

Chapter 590

Multiple Contrasts (Simulation)

Introduction

This procedure uses simulation to analyze the power and significance level of two multiple-comparison procedures that perform two-sided hypothesis tests of contrasts of the group means. These are the Dunn-Bonferroni test and the Dunn-Welch test. For each scenario, two simulations are run: one estimates the significance level and the other estimates the power.

The term *contrast* refers to a user-defined comparison of the group means. The term *multiple contrasts* refers to a set of such comparisons. An additional restriction imposed is that the contrast coefficients to sum to zero.

When several contrasts are tested, the interpretation of the results is more complex because of the problem of *multiplicity*. *Multiplicity* here refers to the fact that the probability of making at least one incorrect decision increases as the number of statistical tests increases. Methods for testing *multiple contrasts* have been developed to account for this multiplicity.

Error Rates

When dealing with several simultaneous statistical tests, both individual-wise and experiment wise error rates should be considered.

1. **Comparison-wise error rate.** This is the probability of a type-I error (rejecting a true H₀) for a particular test. In the case of the five-group design, there are ten possible comparison-wise error rates, one for each of the ten possible pairs. We will denote this error rate α_c .
2. **Experiment-wise (or family-wise) error rate.** This is the probability (FWER) of making one or more type-I errors in the set (family) of comparisons. We will denote this error rate α_f .

The relationship between these two error rates when the tests are independent is given by

$$\alpha_f = 1 - (1 - \alpha_c)^C$$

where C is the total number of contrasts. For example, if α_c is 0.05 and C is 10, α_f is 0.401. There is about a 40% chance that at least one of the ten contrasts will be concluded to be non-zero when in fact they are not. When the tests are correlated, as they might be among a set of contrasts, the above formula provides an upper bound to the family-wise error rate.

The techniques described below provide control for α_f rather than α_c .

Technical Details

The One-Way Analysis of Variance Design

The discussion that follows is based on the common one-way analysis of variance design which may be summarized as follows. Suppose the responses Y_{ij} in k groups each follow a normal distribution with means $\mu_1, \mu_2, \dots, \mu_k$ and unknown variance σ^2 . Let n_1, n_2, \dots, n_k denote the number of subjects in each group. The control group is assumed to be group one.

The analysis of these responses is based on the sample means

$$\hat{\mu}_i = \bar{Y}_i = \sum_{j=1}^{n_i} \frac{Y_{ij}}{n_i}$$

and the pooled sample variance

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2}{\sum_{i=1}^k (n_i - 1)}$$

The F test is the usual method of analysis of the data from such a design, testing whether all of the means are equal. However, a significant F test does not indicate which of the groups are different, only that at least one is different. The analyst is left with the problem of determining which of the groups are different and by how much.

To determine the mean differences that are most important, the researcher may specify a set of contrasts. These contrasts, called *a priori* or *planned*, contrasts should be specified before the experimental results are viewed.

The Dunn-Bonferroni procedure and the Dunn-Welch procedure have been developed to test these planned contrasts. The calculations associated with each of these tests are given below.

Contrasts

A contrast of the means is a stated difference among the means. The difference is constructed so that it represents a useful hypothesis. For example, suppose there are four groups, the first of which is a control group. It might be of interest to determine which treatments are statistically different from the control. That is, the differences $\mu_2 - \mu_1$, $\mu_3 - \mu_1$, and $\mu_4 - \mu_1$ would be tested to determine if they are non-zero.

Contrasts are often simple differences between two means. However, they may involve more than just two means. For example, suppose the first two groups receive one treatment and the second two groups receive another treatment. The contrast (difference) that would be tested is $(\mu_1 + \mu_2) - (\mu_3 + \mu_4)$.

Every contrast can be represented by the list of contrast coefficients: the values by which the means are multiplied. Here are some examples of contrasts that might be of interest when the experiment involves four groups.

Multiple Contrasts (Simulation)

<u>Difference</u>	<u>Coefficients</u>
$\mu_2 - \mu_1$	-1, 1, 0, 0
$\mu_3 - \mu_1$	-1, 0, 1, 0
$(\mu_1 + \mu_2) - (\mu_3 + \mu_4)$	1, 1, -1, -1
$(\mu_1 + \mu_1 + \mu_1) - (\mu_2 + \mu_3 + \mu_4)$	3, -1, -1, -1
$(\mu_4 + \mu_4) - (\mu_2 + \mu_3)$	0, -1, -1, 2

Note that in each case, the coefficients sum to zero. This makes it possible to test whether the quantity is different from zero.

A lot is written about *orthogonal contrasts* which have the property that the sum of the products of corresponding coefficients is zero. For example, the sum of the products of the last two contrasts given above is $0(3) + (-1)(-1) + (-1)(-1) + (2)(-1) = 0 + 1 + 1 - 2 = 0$, so these two contrasts are orthogonal. However, the first two contrasts are not orthogonal since $(-1)(-1) + (1)(0) + (0)(1) + (0)(0) = 1 + 0 + 0 + 0 = 1$ (not zero). Orthogonal contrasts have nice properties when the sample sizes are equal. Unfortunately, they lose those properties when the group sample sizes are unequal or when the data are not normally distributed.

The procedures described in this chapter do not require that the contrasts be orthogonal. Instead, you should focus on defining a set of contrasts that answer the research questions of interest.

