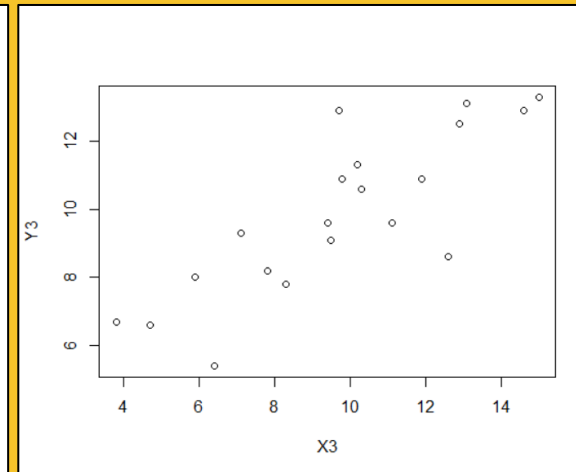
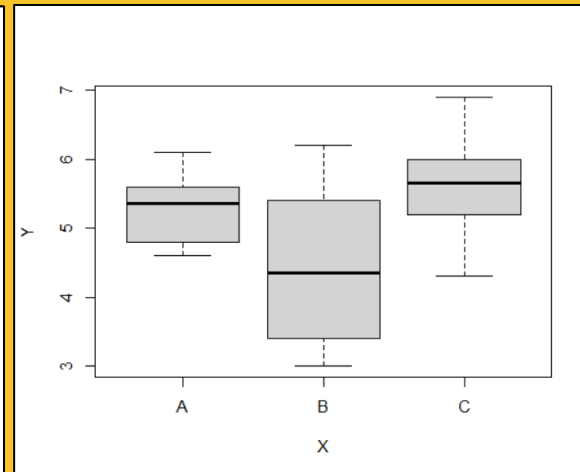
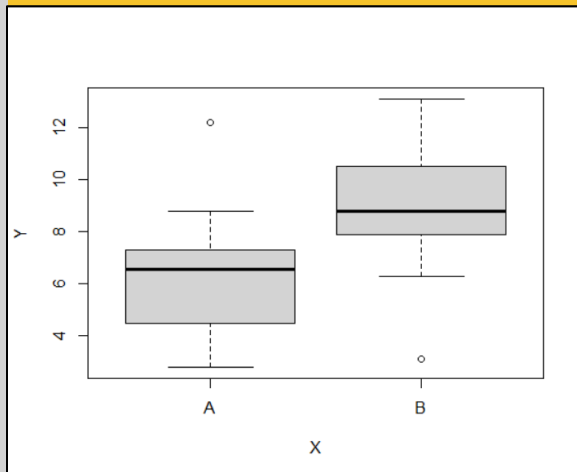
- In a basic power analysis, there are four components [2]:
 - Alpha (α) = risk of a Type I error (false positive)
 - Power = 1 - risk of a Type II error (false negative)
 - Sample size (n) = number of samples needed to detect the effect
 - Effect size = magnitude of the effect under the alternative hypothesis
- Alpha and Power have defaults:
 - Alpha = 0.05
 - Power = 0.80
- Effect size is the major unknown:
 - Can be estimated
 - Can be guessed

3. The Challenge of Effect Size.

- **Effect size matters the most for sample size**
 - The higher the effect size, the fewer samples needed
- **Calculating Effect size:**
 - Prior studies (pilot, trial, etc.)
 - Similar studies (same or related field)
 - Other pertinent background information
- **Guessing Effect size:**
 - Rules of thumb for small, medium, large effects
 - Enter plausible values to effect size equation

4. Basic Comparisons.

- Comparing 2 groups (T-test)
- Comparing 3+ groups (ANOVA)
- Correlation
- Proportions



5. Guessing Effect Size.

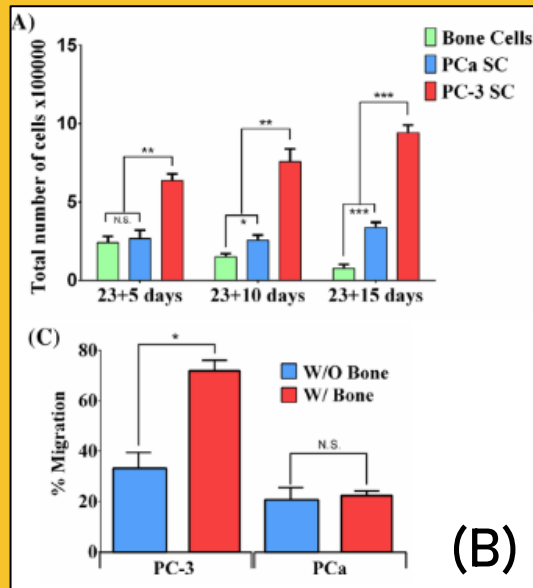
- Rule of thumb values [3,4]
- Enter plausible equation values
 - T-test: Two-fold difference, SD 1/2 of mean

