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 ~ *N*[*μ*, *σ*] means *X* is distributed normally with mean *μ* and standard deviation *σ*.

Difference in Mean Changes, δ

Test statistical hypothesis about

$$\delta = \delta 2 - \delta 1$$

where

δ1 = *μ12* - *μ11* (change in group 1)

δ2 = *μ22* - *μ21* (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 $\sigma_{\bar{X}} = \sigma_{Diff}\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$
3. Calculate $z_p = \frac{z_\alpha \sigma_{\bar{X}} - \delta}{\sigma_{\bar{X}}}$
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 $\sigma_{\bar{X}} = \sigma_{Diff}\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$.
3. Calculate the noncentrality parameter, $\lambda = \frac{\delta}{\sigma_{\bar{X}}}$.
4. Calculate $t_p = \frac{t_\alpha \sigma_{\bar{X}} - \delta}{\sigma_{\bar{X}}} + \lambda$.
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.

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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

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**

Output

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

Numeric Reports

Numeric Results for Comparing Mean Change

Solve For: Power
Design Type: Repeated Measures
Test Type: T-Test
Hypotheses: H0: δ = 0 vs. H1: δ ≠ 0

	Sample Size			Difference Between Group Mean Changes	Standard Deviation of Measurements		Correlation Between Measurements	Standard Deviation of Paired Differences	
Power	N1	N2	N	δ	Time 1 σ(T1)	Time 2 σ(T2)	ρ	σ(Diff)	Alpha
0.09599	10	10	20	4	16	14	0.6	13.535	0.05
0.14145	10	10	20	4	16	14	0.8	9.675	0.05
0.20308	30	30	60	4	16	14	0.6	13.535	0.05
0.35026	30	30	60	4	16	14	0.8	9.675	0.05
0.30998	50	50	100	4	16	14	0.6	13.535	0.05
0.53469	50	50	100	4	16	14	0.8	9.675	0.05
0.41158	70	70	140	4	16	14	0.6	13.535	0.05
0.68046	70	70	140	4	16	14	0.8	9.675	0.05
0.50475	90	90	180	4	16	14	0.6	13.535	0.05
0.78772	90	90	180	4	16	14	0.8	9.675	0.05
0.58788	110	110	220	4	16	14	0.6	13.535	0.05
0.86274	110	110	220	4	16	14	0.8	9.675	0.05
0.66049	130	130	260	4	16	14	0.6	13.535	0.05
0.91323	130	130	260	4	16	14	0.8	9.675	0.05
0.72278	150	150	300	4	16	14	0.6	13.535	0.05
0.94620	150	150	300	4	16	14	0.8	9.675	0.05
0.77545	170	170	340	4	16	14	0.6	13.535	0.05
0.96719	170	170	340	4	16	14	0.8	9.675	0.05
0.81942	190	190	380	4	16	14	0.6	13.535	0.05
0.98028	190	190	380	4	16	14	0.8	9.675	0.05

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

Summary Statements

A parallel two-group repeated-measures design with 2 measurements on each subject (e.g., pre-score and post-score) will be used to test whether the Group 1 pre-/post- mean difference is different from the Group 2 pre-/post- mean difference (H0: δ = 0 versus H1: δ ≠ 0). The comparison will be made using a two-sided, two-sample t-test, based on the difference of the means of pre-/post- differences of each group, with a Type I error rate (α) of 0.05. The standard deviation of measurements at the first time point is assumed to be 16, the standard deviation of measurements at the second time point is assumed to be 14, and the correlation between measurement pairs is assumed to be 0.6. To detect a difference (of differences) of 4, with sample sizes of 10 for Group 1 and 10 for Group 2, the power is 0.09599.

Dropout-Inflated Sample Size

	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
Dropout Rate	**N1**	**N2**	**N**	**N1'**	**N2'**	**N'**	**D1**	**D2**	**D**
20%	10	10	20	13	13	26	3	3	6
20%	30	30	60	38	38	76	8	8	16
20%	50	50	100	63	63	126	13	13	26
20%	70	70	140	88	88	176	18	18	36
20%	90	90	180	113	113	226	23	23	46
20%	110	110	220	138	138	276	28	28	56
20%	130	130	260	163	163	326	33	33	66
20%	150	150	300	188	188	376	38	38	76
20%	170	170	340	213	213	426	43	43	86
20%	190	190	380	238	238	476	48	48	96

Dropout Rate — The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N1, N2, and N — The evaluable sample sizes at which power is computed (as entered by the user). If N1 and N2 subjects are evaluated out of the N1' and N2' subjects that are enrolled in the study, the design will achieve the stated power.
N1', N2', and N' — The number of subjects that should be enrolled in the study in order to obtain N1, N2, and N evaluable subjects, based on the assumed dropout rate. N1' and N2' are calculated by inflating N1 and N2 using the formulas N1' = N1 / (1 - DR) and N2' = N2 / (1 - DR), with N1' and N2' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D1, D2, and D — The expected number of dropouts. D1 = N1' - N1, D2 = N2' - N2, and D = D1 + D2.

