Chapter 432

Tests for Two Groups of Pre-Post Scores

Introduction

This module calculates the power for testing the interaction in a 2-by-2 repeated measures design. This particular repeated measures design is one in which subjects are observed twice over time, as is the case in a pre, post design. Measurements are taken at two, pre-determined time intervals. It is important that the time interval remains constant from subject to subject.

In this case, the test of the interaction compares the average change in measurement for group 1 with that of group 2. For example, if the time points surround the application of a treatment, then the interaction test determines if the change from one time period to the other is the same for both groups.

It turns out that the data may be analyzed using a two-sample t-test on the paired differences. If assumptions about the other features of the two groups are met (such as that the paired differences are normally distributed and their variances are equal), the two-sample *t* test can be used to compare the means of random samples drawn from these two populations.

The formulas used in this chapter are partially based on Rosner (2011).

Technical Details

Suppose a repeated measures design has N_1 subjects in group 1 and N_2 subjects in group 2. Each subject is measured once and then again after a specified time has elapsed. If the difference between the two responses for each subject is of primary interest, the design can be collapsed to a two-sample, parallel-group design. The usual two-sample t-test may be used to compare the average difference of each group.

To compute an estimate of power or sample size, the variance of the differences is needed. This could be found using previous studies or pilot data. However, it may be easier to compute the needed variance using the following formula based on more readily available quantities

$$\sigma_{Diff}^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 \,.$$

Here σ_1^2 is the variance of the first measurement, σ_2^2 is the variance of the second measurement, and ρ is the correlation between the two measurements.

Once you have decided to use the differences, you may base all power and sample size estimates on the usual two-sample t-test (or z-test approximation for larger sample sizes).

Data Model for Measurements X(group, time)

Time 1 Measurements	**Time 2 Measurements**
Group 1: X11 ~ N[μ11, σ(T1)]	Group 1: X12 ~ N[μ12, σ(T2)]
Group 2: X21 ~ N[μ21, σ(T1)]	Group 2: X22 ~ N[μ22, σ(T2)]

where

$X \sim N[\mu, \sigma]$ means X is distributed normally with mean μ and standard deviation σ.

Difference in Mean Changes, δ

Test statistical hypothesis about

$$\delta = \delta2 - \delta1$$

where

$$\delta1 = \mu12 - \mu11 \text{ (change in group 1)}$$
$$\delta2 = \mu22 - \mu21 \text{ (change in group 2)}$$

Case 1 – Standard Deviations Known and Equal (Z-Test)

When $\sigma_1 = \sigma_2 = \sigma$ and are known, the power of the *t* test is calculated as follows for a directional alternative (one-tailed test) in which $\delta > 0$. Since the standard deviations are rarely known, this case usually used when quick results based on the normal approximation to the t-test is desired. This is the case presented in Rosner (2011).

1. Find z_α such that $1 - \Phi(z_\alpha) = \alpha$, where $\Phi(x)$ is the area under the standardized normal curve to the left of *x*.

2. Calculate $\sigma_{\bar{X}} = \sigma_{Diff}\sqrt{\dfrac{1}{N_1} + \dfrac{1}{N_2}}$

3. Calculate $z_p = \dfrac{z_\alpha \sigma_{\bar{X}} - \delta}{\sigma_{\bar{X}}}$

4. Calculate $Power = 1 - \Phi(z_p)$.

Case 2 – Standard Deviations Unknown and Equal (T-Test)

When $\sigma_1 = \sigma_2 = \sigma$ and are unknown, the power of the *t* test is calculated as follows for a directional alternative (one-tailed test) in which $\delta > 0$.

1. Find t_α such that $1 - T_{df}(t_\alpha) = \alpha$, where $T_{df}(t_\alpha)$ is the area under a central-*t* curve to the left of *x* and $df = N_1 + N_2 - 2$.