$$d = \frac{10 - 5}{5} = 1$$

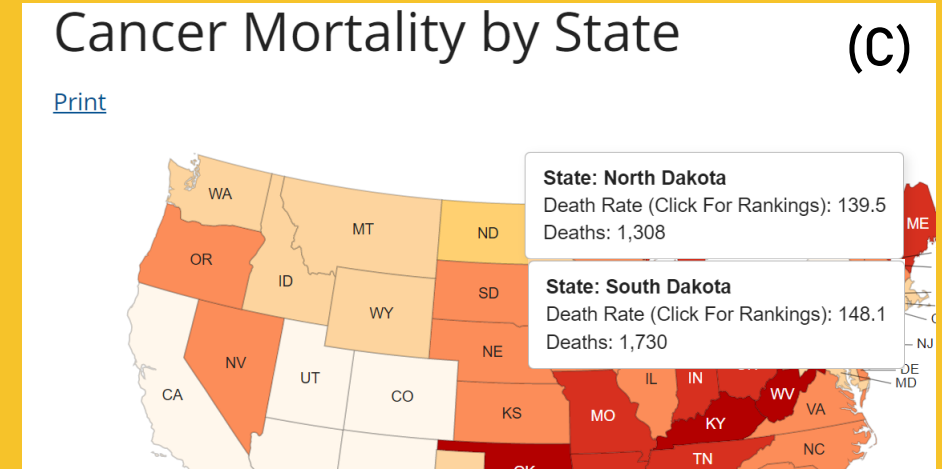
| | Effect Size | | |
|-----------------|-------------|--------|-------|
| Test | Small | Medium | Large |
| T-test (d) | 0.20 | 0.50 | 0.80 |
| ANOVA (f) | 0.10 | 0.25 | 0.40 |
| Correlation (R) | 0.10 | 0.30 | 0.50 |

6. Calculating Effect Size.

- Raw data (A)
- Literature search (B)
- Summary statistics (C)

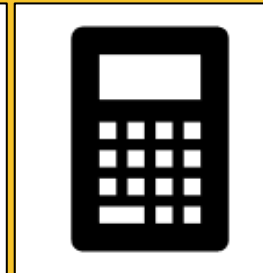


| | Treat | Control |
|------|-------|---------|
| | 6.8 | 4.7 |
| | 3.2 | 3.3 |
| | 3.4 | 5 |
| | 5.1 | 2.5 |
| | 7.9 | 2.4 |
| | 7.4 | 1.9 |
| | 8.8 | 0.6 |
| | 4.4 | 2.6 |
| | 5.1 | 7.2 |
| | 8.2 | 5 |
| mean | 6 | 3.5 |
| SD | 2 | 1.9 |



7. Tools to Use.

- By hand with equations (!!!)
- G*Power[5]
- Online calculator
 - Powerandsamplesize.com[6]
 - GIGAcaculator[7]
 - Statistics Kingdom[8]
 - Sample-size.net[9]
- Software
 - R: package 'pwr'
 - SPSS -> Analyze -> Power Analysis



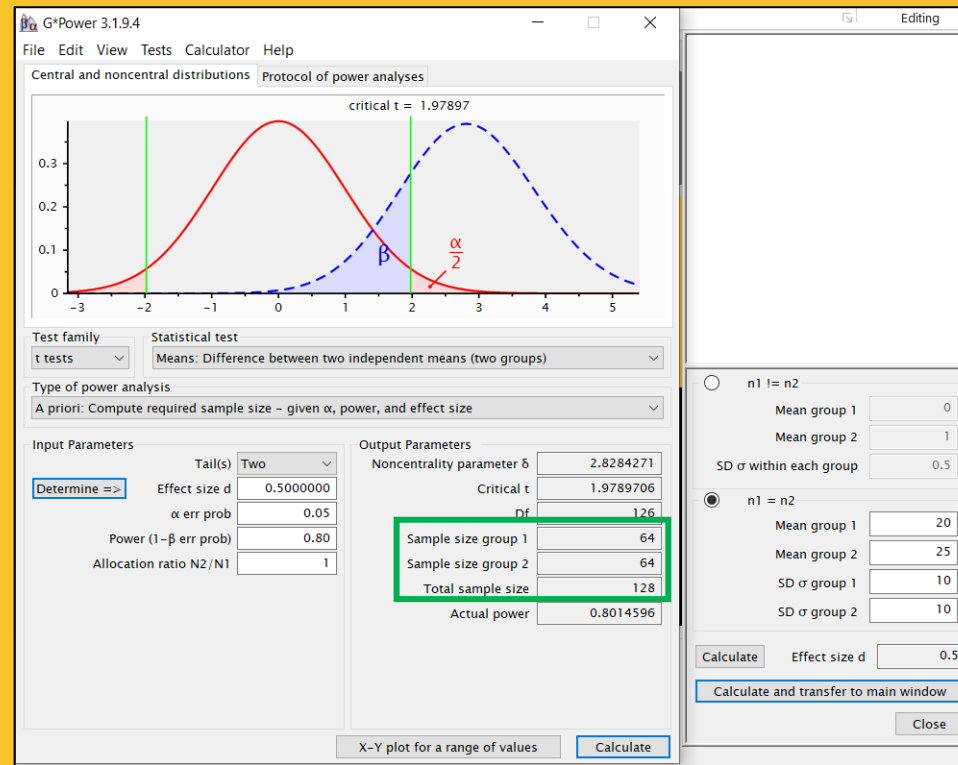
8. T-test Example.

