Chapter 513

Superiority by a Margin Tests for the Ratio of Two Means in a 2x2 Cross-Over Design

Introduction

This procedure calculates power and sample size of statistical tests for non-unity null tests from a 2x2 cross-over design. This routine deals with the case in which the statistical hypotheses are expressed in terms mean ratios rather than mean differences.

The details of testing the non-unity null of two treatments using data from a 2x2 cross-over design are given in another chapter and they will not be repeated here. If the logarithms of the responses can be assumed to follow the normal distribution, hypotheses about non-unity null hypotheses stated in terms of the ratio can be transformed into hypotheses about the difference. The details of this analysis are given in Julious (2004). They will only be summarized here.

Non-Inferiority Testing Using Ratios

It will be convenient to adopt the following specialized notation for the discussion of these tests.

Parameter	PASS Input/Output	Interpretation
μ_T	Not used	*Treatment mean*. This is the treatment mean.
μ_R	Not used	*Reference mean*. This is the mean of a reference population.
M_s	SM	*Margin of superiority*. This is a tolerance value that defines the magnitude of difference that is required for practical importance. This may be thought of as the smallest difference from the reference that is considered to be different.
ϕ	R1	*True ratio*. This is the value of $\phi = \mu_T / \mu_R$ at which the power is calculated.

Note that the actual values of μ_T and μ_R are not needed. Only their ratio is needed for power and sample size calculations.

The null hypothesis of non-superiority is

$$\mathrm{H}_0: \phi \le \phi_L \text{ where } \phi_L > 1.$$

and the alternative hypothesis of superiority is

$$\mathrm{H}_1: \phi > \phi_L$$

Log Transformation

In many cases, hypotheses stated in terms of ratios are more convenient than hypotheses stated in terms of differences. This is because ratios can be interpreted as scale-less percentages, but differences must be interpreted as actual amounts in their original scale. Hence, it has become a common practice to take the following steps in hypothesis testing.

1. State the statistical hypotheses in terms of ratios.
2. Transform these into hypotheses about differences by taking logarithms.
3. Analyze the logged data—that is, do the analysis in terms of the difference.
4. Draw the conclusion in terms of the ratio.

The details of step 2 for the null hypothesis are as follows.

$$\phi_L \le \phi$$

$$\Rightarrow \phi_L \le \left\{\frac{\mu_T}{\mu_R}\right\}$$

$$\Rightarrow \ln(\phi_L) \le \left\{\ln(\mu_T) - \ln(\mu_R)\right\}$$

Thus, a hypothesis about the ratio of the means on the original scale can be translated into a hypothesis about the difference of two means on the logged scale.

Coefficient of Variation

The coefficient of variation (COV) is the ratio of the standard deviation to the mean. This parameter is used to represent the variation in the data because of a unique relationship that it has in the case of log-normal data.

Suppose the variable X is the logarithm of the original variable Y. That is, $X = \ln(Y)$ and $Y = \exp(X)$. Label the mean and variance of X as μ_X and σ_X^2, respectively. Similarly, label the mean and variance of Y as μ_Y and σ_Y^2, respectively. If X is normally distributed, then Y is log-normally distributed. Julious (2004) presents the following well-known relationships between these two variables

$$\mu_Y = \left(e^{\mu_X + \frac{\sigma_X^2}{2}}\right)$$

$$\sigma_Y^2 = \mu_Y^2\left(e^{\sigma_X^2} - 1\right)$$

From this relationship, the coefficient of variation of *Y* can be expressed as

$$
\begin{aligned}
COV_Y &= \frac{\sqrt{\mu_Y^2\left(e^{\sigma_X^2}-1\right)}}{\mu_Y} \\
&= \sqrt{e^{\sigma_X^2}-1} \\
&= \sqrt{e^{\sigma_w^2}-1}
\end{aligned}
$$

where σ_w^2 is the within mean square error from the analysis of variance of the logged data. Solving this relationship for σ_X^2, the standard deviation of *X* can be stated in terms of the coefficient of variation of *Y*. This equation is

$$\sigma_X = \sqrt{\ln\left(COV_Y^2+1\right)}$$

Similarly, the mean of *X* is

$$\mu_X = \frac{\mu_Y}{\ln\left(COV_Y^2+1\right)}$$

Thus, the hypotheses can be stated in the original (Y) scale and then power can be analyzed in the transformed (X) scale.

Power Calculation

As is shown above, the hypotheses can be stated in the original (Y) scale using ratios or the logged (X) scale using differences. Either way, the power and sample size calculations are made using the formulas for testing the equivalence of the difference in two means. These formulas are presented in another chapter and are not duplicated here.

