

What is the Power Analysis for a Two-Group Independent Sample t-test in SAS Data Analysis Examples?

Authored by
stats writer

June 29, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the Power Analysis for a Two-Group Independent Sample t-test in SAS Data Analysis Examples?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=159230>

Power analysis is a statistical method used to determine the minimum sample size needed for a study to have a high probability of detecting a true effect. In the context of a two-group independent sample t-test in SAS data analysis, power analysis is used to calculate the necessary sample size to achieve a desired level of statistical power. This involves considering factors such as the effect size, significance level, and desired power level. By conducting a power analysis, researchers can ensure that their study has enough participants to detect a significant difference between the two groups being compared. This allows for more accurate and reliable results in statistical analysis using SAS software.

Power Analysis for Two-group Independent sample t-test | SAS Data Analysis Examples

Examples

Example 1. A clinical dietician wants to compare two different diets, A and B, for diabetic patients. She hypothesizes that diet A (Group 1) will be better than diet B (Group 2), in terms of lower blood glucose. She plans to get a random sample of diabetic patients and randomly assign them to one of the two diets. At the end of the experiment, which lasts 6 weeks, a fasting blood glucose test will be conducted on each patient. She also expects that the average difference in blood glucose measure between the two group will be about 10 mg/dl. Furthermore,

she also assumes the standard deviation of blood glucose distribution for diet A to be 15 and the standard deviation for diet B to be 17. The dietician wants to know the number of subjects needed in each group assuming equal sized groups.

Example 2. An audiologist wanted to study the effect of gender on the response time to a certain sound frequency. He suspected that men were better at detecting this type of sound than were women. He took a random sample of 20 male and 20 female subjects for this experiment. Each subject was given a button to press when he/she heard the sound. The audiologist then measured the response time - the time between the sound was emitted and the time the button was pressed. Now, he wants to know what the statistical power is based on his total of 40 subjects to detect the gender difference.

Prelude to The Power Analysis

There are two different aspects of power analysis. One is to calculate the necessary sample size for a specified power as in Example 1. The other aspect is to calculate the power when given a specific sample size as in Example 2. Technically, power is the probability of rejecting the null hypothesis when the specific alternative hypothesis is true.

For the power analyses below, we are going to focus on Example 1, calculating the sample size for a given statistical power of testing the difference in the effect of diet A and diet B. Notice the assumptions that the dietician has made in order to perform the power analysis. Here is the information we have to know or have to assume in order to perform the power analysis:

Notice that in the first example, the dietician didn't specify the mean for each group, instead she only specified the difference of the two means. This is

because that she is only interested in the difference, and it does not matter what the means are as long as the difference is the same.

Power Analysis

In SAS, it is fairly straightforward to perform power analysis for comparing means. For example, we can use proc power of SAS for our calculation as shown below. We first specify the two means, the mean for Group 1 (diet A) and the mean for Group 2 (diet B). Since what really matters is the difference, instead of means for each group, we can enter a mean of zero for Group 1 and 10 for the mean of Group 2, so that the difference in means will be 10. Next, we need to specify the pooled standard deviation, which is the square root of the average of the two standard deviations squared (i.e., variances). In this case, it is $\sqrt{((15^2 + 17^2)/2)} = 16.03$. The default significance level (alpha level) is .05.

For this example we will set the power to be at .8.

```
proc power;  
twosamplemeans test=diff  
groupmeans = 0 | 10  
stddev = 16.03  
npergroup = .  
power = 0.8;  
run;
```

Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal

Method Exact

Group 1 Mean 0

Group 2 Mean 10

Standard Deviation 16.03

Nominal Power 0.8

Number of Sides 2

Null Difference 0

Alpha 0.05

Computed N Per Group

Actual N Per Power Group

0.807 42

The calculation results indicate that we need 42 subjects for diet A and another 42 subject for diet B in our sample in order the effect. Now, let's use another pair of means with the same difference. As we have discussed earlier, the results should be the same, and they are.

```
proc power;  
twosamplemeans test=diff  
groupmeans = 5 | 15  
stddev = 16.03  
npergroup = .  
power = 0.8;  
run;
```

Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal

Method Exact

Group 1 Mean 5

Group 2 Mean 15

Standard Deviation 16.03

Nominal Power 0.8

Number of Sides 2

Null Difference 0

Alpha 0.05

Computed N Per Group

**Actual N Per
Power Group**

0.807 42

Now the dietician may feel that a total sample size of 84 subjects is beyond her budget. One way of reducing the sample size is to increase the Type I error rate, or the alpha level. Let's say instead of using alpha level of .05 we will use .07. Then our sample size will reduce by 4 for each group as shown below.

```
proc power;  
twosamplemeans test=diff  
groupmeans = 0 | 10  
stddev = 16.03  
npergroup = .  
alpha = .07  
power = 0.8;  
run;
```

Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal

Method Exact

Alpha 0.07

Group 1 Mean 0

Group 2 Mean 10

Standard Deviation 16.03

Nominal Power 0.8

Number of Sides 2

Null Difference 0

Computed N Per Group

Actual N Per Power Group

0.810 38

Now suppose the dietician can only collect data on 60 subjects with 30 in each group. What will the statistical power for her t-test be with respect to alpha level of .05?

```
proc power;  
twosamplemeans test=diff  
groupmeans = 0 | 10  
stddev = 16.03  
npergroup = 30  
power = .;  
run;
```

Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal

Method Exact

Group 1 Mean 0

Group 2 Mean 10

Standard Deviation 16.03

Sample Size Per Group 30

Number of Sides 2

Null Difference 0

Alpha 0.05

Computed Power

Power

0.661

What if she actually collected her data on 60 subjects but with 40 on diet A and 20 on diet B instead of equal sample sizes in the groups?

```
proc power;
```

```
twosamplemeans test=diff
```

```
groupmeans = 0 | 10
```

```
stddev = 16.03
```

```
groupns = (40 20)
```

```
power = .;
```

run;

Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal

Method Exact

Group 1 Mean 0

Group 2 Mean 10

Standard Deviation 16.03

Group 1 Sample Size 40

Group 2 Sample Size 20

Number of Sides 2

Null Difference 0

Alpha 0.05

Computed Power

Power

0.610

As you can see the power goes down from .661 to .610 even though the total number of subjects is the same. This is why we always say that a balanced design is more efficient.

