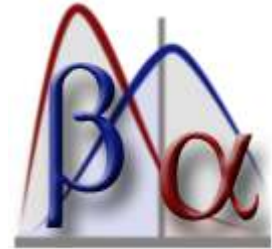G*Power 3 is a wonderful free resource for power analysis. This program provides power analyses for tests that use F, t, chi-square, or z distributions plus various 'exact' distributions for nonparametric applications.

Power analysis involves four interdependent concepts that can be remembered with the mnemonic BEAN:

> **B** = beta error, where power = (1 – Beta error)
> **E** = effect size
> **A** = alpha error rate
> **N** = sample size

G*Power can compute any one of these four values given the other three.

Download the free program:

G*Power 3 was written by Franz Faul of the University of Kiel, Germany. Download from
http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/download-and-register

Options:

The program is loaded with options for power analyses with F, t, z, chi-square, and exact distributions. Within each there are choices for the statistical test. For example, within F are tests for ANOVA, MANOVA, Hotellings T, and Multiple Regression.

Example of power analysis for a One-Way ANOVA:

> **Suppose we wish to compare reading readiness scores for three populations of children. Previous research shows the standard deviation of our test within groups to be about 10. Based on our knowledge of this test and the populations under study, we consider a difference of 5 points between two population means to be the smallest effect size of practical importance. We would like to estimate the number of cases we need to sample to attain power = .80 with alpha set at .05.**

Thus, we wish to compute N, and we must provide information on B, E and A. We specify B by choosing power = .80 and we specify A by choosing alpha = .05. E (effect size) is trickier.

A measure of effect size is Cohen's f = $\phi' = \dfrac{\sigma_T}{\sigma_E} = \sqrt{\dfrac{\sum \alpha_j^2}{k\sigma_\varepsilon^2}}$. Note: α_j is not the critical value α

that we use to evaluate our calculated *p*-value but rather it is the difference between the population mean for group *j* and the grand mean.

Alternate notation for the same computation: f = $\phi' = \sqrt{\dfrac{\sum(\mu_j - \mu)^2 / k}{\sigma_\varepsilon^2}}$.

Application of G*Power to select sample size for One-Way ANOVA

1. Open G*Power.
2. Under Test family, select **F tests**.
3. Under Statistical test, select **ANOVA: Fixed effects, omnibus, one-way**.
4. Under Type of power analysis, select **A priori: Compute required sample size**
5. Set **α error prob** to **0.05, Power** to **0.80,** and **Number of groups** to **3.**
6. Under Input Parameters, click **Determine =>** to open a computation window.
7. Select procedure as **Effect size from means**.
8. Set Number of groups = **3** and SD σ within each group as **10**.
9. To specify a maximum difference between means of 5 points, we can enter **100** for the Mean of Group 1 and **105** for the Mean of Group 3. The Mean for Group 2 could then be any value between 100 and 105, but the most conservative choice is at the mean of the two extreme groups, **102.5**. This will result in a larger value for N than any other choice. The size of the groups can be modified if we wish to insist on a desired ratio between sample sizes. However, most often we desire equal groups, so we can keep the default values of 5.
10. Press **Calculate and transfer to main window** and we see f = .2041241 is shown in both windows.
11. Click **Calculate** in the main window.
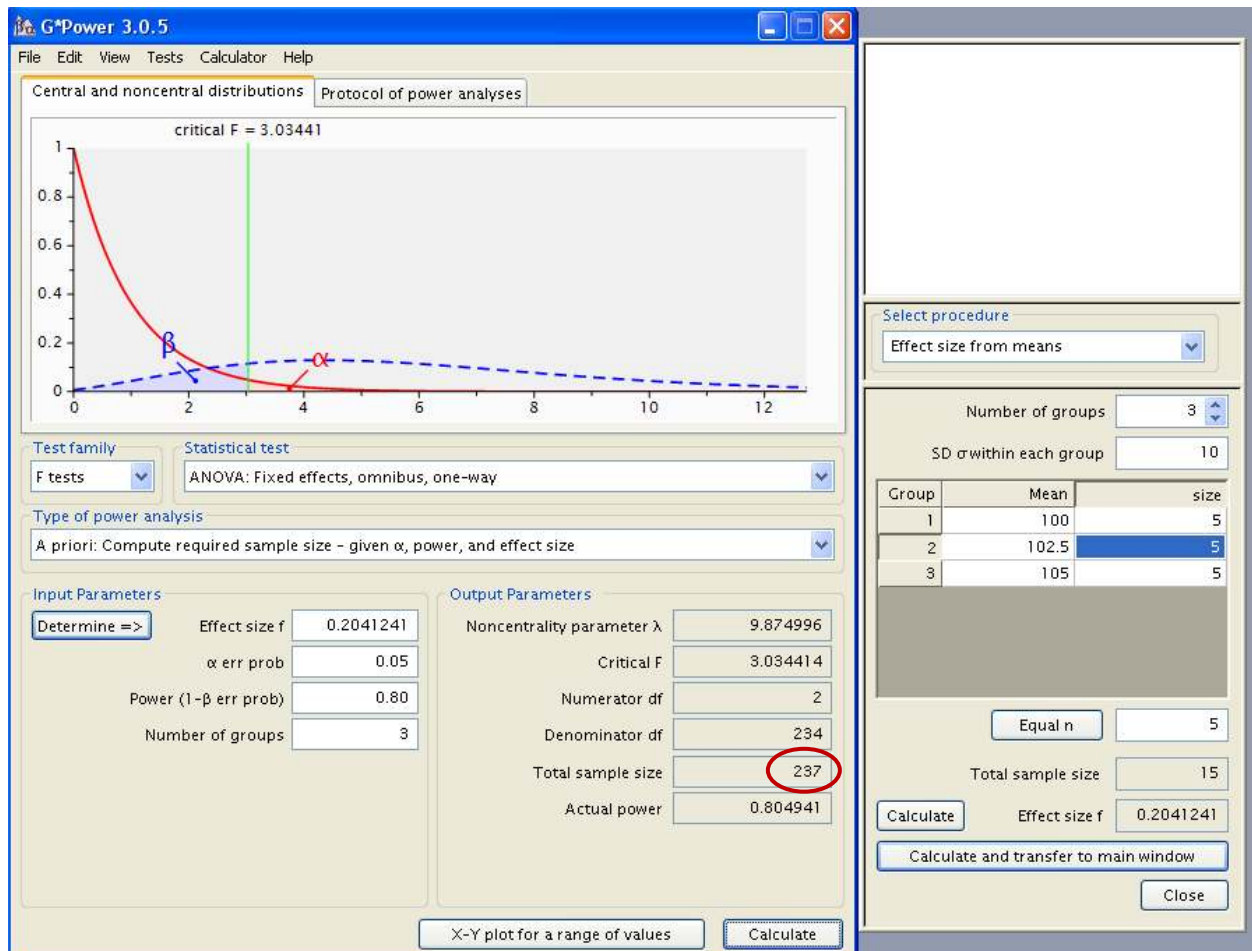
12. The Total sample size is computed to be 237, or about 77 per group.