- Data:

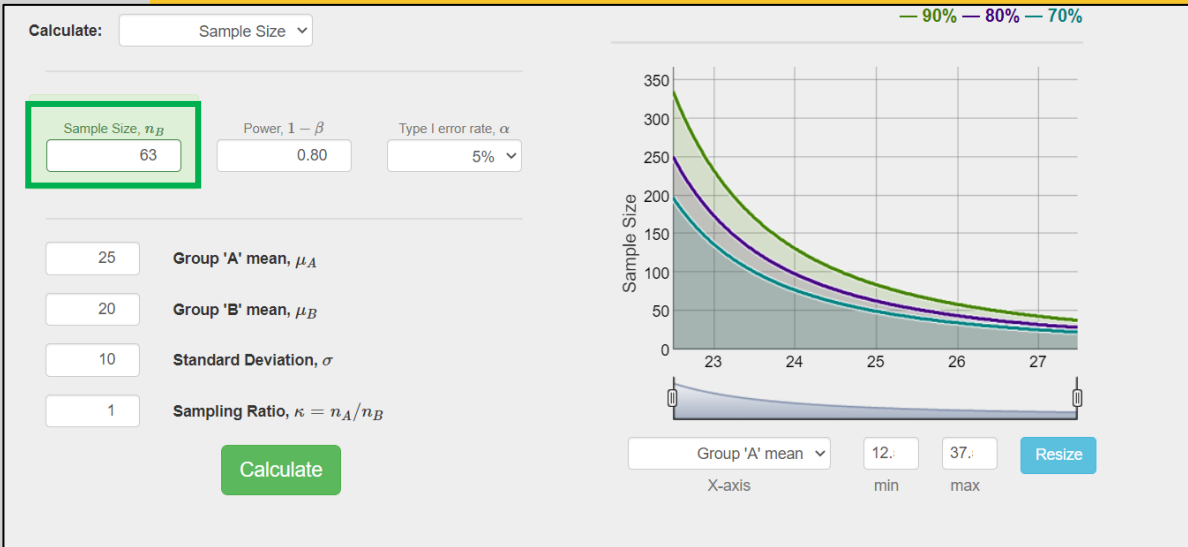
| | Treat | Control |
|------|-------|---------|
| mean | 20 | 25 |
| std | 10 | 10 |

- Method:

- G*Power
- Online calculator
- R
- SPSS



8. T-test Example.



```
R 4.2.1 · ~/    
> d <-(25-20)/10  
> d  
[1] 0.5  
> pwr.t.test(d=d, sig.level=0.05 , power=0.8, type=c("two.sample"))
```

Two-sample t test power calculation

```
n = 63.76561  
d = 0.5  
sig.level = 0.05  
power = 0.8  
alternative = two.sided
```

NOTE: n is number in *each* group

8. T-test Example.

Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Power Analysis > Means > One-Sample T Test
 Meta Analysis > Proportions > Paired-Samples T Test
 Reports > Correlations > Independent-Samples T Test
 Descriptive Statistics > Regression > One-Way ANOVA

Visible: 0 of 0 Variables

Power Analysis: Independent-Sample Means

Test Assumptions

Estimate: Sample size

Single power value: 0.80

Grid power values: Grid

Grid values: None selected

Group size ratio: 1

Specify: Hypothesized Values

Population mean difference: (empty)

Population mean for group 1: 20 and group 2: 25

Population standard deviations are

Equal for two groups

Pooled standard deviation: 10

Not equal for two groups

Standard deviation for group 1: 10 and group 2: 10

Test Direction

Nondirectional (two-sided) analysis

Directional (one-sided) analysis

Significance level: 0.05

OK Paste Reset Cancel Help

Power Analysis Table

| Test for Mean Difference ^a | N | | Actual Power ^b | Power | Test Assumptions | | |
|---------------------------------------|----|----|---------------------------|-------|------------------------|-------------|------|
| | N1 | N2 | | | Std. Dev. ^c | Effect Size | Sig. |
| | 64 | 64 | .801 | .8 | 10 | .500 | .05 |

a. Two-sided test.
 b. Based on noncentral t-distribution.
 c. Group variances are assumed to be equal.