Dunn-Bonferroni Test

Dunn (1964) developed a procedure for simultaneously testing several contrasts. This method is also discussed in Kirk (1982) pages 106 to 109. The method consists of testing each contrast with Student's t distribution with degrees of freedom equal to $N-k$ with a Bonferroni adjustment of the alpha value. That is, the alpha value is divided by C , the number of contrasts simultaneously tested.

The test rejects H_0 if

$$\frac{|\sum_{i=1}^k c_i \bar{Y}_i|}{\sqrt{\hat{\sigma}^2 \left(\sum_{i=1}^k \frac{c_i^2}{n_i} \right)}} \geq |t_{1-\alpha/(2C), N-k}|$$

Note that this is a two-sided test of the hypothesis that $\sum_{i=1}^k c_i \mu_i = 0$ where $\sum_{i=1}^k c_i = 0$.

Dunn-Welch Test

Dunn (1964) developed a procedure for simultaneously testing several contrasts. This method, using Welch's (1947) modification for the unequal variances, is discussed in Kirk (1982) pages 100, 101, 106 - 109. The method consists of testing each contrast with Student's t distribution with degrees of freedom given below with a Bonferroni adjustment of the alpha value. That is, the alpha value is divided by C , the number of contrasts simultaneously tested.

The two-sided test statistic rejects H_0 if

$$\frac{|\sum_{i=1}^k c_i \bar{Y}_i|}{\sqrt{\sum_{i=1}^k \frac{c_i^2 \hat{\sigma}_i^2}{n_i}}} \geq |t_{1-\alpha/(2C), v'}|$$

where

$$v' = \frac{\left(\sum_{i=1}^k \frac{c_i^2 \hat{\sigma}_i^2}{n_i} \right)^2}{\sum_{i=1}^k \frac{c_i^4 \hat{\sigma}_i^4}{n_i^2 (n_i - 1)}}$$

Definition of Power for Multiple Contrasts

The notion of power is well-defined for individual tests. Power is the probability of rejecting a false null hypothesis. However, this definition does not extend directly when there are a number of simultaneous tests. The two definitions that we emphasize in **PASS**, where recommended by Ramsey (1978). They are *any-contrast power* and *all-contrasts power*. Other design characteristics, such as *average-contrast power* and *false-discovery rate*, are important to consider. However, our review of the statistical literature resulted in our focus on these two definitions of power.

Any-Contrast Power

Any-contrast power is the probability of detecting at least one of the contrasts that are actually non-zero.

All-Contrasts Power

All-contrast power is the probability of detecting all of the contrasts that are actually non-zero.

Simulation Details

Computer simulation allows us to estimate the power and significance level that is actually achieved by a test procedure in situations that are not mathematically tractable. Computer simulation was once limited to mainframe computers. But, in recent years, as computer speeds have increased, simulation studies can be completed on desktop and laptop computers in a reasonable period of time.

The steps to a simulation study are

1. Specify how each test is to be carried out. This includes indicating how the test statistic is calculated and how the significance level is specified.
2. Generate random samples from the distributions specified by the alternative hypothesis. Calculate the test statistics from the simulated data and determine if the null hypothesis is accepted or rejected. The number rejected is used to calculate the power of each test.
3. Generate random samples from the distributions specified by the null hypothesis. Calculate each test statistic from the simulated data and determine if the null hypothesis is accepted or rejected. The number rejected is used to calculate the significance level of each test.
4. Repeat steps 2 and 3 several thousand times, tabulating the number of times the simulated data leads to a rejection of the null hypothesis. The power is the proportion of simulated samples in step 2 that lead to rejection. The significance level is the proportion of simulated samples in step 3 that lead to rejection.

Generating Random Distributions

Two methods are available in **PASS** to simulate random samples. The first method generates the random variates directly, one value at a time. The second method generates a large pool (over 10,000) of random values and then draws the random numbers from this pool. This second method can cut the running time of the simulation by 70%!

As mentioned above, the second method begins by generating a large pool of random numbers from the specified distributions. Each of these pools is evaluated to determine if its mean is within a small relative tolerance (0.0001) of the target mean. If the actual mean is not within the tolerance of the target mean, individual members of the population are replaced with new random numbers if the new random number moves the mean towards its target. Only a few hundred such swaps are required to bring the actual mean to within tolerance of the target mean. This population is then sampled with replacement using the uniform distribution. We have found that this method works well as long as the size of the pool is the maximum of twice the number of simulated samples desired and 10,000.

Example 1 – Power at Various Sample Sizes

A study is being planned to find the threshold level of a certain drug. Below this threshold level, the response has little change. Once the threshold level is reached, there is a sizeable jump in the mean response rate. Little change in the response occurs as the drug level is increased above the threshold. Scientists believe that the threshold level is between 3 and 7—their best estimate, based on previous studies, is 5. Previous studies have shown that the standard deviation within a group is 3.0.

In order to find the threshold, they design a study with five levels: 3.0, 4.0, 5.0, 6.0, and 7.0. Since there is no trend in the mean value (only a sudden shift) as the dose level is increased, they decide to test the following hypotheses:

<u>Difference</u>	<u>Coefficients</u>
$\mu_2 - \mu_1$	-1, 1, 0, 0, 0
$\mu_3 - \mu_2$	0, -1, 1, 0, 0
$\mu_4 - \mu_3$	0, 0, -1, 1, 0
$\mu_5 - \mu_4$	0, 0, 0, -1, 1

Notice that this set of hypotheses answers the question directly. An overall F-test would test the hypothesis that at least one mean is different, but it would not indicate which is different. The question might be settled by considering all possible pairs, but there are ten pairs, so ten hypothesis tests would have to be considered instead of only four—decreasing the power.

