What is the process for conducting a Power Analysis for a Paired Sample t-test in SAS for data analysis?

Authored by
stats writer

June 29, 2024

The process for conducting a Power Analysis for a Paired Sample t-test in SAS involves several steps. First, the researcher must identify the research question and the specific hypothesis being tested. Then, the sample size and desired level of statistical power must be determined. Next, the researcher must collect the necessary data and input it into the SAS software. The software will then calculate the effect size, which is a measure of the difference between the two paired groups. Using this information, the researcher can determine the power of the test and whether the sample size is sufficient to detect the expected effect. If the power is low, the researcher may need to adjust the sample size or reconsider the research question. Finally, the results of the Power Analysis can be used to inform the interpretation of the t-test results and the overall conclusions of the study. Overall, the Power Analysis for a Paired Sample t-test in SAS is a crucial step in ensuring the validity and reliability of the data analysis.

Power Analysis for Paired Sample t-test | SAS Data Analysis Examples

Examples

Example 1. A company markets an eight week long weight loss program and claims that at the end of the program on average a participant will have lost 5 pounds. On the other hand, you have studied the program and you believe that their program is scientifically unsound and shouldn't work at all. With some limited funding at hand, you want test the hypothesis that the weight loss program does not help people lose weight. Your plan is to get a random sample of

people and put them on the program. You will measure their weight at the beginning of the program and then measure their weight again at the end of the program. Based on some previous research, you believe that the standard deviation of the weight before and after the program remains the same and will be 5 pounds. You now want to know how many people you should enroll in the program to test your hypothesis.

Example 2. A human factors researcher wants to study the difference between dominant hand and the non-dominant hand in terms of manual dexterity. She designs an experiment where each subject would place 10 small beads on the table in a bowl, once with the dominant hand and once with the non-dominant hand. She measured the number seconds needed in each round to complete the task. She has also decided that the order in which the two hands are measured should be counter balanced. She

expects that the average difference in time would be 5 seconds with the dominant hand being more efficient and the standard deviation for each hand to be fairly equal and equal to 10. She collects her data on a sample of 35 subjects. The question is, what is the statistical power of her design with an N of 35 to detect the difference in the magnitude of 5 seconds.

Prelude to the Power Analysis

In both of the examples, there are two measures on each subject, and we are interested in the mean of the difference of the two measures. This can be done with a t-test for paired samples (dependent samples). In a power analysis, there are always a pair of hypotheses: a specific null hypothesis and a specific alternative hypothesis. For instance, in Example 1, the null hypothesis is that the mean weight loss is 5 pounds and the alternative is zero pounds. In Example 2, the null hypothesis is that mean difference is zero seconds and the

alternative hypothesis is that the mean difference is 5 seconds.

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

Both of these calculations depend on the Type I error rate, the significance level. The significance level (called alpha), or the Type I error rate, is the probability of rejecting H0 when it is actually true. The smaller the Type I error rate, the larger the sample size required for the same power. Likewise, the smaller the Type I error rate, the smaller the power for the same sample size. This is the trade-off between the reliability and sensitivity of

the test.

Power Analysis

In SAS, it is fairly straightforward to perform a power analysis for
the paired sample t-test using proc power.

For the calculation of Example 1, we can set the power at different levels
and calculate the sample size for each level. We will specify the difference in
means, which is 5-0 = 5, and the standard deviation for either before or after the program which in this example are assumed to be equal to 5. One
thing that SAS requires is the correlation between the two measures, pre and
post. In this example, we don't know the magnitude of the correlation of the pre
and post measures, we will set it to be .5, a medium strength of correlation.
This way, the standard deviation can be considered to be the pooled standard
deviation from the standard deviation of the two measures. We set the power

level from .6 to .9 and look for sample size for each level of power.

```
proc power;
pairedmeans test=diff
meandiff = 5
std = 5
corr = .5
npairs = .
power = 0.6 to .9 by .1;
run;
```

Paired t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal
Method Exact
Mean Difference 5
Standard Deviation 5
Correlation 0.5
Number of Sides 2
Null Difference 0
Alpha 0.05

Computed N Pairs

Nominal Actual N
Index Power Power Pairs

1 0.6 0.600 7
2 0.7 0.748 9
3 0.8 0.803 10
4 0.9 0.911 13

Next, let's change the level of significance to .01 with a power of .911. What does this mean for our sample size calculation?

```
proc power;
pairedmeans test=diff
meandiff = 5
std = 5
corr = .5
npairs = .
alpha =.01
power = 0.911;
run;
```

Paired t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal
Method Exact
Alpha 0.01
Mean Difference 5
Standard Deviation 5
Correlation 0.5
Nominal Power 0.911
Number of Sides 2
Null Difference 0

Computed N Pairs

Actual N
Power Pairs

0.915 19

As you can see, the sample size goes up from 13 to 19 for specified power of .911
when alpha drops from .05 to .01.