9. ANOVA Example.

- **Data:**

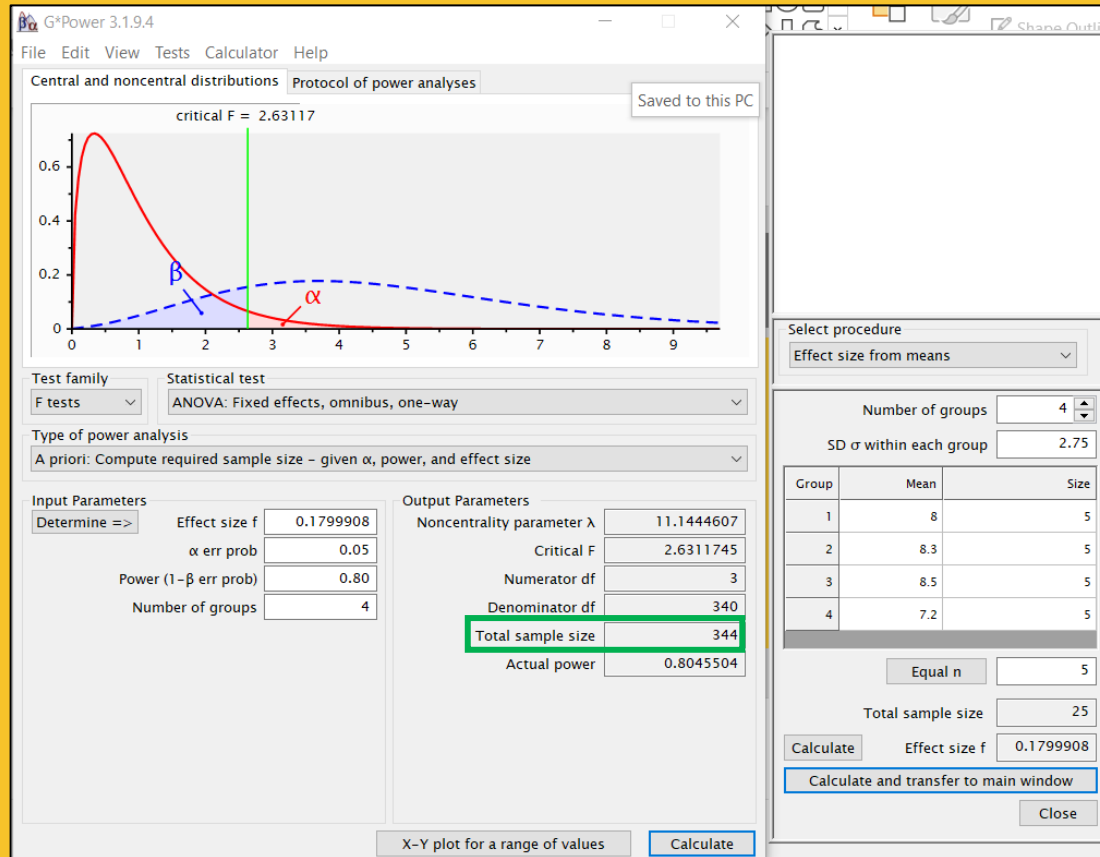
| | A | B | C | D |
|------|-----|-----|-----|-----|
| mean | 8.0 | 8.3 | 8.5 | 7.2 |
| std | 2.1 | 3.2 | 2.2 | 3.5 |

| [10, 11] | Df | SS | Mean | F | p |
|-----------|----|-------|-------|-------|-------|
| X var | 3 | 6.5 | 2.163 | 0.211 | 0.888 |
| Residuals | 36 | 369.2 | 2.2 | | |

- **Method:**

- G*Power
- ~~Online calculator~~
- R
- SPSS

9. ANOVA Example.



```
R 4.2.1 · ~/
```

```
> n2 <- 6.5/(6.5+369.2)
> n2
[1] 0.01730104
> f <-sqrt(n2/(1-n2))
> f
[1] 0.1326862
> pwr.anova.test(k=4 , f=f, sig.level=0.05, power=0.8)

Balanced one-way analysis of variance power calculation

      k = 4
      n = 155.7967 Using f calculated from eta squared ( $\eta^2$ )
      r = 0.1326862
sig.level = 0.05
power = 0.8

NOTE: n is number in each group

>
> f.2 <-0.179
> pwr.anova.test(k=4 , f=f.2, sig.level=0.05, power=0.8)

Balanced one-way analysis of variance power calculation

      k = 4
      n = 86.05041 Using f calculated by G*Power
      f = 0.179
sig.level = 0.05
power = 0.8

NOTE: n is number in each group
```

9. ANOVA Example.

Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Power Analysis > Means > One-Sample T Test
Meta Analysis > Proportions > Paired-Samples T Test
Reports > Correlations > Independent-Samples T Test
Descriptive Statistics > Regression > One-Way ANOVA

