How to drive better research outcomes with Power Analysis

Rick Marcantonio
Quality Assurance Engineer, SPSS Statistics

IBM SPSS Statistics
Accelerate research and analysis with the leading statistical software solution
Power Analysis

• What is it?
• When do I use it?
• Where do I find it?
• Why use it?
What is Power Analysis?


Used when performing a statistical test
- IBM SPSS Statistics offers power analysis for a variety of tests: Means, proportions, correlations, and regression

Statistically significant, not necessarily clinically meaningful
Types of Errors

Statistical tests carry with them some amount of uncertainty, quantified in probabilities. Two of these concern errors we can make in test interpretation:

1. Type I error (alpha, or α): The probability that we declare a difference significant when it really isn’t.

2. Type II error (beta, or β): The probability that we declare a difference not statistically significant when it really is.
The null hypothesis should be rejected

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>The researcher</td>
<td>No</td>
<td>1 - $\alpha$</td>
</tr>
<tr>
<td>rejects the</td>
<td></td>
<td></td>
</tr>
<tr>
<td>null hypothesis</td>
<td>Yes</td>
<td>$\alpha$</td>
</tr>
</tbody>
</table>
What affects the power of a test?

1. Alpha error level ($\alpha$)
   - as $\alpha \uparrow$, $\beta \downarrow$
   - Settling for $\alpha = .05$ rather than .01 or even lower
   - Hypothesizing a direction (one- vs. two-tailed)
What affects the power of a test?

2. The effect size (ES)

- Distance between $H_0$ and $H_A$
- Degree to which the phenomenon is present in the population
- A specific value (or power cannot be determined)
What affects the power of a test?

ES components vary by test. For example:

Independent t-test: \( ES = \frac{(\mu_A - \mu_0)}{\sigma} \)

Difference between proportions: \( ES = \phi_A - \phi_0 \)

Correlation: \( ES = \rho_A \)
When do I use Power Analysis?

**Before data are collected**
- Data collection is expensive (both time and money)
- Good research uses only the subjects it needs

**After consideration of**
- The null and alternative hypotheses
- Desired alpha and beta levels
- One vs. two-tailed test
Where do I find Power Analysis?

Power Analysis appears as the top-most menu selection under Analyze, because it should be done before collecting and analyzing data.

That is also why the POWER procedures produce results without the need for data in the Data Editor, unlike almost every other procedure.
Example: Finding N given ES and Power

How many subjects do I need to test the hypothesis that the means of my control and treatment groups differ?

Power Analysis  >  Means  >  Independent-Samples T Test
How many subjects do I need to test the hypothesis that the means of my control and treatment groups differ?

Example: Finding N given ES and Power

[SPSS Image]
Example: Finding N given ES and Power

How many subjects do I need to test the hypothesis that the means of my control and treatment groups differ?
Example: Finding N given ES and Power

How many subjects do I need to test the hypothesis that the means of my control and treatment groups differ?

Enter the hypothesized means for the groups (or the difference between them)

Enter the (estimated) standard deviation

(These are needed to calculate the effect size)
Example: Finding N given ES and Power

How many subjects do I need to test the hypothesis that the means of my control and treatment groups differ?

Can you hypothesize a direction for alpha error?

Enter the highest alpha value you are willing to accept

Finally, click and select

Test Direction

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

Significance level: 0.05

Two-Dimensional Plot

- Power estimation versus sample size
Given the effect size and alpha entered...

...we need $N \geq 14$ per group for power $\geq .80$
The Power Curve

A graphical representation of the Power Analysis Table
But if we cut the ES in half...

We need \( N \geq 51 \) per grp for power \( \geq .80 \)

### Power Analysis – Independent Sample Means

<table>
<thead>
<tr>
<th>Test for Mean Difference(^a)</th>
<th>N1</th>
<th>N2</th>
<th>Actual Power(^b)</th>
<th>Power</th>
<th>Std. Dev.(^c)</th>
<th>Effect Size</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>.129</td>
<td>.100</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>.155</td>
<td>.150</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>6</td>
<td>.201</td>
<td>.200</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>9</td>
<td>.265</td>
<td>.250</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>11</td>
<td>.304</td>
<td>.300</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>14</td>
<td>.361</td>
<td>.350</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>7</td>
<td>17</td>
<td>17</td>
<td>.414</td>
<td>.400</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>8</td>
<td>20</td>
<td>20</td>
<td>.463</td>
<td>.450</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>9</td>
<td>23</td>
<td>23</td>
<td>.510</td>
<td>.500</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>10</td>
<td>26</td>
<td>26</td>
<td>.553</td>
<td>.550</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>11</td>
<td>30</td>
<td>30</td>
<td>.606</td>
<td>.600</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>12</td>
<td>34</td>
<td>34</td>
<td>.654</td>
<td>.650</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>13</td>
<td>39</td>
<td>39</td>
<td>.707</td>
<td>.700</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>14</td>
<td>44</td>
<td>44</td>
<td>.752</td>
<td>.750</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>15</td>
<td>51</td>
<td>51</td>
<td>.806</td>
<td>.800</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>16</td>
<td>59</td>
<td>59</td>
<td>.854</td>
<td>.850</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
<tr>
<td>17</td>
<td>70</td>
<td>70</td>
<td>.903</td>
<td>.900</td>
<td>10</td>
<td>-.500</td>
<td>.05</td>
</tr>
</tbody>
</table>

- \(^a\) One–sided test.
- \(^b\) Based on noncentral t–distribution.
- \(^c\) Group variances are assumed to be equal.
How Effect Size Affects Power
There are times when a researcher knows (or has a very good guess) how many cases are available; what they do not know is, what power can they expect given that sample size?