Description of the G*Power graph.

The solid red descending curve on the left in the graph shows the F distribution that we would be sampling from if the null hypothesis (no difference in population means) were true, with the critical value cutting off an area corresponding to alpha on the right tail. The dashed curve is the non-central F distribution, corresponding to the F distribution that we are actually sampling from if the effect size is as specified by f and the sample is large enough to provide the desired power level. The area under the dashed curve to the right of the critical value corresponds to power.

Computation of effect size.

Effect size = f = $\phi' = \sqrt{\dfrac{\sum (\mu_j - \mu)^2 / k}{\sigma_\varepsilon^2}}$. In our example, we determined that a difference of 5 points between the largest and smallest means is the smallest difference that is meaningful. Thus, we set the population means at 100, 102.5, and 105. The grand mean is 102.5, so the deviations from the grand mean are -2.5, 0, and 2.5. The sum of squares of these deviations is 6.25 + 0 + 6.25 = 12.5. The number of groups (k) = 3 and the within groups variance is 10^2 = 100. This gives f = square root of (12.5/300) = .2041241, as shown by G*Power.

Specifying an effect size

A common error is to set the effect size at a value that is expected in the population rather than the minimum value that is considered meaningful. Failure to attain statistical significance does not necessarily mean that the study was under-powered. The sample size may have been large enough to have adequate power to detect a meaningful effect if it existed.

An estimate of power based on an observed study is called 'achieved power' or 'empirical power.' This estimate of power generally is not very useful because it is redundant with the test of statistical significance – if we attained significance, achieved power is high, but if we failed to attain statistical significance, the achieved power is low. However, the important question is whether we had adequate power to detect a meaningful effect. It may be useful to compute the minimum effect size for which we did have adequate power given our sample size, etc.

Note on 'Blob' tests

When we have three or more groups in ANOVA, the overall F-test is a 'blob' test that doesn't indicate the location of an observed significant difference. An overall F-test can be used as a screening device to help control possible inflation of alpha error, but it is not very useful for interpreting findings in detail. In general, it is better practice to focus on comparisons with df=1 in the numerator of the F-test. Tests of contrasts with df=1 in the numerator of the F test, such as comparisons of two groups at a time, allow the researcher to describe observed effects more clearly.

Additional plots and tables

Press X-Y plot for a range of values to generate graphs showing sample size as a function of power for various effect sizes. The user has control over the range of values to be evaluated, size of steps, etc. The same information can be shown in a table.

Example of power analysis for multiple regression with G*Power:

> **Suppose we plan a hierarchical regression analysis with three control variables followed by two treatment variables. We wish to have a 90% chance of attaining statistical significance with alpha = .05 if the population value of R^2 change is large enough to add .10 beyond an anticipated R^2 of .30 produced by the first three control variables. We compute this population effect size to be f^2 = (R^2 change) / (1 – cumulative R^2) = .1 / (1- (.3+.1)) = .1/.6 = .166667. This is close to the value of .15 that Cohen (1988) called a 'medium' effect size.**

Open G*Power.
Under Test family, select **F tests**.
Under Statistical test, select **Linear Multiple Regression: Fixed model, R^2 increase**.
Under Type of power analysis, select **A priori: Compute required sample size**
Under Input Parameters, enter the chosen values:
f^2 = .166667; α = .05; Power = .90; Numerator df = 2; Number of predictors = 5

Click **Calculate**. The program gives the Total sample size = 80. Power will be greater than 90% if the population effect size is larger than .166667.

The graph shows the null distribution that F is sampled from if the null hypothesis is true (solid curve on left) and the alternate distribution that F is sampled from if the effect size is .166667 and N=80 with alpha = .05 (low dashed-line distribution). If you are not sure how to compute the effect size, the program can help determine that value. The output includes the noncentrality parameter (a measure of effect size), critical F, df for the error term (df=74), Total sample size (N=80), and the actual power (.903).

Pressing **X-Y plot** will generate a plot showing sample size on Y as a function of power on X, with separate lines for alternative effect sizes.