Power Analysis: One-way ANOVA

Test Assumptions

Estimate: Sample size v

Single power value: 0.8 Grid power values: Grid

Grid values: None selected

Specify: Hypothesized Values v

Pooled population standard deviation: 2.75

| | Group sizes | Group weights | Group means |
|-------------------------------------|-------------|---------------|-------------|
| <input checked="" type="checkbox"/> | 1 | 8 | |
| <input checked="" type="checkbox"/> | 1 | 8.3 | |
| <input checked="" type="checkbox"/> | 1 | 8.5 | |
| <input checked="" type="checkbox"/> | 1 | 7.2 | |

Total sample size: A minimum of two groups are required

Significance level: 0.05

OK Paste Reset Cancel Help

Power Analysis - One-way ANOVA

| Overall Test ^a | N ^b | Actual Power ^c | Power | Test Assumptions | | |
|---------------------------|----------------|---------------------------|-------|------------------|--------------------------|------|
| | | | | Std. Dev. | Effect Size ^d | Sig. |
| | 344 | .805 | .8 | 2.75 | .208 | .05 |

a. Test the null hypothesis that population mean is the same for all groups.
b. Total sample size across groups.
c. Based on noncentral F-distribution.
d. Effect size measured by the root-mean-square standardized effect.

Group Size Allocation for Overall Test

| N | |
|---------|-----|
| Group 1 | 86 |
| Group 2 | 86 |
| Group 3 | 86 |
| Group 4 | 86 |
| Overall | 344 |

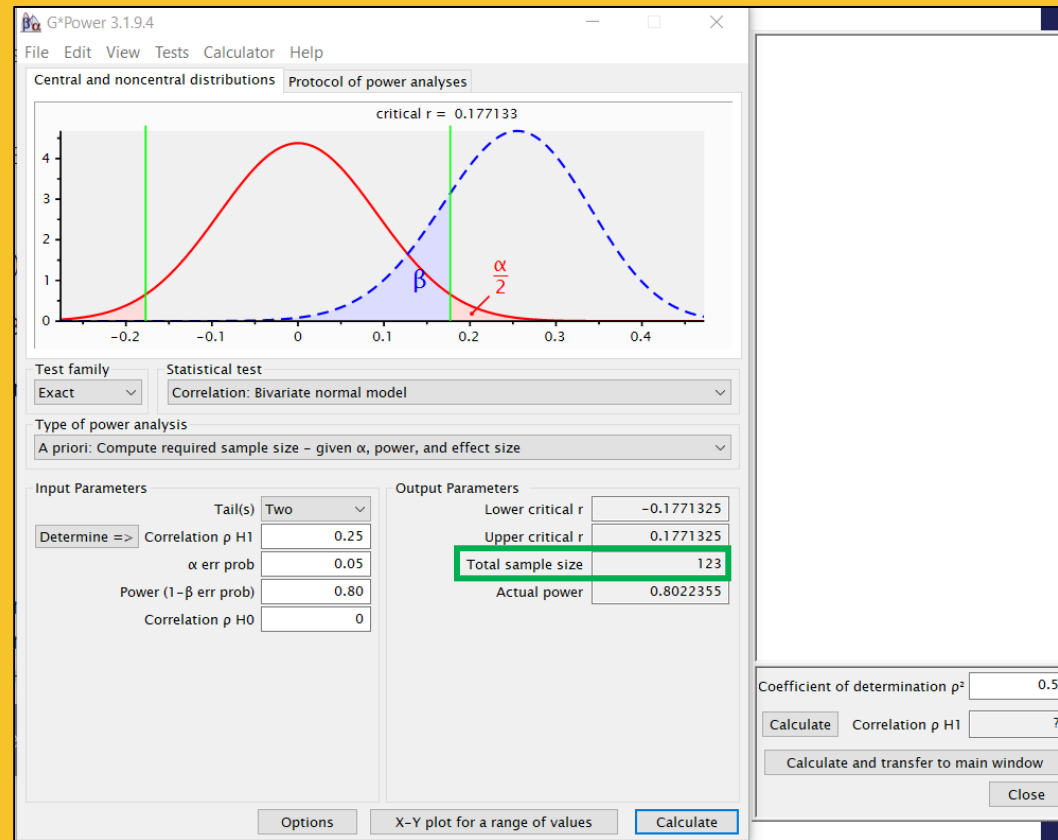
10. Correlation Example.

• Data:

| | H1 | H0 |
|-------------|------|------|
| Correlation | 0.25 | 0.00 |

• Method:

- G*Power
- Online calculator
- R
- SPSS



10. Correlation Example.

Correlation sample size

Total sample size required to determine whether a correlation coefficient differs from zero.

