Chapter 589

One-Way Analysis of Variance Allowing Unequal Variances

Introduction

This procedure computes power and sample size of tests of the equality of multiple means which are analyzed using the Welch test which is recommended when the group variances are not equal. Note that the Welch test uses an adjusted degrees of freedom to compensate for the unequal variances. The results in this chapter come from Jan and Shieh (2014, 2019).

Technical Details

Suppose *G* groups each have a normal distribution and with means $\mu_1, \mu_2, \ldots, \mu_G$ and standard deviations $\sigma_1, \sigma_2, \ldots, \sigma_G$. Let $n_1, n_2, \ldots, n_G$ denote the sample size of each group and let *N* denote the total sample size of all groups. The statistical hypotheses that are tested is

$$H_0\colon \omega^2 \leq 0 \quad \text{versus} \quad H_1\colon \omega^2 > 0$$

Here we let ω^2 represent the variation in the standardized means.

Test Statistic

ANOVA F-Test

Assuming homogeneity of variance among the groups, the most popular procedure for analyzing a set of G means is the ANOVA F-Test which is calculated as follows.

$$F^* = \frac{SSM/(G-1)}{SSE/(N-G)}$$

where *SSM* is the sum of squares of treatment means, *SSE* is the sum of squares of error, and *N* is the total sample size.

Welch's Test

If variance heterogeneity is suspected, a common approach is to use Welch's procedure which is calculated as follows.

$$W = \frac{\sum_{i=1}^{G} W_i(\bar{X}_i - \bar{X})/(G-1)}{1 + 2(G-2)Q/(G^2-1)}$$

where

$W_i = n_i/S_i^2$, $S_i^2 = \sum_{j=1}^{n_i}\left(X_{ij} - \bar{X}_i\right)^2/(n_i - 1)$, $\bar{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i$, $\bar{X} = \sum_{i=1}^{G} W_i\bar{X}_i/U$, $U = \sum_{i=1}^{G} W_i$, and

$Q = \sum_{i=1}^{G}\left(1 - \frac{W_i}{U}\right)^2/(n_i - 1)$.

Under the null hypothesis, Welch (1951) gave the approximate distribution of *W* as an *F* distribution with degrees of freedom *G* – 1 and *v*, where

$$v = \frac{G^2 - 1}{3Q}$$

Jan and Shieh's Extension of Welch's Test

Jan and Shieh (2019) proposed testing the *G* group means in the face of variance heterogeneity using the hypotheses

$$H_0: \omega^2 \leq 0 \quad \text{versus} \quad H_1: \omega^2 > 0$$

where

$\omega^2 = \sum_{i=1}^{G} w_i(\mu_i - \mu^*)^2$, $w_i = \frac{n_i}{N\sigma_i^2}$, $\mu^* = \sum_{i=1}^{G}\frac{w_i\mu_i}{v}$, and $v = \sum_{i=1}^{G} w_i$.

Under H_0, *W* is assumed to follow the F distribution $W \sim F_{G-1,v}$.

The null hypothesis is rejected at the significance level α if $W > F_{1-\alpha,G-1,v}$.

Power

The power function of the extended Welch's test computed at a particular set of means, $\mu_{11}, \mu_{12}, \ldots, \mu_{1G}$, is given by

$$\text{Power} = \Pr\left[F'_{G-1,N-G,\Omega_1} > F_{1-\alpha,G-1,\eta}\right]$$

where

$\Omega_1 = N\omega_1^2$, $\eta = (G^2-1)/(3\tau)$, $\tau = \sum_{i=1}^{G}\left(1-\frac{w_i}{v}\right)^2/(n_i-1)$, $\omega_1^2 = \sum_{i=1}^{G} w_i(\mu_{1i}-\mu_1^*)^2$, and $\mu_1^* = \sum_{i=1}^{G}\frac{w_i\mu_{1i}}{v}$.

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

Example 1 – Finding Sample Size

An experiment is being designed to assess sample size needed for an F test of the four means using the Welch test with a significance level of 0.05 and a power of 0.9. The first group is a control group and the three remaining groups are treatment groups. Previous studies have shown that the standard deviations in the four groups are {5, 4, 3, 4}. The sample sizes will be equal across all groups.

Three sets of alternative treatment means are to be compared: {17, 17, 13, 13}, {17, 16, 14, 13}, and {17, 15, 15, 13}. Note that these three alternative sets of means all have the same range (4) but different configurations.

