G*Power: One-Way Independent Samples ANOVA

See the power analysis done by hand in my document One-Way Independent Samples Analysis of Variance. Here I shall do it with G*Power.

I wish to test the null hypothesis that for GRE-Q, the population means for undergraduates intending to major in social psychology, clinical psychology, and experimental psychology are all equal. I decide that the minimum nontrivial effect size is if each mean differs from the next by 12.2 points (about .12 \( \sigma \)). I expect to have 25 subjects in each group.

How many subjects would be needed to raise power to 80%?

F tests – ANOVA: Fixed effects, omnibus, one–way

Analysis: A priori: Compute required sample size

Input:
- Effect size f = 0.0996126
- \( \alpha \) err prob = 0.05
- Power (1 – \( \beta \) err prob) = 0.80
- Number of groups = 3

Output:
- Noncentrality parameter \( \lambda \) = 9.6746033
- Critical F = 3.0049842
- Numerator df = 2
- Denominator df = 972
- Total sample size = 975
- Actual power = 0.8004415

That was easy, wasn’t it?

If your ANOVA is significant, you will likely want to make pairwise comparisons. Since there are only three groups, Fisher’s procedure would be appropriate. Assuming equal sample sizes, each
group will have $975/3 = 325$ subjects. Although I would get more power using a pooled error term, I shall assume here that I’ll use the more conservative procedure, employing individual error terms. Each comparison will involve 325 subjects in each of the two groups being compared. The difference between means will be 12.2 for two comparisons and 24.4 for one comparison. How much power would I have for the comparisons with the smaller differences between means?

Only 34% power for that comparison. How many subjects would I need to have 80% power?

**t tests** Means: Difference between two independent means (two groups)

**Analysis:** A priori: Compute required sample size

**Input:**
- Tail(s) = Two
- Effect size $d$ = 0.1220000
- $\alpha$ err prob = 0.05
- Power (1-$\beta$ err prob) = 0.80
- Allocation ratio $N_2/N_1$ = 1

**Output:**
- Noncentrality parameter $\delta$ = 2.8033466
- Critical $t$ = 1.9610889
- Df = 2110
- Sample size group 1 = 1056
- Sample size group 2 = 1056
- Total sample size = 2112
- Actual power = 0.8001368

We need 1,056 subjects in each of three groups, that is, a total of 3,168 subjects.

Usually one will not compute the effect size $f$ from group means and standard deviations but rather rely on Cohen’s benchmarks for $f$ (.1 = small, .25 = medium, .40 = large). Suppose that I have 20 subjects in each of four groups and I want to know what my chances are of obtaining significant results if the population effect size is small.
Oh my, that is not a very happy result – only 10% power – if there really is a small effect, then we most likely (90%) will be making a Type II error – the chances of making a Type II error would be nine times greater than the chances of making a correct decision. We could hope that the effect is at least medium. Power would then be greater.

Well, that is better, but we are still more likely to make a Type II error than a correct decision. How many subjects would we need to have 80% power (the usual but not, IMHO, well-advised standard in the USA).
We need 45 subjects per group. Ideally, we would have enough data to have 95% power to reject the null hypothesis even if the population effect size was the smallest value that would be of practical importance ($f = .10$). In that case, $\beta = \alpha$, .05, and if were unable to reject the null, we would be in a good position to argue that the population effect is so small that it might as well be zero.

Holy moly, we would need a total of 1,724 subjects to get the desired amount of power, and that does not take into consideration the power for pairwise comparisons. Suppose the ANOVA is significant and I am going to make pairwise comparisons. That will be six comparisons. To cap familywise alpha at .05, I lower the per-comparison alpha to $.05/6 = .0083$. To have 95% power to detect a small difference between two means I need 920 subjects in each group, for a total of 4(920) = 3,680 subjects.
t tests – Means: Difference between two independent means (two groups)

Analysis: A priori: Compute required sample size

Input:
- Tail(s) = Two
- Effect size d = .20
- α err prob = 0.0083
- Power (1 - β err prob) = 0.95
- Allocation ratio N2/N1 = 1

Output:
- Noncentrality parameter δ = 4.2895221
- Critical t = 2.6424796
- Df = 1838
- Sample size group 1 = 920
- Sample size group 2 = 920
- Total sample size = 1840
- Actual power = 0.9501016

The Delta ANOVA Effect Size Statistic

Cohen actually defined ANOVA effect size in two different ways. The f effect size statistic, used by G*Power, is the standardized average dispersion among the group means. Cohen also proposed the delta (δ) ANOVA effect size statistic, which is the difference between the largest and smallest population means divided by the within-population standard deviation. Cohen’s benchmarks for delta are .25 = small, .75 = medium, and 1.25 = large. Michael Friendly has an online app that will do power analysis using this effect size statistic. Here I illustrate using it for determining what power would be with three groups, small/medium/large effect sizes, and various sample sizes.

Power analysis for ANOVA designs

3 x 1 layout Ha: T1=GM-Delta/2, T2=T3=...=Tk-1=GM, Tk=GM+Delta/2 tested at Alpha= 0.050

<table>
<thead>
<tr>
<th>DELTA (in units of sigma=Std. Dev.)</th>
<th>N</th>
<th>0.250</th>
<th>0.750</th>
<th>1.250</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>.052</td>
<td>.068</td>
<td>.101</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.054</td>
<td>.091</td>
<td>.173</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.056</td>
<td>.116</td>
<td>.249</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>.059</td>
<td>.141</td>
<td>.325</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>.061</td>
<td>.167</td>
<td>.400</td>
</tr>
</tbody>
</table>
$N$ is the number of scores in each group. With 25 cases in each of 3 groups, power is 11% for a small effect, 64% for a medium effect, and 98% for a large effect.

Links
- [Cohen's d, Cohen's $f$, and $\eta^2$](#)
- [Karl Wuensch's Statistics Lessons](#)
- [Internet Resources for Power Analysis](#)
- [UCLA Power Lesson](#)

Karl L. Wuensch  
Dept. of Psychology  
East Carolina University  
Greenville, NC  USA  

March, 2017