Researchers want to detect a shift in the mean as small as 2.0. Hence, they want to study the power when the means are 0.0, 0.0, 2.0, 2.0, 2.0. They want to investigate sample sizes of 10, 30, 50, and 70 subjects per group.

They have no reason to assume that the variance will change a great deal from group to group, so they decide to analyze the data using the Dunn-Bonferroni procedure. They set the FWER to 0.05. Note that, based on these means, only the second of the four contrasts will be significant, so the any-contrast power will be the same as the all-contrast power.

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
MC Procedure.....	Dunn-Bonferroni Test
Simulations	2000
Random Seed.....	3698677 (for Reproducibility)
FWER (Alpha).....	0.05
n (Sample Size Multiplier).....	10 30 50 70
Group Sample Size Pattern	Equal
Set A Grps	2
Set A Group Distribution(s) H0	Normal(M0 S)
Set A Group Distribution(s) H1	Normal(M0 S)

Multiple Contrasts (Simulation)

Set B Grps **3**
 Set B Group Distribution(s)|H0 **Normal(M0 S)**
 Set B Group Distribution(s)|H1 **Normal(M1 S)**
 Equivalence Margin **0.1**
 M0 (Mean|H0) Parameter Value(s) **0**
 M1 (Mean|H1) Parameter Value(s) **2**
 Parameter 1 Label **S**
 Parameter 1 Value(s)..... **3**

Contrasts Tab

Contrasts **Each With Next**

Output

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

Simulation Summary Report

Summary of Simulations for Simultaneously Testing Multiple Contrasts of Groups

Solve For: **Power**
 MC Procedure: **Dunn-Bonferroni Test**
 Number of Groups: **5**
 Number of Contrasts: **4**

Scenario	Any-Contrast Power	Sample Size		All-Contrasts Power	Standard Deviation of Group Means Sm H1	Within-Group Standard Deviation SD H1	Family-Wise Error Rate		M0	M1	S
		Average Group n	Total N				Actual FWER	Target FWER			
1	0.153 (0.016) [0.137 0.169]	10	50	0.153 (0.016) [0.137 0.169]	1	3	0.045 (0.009) [0.036 0.054]	0.05	0	2	3
2	0.526 (0.022) [0.504 0.548]	30	150	0.526 (0.022) [0.504 0.548]	1	3	0.048 (0.009) [0.039 0.057]	0.05	0	2	3
3	0.794 (0.018) [0.776 0.811]	50	250	0.794 (0.018) [0.776 0.811]	1	3	0.041 (0.009) [0.032 0.05]	0.05	0	2	3
4	0.919 (0.012) [0.907 0.931]	70	350	0.919 (0.012) [0.907 0.931]	1	3	0.045 (0.009) [0.036 0.054]	0.05	0	2	3

Pool Size: 10000. Simulations: 2000. Run Time: 3.13 seconds.
 Equivalence Margin: 0.1. User-Entered Random Seed: 3698677

H0 The null hypothesis that the contrast of the means is zero.
 H1 The alternative hypothesis that the contrast of the means is not zero.
 Contrast A weighted average of the means in which the weights sum to zero.
 Any-Contrast Power The estimated probability of detecting at least one non-zero contrast. The second row provides the precision and a 95% confidence interval for Any-Contrast Power, (Any-Contrast Power Precision) [95% LCL and UCL], based on the size of the simulation.
 n The average of the group sample sizes.
 N The combined sample size of all groups.
 All-Contrasts Power The estimated probability of detecting all non-zero contrasts. The second row provides the precision and a 95% confidence interval for All-Contrasts Power, (All-Contrasts Power Precision) [95% LCL and UCL], based on the size of the simulation.
 Sm|H1 The standard deviation of the group means under H1.
 SD|H1 The pooled, within-group standard deviation under H1.

Multiple Contrasts (Simulation)

FWER	Family-Wise Error Rate. The probability of detecting at least one zero contrast assuming H0.
Actual FWER	The Family-Wise Error Rate estimated by the alpha simulation. The second row provides the precision and a 95% confidence interval for FWER, (FWER Precision) [95% LCL and UCL], based on the size of the simulation.
Target FWER	The user-specified Family-Wise Error Rate.
M0, M1, S, etc.	The value(s) of the user-specified distribution parameters.

Summary Statements

A one-way design with 5 groups will be used to test each contrast of the group means (as defined in the contrast coefficients section) against zero. The comparisons will be made using Dunn-Bonferroni contrast tests with a target overall (family-wise) Type I error rate (α) of 0.05. Based on 2000 simulations of the null distributions: N(M0 S); N(M0 S); N(M0 S); N(M0 S); and N(M0 S), and of the alternative distributions: N(M0 S); N(M0 S); N(M1 S); N(M1 S); and N(M1 S) (where M0 = 0, M1 = 2, and S = 3), with an average group sample size of 10 (for a total of 50 subjects), the any-contrast power (probability of detecting at least one of the contrasts) is 0.153, and the all-contrast power (probability of detecting all of the contrasts) is 0.153. (Additional details: The standard deviation of the group means under the null hypothesis is 0, and the standard deviation of the group means under the alternative hypothesis is 1. The average of the within-group standard deviations, assuming the alternative hypothesis distributions, is 3. The actual family-wise Type I error rate (α), based on the null hypothesis distribution simulations, is 0.045.)