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 .... **Sample Size**
Power.. **0.90**
Alpha.. **0.05**
G (Number of Groups) **4**
Group Allocation Input Type .... **Equal (n1 = ⋯ = nG)**
μ1 Input Type.. **Enter Columns Containing Sets of μ1's**
Columns Containing Sets of μ1's................... **1-3**
σ Input Type.. **Enter σ (Group Standard Deviations)**
σ (Group Standard Deviations)...................... **5 4 3 4**

Input Spreadsheet Data

Row	C1	C2	C3
1	17	17	17
2	17	16	15
3	13	14	15
4	13	13	13

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size
Number of Groups: 4

	Sample Size		Group	Standard	Standard	
	Total	Group	Means|H1 Pattern	Deviation Pattern	Deviation of Standardized Means	
Power	N	ni	μ1	σ	ω1	Alpha
0.90968	64	16	C1(1)	σ(1)	0.508	0.05
0.90619	112	28	C2(2)	σ(1)	0.371	0.05
0.90006	148	37	C3(3)	σ(1)	0.317	0.05

Item	Values
C1(1)	17, 17, 13, 13
C2(2)	17, 16, 14, 13
C3(3)	17, 15, 15, 13
σ(1)	5, 4, 3, 4

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N The total number of subjects in the study. N is the sum of the group sample sizes.
μ1 The Group Means|H1 is the column name and set number of the group means under the alternative hypothesis. This is the set of means at which the power is calculated.
σ The set name and number of the group standard deviations.
ω1 The standard deviation of the standardized means (μ1i / σi) assumed by the alternative hypothesis, H1.
Alpha The probability of rejecting a true null hypothesis.

Group Sample Size Details

n	N	Group Sample Sizes	Group Allocation Proportions
n(1)	64	16, 16, 16, 16	0.25, 0.25, 0.25, 0.25
n(2)	112	28, 28, 28, 28	0.25, 0.25, 0.25, 0.25
n(3)	148	37, 37, 37, 37	0.25, 0.25, 0.25, 0.25

Summary Statements

A one-way ANOVA design with 4 groups will be used to test whether there is a difference among the 4 group means. The comparison will be made using a Welch's unequal variance F-test with a Type I error rate (α) of 0.05. The within-group standard deviations of responses for the 4 groups are assumed to be 5, 4, 3, 4. To detect the means 17, 17, 13, 13 (standard deviation of group means = 0.508), with 90% power, the needed group sample sizes are 16, 16, 16, 16 (for a total of 64 subjects).

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	64	80	16
20%	112	140	28
20%	148	185	37

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.
N The evaluable sample size at which power is computed. If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated power.
N' The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects, based on the assumed dropout rate. After solving for N, N' is calculated by inflating N using the formula N' = N / (1 - DR), with N' 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.)
D The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 80 subjects should be enrolled to obtain a final sample size of 64 subjects.