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 the parameters associated with this test such as the means, sample sizes, alpha, and power.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. Under most situations, you will select either *Power* for a power analysis or *Sample Size* for sample size determination.

Test

Higher Means Are

This option defines whether higher values of the response variable are to be considered better or worse.

The choice here determines the direction of the test.

If Higher Means Are Better the null hypothesis is R ≤ 1+SM and the alternative hypothesis is R > 1+SM. If Higher Means Are Worse the null hypothesis is R ≥ 1-SM and the alternative hypothesis is R < 1-SM.

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 non-superiority when in fact the treatment mean is superior.

Values must be between zero and one. Historically, the value of 0.20 was often used for beta. Recently, the standard has shifted to 0.10.

Power is defined as one minus beta. Power is equal to the probability of rejecting a false null hypothesis. Hence, specifying the beta error level also specifies the power level. For example, if you specify beta values of 0.05, 0.10, and 0.20, you are specifying the corresponding power values of 0.95, 0.90, and 0.80, respectively.

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. In this procedure, a type-I error occurs when rejecting the null hypothesis of non-superiority when in fact the treatment group is not superior to the reference group.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Sample Size

N (Total Sample Size)

This option specifies one or more values of the sample size, the number of individuals in the study (total subjects in both sequences). This value must be an integer greater than one.

When N is even, it is split evenly between the two sequences. When N is odd, the first sequence has one more subject than the second sequence.

Note that you may enter a list of values using the syntax *50,100,150,200,250* or *50 to 250 by 50*.

Effect Size – Ratios

SM (Superiority Margin)

This is the magnitude of the margin of superiority. It must be entered as a positive number.

When higher means are better, this value is the distance above one that is required for the mean ratio (Treatment Mean / Reference Mean) to be considered superior. When higher means are worse, this value is the distance below one that is required for the mean ratio (Treatment Mean / Reference Mean) to be considered superior.

R1 (True Ratio)

This is the value of the ratio of the two means (Treatment Mean / Reference Mean) at which the power is to be calculated.

When higher means are better, this value should be greater than 1+SM. When higher means are worse, this value should be less than 1-SM.

Effect Size – Coefficient of Variation

COV (Coefficient of Variation)

The coefficient of variation is the ratio of the standard deviation and the mean (SD/Mean). It is used to specify the variability (standard deviation). Note that this COV is defined on the original (not logarithmic) scale. This value must be determined from past experience or from a pilot study.

To be clear, consider the following definition. Suppose data on a response variable Y are collected. This procedure assumes that the values of X = Ln(Y) are analyzed using an appropriate ANOVA procedure. Thus, there are two sets of means and standard deviations: those of X labelled MX and SX and those of Y labelled MY and SY. The COV entered here is the COV of Y = SY/MY. For log-normal data, the following relationship exists: COV(Y) = SQR(Exp(SX*SX)-1) where SX is the square root of the within mean square error in the ANOVA table of the log-transformed values.

Example 1 – Finding Power

A company has developed a generic drug for treating rheumatism and wants to show that it is superior to the standard drug by a certain amount. A 2x2 cross-over design will be used to test the superiority of the treatment drug to the reference drug.

Researchers have decided to set the margin of superiority to 0.20. Past experience leads the researchers to set the COV to 1.50. The significance level is 0.05. The power will be computed assuming that the true ratio is 1.40. Sample sizes between 50 and 550 will be included in the analysis.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Superiority by a Margin Tests for the Ratio of Two Means in a 2x2 Cross-Over Design** procedure window by expanding **Means**, then **Cross-Over (2x2) Design**, then clicking on **Superiority by a Margin**, and then clicking on **Superiority by a Margin Tests for the Ratio of Two Means in a 2x2 Cross-Over Design**. 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**
Higher Means Are	**Better**
Alpha	**0.05**
N (Total Sample Size)	**50 to 550 by 100**
SM (Superiority Margin)	**0.20**
R1 (True Ratio)	**1.40**
COV (Coefficient of Variation)	**1.50**

Annotated Output

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

Numeric Results

Numeric Results for Superiority Ratio Test (H0: R ≤ 1 + SM; H1: R > 1 + SM)
Higher Means are Better