Dropout-Inflated Sample Size

Average Group Sample Size n	Group	Dropout Rate	Sample Size Ni	Dropout-Inflated Enrollment Sample Size Ni'	Expected Number of Dropouts Di
10	1 - 5	20%	10	13	3
	Total		50	65	15
30	1 - 5	20%	30	38	8
	Total		150	190	40
50	1 - 5	20%	50	63	13
	Total		250	315	65
70	1 - 5	20%	70	88	18
	Total		350	440	90

n	The average group sample size.
Group	Lists the group numbers.
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.
Ni	The evaluable sample size for each group at which power is computed (as entered by the user). If Ni subjects are evaluated out of the Ni' subjects that are enrolled in the study, the design will achieve the stated power.
Ni'	The number of subjects that should be enrolled in each group in order to obtain Ni evaluable subjects, based on the assumed dropout rate. Ni' is calculated by inflating Ni using the formula $Ni' = Ni / (1 - DR)$, with Ni' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lohknygina, Y. (2018) pages 32-33.)
Di	The expected number of dropouts in each group. $Di = Ni' - Ni$.

Multiple Contrasts (Simulation)

Dropout Summary Statements

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

.

.

.

References

Devroye, Luc. 1986. Non-Uniform Random Variate Generation. Springer-Verlag. New York.

Dunn, O. J. 1964. 'Multiple comparisons using rank sums', *Technometrics*, Volume 6, pages 241-252.

Hsu, Jason. 1996. *Multiple Comparisons: Theory and Methods*. Chapman & Hall. London.

Kirk, Roger E. 1982. *Experimental Design: Procedures for the Behavioral Sciences*. Brooks/Cole. Pacific Grove, California.

Matsumoto, M. and Nishimura, T. 1998. 'Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator.' *ACM Trans. On Modeling and Computer Simulations*.

This report shows that a group sample size of about 50 will be needed to achieve 80% power or about 70 for 90% power.

Any-Contrast Power

This is the probability of detecting any of the significant contrasts. This value is estimated by the simulation using the H1 distributions.

Note that a precision value (half the width of its confidence interval) and a confidence interval are shown on the line below this row. These values provide the precision of the estimated power.

All-Contrast Power

This is the probability of detecting all of the significant contrasts. This value is estimated by the simulation using the H1 distributions.

Note that a precision value (half the width of its confidence interval) and a confidence interval are shown on the line below this row. These values provide the precision of the estimated power.

Group Sample Size n

This is the average of the individual group sample sizes.

Total Sample Size N

This is the total sample size of the study.

Standard Deviation of Group Means $S_m | H_1$

This is the standard deviation of the hypothesized means of the alternative distributions. Under the null hypothesis this value is zero. It represents the magnitude of the difference among the means. It is roughly equal to the average difference between the group means and the overall mean.

Note that the effect size is the ratio of $S_m | H_1$ and $SD | H_1$.

Multiple Contrasts (Simulation)

Within-Group Standard Deviation SD | H1

This is the within-group standard deviation calculated from samples from the alternative distributions.

Actual FWER

This is the value of FWER (family-wise error rate) estimated by the simulation using the H0 distributions. It should be compared with the Target FWER to determine if the test procedure is accurate.

Note that a precision value (half the width of its confidence interval) and a confidence interval are shown on the line below this row. These values provide the precision of the Actual FWER.

Target FWER

This is the target value of FWER that was set by the user.

M0

This is the value entered for M0, the group means under H0.

M1

This is the value entered for M1, the group means under H1.

S

This is the value entered for S, the standard deviation.

Error Rate Summary from H0 (Alpha) Simulations

Error Rate Summary from H0 (Alpha) Simulations

MC Procedure: Dunn-Bonferroni Test

Number of Groups: 5

Number of Contrasts: 4

Scenario	Total Sample Size	Type I Errors		Family-Wise Error Rate		Individual Contrast Alpha		
		Average Number	Proportion	Actual	Target	Average	Minimum	Maximum
1	50	0.051	0.013	0.045	0.05	0.013	0.007	0.015
2	150	0.053	0.013	0.048	0.05	0.013	0.011	0.016
3	250	0.046	0.011	0.041	0.05	0.011	0.011	0.013
4	350	0.050	0.012	0.045	0.05	0.012	0.010	0.014

- Contrast H0: A weighted average of the means in which the weights sum to zero.
- Contrast H1: The null hypothesis that a particular contrast of the means is zero.
- Type I Error: The alternative hypothesis that a particular contrast of the means is not zero.
- Family: Failing to reject the hypothesis that a contrast is zero when it is.
- Scenario: The set of all contrasts.
- Identification number that signifies a set of options. This is used across the various reports.
- Total Sample Size: The total number of subjects in all groups.
- Average Number of Type I Errors: The average number of Type I errors per family.
- Proportion of Type I Errors: The probability of falsely detecting at least one zero-valued contrast under H0.
- FWER: Family-wise error rate. The family consists of the set of all contrasts.
- Actual FWER: The computed value of the family-wise error rate.
- Target FWER: The specified value of the family-wise error rate.

Multiple Contrasts (Simulation)

Average Indiv. Cont. Alpha	The average of the individual contrast alphas.
Minimum Indiv. Cont. Alpha	The minimum of the individual contrast alphas.
Maximum Indiv. Cont. Alpha	The maximum of the individual contrast alphas.

This report shows the results of the H0 simulation. This simulation uses the H0 settings for each group. Its main purpose is to provide an estimate of the FWER.

Scenario

Identification number that signifies a set of options. This is used across the various reports.

Total Sample Size

Identification number that signifies a set of options. This is used across the various reports.