Instructions: Enter parameters in the **green** cells. Answers will appear in the **blue** box below.

α (two-tailed) = Threshold probability for rejecting the null hypothesis. Type I error rate.
 β = Probability of failing to reject the null hypothesis under the alternative hypothesis. Type II error rate.
 r = The expected correlation coefficient.

Calculate

The standard normal deviate for $\alpha = Z_\alpha = 1.9600$

The standard normal deviate for $\beta = Z_\beta = 0.8416$

$C = 0.5 * \ln[(1+r)/(1-r)] = 0.2554$

Total sample size = $N = [(Z_\alpha + Z_\beta)/C]^2 + 3 = 123$

R 4.2.1 · ~/ ↻

```
> pwr.r.test(r=0.25, sig.level=0.05, power=0.80)
```

approximate correlation power calculation (arctangh transformation)

n = 122.4466

r = 0.25

sig.level = 0.05

power = 0.8

alternative = two.sided

10. Correlation Example.

The screenshot shows the IBM SPSS Statistics Data Editor interface. The 'Analyze' menu is open, and the 'Correlations' sub-menu is selected, showing options for 'Pearson Product-Moment', 'Spearman Rank-Order', and 'Partial'. The 'Power Analysis: Pearson Correlation' dialog box is open, displaying the following settings:

- Test Assumptions: Estimate: Sample size
- Single power value: .8
- Grid power values: Grid (Grid values: None selected)
- Pearson correlation parameter: 0.25
- Null value: 0
- Use bias-correction formula in the power estimation:
- Test Direction: Nondirectional (two-sided) analysis
- Significance level: 0.05

Power Analysis Table

| | N | Actual Power ^b | Power | Test Assumptions | | |
|----------------------------------|-----|---------------------------|-------|------------------|-------------|------|
| | | | | Null | Alternative | Sig. |
| Pearson Correlation ^a | 123 | .802 | .8 | 0 | .25 | .05 |

a. Two-sided test.
b. Based on Fisher's z-transformation and normal approximation with bias adjustment.

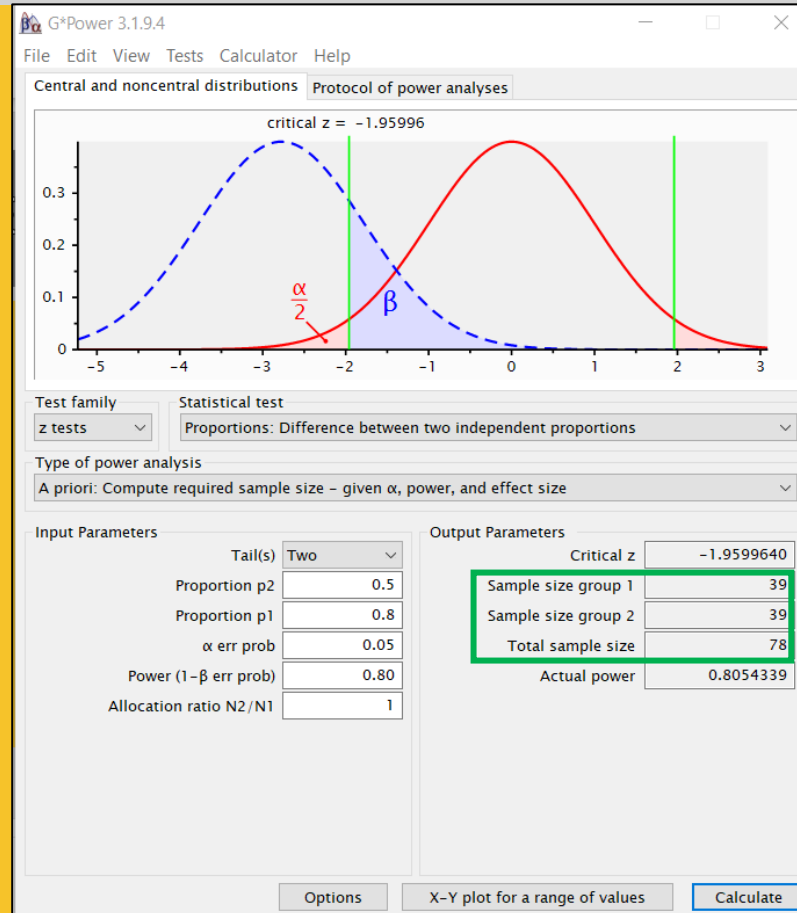
11. Proportion Example.

- Data:

| | Treat | Control |
|------------|-------|---------|
| proportion | 0.50 | 0.80 |

- Method:

- G*Power
- Online calculator
- R
- SPSS



11. Proportion Example.

Calculate h effect size calculator

P₁
0.8

P₂
0.5

Calculate h

Sample:
Two samples

Statistical power:
0.8

h effect size:
0.643501108793284

Expected h:
Calculate h

Rounding:
4

Calculate Clear

How to do with R?

