Chapter 596

2x2 Factorial Analysis of Variance Allowing Unequal Variances

Introduction

This procedure computes power and sample size of tests of the means from a 2x2 factorial design which will be analyzed using the Welch-Satterthwaite t-test. This method is recommended when the four group variances are not equal. The results in this chapter come from Jan and Shieh (2016).

The 2x2 factorial design may be used when tests of two factors and their interaction are desired. This procedure allows both the group variances and the sample sizes to be unequal. It also allows the tests to be made in the presence of a non-zero null distribution.

Technical Details

Suppose data from the four groups of a 2x2 factorial design each have a normal distribution with means $\mu_{1,1}, \mu_{1,2}, \mu_{2,1}, \mu_{2,2}$, standard deviations $\sigma_{1,1}, \sigma_{1,2}, \sigma_{2,1}, \sigma_{2,2}$, and sample sizes $n_{1,1}, n_{1,2}, n_{2,1}, n_{2,2}$. Let N denote the total sample size of all groups. The common ANOVA model for this design is

$$Y_{i,j,k} = \mu_0 + A_i + B_j + AB_{i,j} + \varepsilon_{i,j,k}, \text{where } i = 1,2 \text{ and } j = 1,2.$$

Since the variances are unequal, the usual F-test from the ANOVA cannot be used.is used to show that various comparisons of the means are significantly different from each other. Since each of the three terms that may need to be tested only have one degree of freedom, the results can be analyzed using appropriate contrasts. This procedure provides results for contrasts used in an ANOVA design when the variances are not necessarily equal and a non-zero null value is known.

A contrast of the means may be defined as

$$\delta = \sum_{i=1}^{2} \sum_{j=1}^{2} c_{i,j} \mu_{i,j}$$

Here, the $c_{i,j}$ are the known contrast coefficients. A unbiased estimate of δ is obtained by replacing the population means by the corresponding sample means.

The three terms in the ANOVA model may be tested using one of the following three sets of contrast coefficients:

$\boldsymbol{c_t}$	$\boldsymbol{c_{1,1}}$	$\boldsymbol{c_{1,2}}$	$\boldsymbol{c_{2,1}}$	$\boldsymbol{c_{2,2}}$
c_A	-0.5	-0.5	0.5	0.5
c_B	-0.5	0.5	-0.5	0.5
c_{AB}	0.5	-0.5	-0.5	0.5

A hypothesis test of each term uses the appropriate version of c_t. The null and alternative hypotheses for each model term are given by $\mathrm{H}_0\colon \delta = \delta_0$ versus $\mathrm{H}_1\colon \delta \neq \delta_0$. Here, δ_0 is the value of the contrast for the means at the null hypothesis, H0.

The corresponding test statistics are given by

$$T_t = \frac{\hat{\delta} - \delta_0}{\widehat{\sigma(\delta)}}$$

where

$$\delta_0 = \sum_{i=1}^{2}\sum_{j=1}^{2} c_{i,j}\mu_{0,i,j}$$

$$\hat{\delta} = \sum_{i=1}^{2}\sum_{j=1}^{2} c_{i,j}\bar{Y}_{i,j}$$

$$\widehat{\sigma^2(\delta)} = \sum_{i=1}^{2}\sum_{j=1}^{2} c_{i,j}^2 S_{i,j}^2/n_{i,j}$$

$$\bar{Y}_{i,j} = \sum_{k=1}^{n_{i,j}} Y_{i,j,k}/n_{i,j}$$

$$S_{i,j}^2 = \sum_{k=1}^{n_{i,j}} \left(Y_{i,j,k} - \bar{Y}_{i,j}\right)^2/\left(n_{i,j} - 1\right)$$

Here $\widehat{\sigma^2(\delta)}$ is an estimator of $\sigma^2(\delta) = \mathrm{Var}(\hat{\delta})$ where

$$\mathrm{Var}(\hat{\delta}) = \sum_{i=1}^{2}\sum_{j=1}^{2} c_{i,j}^2 \sigma_{i,j}^2/n_{i,j}$$