Average Number of Type I Errors

This is the average number of Type I errors (false detections) per set (family).

Proportion of Contrasts with Type I Errors

This is the proportion of Type I errors (false detections) among all tests that were conducted.

Actual FWER

This is the proportion of the H0 simulations in which at least one Type I error occurred. This is an abbreviation of family-wise error rate.

Target FWER

This is the target value of FWER that was input.

Average Individual Contrast Alpha

Alpha is the probability of rejecting H0 when H0 is true. It is a characteristic of an individual test. This is the average individual alpha value over all individual contrasts.

Minimum Individual Contrast Alpha

This is the minimum of the individual contrast alphas.

Maximum Individual Contrast Alpha

This is the maximum of the individual contrast alphas.

Multiple Contrasts (Simulation)

Error Rate Summary from H1 (Power) Simulation

Error Rate Summary from H1 (Power) Simulation

Multiple-Comparison Procedure: Dunn-Bonferroni Test
 Number of Groups: 5
 Number of Contrasts: 4

Scenario	Number of Contrasts		Average Number		Proportion of Contrasts with Errors	Proportion of Contrasts that Are				Power of Non-Zero Contrasts		Individual Contrast Power		
	Not Zero		False Pos.	False Neg.		Zero but Detected	Non-Zero but Undetected	Detected but Zero	Undetected but Non-Zero	All	Any	Ave	Min	Max
	Zero	Zero												
1	3	1	0.04	0.85	0.223	0.015	0.847	0.223	0.223	0.153	0.153	0.049	0.012	0.153
2	3	1	0.03	0.47	0.127	0.011	0.474	0.059	0.138	0.526	0.526	0.140	0.008	0.526
3	3	1	0.03	0.21	0.060	0.011	0.207	0.041	0.065	0.794	0.794	0.207	0.011	0.794
4	3	1	0.03	0.08	0.028	0.010	0.081	0.033	0.027	0.919	0.919	0.238	0.008	0.919

- Contrast H0: A weighted average of the means in which the weights sum to zero.
- Contrast H1: The null hypothesis that a particular contrast of the means is zero.
- Family: The alternative hypothesis that a particular contrast of the means is not zero.
- Scenario: The set of all contrasts in the study.
- Zero Contrasts: An identification number that signifies a set of options. This is used across the various reports.
- Non-Zero Contrasts: The number of contrasts that are zero under H1.
- False Pos.: The number of contrasts that are not zero under H1.
- False Neg.: Zero contrasts that were falsely concluded as being non-zero.
- Proportion Errors: Non-zero contrasts that were falsely not detected.
- Zero, Detected: The proportion of contrasts that exhibit either a Type I or Type II error.
- Non-Zero, Undetected: The proportion of contrasts that were zero but were concluded as being non-zero.
- Detected, Zero: The proportion of contrasts that were non-zero but were concluded as being zero.
- Undetected, Non-Zero: The proportion of contrasts that were concluded as non-zero but were actually zero.
- Power of All Contrasts: The proportion of contrasts that were concluded as zero but were actually non-zero.
- Power of Any Contrasts: This is the probability of detecting all non-zero contrasts in the H1 simulations.
- Average Indiv. Cont. Power: This is the probability of detecting any non-zero contrasts in the H1 simulations.
- Minimum Indiv. Cont. Power: The average of the individual contrast powers.
- Maximum Indiv. Cont. Power: The minimum of the individual contrast powers.
- Maximum Indiv. Cont. Power: The maximum of the individual contrast powers.

This report shows the results of the H1 simulation. This simulation uses the H1 settings for each group. Its main purpose is to provide an estimate of the power.

Number of Contrasts that are Zero

The number of contrasts that were zero under H1.

Number of Contrasts that are Not Zero

The number of contrasts that were non-zero under H1.

Average Number of False Positives

This is the average number of zero contrasts that were declared as being non-zero by the testing procedure. A *false positive* is a Type I (alpha) error.

Average Number of False Negatives

This is the average number of non-zero contrasts that were not declared as being non-zero by the testing procedure. A *false negative* is a type-2 (beta) error.

Proportion of Contrasts with Errors

This is the proportion of Type I and Type II errors.

Multiple Contrasts (Simulation)

Proportion of Contrasts that are Zero but Detected

This is the proportion of the zero contrasts in the H1 simulations that were declared as non-zero.

Proportion of Contrasts that are Non-Zero but Undetected

This is the proportion of non-zero contrasts in the H1 simulations that were not declared as being non-zero.

Proportion of Contrasts that are Detected but Zero (FDR)

This is the proportion of all detected contrasts in the H1 simulations that were actually zero. This is often called the *false discovery rate*.

Proportion of Contrasts that Are Undetected but Non-Zero

This is the proportion of undetected contrasts in the H1 simulations that were actually non-zero.

Power of Any Non-Zero Contrasts

This is the probability of detecting all non-zero contrasts in the H1 simulation.

Power of All Non-Zero Contrasts

This is the probability of detecting any non-zero contrasts in the H1 simulation.

Mean Individual Contrast Power

The average of the individual contrast powers from the H1 simulations.

Minimum Individual Contrast Power

The minimum of the individual contrast powers from the H1 simulations.

Maximum Individual Contrast Power

The maximum of the individual contrast powers from the H1 simulations.