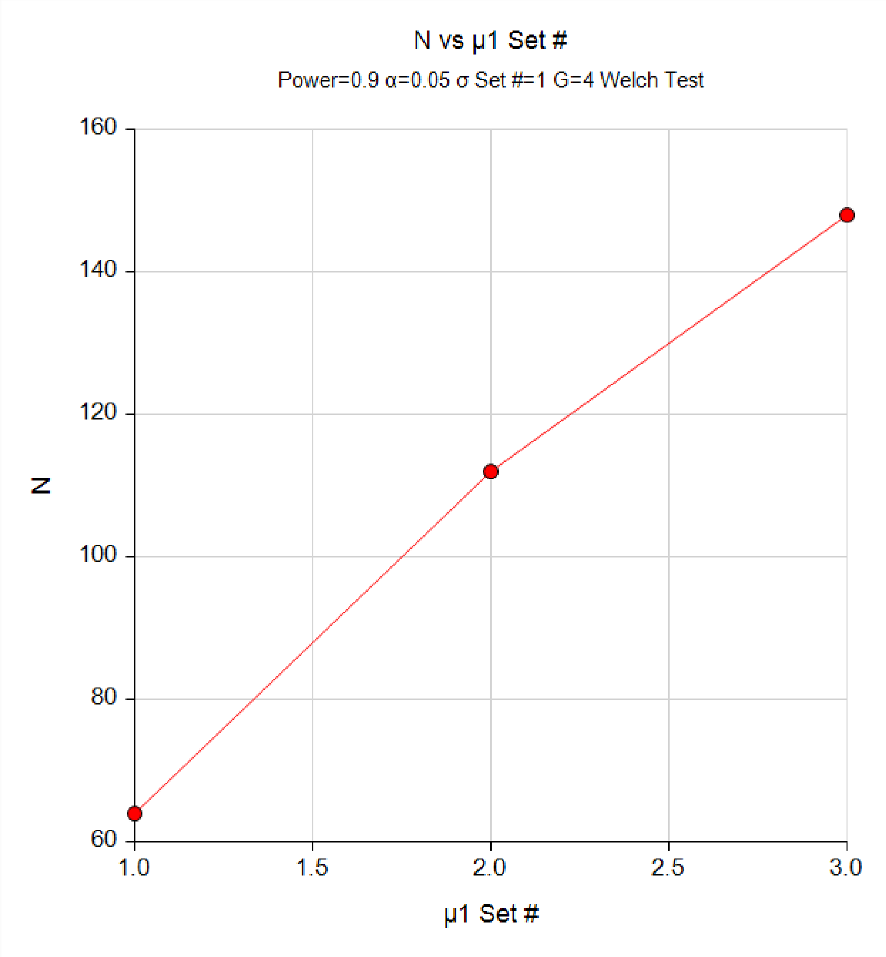
References

Jan, S-L and Shieh, G. 2019. 'On the Extended Welch Test for Assessing Equivalence of Standardized Means'. Statistics in Biopharmaceutical Research. DOI:10.1080/19466315.2019.1654915
Shieh, G. 2018. 'On Detecting a Minimal Important Difference among Standardized Means'. Current Psychology, Vol 37, Pages 640-647. Doi: 10.1007/s12144-016-9549-5
Jan, S-L and Shieh, G. 2014. 'Sample size determinations for Welch's test in one-way heteroscedastic ANOVA'. British Journal of Mathematical and Statistical Psychology. Vol. 67, Pages 72-93.
Shieh, G. and Jan, S-L. 2013. 'Determining Sample Size With a Given Range of Mean Effects in One-Way Heteroscedastic Analysis of Variance'. The Journal of Experimental Education, Vol 81(3), Pages 281-294.
Cohen, Jacob. 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates. Hillsdale, New Jersey.
Welch, B.L. 1951. 'On the Comparison of Several Mean Values: An Alternative Approach'. Biometrika,38,330-336.

This report shows the numeric results of this study.

Plots Section

Plots



This plot gives a visual presentation of the results in the Numeric Report. Note that the horizontal axis plots the indices {1, 2, 3} of μ1(1), μ1(2), and μ1(3). That is, the horizontal value "1", is the index of the set μ1(1), and so on.

Example 2 – Validation using Jan and Shieh (2014)

Jan and Shieh (2014) page 79 presents an example in which alpha = 0.05, G = 4, the sample sizes are {10, 20, 30, 40}, the standard deviations are {1, 2, 3, 4}, and the alternative means are {1, 0, 0, -1}. The resulting power is given as 0.7129.

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**
Alpha.. **0.05**
G (Number of Groups) **4**
Group Allocation Input Type .... **Enter n (Group Sample Sizes)**
n (Group Sample Sizes)................................ **10 20 30 40**
μ1 Input Type... **Enter μ1 (Group Means|H1)**
μ1 (Group Means|H1) .... **1 0 0 -1**
σ Input Type... **Enter σ (Group Standard Deviations)**
σ (Group Standard Deviations)...................... **1 2 3 4**

Output

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

Numeric Results

Solve For: Power
Number of Groups: 4

	Sample Size					
Power	Total N	Group n	Group Means|H1 Pattern μ1	Standard Deviation Pattern σ	Standard Deviation of Standardized Means ω1	Alpha
0.71286	100	n(1)	μ1(1)	σ(1)	0.313	0.05

Item	Values
n(1)	10, 20, 30, 40
μ1(1)	1, 0, 0, -1
σ(1)	1, 2, 3, 4

Group Sample Size Details

n	N	Group Sample Sizes	Group Allocation Proportions
n(1)	100	10, 20, 30, 40	0.1, 0.2, 0.3, 0.4

PASS also found the power to be 0.71286. The procedure is validated.