Under the null hypothesis, Satterthwaite (1946) and Welch (1947) showed that *T* is approximately distributed as a Student's *t* with *v* degrees of freedom where

$$v = \frac{\left(\sum_{i=1}^{2}\sum_{j=1}^{2} c_{i,j}^2\sigma_{i,j}^2/n_{i,j}\right)^2}{\sum_{i=1}^{2}\sum_{j=1}^{2} c_{i,j}^4\sigma_{i,j}^4/\left[n_{i,j}^2\left(n_{i,j}-1\right)\right]}$$

This value, *v*, is estimated as

$$\hat{v} = \frac{\left(\sum_{i=1}^{2}\sum_{j=1}^{2} c_{i,j}^2 S_{i,j}^2/n_{i,j}\right)^2}{\sum_{i=1}^{2}\sum_{j=1}^{2} c_{i,j}^4 S_{i,j}^4/\left[n_{i,j}^2\left(n_{i,j}-1\right)\right]}$$

The test statistic T has the approximate distribution

$$T \sim t_{\hat{v}}$$

The two-sided, Welch-Satterthwaite test rejects H_0 at a significance level α if $|T| > t_{\hat{v},1-\frac{\alpha}{2}}$.

Power

Shieh and Jan (2015) noted that T has the general approximate distribution

$$T \sim t_{\hat{v},\Delta}$$

where $t_{\hat{v},\Delta}$ is a noncentral t with $\hat{v}$ degrees of freedom and noncentrality parameter Δ. Here, Δ is defined as

$$\Delta = \frac{\delta_1 - \delta_0}{\sigma(\delta)}$$

where

$$\delta_1 = \sum_{i=1}^{2}\sum_{j=1}^{2} c_{i,j}\mu_{1,i,j}$$

is the value of the contrast for the means at the alternative hypothesis, H1.

Hence, the power can be approximated as

$$Power = P\left(\left|t_{v,\Delta}\right| > t_{v,1-\frac{\alpha}{2}}\right)$$

When a sample size is desired, it can be determined using a standard binary search algorithm.

Example 1 – Finding Sample Size

Suppose an experiment is being designed to assess the sample size needed for a 2x2 design that will be analyzed with the extended Welch test at a significance level of 0.05 and a power of 0.9. The null means are all 0. The alternative means are {48, 62, 66, 64}. The standard deviations are {3, 5, 4, 6}. Two types of group allocations will be used: equal (Eq) and proportional to σ (SD).

The following table shows the setting each of the parameters.

A	B	μ1	σ	Eq	SD
1	1	48	3	1	3
1	2	62	5	1	5
2	1	66	4	1	4
2	2	64	6	1	6

The group allocation values are entered into the spreadsheet as follows.

Eq	SD
1	3
1	5
1	4
1	6

To better understand the pattern of means that we have chosen, press the ***Sm*** button to the right of the ***μ1* (*H1 Group Means*)** section. This will bring up the ***Standard Deviation of Means Calculator*** window. Enter the four means in the ***Input*** table at the top, two per row. When you are finished, the ***Input*** table should appear as follows:

	B1	**B2**
A1	48	62
A2	66	64

The ***Output*** table shows the ***Effects*** for each term. Note that the ***Effects*** of factor A are -5 and 5, those of factor B are -3 and 3, and those of the interaction are -4 and 4 (shown in the 2-by-2 table in the upper left of the table).

Notice that the ***Standard Deviation of Means*** **(*Sm*)** report at the bottom presents the standard deviations of each of the terms: Sm(A) = 5, Sm(B) = 3, and Sm(AB) = 4. Note that these values are equal to the absolute values of the ***Effects*** shown in the ***Output*** table. This will always be true for 2-by-2 tables.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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	**Sample Size**
Alternative Hypothesis	**Two-Sided (H1: δ ≠ δ0)**
Power	**0.90**
Alpha	**0.05**
Group Allocation Input Type	**Enter Columns of Allocation Patterns**
Columns of Allocation Patterns	**1-2**

Model Terms.. **A, B, AB**
μ0 Input Type.. **μ0 (H0 Group Means)**
μ0(1,1) **0**
μ0(1,2) **0**
μ0(2,1) **0**
μ0(2,2) **0**
μ1 Input Type.. **Enter μ1 (H1 Group Means)**
μ1(1,1) **48**
μ1(1,2) **62**
μ1(2,1) **66**
μ1(2,2) **64**
σ Input Type... **σ (Group Standard Deviations)**
σ(1,1) **3**
σ(1,2) **5**
σ(2,1) **4**
σ(2,2) **6**