2. Calculate $\sigma_{\bar{X}} = \sigma_{Diff}\sqrt{\dfrac{1}{N_1} + \dfrac{1}{N_2}}$.

3. Calculate the noncentrality parameter $\lambda = \dfrac{\delta}{\sigma_{\overline{X}}}$.

4. Calculate $t_p = \dfrac{t_\alpha \sigma_{\overline{X}} - \delta}{\sigma_{\overline{X}}} + \lambda$.

5. Calculate: Power $= 1 - T'_{df,\lambda}(t_p)$, where $T'_{df,\lambda}(x)$ is the area to the left of x under a noncentral-t curve with degrees of freedom df and noncentrality parameter λ.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. In most situations, you will select either *Power* or *Sample Size (N1)*.

Select *Sample Size (N1)* when you want to calculate the sample size needed to achieve a given power and alpha level.

Select *Power* when you want to calculate the power of an experiment.

Test

Test Type

Specify the type of test that will be used to analyze the data from this experiment. Since the "T-Test" is more accurate, it should be used in most situations.

You might select "Z-Test" to match an example in an elementary textbook.

Alternative Hypothesis

This option specifies the type of test by specifying the direction of the alternative. Since the null hypothesis is the opposite of the alternative hypothesis, specifying the alternative is all that is needed. Usually, the two-tailed ($\neq$) option is selected.

The options containing only $<$ or $>$ are one-tailed tests. When you choose one of these, you must be sure that the values of the parameters match the condition.

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal means when in fact the means are different.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

Alpha is the probability of obtaining a FALSE POSITIVE on the statistical test. That is, it is the probability of rejecting a true null hypothesis. The null hypothesis is usually that the parameters of interest (means, proportions, etc.) are all equal. Since Alpha is a probability, it is bounded by 0 and 1. Commonly, it is between 0.001 and 0.250. Alpha is often set to 0.05 for two-sided tests and to 0.025 for one-sided tests (such as non-inferiority tests).

You can enter a list of values such as '.05 .10 .15' or '.05 to .15 by .01.'

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on *N1* or *N2* or both.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter N2, solve for N1**

  Select this option when you wish to fix *N2* at some value (or values), and then solve only for *N1*. Please note that for some values of *N2*, there may not be a value of *N1* that is large enough to obtain the desired power.

- **Enter R = N2/N1, solve for N1 and N2**

  For this choice, you set a value for the ratio of *N2* to *N1*, and then PASS determines the needed *N1* and *N2*, with this ratio, to obtain the desired power. An equivalent representation of the ratio, *R*, is

  $$N2 = R * N1.$$

- **Enter percentage in Group 1, solve for N1 and N2**

  For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed *N1* and *N2* with this percentage to obtain the desired power.

N2 (Sample Size, Group 2)

This option is displayed if Group Allocation = "Enter N2, solve for N1"

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter R = N2/N1, solve for N1 and N2."

R is the ratio of *N2* to *N1*. That is,

$$R = N2 / N1.$$

Use this value to fix the ratio of *N2* to *N1* while solving for *N1* and *N2*. Only sample size combinations with this ratio are considered.

N2 is related to *N1* by the formula:

$$N2 = [R \times N1],$$

where the value *[Y]* is the next integer $\geq Y$.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1 (e.g., *N1* = 10 and *N2* = 20, or *N1* = 50 and *N2* = 100).

R must be greater than 0. If *R* < 1, then *N2* will be less than *N1*; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for N1 and N2."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for *N1* and *N2*. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When <u>Not</u> Solving for Sample Size)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal (N1 = N2)**

    This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter N1 and N2 individually**

    This choice permits you to enter different values for *N1* and *N2*.

- **Enter N1 and R, where N2 = R * N1**

    Choose this option to specify a value (or values) for *N1*, and obtain *N2* as a ratio (multiple) of *N1*.

- **Enter total sample size and percentage in Group 1**

    Choose this option to specify a value (or values) for the total sample size (*N*), obtain *N1* as a percentage of *N*, and then *N2* as *N - N1*.

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal (N1 = N2)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for *N1*, and also the value for *N2*.