Multiple Contrasts (Simulation)

Detailed Model Reports

Detailed Model Report for Scenario 1

Target FWER = 0.05, M0 = 0, M1 = 2, S = 3
 Multiple-Comparison Procedure: Dunn-Bonferroni Test
 Number of Groups: 5

Hypothesis Type	Groups	Group Labels	n	N	Group Mean	Standard Deviation of Group Means Sm	Average Group Standard Deviation Ave(SD)	Simulation Model
H0	1-2	A1-A2	10	50	0		3	N(M0 S)
H0	3-5	B1-B3	10	50	0		3	N(M0 S)
H0	All					0	3	
H1	1-2	A1-A2	10	50	0		3	N(M0 S)
H1	3-5	B1-B3	10	50	2		3	N(M1 S)
H1	All					1	3	

Hypothesis Type H0 (null) for the alpha simulation and H1 (alternative) for the power simulation.
 Groups The groups reported on this line. "All" is used for the average of all groups.
 Group Labels The labels that are used in the individual alpha-level reports.
 n n is the average group size.
 N N is the combined sample size.
 Sm The standard deviation of the means. This value is tested with the F-test.
 Ave(SD) The average within-group standard deviation of all groups reported on this line.
 Simulation Model The distribution used to simulate data for the groups reported on this line.

(More Reports Follow)

This report shows details of each row of the previous reports.

Hypothesis Type

This indicates which simulation is being reported on each row. H0 represents the null simulation and H1 represents the alternative simulation.

Groups

Each group in the simulation is assigned a number. This item shows the arbitrary group number that was assigned.

Group Labels

These are the labels that were used in the individual alpha-level reports.

n

n is the average sample size of the groups.

N

N is the total sample size across all groups.

Multiple Contrasts (Simulation)

Group Mean

These are the means of the individual groups as specified for the H0 and H1 simulations.

SM

The standard deviation of the means in this set of simulations.

Ave(SD).

This is the average within-group standard deviation of those groups reported on each line. Note that it is calculated from the simulated data.

Simulation Model

This is the distribution that was used to simulate data for the groups reported on each line.

List of Contrast Coefficients**List of Contrast Coefficients**

Contrasts	Groups				
	A1	A2	B1	B2	B3
Con1	-1	1	0	0	0
Con2	0	-1	1	0	0
Con3	0	0	-1	1	0
Con4	0	0	0	-1	1

The contrasts are shown down the rows. The groups are shown across the columns.
The coefficients (weights) are shown as the body of the table.

This report shows values of the contrast coefficients so you can double-check that they are what was intended.

Probability of Rejecting Individual Contrasts**Probability of Rejecting Individual Contrasts. Simulation No. 1**

Contrasts	Alpha	Power
Con1	0.015	0.012
Con2	0.007	0.153
Con3	0.015	0.019
Con4	0.014	0.014

Alpha: probability of rejecting hypothesis that contrast is zero under alpha (H0) simulation.
Power: probability of rejecting hypothesis that contrast is zero under power (H1) simulation.

(More Reports Follow)

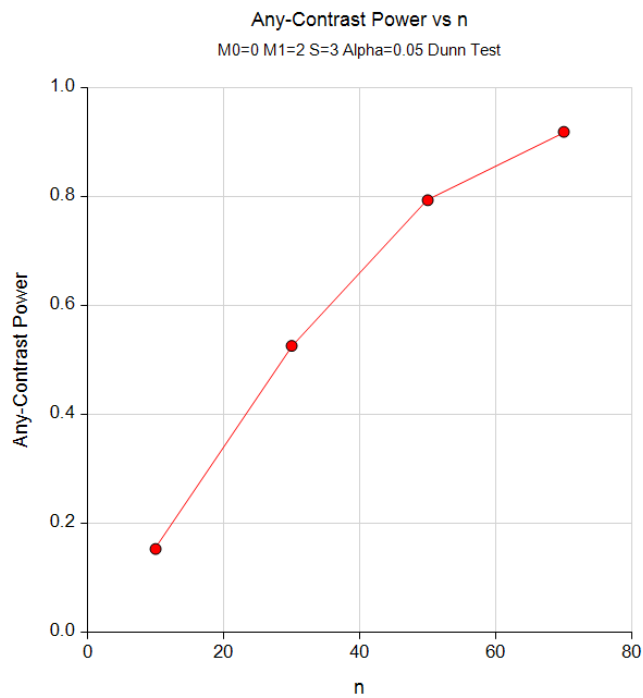
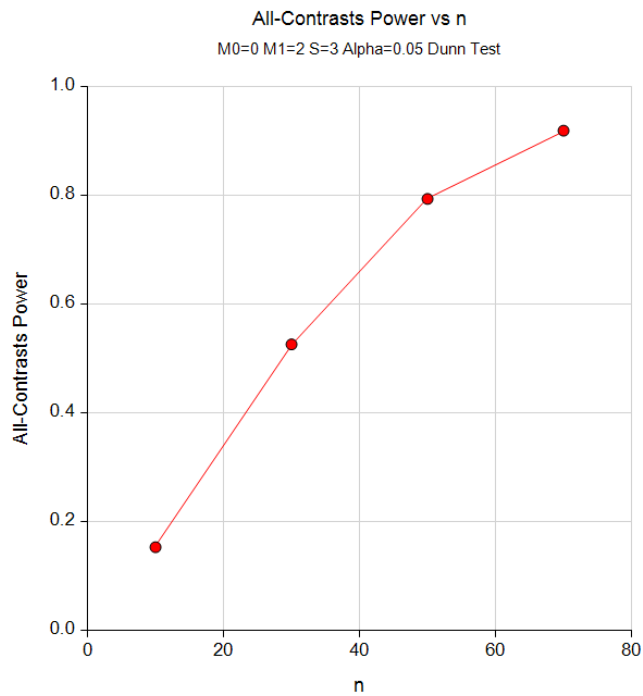
This report shows alpha and individual power for each contrast for each simulation that was run.