As we have discussed before, what really matters in the calculation of power or sample size is the difference of the means over the pooled standard deviation. This is a measure of effect size. Let's now look at how the effect size affects the sample size assuming a given sample power. We can simply assume the difference in means and set the standard deviation to be 1.

```
proc power;  
twosamplemeans test=diff  
meandiff = .2 to 1.2 by .2  
stddev = 1  
power = .8  
npergroup = . ;  
run;
```

Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal

Method Exact

Standard Deviation 1

Nominal Power 0.8

Number of Sides 2

Null Difference 0

Alpha 0.05

Computed N Per Group

Mean Actual N Per

Index Diff Power Group

1 0.2 0.801 394

2 0.4 0.804 100

3 0.6 0.804 45

4 0.8 0.807 26

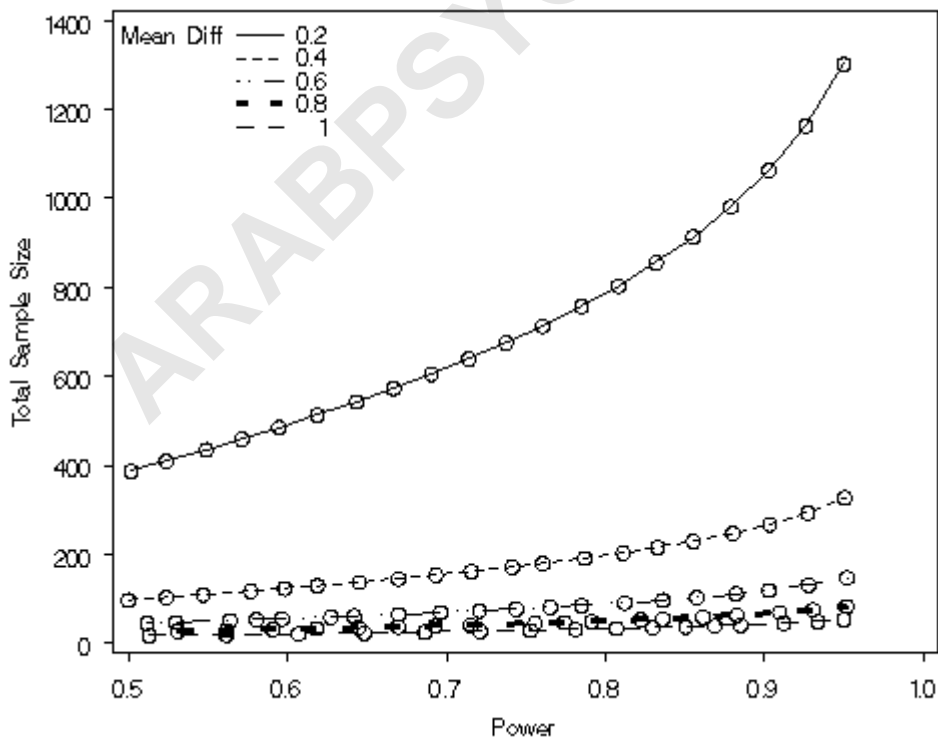
5 1.0 0.807 17

6 1.2 0.802 12

We see that as the effect size goes down the sample

size goes up very quickly. We can display this information in a plot.

```
proc power plotonly;  
twosamplemeans test=diff  
meandiff = .2 to 1 by .2  
stddev = 1  
power = .9  
ntotal = .;  
plot x = power min= .5 max=.95  
key=byfeature(pos=inset);  
run;
```



It shows that if the effect size is tiny, such .2 then we need fairly large sample size. For example, it looks like that we need a sample size of about 700 just for a reasonable power such .75 to detect an effect size of .2.

Discussion

An important technical assumption is the normality assumption. If the distribution is skewed, then a small sample size may not have the power shown in the results, because the value in the results is calculated using the method based on the normality assumption. We have seen that in order to compute the power or the sample size, we have to make a number of assumptions. These assumptions are used not only for the purpose of calculation, but are also used in the actual t-test itself. So one important side benefit of performing power analysis is to help us to better understand our designs

and our hypotheses.

We have seen in the power calculation process that what matters in the two-independent sample t-test is the difference in the means and the standard deviations for the two groups. This leads to the concept of effect size. In this case, the effect size will be the difference in means over the pooled standard deviation. The larger the effect size, the larger the power for a given sample size. Or, the larger the effect size, the smaller sample size needed to achieve the same power. So, a good estimate of effect size is the key to a good power analysis. But it is not always an easy task to determine the effect size. Good estimates of effect size come from the existing literature or from pilot studies.

See Also

D. Moore and G. McCabe, Introduction to the Practice

of Statistics, Third Edition, Section 6.4

ARABPSYCHOLOGY.COM