The Sample Size Per Group must be ≥ 2. You can enter a single value or a series of values.

N1 (Sample Size, Group 1)

*This option is displayed if Group Allocation = "Enter N1 and N2 individually" or "Enter N1 and R, where N2 = R * N1."*

N1 is the number of items or individuals sampled from the Group 1 population.

N1 must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter N1 and N2 individually."

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

*This option is displayed only if Group Allocation = "Enter N1 and R, where N2 = R * N1."*

R is the ratio of *N2* to *N1*. That is,

$$R = N2/N1$$

Use this value to obtain *N2* as a multiple (or proportion) of *N1*.

N2 is calculated from *N1* using the formula:

$$N2=[R \ x \ N1],$$

where the value *[Y]* is the next integer ≥ *Y*.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If *R* < 1, then *N2* will be less than N1; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines *N1* and *N2*.

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.

You can enter a single value or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Effect Size – Means

δ (Difference Between Group Mean Changes)

Enter a value for the hypothesized difference between the group mean changes under the alternative hypothesis. Under the null hypothesis, this value is zero. You may enter a range of values such as '10, 20, 30' or '0 to 100 by 25.'

Effect Size – Standard Deviations

σ(T1) (SD of Measurements at Time 1)

Enter an estimate of the standard deviation of measurements made at time 1 (baseline). The standard deviations at time 1 are assumed to be equal for both groups.

σ(T2) (SD of Measurements at Time 2)

Enter an estimate of the standard deviation of measurements made at time 1 (follow-up). The standard deviations at time 2 are assumed to be equal for both groups.

Effect Size – Correlation

ρ (Correlation Between Measurements)

ρ is the correlation between the first and second measurements made on a particular subject. In a 2-by-2 repeated measures design, each subject is measured twice: first at the baseline, and again at the follow-up time.

Remember that ρ is the square root of the R-squared between these two measurements.

Example 1 – Power after a Study

Suppose you are planning a repeated measures (longitudinal) study to compare blood pressures in a group receiving a special drug to a control group not receiving any drug. From previous studies, you know that the standard deviation of the control group is 16 at baseline and 14 at follow-up two days later. Since the two measurements are only two days apart, you decide to use rather high correlation valued of 0.6 and 0.8. You would like to determine the power to detect a difference of 4 for group sample sizes of 10 to 200 in increments of 20 when alpha is 0.05 for a two-sided t-test based on the differences.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Groups of Pre-Post Scores** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Repeated Measures**, and then clicking on **Tests for Two Groups of Pre-Post Scores**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

Option	**Value**
Design Tab	
Solve For	**Power**
Test Type...	**T-Test**
Alternative Hypothesis	**Ha: δ ≠ 0 (Two-Sided)**
Alpha...	**0.05**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**10 to 200 by 20**
δ..	**4**
σ(T1)	**16**
σ(T2)	**14**
ρ..	**0.6 0.8**

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results for Comparing Mean Change in a Repeated Measures Design
Null Hypothesis: $\delta = 0$. Alternative Hypothesis: $\delta \neq 0$
Results assuming a t-test.