In this example, only the second contrast was non-zero, so that is the only one which has large values for the power.

Multiple Contrasts (Simulation)

Plots Section

Plots



These plots give a visual presentation of the all-contrasts power values and the any-contrast power values.

Example 2 – Comparative Results

Continuing with Example 1, the researchers want to study the characteristics of alternative multiple contrast procedures.

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**
 MC Procedure..... **Dunn-Bonferroni Test**
 Simulations **2000**
 Random Seed..... **3764272** (for Reproducibility)
 FWER (Alpha)..... **0.05**
 n (Sample Size Multiplier) **10 30 50 70**
 Group Sample Size Pattern **Equal**
 Set A Grps **2**
 Set A Group Distribution(s)|H0 **Normal(M0 S)**
 Set A Group Distribution(s)|H1 **Normal(M0 S)**
 Set B Grps **3**
 Set B Group Distribution(s)|H0 **Normal(M0 S)**
 Set B Group Distribution(s)|H1 **Normal(M1 S)**
 Set C Grps **0**
 Equivalence Margin **0.1**
 M0 (Mean|H0) Parameter Value(s) **0**
 M1 (Mean|H1) Parameter Value(s) **2**
 Parameter 1 Label **S**
 Parameter 1 Value(s)..... **3**

Contrasts Tab

Contrasts **Each With Next**

Reports Tab

Show Comparative Reports **Checked**

Comparative Plots Tab

Comparative All-Contrasts Power Plot..... **Checked**

Comparative Any-Contrast Power Plot **Checked**

Multiple Contrasts (Simulation)

Output

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

Power Comparison for Simultaneously Testing Multiple Contrasts

Number of Groups: 5
Number of Contrasts: 4

Scenario	Total Sample Size	Target Alpha	Power			
			All-Contrasts		Any-Contrast	
			Dunn-Bonferroni	Dunn-Welch	Dunn-Bonferroni	Dunn-Welch
1	50	0.05	0.153	0.113	0.153	0.113
2	150	0.05	0.524	0.504	0.524	0.504
3	250	0.05	0.785	0.777	0.785	0.777
4	350	0.05	0.928	0.923	0.928	0.923

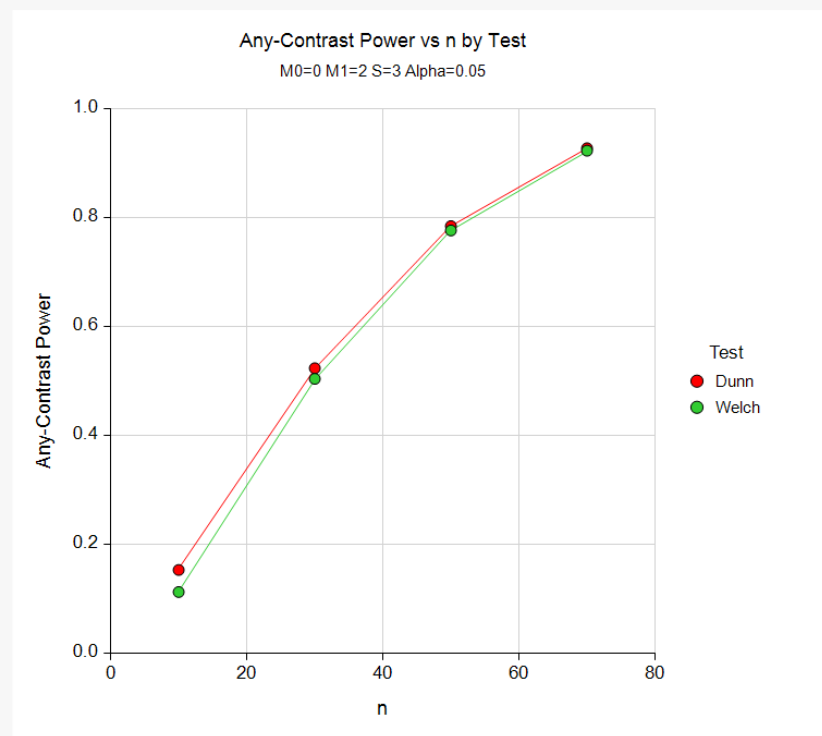
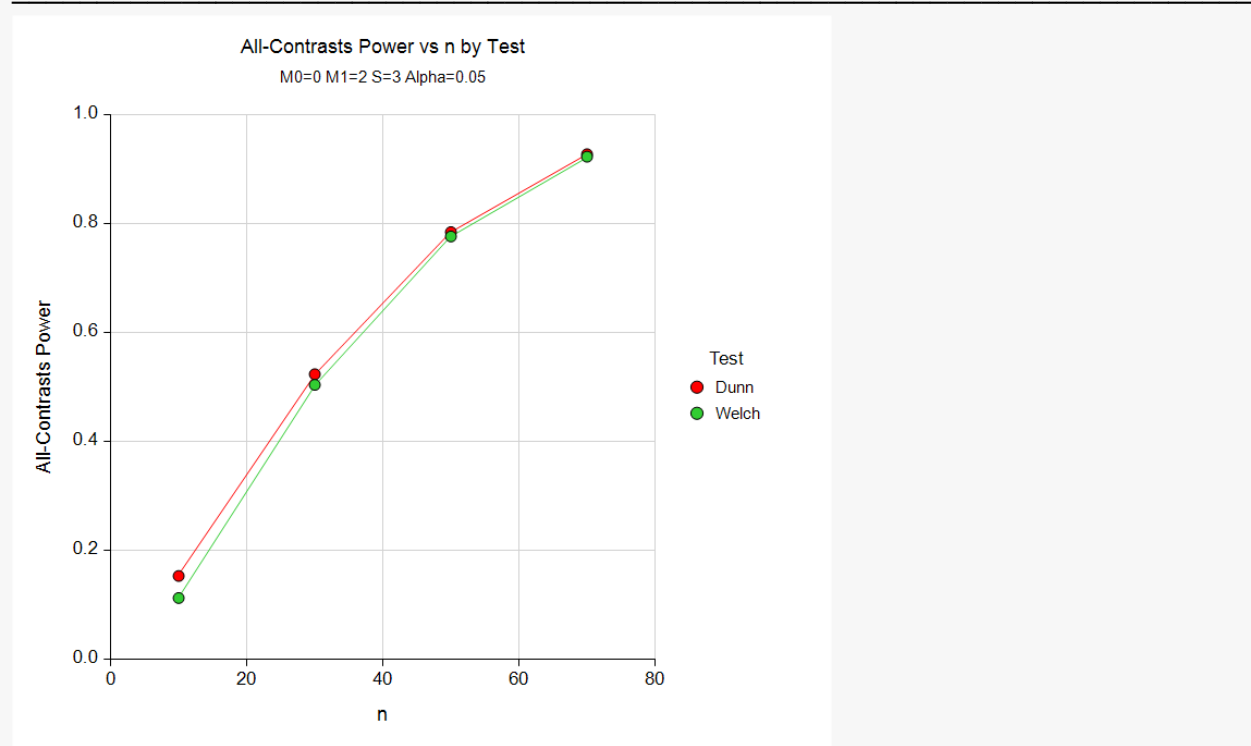
Family-Wise Error Rate Comparison for Simultaneously Testing Multiple Contrasts