Input Spreadsheet Data

Row	Eq	SD
1	1	3
2	1	5
3	1	4
4	1	6

Annotated Output

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

Numeric Results

Numeric Results
Model Y(i,j,k) = μ + A(i) + B(j) + AB(i,j) + e(i,j,k), i = 1,2, j = 1,2
Hypotheses H0: δ = δ0 vs. H1: δ ≠ δ0

		Sample Size			Group Means		Diff's Among Means		Group	SE of		
Term	Power	Total N	Alloc r	Grp n	H0 μ0	H1 μ1	H0 δ0	H1 δ1	SD's σ	δ's σ(δ)	NCP Δ	Alpha
1: A	0.97150	16	Eq(1)	n(1)	μ0(1)	μ1(1)	0	10	σ(1)	2.318	4.313	0.05
2: B	0.90184	28	Eq(1)	n(2)	μ0(1)	μ1(1)	0	6	σ(1)	1.753	3.424	0.05
3: AB	0.94549	20	Eq(1)	n(3)	μ0(1)	μ1(1)	0	-8	σ(1)	2.074	-3.858	0.05
1: A	0.91419	12	SD(2)	n(4)	μ0(1)	μ1(1)	0	10	σ(1)	2.606	3.837	0.05
2: B	0.91081	27	SD(2)	n(5)	μ0(1)	μ1(1)	0	6	σ(1)	1.735	3.458	0.05
3: AB	0.93828	18	SD(2)	n(6)	μ0(1)	μ1(1)	0	-8	σ(1)	2.121	-3.771	0.05

Value Lists

Name	(1,1)	(1,2)	(2,1)	(2,2)
C_A*	-0.5	-0.5	0.5	0.5
C_B*	-0.5	0.5	-0.5	0.5
C_AB*	0.5	-0.5	-0.5	0.5
Eq(1)	1	1	1	1
SD(2)	3	5	4	6
μ0(1)	0	0	0	0
μ1(1)	48	62	66	64
σ(1)	3	5	4	6

*C_A is the contrast for testing A. C_B is the contrast for testing B. C_AB is the contrast for testing AB.

Group Sample Size Details

		Group Sample Sizes				Group Allocation Proportions			
n	N	n(1,1)	n(1,2)	n(2,1)	n(2,2)	p(1,1)	p(1,2)	p(2,1)	p(2,2)
n(1)	16	4	4	4	4	0.250	0.250	0.250	0.250
n(2)	28	7	7	7	7	0.250	0.250	0.250	0.250
n(3)	20	5	5	5	5	0.250	0.250	0.250	0.250
n(4)	12	2	3	3	4	0.167	0.250	0.250	0.333
n(5)	27	4	8	6	9	0.148	0.296	0.222	0.333
n(6)	18	3	5	4	6	0.167	0.278	0.222	0.333

References

Jan, S.L. and Shieh, G. 2016. 'A systematic approach to designing statistically powerful heteroscedastic 2 x 2 factorial studies while minimizing financial costs.' BMC Medical Research Methodology, 16:114.
Kirk, Roger E. 2013. Experimental Design: Procedures for the Behavioral Sciences, 4th Edition. Sage. Washington, D.C.
Luh, W.M. and Guo, J.H. 2016. 'Allocating sample sizes to reduce budget for fixed-effect 2 x 2 heterogeneous analysis of variance.' Journal of Experimental Education, 84:197-211.
Satterthwaite, F.E. 1946. 'An approximate distribution of estimate of variance components,' Biometric Bulletin, 2:110-114.
Shieh, G. and Jan, S-L. 2015. 'Power and sample size calculations for testing linear combinations of group means under variance heterogeneity with applications to meta and moderation analysis'. Psicologica, 36:367-390.
Welch, B.L. 1951. 'On the Comparison of Several Mean Values: An Alternative Approach'. Biometrika,38,330-336.