Power	N1	N2	N	δ	σ(T1)	σ(T2)	ρ	σ(Diff)	Alpha
0.09599	10	10	20	4.0	16.0	14.0	0.600	13.535	0.050
0.14145	10	10	20	4.0	16.0	14.0	0.800	9.675	0.050
0.20308	30	30	60	4.0	16.0	14.0	0.600	13.535	0.050
0.35026	30	30	60	4.0	16.0	14.0	0.800	9.675	0.050
0.30998	50	50	100	4.0	16.0	14.0	0.600	13.535	0.050
0.53469	50	50	100	4.0	16.0	14.0	0.800	9.675	0.050
0.41158	70	70	140	4.0	16.0	14.0	0.600	13.535	0.050
0.68046	70	70	140	4.0	16.0	14.0	0.800	9.675	0.050
0.50475	90	90	180	4.0	16.0	14.0	0.600	13.535	0.050
0.78772	90	90	180	4.0	16.0	14.0	0.800	9.675	0.050
0.58788	110	110	220	4.0	16.0	14.0	0.600	13.535	0.050
0.86274	110	110	220	4.0	16.0	14.0	0.800	9.675	0.050
0.66049	130	130	260	4.0	16.0	14.0	0.600	13.535	0.050
0.91323	130	130	260	4.0	16.0	14.0	0.800	9.675	0.050
0.72278	150	150	300	4.0	16.0	14.0	0.600	13.535	0.050
0.94620	150	150	300	4.0	16.0	14.0	0.800	9.675	0.050
0.77545	170	170	340	4.0	16.0	14.0	0.600	13.535	0.050
0.96719	170	170	340	4.0	16.0	14.0	0.800	9.675	0.050
0.81942	190	190	380	4.0	16.0	14.0	0.600	13.535	0.050
0.98028	190	190	380	4.0	16.0	14.0	0.800	9.675	0.050

References
Rosner, Barnard. 2011. Fundamentals of Biostatistics, Seventh Edition. Brooks/Cole. Boston, MA.

Report Definitions
Power is the probability of rejecting a false null hypothesis.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
δ is mean change of group 2 minus the mean change of group 1 assuming the alternative hypothesis.
σ(T1) and σ(T2) are the standard deviations of measurements at time 1 and time 2, respectively.
ρ is correlation between a pair of observations made on the same subject. σ(Diff) is the standard deviation of
   the paired differences, assumed equal for the two groups.
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements
A 2-by-2 repeated measures design consists of two groups of subjects, each measured at two time
points. In this case, the primary goal of the study is to compare the change across time in
group 1 to the change across time in group 2. Sample sizes of 10 in group 1 and 10 in group 2
achieve 0% power to detect a difference in mean changes of 4.0 with a standard deviation of
16.0 at the first time point, a standard deviation of 14.0 at the second time point, and a
correlation between measurement pairs of 0.600. The significance level (alpha) is 0.05000 using
a two-sided, two-sample t-test.

This report shows the values of each of the parameters, one scenario per row.