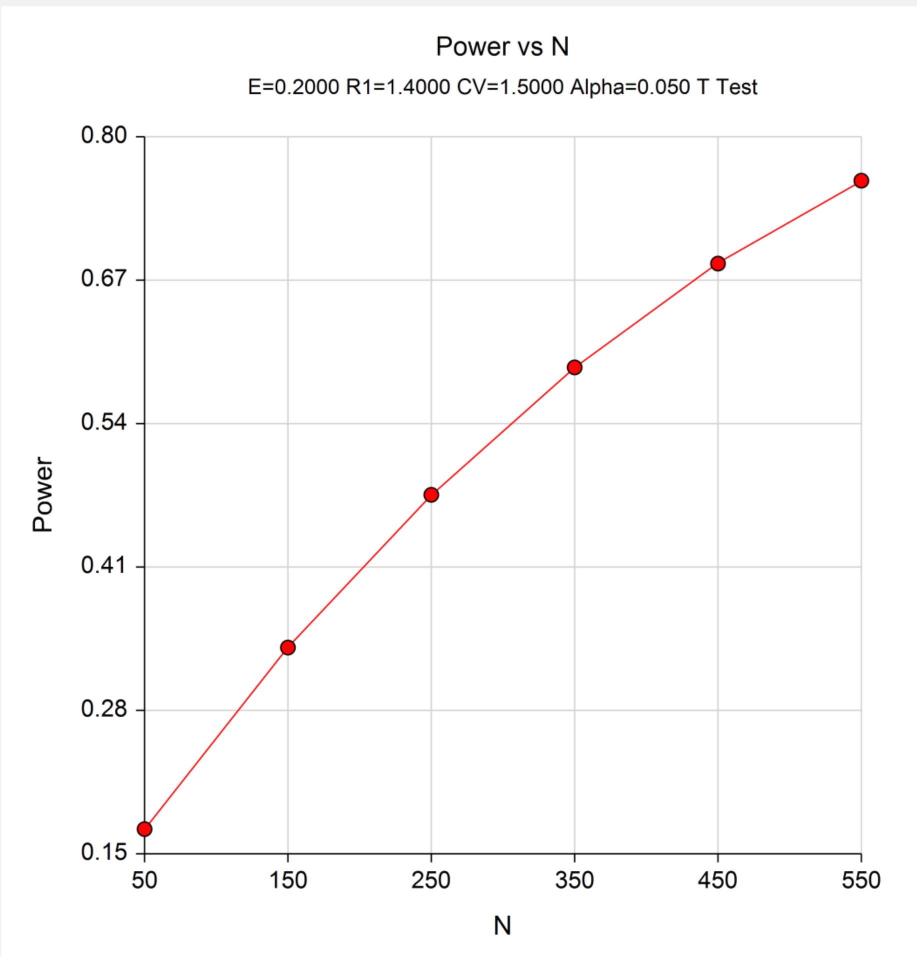
Power	N	Superiority Margin (SM)	Superiority Bound (SB)	Actual Ratio (R1)	Significance Level (Alpha)	Beta	COV
0.1724	50	0.2000	1.2000	1.4000	0.0500	0.8276	1.5000
0.3369	150	0.2000	1.2000	1.4000	0.0500	0.6631	1.5000
0.4754	250	0.2000	1.2000	1.4000	0.0500	0.5246	1.5000
0.5909	350	0.2000	1.2000	1.4000	0.0500	0.4091	1.5000
0.6850	450	0.2000	1.2000	1.4000	0.0500	0.3150	1.5000
0.7602	550	0.2000	1.2000	1.4000	0.0500	0.2398	1.5000

Report Definitions
H0 (null hypothesis) is R ≤ 1 + SM, where R = Treatment Mean / Reference Mean.
H1 (alternative hypothesis) is R > 1 + SM.
Power is the probability of rejecting H0 when it is false.
N is the total sample size drawn from all sequences. The sample is divided equally among sequences.
SM is the magnitude of the margin of superiority. Since higher means are better, this value is positive and is the distance above one that is required to be considered superior.
SB is the corresponding bound to the superiority margin, and equals 1 + SM.
R1 is the mean ratio (treatment/reference) at which the power is computed.
Alpha is the probability of falsely rejecting H0.
Beta is the probability of not rejecting H0 when it is false.
COV is the coefficient of variation on the original scale.

Summary Statements
A total sample size of 50 achieves 17% power to detect superiority using a one-sided t-test when the margin of superiority is 0.2000, the true mean ratio is 1.4000, the significance level is 0.0500, and the coefficient of variation on the original, unlogged scale is 1.5000. A 2x2 cross-over design with an equal number in each sequence is used.

This report shows the power for the indicated scenarios. Note that if they want 80% power, they will require a sample of more than 450 subjects.

Plot Section



This plot shows the power versus the sample size.

Example 2 – Validation

This procedure uses the same mechanics as the Non-Inferiority Tests for the Ratio of Two Means in a 2x2 Cross-Over Design procedure. We refer the user to Example 2 of Chapter 515 for the validation.