The sample size **38** for each group, will gain the power of **0.800943**. (n₁ = n₂ = 38)

```
R 4.2.1 · ~ / ↻  
> h <- 2*asin(sqrt(0.8))-2*asin(sqrt(0.5))  
> h  
[1] 0.6435011  
> pwr.2p.test(h=h, sig.level=0.05, power=0.8, alternative=c("two.sided"))  
  
Difference of proportion power calculation for binomial distribution (arcsine transformation)  
  
      h = 0.6435011  
      n = 37.90862  
sig.level = 0.05  
power = 0.8  
alternative = two.sided  
  
NOTE: same sample sizes
```

11. Proportion Example.

Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Power Analysis > Means > Proportions > One-Sample Binomial Test
 Proportions > Related-Samples Binomial Test
 Regression > Independent-Samples Binomial Test

Visible: 0 of 0 Variables

Power Analysis: Independent-Sample Proportions

Test Assumptions

Estimate: Sample size

Single power value: 0.80 Grid power values: Grid
 Grid values: None selected

Group size ratio: 1

Proportion parameters for group 1: 0.50 and group 2: 0.80

Significance level: 0.05

Test Method

Chi-squared test
 Standard deviation is pooled
 Apply continuity correction
 I-test
 Standard deviation is pooled
 Likelihood ratio test
 Fisher's exact test

Estimation Method

Normal approximation
 Binomial enumeration
 Time limit: 5 minutes

Test Direction

Nondirectional (two-sided) analysis
 Directional (one-sided) analysis

OK Paste Reset Cancel Help

Data View Variable View

Independent-Samples Binomial Test

IBM SPSS Statistics Processor is ready Unicode:ON

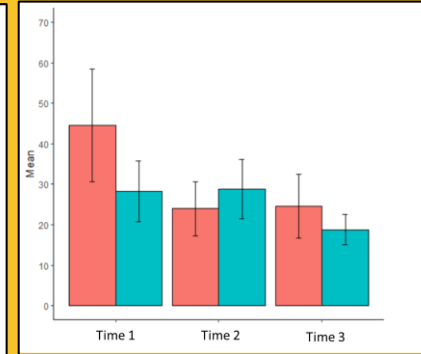
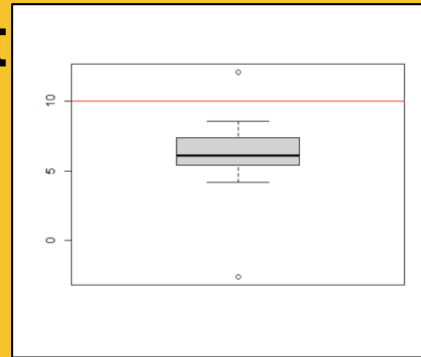
Power Analysis Table

| Test for Proportion Difference ^a | N | | Actual Power ^b | Power | Test Assumptions | | | |
|---|----|----|---------------------------|-------|------------------|------------|------------|------|
| | N1 | N2 | | | Risk Difference | Risk Ratio | Odds Ratio | Sig. |
| Test for Proportion Difference ^a | 39 | 39 | .805 | .8 | -.300 | .625 | .250 | .05 |

a. Two-sided test using large-sample approximation.
 b. The estimation of power is based on the Pearson Chi-Square test and the pooled standard deviation.

12. Additional Designs.

- 1-sample or Paired T-test
- Repeated Measures
- Regression
- Chi-squared
- ...



| Expected | Observed |
|----------|----------|
| 9 | 56 |
| 3 | 18 |
| 3 | 7 |
| 1 | 7 |

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \epsilon$$

“A Power Analysis for every analysis”

13. Practical Limitations.

- Understanding what goes in.
- Simulation sanity check. (A)
- Realistic patient/sample number.
- Balance of art and science.



(A)

| | |
|--------------------|----|
| Mean A | 30 |
| Mean B | 22 |
| Std | 12 |
| $d = (30-22) / 12$ | |
| $d = 0.667$ | |

```
> pwr.t.test(d=0.667, sig.level=0.05, power=0.8, type=c("two.sample"))

Two-sample t test power calculation

      n = 36.27042
      d = 0.667
sig.level = 0.05
  power = 0.8
alternative = two.sided

NOTE: n is number in *each* group
```

```
> Treat<-c(rep("A",37), rep("B",37))
> Response<-round(c(rnorm(37, mean=30, sd=12), rnorm(37, mean=22, sd=12)),1)
> ss1 <-data.frame(Treat=Treat, Response=Response)
> t.test(Response~Treat, data=ss1)

Welch Two Sample t-test

data: Response by Treat
t = 2.9245, df = 69.914, p-value = 0.004647
alternative hypothesis: true difference in means between group A and group B is not equal to 0
95 percent confidence interval:
 2.942929 15.565179
sample estimates:
mean in group A mean in group B
 31.42432      22.17027
```