Tests for Two Groups of Pre-Post Scores

Plots Section
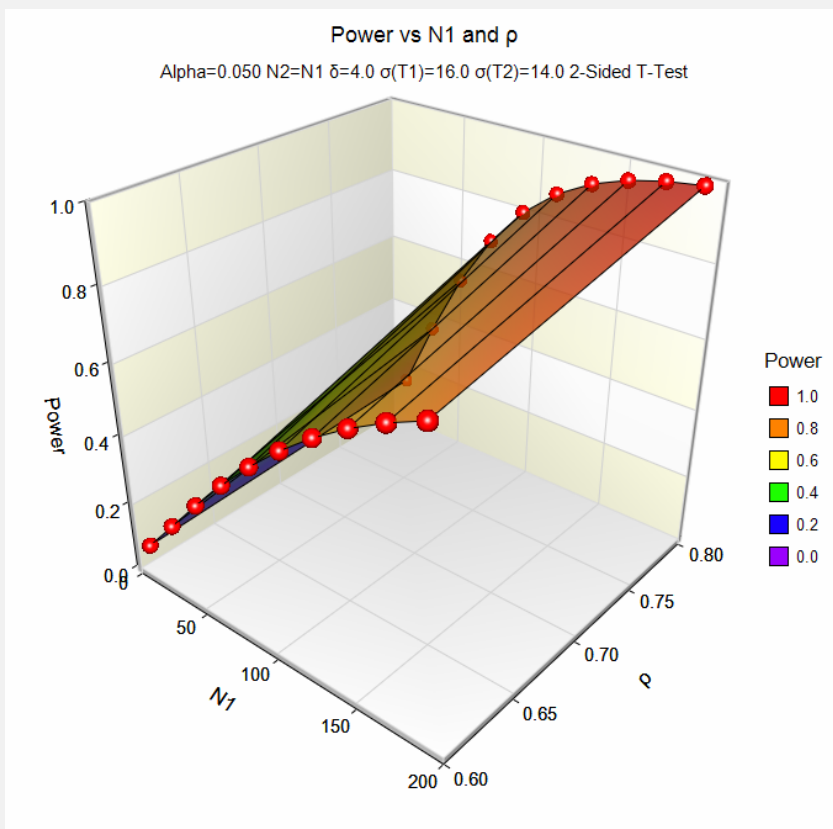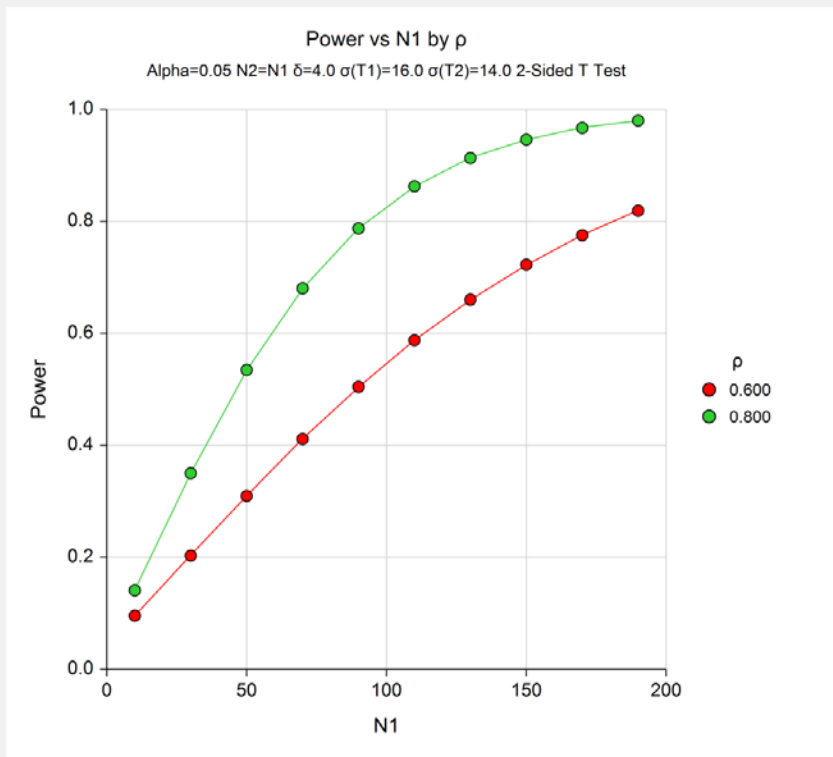
This plot shows the relationship between sample size and power for each value of ρ.

Example 2 – Validation using Rosner

Rosner (2011) page 306 calculates the power for a longitudinal study in which the standard deviations are both 15, the correlation is 0.7, the sample sizes are both 75, the mean change is 5, and a two-sided z-test is used with an alpha of 0.05. Rosner computes the power as 0.75.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Groups of Pre-Post Scores** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Repeated Measures**, and then clicking on **Tests for Two Groups of Pre-Post Scores**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Power**
Test Type...	**Z-Test**
Alternative Hypothesis	**Ha: δ ≠ 0 (Two-Sided)**
Alpha...	**0.05**
N1 (Sample Size Group 1)	**75**
N2 (Sample Size Group 2)	**Use R**
R (Sample Allocation Ratio)	**1.0**
δ..	**5**
σ(T1)	**15**
σ(T2)	**σ(T1)**
ρ..	**0.7**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results for Comparing Mean Change in a Repeated Measures Design
Null Hypothesis: δ = 0. Alternative Hypothesis: δ ≠ 0
Results assuming a z-test.

Power	N1	N2	N	δ	σ(T1)	σ(T2)	ρ	σ(Diff)	Alpha
0.75025	75	75	150	5.0	15.0	15.0	0.700	11.619	0.050

Note that the results of PASS match those of Rosner.