Report Definitions

Term is the number and name of the model term being reported on this report line. Each term is associated with a different array of contrast coefficients: C_A, C_B, or C_AB. Note that the term number is the number that is plotted along the horizontal axis, in the legend, or in the plot title.
Power is the probability of rejecting a false non-zero null hypothesis in favor of the alternative hypothesis.
N, the Total Sample Size, is the total number of subjects in the study found by summing the group sample sizes.
r, the Group Allocation Set, is the name and number of the set containing the Group Allocation Pattern {r(1,1), r(1,2), r(2,1), r(2,2)}. The pattern values are rescaled to sum to one and thus become the Group Allocation Proportions.
n, the Group Sample Size, is the name and number of the set containing the sample size of each group.
μ0, the H0 Group Means, is the name and number of the set containing the group means under H0. Note that δ0 = μ0'C, where C is the appropriate array of contrast coefficients.
μ1, the H1 Group Means, is the name and number of the set containing the group means under H1. This is the set of means at which the power is calculated using δ1 = μ1'C.
δ0, the Diff's Among Means|H0, is the dot product of μ0 and C, where C is the appropriate set of contrast coefficients (C = C_A, C_B, or C_AB). The dot product is the sum of the products of the corresponding entries of the two sets of numbers. Note that you must have δ0 ≠ δ1. Also note that when δ0 ≠ 0, this test is called a 'non-zero null' test.
δ1, the Diff's Among Means|H1, is the dot product of μ1 and C, where C is the appropriate set of contrast coefficients (C = C_A, C_B, or C_AB). The dot product is the sum of the products of the corresponding entries of the two sets of numbers. Thus, δ1 is the difference between the means at the two levels of the term. Note that (δ1)/2 is sometimes called the 'effect' of with this term. Also note that you must have δ0 ≠ δ1.
σ, the Group Standard Deviations, is the name and number of the set containing the standard deviation of each group.
σ(δ), the standard error of the δ's, is used in the calculation of Δ. Note that $\sigma(\delta)^2 = \Sigma i\{\Sigma j[c(i,j)\ \sigma(i,j)]^2 / n(i,j)\}$, where the c(i,j) are the individual contrast coefficients of the term being reported on this row.
Δ, the NCP (noncentrality parameter), is used with the adjusted noncentral t-distribution to calculate the power. Note that Δ = (δ1 - δ0) / σ(δ).
Alpha is the significance level of the test: the probability of rejecting H0 when it is actually true.