Number of Groups: 5
Number of Contrasts: 4

Scenario	Total Sample Size	Family-Wise Error Rate		
		Target	Dunn-Bonferroni	Dunn-Welch
1	50	0.05	0.047	0.038
2	150	0.05	0.047	0.045
3	250	0.05	0.039	0.042
4	350	0.05	0.047	0.047

Multiple Contrasts (Simulation)

Plots



These reports show the power and FWER of both multiple contrast procedures. In these simulations of groups from the normal distributions with equal variances, there is little difference in the power of the two procedures.

Example 3 – Validation

We could not find an article that gives power values for this test, so we decided to validate the procedure by comparing its results to those of the one-way ANOVA procedure which allows a single contrast to be tested. Using the settings of Example 1 and using the contrast '0, -1, 1, 0, 0', we obtained the following powers: 0.3085, 0.7274, 0.9131, and 0.9758.

For reproducibility, we'll use a random seed of 5226632.

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 3** 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
MC Procedure.....	Dunn-Bonferroni Test
Simulations	2000
Random Seed.....	5226632 (for Reproducibility)
FWER (Alpha).....	0.05
n (Sample Size Multiplier).....	10 30 50 70
Group Sample Size Pattern	Equal
Set A Grps	2
Set A Group Distribution(s) H0	N(M0 S)
Set A Group Distribution(s) H1	N(M0 S)
Set B Grps	3
Set B Group Distribution(s) H0	N(M0 S)
Set B Group Distribution(s) H1	N(M1 S)
Set C Grps	0
Equivalence Margin	0.1
M0 (Mean H0) Parameter Value(s)	0
M1 (Mean H1) Parameter Value(s)	2
Parameter 1 Label	S
Parameter 1 Value(s).....	3

Contrasts Tab

Contrasts	0 -1 1 0 0
-----------------	-------------------

Multiple Contrasts (Simulation)

Output

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

Summary of Simulations for Simultaneously Testing Multiple Contrasts of Groups

Solve For: Power
 MC Procedure: Dunn-Bonferroni Test
 Number of Groups: 5
 Number of Contrasts: 1

Scenario	Any-Contrast Power	Sample Size		All-Contrasts Power	Standard Deviation of Group Means Sm H1	Within-Group Standard Deviation SD H1	Family-Wise Error Rate		M0	M1	S
		Average Group n	Total N				Actual FWER	Target FWER			
1	0.315 (0.02) [0.294 0.335]	10	50	0.315 (0.02) [0.294 0.335]	1	3	0.053 (0.01) [0.043 0.063]	0.05	0	2	3
2	0.730 (0.019) [0.711 0.749]	30	150	0.730 (0.019) [0.711 0.749]	1	3	0.056 (0.01) [0.046 0.066]	0.05	0	2	3
3	0.907 (0.013) [0.894 0.92]	50	250	0.907 (0.013) [0.894 0.92]	1	3	0.044 (0.009) [0.035 0.053]	0.05	0	2	3
4	0.975 (0.007) [0.968 0.982]	70	350	0.975 (0.007) [0.968 0.982]	1	3	0.053 (0.01) [0.043 0.063]	0.05	0	2	3

Pool Size: 10000. Simulations: 2000. Run Time: 3.45 seconds.
 Equivalence Margin: 0.1. User-Entered Random Seed: 5226632

In each case, the result matches the analytic value reasonably well.