Example: Finding Power given ES and N
There are times when a researcher knows (or has a very good guess) how many cases are available; what they do not know is, what power can they expect given that sample size?

Example: Finding Power given ES and N
Given that the researcher hypothesizes $r = .35$ and has $N = 100$, statistical power is very high against a null hypothesis of 0.

<table>
<thead>
<tr>
<th>Power Analysis – Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Power Analysis Table</strong></td>
</tr>
<tr>
<td><strong>Power^b</strong></td>
</tr>
<tr>
<td>Pearson Correlation^a</td>
</tr>
<tr>
<td>.976</td>
</tr>
<tr>
<td>.35</td>
</tr>
</tbody>
</table>

a. One-sided test.

b. Based on Fisher’s z-transformation and normal approximation with bias adjustment.

Now, suppose the researcher wants to test the same alternative hypothesis value against different null hypotheses, say, $r = .10$ or $= .20$?
Example: Finding Power given ES and N

### Power Analysis – Pearson Correlation

**Power Analysis Table**

<table>
<thead>
<tr>
<th>Pearson Correlation&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Power&lt;sup&gt;b&lt;/sup&gt;</th>
<th>N</th>
<th>Test Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.836</td>
<td>100</td>
<td>Null</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Alternative</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sig.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.05</td>
</tr>
</tbody>
</table>

a. One-sided test.

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

### Power Analysis – Pearson Correlation

**Power Analysis Table**

<table>
<thead>
<tr>
<th>Pearson Correlation&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Power&lt;sup&gt;b&lt;/sup&gt;</th>
<th>N</th>
<th>Test Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.486</td>
<td>100</td>
<td>Null</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Alternative</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sig.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.05</td>
</tr>
</tbody>
</table>

a. One-sided test.

b. Based on Fisher's z-transformation and normal approximation with bias adjustment.
Note again how power changes as effect size changes

\[ N = 100; \ r_A = .35 \]
Two questions from a test are being examined. 100 students completed both questions. Previous research has shown that 60% of students on average get the first question correct while 70% get the second one correct. We suspect the difference is significant but want to know the power to correctly reject the null hypothesis of no difference.

First, we need to find the correlation between the responses.

```
DATA LIST FREE /Item1 Item2 (2A4) N (F3).
BEGIN DATA.
Pass Fail  5
Pass Pass 55
Fail Fail 25
Fail Pass 15
END DATA.
WEIGHT BY N.
CROSSTABS /TABLES=Item1 BY Item2 /STATISTICS=PHI.
```

<table>
<thead>
<tr>
<th>Symmetric Measures</th>
<th>Value</th>
<th>Approximate Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal by Nominal Phi</td>
<td>.579</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Cramer's V</td>
<td>.579</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
Example: Finding Power for Correlated Proportions
Example: Finding Power for Correlated Proportions

Power Analysis Table

<table>
<thead>
<tr>
<th>Test for Proportion Difference</th>
<th>Power</th>
<th>Risk Difference</th>
<th>Test Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. One-sided test.</td>
<td>.675</td>
<td>-.100</td>
<td>Risk Ratio: .333</td>
</tr>
<tr>
<td>b. Based on binomial enumeration.</td>
<td>.298</td>
<td>Sig.: .05</td>
<td></td>
</tr>
</tbody>
</table>
Summary and Future Directions

• Consider Power Analysis when in the planning stages, before data collection
• There is power, effect size, and number of cases; once any two of those are fixed, the 3rd is automatically determined
• Experiment with a range of effect sizes and either sample size (if calculating power) or power level (if calculating how many cases are needed)
• Power Analysis in future versions of IBM SPSS Statistics will allow you to:
  1. Determine the sample size necessary to achieve a specified distance from the mean to the confidence limit(s) for a CI level of your choosing;
  2. Directly input the effect size, rather than its component parts
Resources

• Explore other Tech talks: [https://ibm.biz/spsstechtalk](https://ibm.biz/spsstechtalk)
• Blog: What’s new with SPSS Statistics 28
• GradPack (Students / Faculty): Academic discounts
• Take the no cost 30-day trial
• Interactive Product Tour: Start your tour here
• Join the SPSS Statistics Community: Get connected
• Use cases: How industry experts use SPSS Statistics
• Blog: Learning resources for SPSS Statistics
Next steps

Purchase SPSS Statistics
ibm.com/spssstatistics

Try the Free Edition
ibm.biz/SPSSStats
Trademarks and notes

• © IBM Corporation 2021

• IBM, the IBM logo, ibm.com, and SPSS are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (“®” or “™”), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml.

• Other company, product, and service names may be trademarks or service marks of others.

• References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.