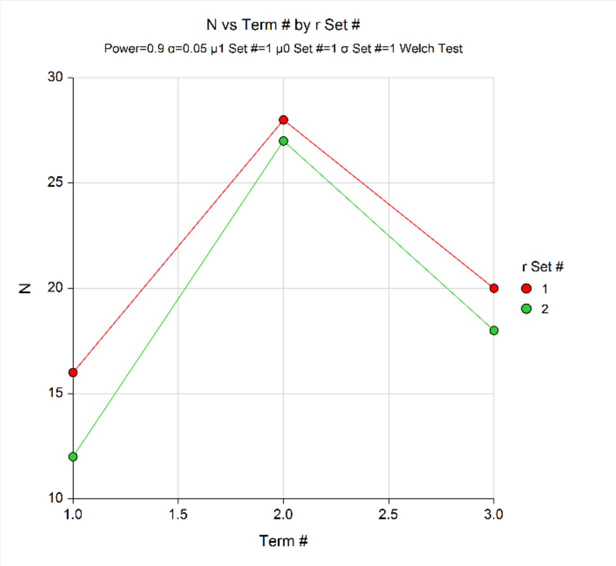
Summary Statements

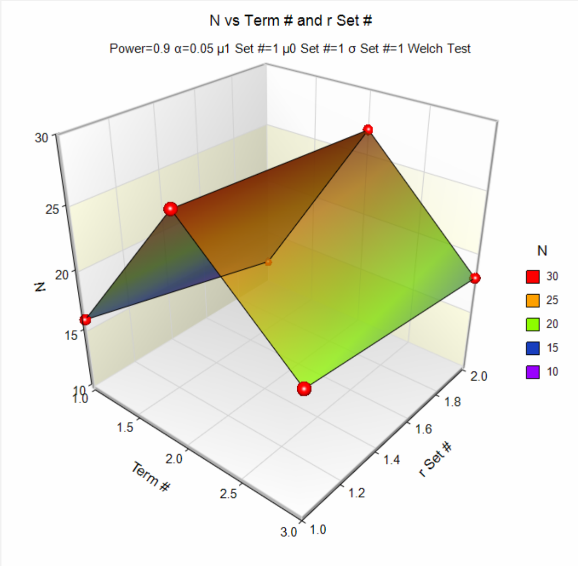
In a non-zero null, 2x2 factorial ANOVA study that allows for unequal group variances, a sample of 16 subjects, divided among four groups, achieves a power of 97% for testing the main-effect A. This power assumes the data will be analyzed with Welch's adjusted degrees of freedom t-test at a significance level of 0.05. The group subject counts are 4, 4, 4, 4. The group means under the null hypothesis are 0, 0, 0, 0. The group means under the alternative hypothesis are 48, 62, 66, 64. The term contrast coefficients are -0.5,-0.5,0.5,0.5. The group standard deviations are 3, 5, 4, 6. The value of the contrast applied to the means under the null hypothesis is 0. The value of the contrast applied to the means under the alternative hypothesis is 10. The noncentrality parameter is 4.313.

This report shows the numeric results of this study along with the report definitions.

Chart Section

Chart Section





These plots give a visual presentation of the results in the Numeric Report.

Note that the three values of the Term # axis are 1, 2, and 3. These values correspond to the terms A, B, and AB.

Example 2 – Validation using Jan and Shieh (2016)

Jan and Shieh (2016) page 6, Table 1, presents an example in which alpha = 0.05, the sample sizes are {16, 14, 7, 15}, the standard deviations are {0.83, 0.72, 0.34, 0.77}, the null means are {0, 0, 0, 0}, and the alternative means are {1.23, 0.42, 0.13, 0.38}. The resulting power for testing the interaction is given as 0.8038.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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**
Alternative Hypothesis	**Two-Sided (H1: δ ≠ δ0)**
Alpha	**0.05**
Group Allocation Input Type	**Enter n (Group Sample Sizes)**
n(1,1)	**16**
n(1,2)	**14**
n(2,1)	**7**
n(2,2)	**15**
Model Terms	**AB**
μ0 Input Type	**μ0 (H0 Group Means)**
μ0(1,1)	**0**
μ0(1,2)	**0**
μ0(2,1)	**0**
μ0(2,2)	**0**
μ1 Input Type	**Enter μ1 (H1 Group Means)**
μ1(1,1)	**1.23**
μ1(1,2)	**0.42**
μ1(2,1)	**0.13**
μ1(2,2)	**0.38**
σ Input Type	**σ (Group Standard Deviations)**
σ(1,1)	**0.83**
σ(1,2)	**0.72**
σ(2,1)	**0.34**
σ(2,2)	**0.77**

Output

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

Numeric Results

Numeric Results

Model Y(i,j,k) = μ + A(i) + B(j) + AB(i,j) + e(i,j,k), i = 1,2, j = 1,2
Hypotheses H0: δ = δ0 vs. H1: δ ≠ δ0

		Sample Size		Group Means		Diff's Among Means		Group	SE of		
		Total	Grp	H0	H1	H0	H1	SD's	δ's	NCP	
Term	Power	N	n	μ0	μ1	δ0	δ1	σ	σ(δ)	Δ	Alpha
1: AB	0.80376	52	n(1)	μ0(1)	μ1(1)	0	0.53	σ(1)	0.184	2.873	0.05

Value Lists

Name	(1,1)	(1,2)	(2,1)	(2,2)
C_AB*	0.5	-0.5	-0.5	0.5
n(1)	16	14	7	15
μ0(1)	0	0	0	0
μ1(1)	1.23	0.42	0.13	0.38
σ(1)	0.83	0.72	0.34	0.77

*C_AB is the contrast for testing AB.

Group Sample Size Details

		Group Sample Sizes				Group Allocation Proportions			
n	N	n(1,1)	n(1,2)	n(2,1)	n(2,2)	p(1,1)	p(1,2)	p(2,1)	p(2,2)
n(1)	52	16	14	7	15	0.308	0.269	0.135	0.288

PASS found the power to be 0.80376 which rounds to 0.8038. The procedure is validated.