14. Where to Next?

- More resources

- [Power Analysis in G*Power](#) (videos and slides)
- [Power Analysis in R](#) (videos and slides)
- [Power Analysis in R: GLMMs](#) (videos and slides)
- [Advanced Power Analysis: Into the Weeds](#) (video and slides)

- How to practice

- Design mock experiments
- Calculate/guess effect size
- Get comfortable using calculator/software

- Handy handouts [12]

- Effect sizes for common designs

DACCOTA Statistical Resources

The Biostatistics, Epidemiology, and Research Design Core (BERDC) offers a variety of statistical resources from both our core and other biostatistical cores.

| Effect Sizes for common designs | | | |
|------------------------------------|----------------------------------|--|--|
| Statistical Test | Effect size | Equation | Rule of thumb for effect sizes |
| 1 sample t-test | Cohen's d | $d = (\text{mean} - \text{constant}) / SD$ | small=0.20, medium=0.50, large=0.80 |
| 2 sample t-test | Cohen's d | $d = (\text{mean}_1 - \text{mean}_2) / SD_{\text{pooled}}$ | small=0.20, medium=0.50, large=0.80 |
| Paired t-test | Cohen's d | $d = (\text{mean}_1 - \text{mean}_2) / SD_{\text{pooled}}$ | small=0.20, medium=0.50, large=0.80 |
| 1-Way ANOVA | Eta squared Cohen's f | $\eta^2 = SS_{\text{treatment}} / SS_{\text{total}}$ $f = \sqrt{\eta^2 / (1 - \eta^2)}$ | small=0.01, medium=0.05, large=0.14 small=0.10, medium=0.25, large=0.40 |
| 2-Way ANOVA | Eta squared Cohen's f | $\eta^2 = SS_{\text{treatment}} / SS_{\text{total}}$ $f = \sqrt{\eta^2 / (1 - \eta^2)}$ | small=0.01, medium=0.06, large=0.14 small=0.10, medium=0.25, large=0.40 |
| Repeated Measures ANOVA | Partial Eta squared Cohen's f | Partial $\eta^2 = SS_{\text{effect}} / (SS_{\text{effect}} + SS_{\text{error}})$... | small=0.01, medium=0.06, large=0.14 small=0.10, medium=0.25, large=0.40 |
| 1 proportion test | Cohen's h | $h = 2 * \text{asin}(\sqrt{\text{prop}_1}) - 2 * \text{asin}(\sqrt{\text{prop}_{\text{control}}})$ | small=0.20, medium=0.50, large=0.80 |
| 2 proportions test | Cohen's h | $h = 2 * \text{asin}(\sqrt{\text{prop}_1}) - 2 * \text{asin}(\sqrt{\text{prop}_2})$ | small=0.20, medium=0.50, large=0.80 |
| Chi-squared test | Cohen's w | $w = \sqrt{(\text{prop}_{\text{obs}} - \text{prop}_{\text{exp}})^2 / \text{prop}_{\text{exp}}}$ | small=0.10, medium=0.30, large=0.50 |
| Pearson Correlation | Correlation (R) | ... | small=0.10, medium=0.30, large=0.50 |
| Linear Regression (Entire Model) | F squared | $F^2 = R^2_{\text{model}} / (1 - R^2_{\text{model}})$ | small=0.02, medium=0.15, large=0.35 |
| Linear Regression (Ind. Predictor) | F squared | $F^2 = R^2_{\text{increase}} / (1 - R^2_{\text{increase}})$ | small=0.10, medium=0.30, large=0.50 |

15. References and Acknowledgement.

- [1] <https://stats.oarc.ucla.edu/other/mult-pkg/seminars/intro-power/Others>
- [2] <https://www.coursehero.com/file/p7g5rdk/Power-Analysis-Components-There-are-four-components-to-a-power-analysis-three/>
- [3] <https://www.statology.org/effect-size/>
- [4] <https://www.spss-tutorials.com/effect-size/#anova>
- [5] <https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower>
- [6] <http://powerandsamplesize.com/Calculators/>
- [7] <https://www.gigacalculator.com/calculators/power-sample-size-calculator.php>
- [8] <https://www.statskingdom.com/statistical-power-calculators.html>
- [9] <https://sample-size.net>
- [10] <https://www.statology.org/eta-squared/>
- [11] <https://www.statology.org/partial-eta-squared/>
- [12] <https://docs.google.com/spreadsheets/d/1dqBPqj3VfiHC3oZE4azLypiFOQaeoj9HQ8Z5yj0vybs/edit#gid=0>

The DaCCoTA is supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729.

For the labs that use the Biostatistics, Epidemiology, and Research Design Core in any way, including this Module, please acknowledge us for publications: ***"Research reported in this publication was supported by DaCCoTA (the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729)"***