This
means if we want our test to be more reliable, i.e., not

rejecting the null hypothesis in case it is true, we will need a larger sample size. Remember all the calculation is under the normality assumption. If the distribution is not normal, then 19 subjects are, in general, not enough for this t-test.

Now, let's now turn our calculation around the other way. Let's look at Example 2. In this example, our researcher has already collected data on 35 subjects. How much statistical power does her design have to detect the difference of 5 seconds with standard deviation of each hand equal to 10 seconds?

Again we use the proc power to calculate the power.

```
proc power;
pairedmeans test=diff
meandiff = 5
std = 10
corr = .5
```

npairs = 35

power = .;

run;

Paired t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal

Method Exact

Mean Difference 5

Standard Deviation 10

Correlation 0.5

Number of Pairs 35

Number of Sides 2

Null Difference 0

Alpha 0.05

Computed Power

Power

0.820

This means that the researcher would detect the

difference of 5 seconds about 82 percent of the time. Notice we did this as two-sided test. But since it is believed that our dominant hand is always better than the non-dominant hand, the researcher actually could conduct a one-tailed test. Now, let's recalculate the power for one-tailed paired-sample t-test.

```
proc power;
pairedmeans test=diff
meandiff = 5
std = 10
corr = .5
npairs = 35
sides = 1
power = .;
run;
```

Paired t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal
Method Exact

Number of Sides 1

Mean Difference 5

Standard Deviation 10

Correlation 0.5

Number of Pairs 35

Null Difference 0

Alpha 0.05

Computed Power

Power

0.895

Recall that we set the correlation between the two measures at .5 for all the calculations we have done. Let's take a look at how the strength of correlation affects the sample size.

```
proc power;
pairedmeans test=diff
meandiff = 5
std = 10
corr = -.9 to .9 by .1
```

npairs = .

power = .8;

run;

Paired t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal

Method Exact

Mean Difference 5

Standard Deviation 10

Nominal Power 0.8

Number of Sides 2

Null Difference 0

Alpha 0.05

Computed N Pairs

Actual N

Index Corr Power Pairs

1 -0.9 0.802 122

2 -0.8 0.800 115

3 -0.7 0.801 109

4 -0.6 0.802 103

5 -0.5 0.804 97

6 -0.4 0.801 90

7 -0.3 0.802 84

8 -0.2 0.804 78

9 -0.1 0.806 72

10 -0.0 0.802 65

11 0.1 0.804 59

12 0.2 0.806 53

13 0.3 0.801 46

14 0.4 0.804 40

15 0.5 0.808 34

16 0.6 0.814 28

17 0.7 0.803 21

18 0.8 0.812 15

19 0.9 0.835 9

**We can see clearly that the more positively correlated the two measures are,
the smaller the sample size needs to be.**

**Also, we don't have to assume that the standard deviation is the same for both groups. If we
know the standard deviation for each measure and the**

correlation between the two measures, we can supply that information to proc power.
For instance, the standard deviation for the measure of weight before the program might be smaller than the standard deviation for the measure of weight after the weight-loss program. Let's say, the standard deviation for the first measure (before the program) is 7 and the standard deviation for the second measure (after the program) is 12 and the correlation between the two is .5. We can calculate the sample size in this setting as well.

```
proc power;
pairedmeans test=diff
meandiff = 5
pairedstddevs = (7 12)
corr = .5
npairs = .
power = .6 to .9 by .05;
run;
```

Paired t Test for Mean Difference

Fixed Scenario Elements

Distribution Normal
Method Exact
Mean Difference 5
Standard Deviation 1 7
Standard Deviation 2 12
Correlation 0.5
Number of Sides 2
Null Difference 0
Alpha 0.05

Computed N Pairs

Nominal Actual N
Index Power Power Pairs

1 0.60 0.613 24
2 0.65 0.651 26
3 0.70 0.702 29
4 0.75 0.760 33
5 0.80 0.809 37
6 0.85 0.858 42

7 0.90 0.901 48

Discussion

The other 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. It might not even be a good idea to do a t-test on a small sample to begin with.

What we really need to know is the difference between the two means, not the individual values. In fact, what really matters, is the difference between the means over the pooled standard deviation. We call this the effect size (Cohen's d). It is usually not an easy task to determine the effect size. It usually comes from studying the existing literature or from pilot studies. A good estimate of the effect size is the key to a successful power analysis.

See Also