Dropout Summary Statements

Anticipating a 20% dropout rate, 13 subjects should be enrolled in Group 1, and 13 in Group 2, to obtain final group sample sizes of 10 and 10, respectively.

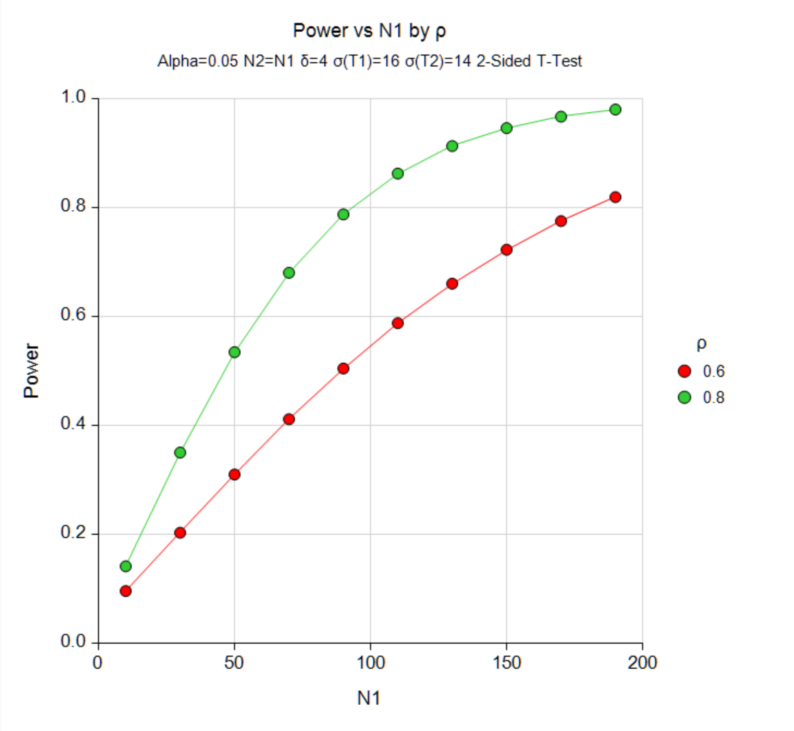
References

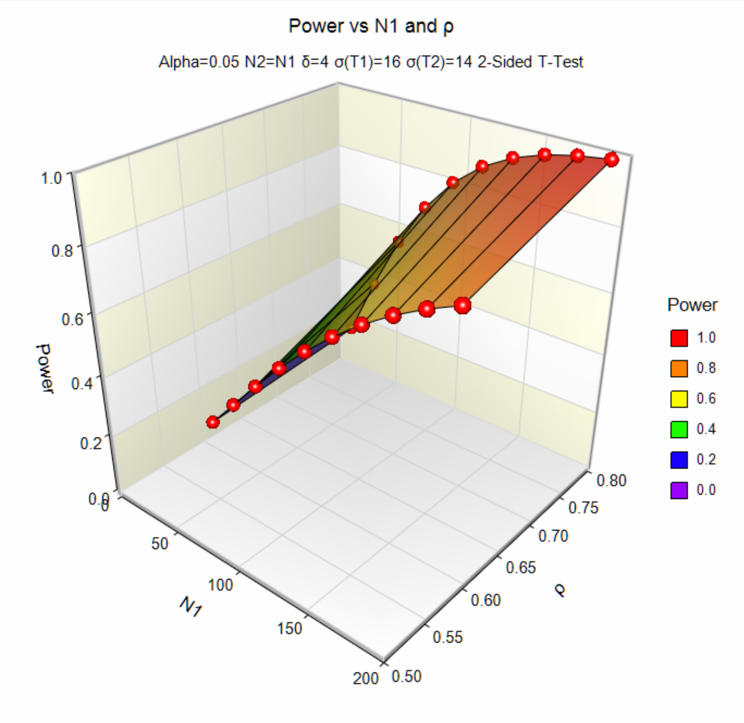
Rosner, Barnard. 2011. Fundamentals of Biostatistics, Seventh Edition. Brooks/Cole. Boston, MA.

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

Plots Section

Plots





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

Example 2 – Validation using Rosner (2011)

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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

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 for Comparing Mean Change

Solve For: Power
Design Type: Repeated Measures
Test Type: Z-Test
Hypotheses: H0: δ = 0 vs. H1: δ ≠ 0

	Sample Size			Difference Between Group Mean Changes	Standard Deviation of Measurements		Correlation Between Measurements	Standard Deviation of Paired Differences	
Power	N1	N2	N	δ	Time 1 σ(T1)	Time 2 σ(T2)	ρ	σ(Diff)	Alpha
0.75025	75	75	150	5	15	15	0.7	11.619	0.05

Note that the results of **PASS** match those of Rosner.