Chapter 526

Tests for the Ratio of Two Means in a Higher-Order Cross-Over Design (Log-Normal Data)

Introduction

This procedure calculates power and sample size for inequality tests which use the ratio of the means of a higher-order cross-over design. Measurements are made on individuals that have been randomly assigned to one of several treatment sequences. Only a brief introduction to the subject will be given here. For a comprehensive discussion on the subject, refer to Chen et al. (1997) and Chow et al. (2003).

Cross-Over Designs

Senn (2002) defines a *cross-over* design as one in which each subject receives all treatments at least once and the object is to study differences among the treatments. The name *cross-over* comes from the most common case in which there are only two treatments. In this case, each subject *crosses over* from one treatment to another. It is assumed that there is a *washout* period between treatments during which the response returns to its baseline value. If this does not occur, there is said to be a *carryover* effect.

A 2x2 cross-over design refers to two periods (and treatments) and two *sequences* (treatment orderings). One sequence of treatments is treatment A followed by treatment B. The other sequence is B and then A. The design includes a washout period between responses to make certain that the effects of the first drug do no carryover to the second. Thus, the groups of subjects in this design are defined by the sequence in which the two treatments are administered, not by the treatments they receive.

Higher-Order Cross-Over Designs

Chen et al. (1997) present sample size results for four cross-over designs that are more complicated than the 2x2 design. Assume that the two treatments are labeled A and B. The available designs are defined by the order and number of times the two treatments are administered.

Balaam's Design

Balaam's design has four sequences with two treatments each. It is popular because it allows the intrasubject variabilities to be estimated. The design is

Sequence	Period 1	Period 2
1	A	A
2	B	B
3	A	B
4	B	A

Two-Sequence Dual Design

This design has two sequences with three periods each. It is popular because it allows the intrasubject variabilities to be estimated. The design is

Sequence	Period 1	Period 2	Period 3
1	A	B	B
2	B	A	A

Four-Period Design with Two Sequences

This design has two sequences of four periods each. The design is

Sequence	Period 1	Period 2	Period 3	Period 4
1	A	B	B	A
2	B	A	A	B

Four-Period Design with Four Sequences

This design has four sequences of four periods each. The design is

Sequence	Period 1	Period 2	Period 3	Period 4
1	A	A	B	B
2	B	B	A	A
1	A	B	B	A
2	B	A	A	B

Advantages of Cross-Over Designs

A comparison of treatments on the same subject is expected to be more precise. The increased precision often translates into a smaller sample size. Also, patient enrollment may be easier to obtain because each patient will receive both treatments.

Disadvantages of Cross-Over Designs

The statistical analysis of a cross-over experiment is more complex than a parallel-group experiment and requires additional assumptions. In a cross-over experiment, it may be difficult to separate the treatment effect from the time effect and the carry-over effect of the previous treatment.

These cross-over designs cannot be used when the treatment (or the measurement of the response) alters the subject permanently. Hence, it cannot be used to compare treatments that are intended to provide a cure.

Because subjects must be measured at least twice, it may be more difficult to keep patients enrolled in the study. This is particularly true when the measurement process is painful, uncomfortable, embarrassing, or time consuming.

The Statistical Hypotheses

A two-sided hypothesis test comparing the means is given by

$$H0: \mu_1/\mu_2 = 1 \text{ versus } H1: \mu_1/\mu_2 \neq 1$$

Rejecting H0 implies that the means are different.

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. Also, taking the logarithm of the data can make skewed data closer to normal. 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.

In step 2, the statistical hypotheses become

$$H0: \ln(\mu_1) - \ln(\mu_2) = 0 \text{ versus } H1: \ln(\mu_1) - \ln(\mu_2) \neq 0$$

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

$$\mu_Y = e^{\mu_X + \frac{\sigma_X^2}{2}}$$

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

From this relationship, the coefficient of variation of Y can be found to be

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

$$= \sqrt{e^{\sigma_X^2} - 1}$$

$$= \sqrt{e^{\sigma_W^2} - 1}$$

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 as

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

Similarly, the mean of X is

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

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

Test Statistic

The presence of a non-zero difference between the means can be tested using an F-test from an appropriate repeated measures model that may include terms for a sequence effect and/or a carry-over effect. The within mean square error from the ANOVA of the logged data provides an estimate of the within-subject variance.

Power Calculation

The power of a two-sided test is given by Chen, Chow, and Li (1997). The basic formula is

$$\text{Power} = T_V\left[\left(\frac{|\ln(\mu_1) - \ln(\mu_2)|}{\sigma_X\sqrt{b/n}}\right) - t_{V,1-\alpha/2}\right]$$

where T represents the cumulative t distribution, V and b depend on the design (shown below), σ_X is defined above, and n is the average number of subjects per sequence. Thus the overall sample size N is given by n times the number of sequences. For a one-sided test, replace α/2 with α in the above formula.

The constants V and b depend on the design as follows:

Design Type	**Parameters (V, b)**
Balaam's Design	$V = 4n - 3$, $b = 2$.
Two-Sequence Dual Design	$V = 4n - 4$, $b = 3/4$.
Four-Period Design with Two Sequences	$V = 6n - 5$, $b = 11/20$.
Four-Period Design with Four Sequences	$V = 12n - 5$, $b = 1/4$.

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.

Select *Sample Size* when you want to calculate the sample size needed to achieve a given power and alpha level. Note that there are two choices for finding *Sample Size*. Select *Sample Size (Equal Per Sequence)* when you want the design to have an equal number of subjects per sequence. Select *Sample Size (Exact)* when you want to find the exact sample size even though the number of subjects cannot be dividing equally among the sequences.

Select *Power* when you want to calculate the power of an experiment.

Test

Design Type

Specify which cross-over design that you want to analyze. These designs allow you to compare two means: a treatment (A) and a reference (B). The designs assume that each subject is measured two or more times (or periods) with ample time in between to wash-out the effect of the treatment.

A sequence is an ordering of the how A and B are applied. For example, 'ABA' means that the subjects are measured three times: first treatment A, then B, then A again. Other possible sequences are BAB, AAB, and BBA. Each design includes several sequences.

The design tries to balance the experiment so that each treatment occurs an equal number of times and after each other treatment an equal number of times.

Alternative Hypothesis

Specify whether your test will be two-sided or one-sided.

Two-Sided

Note that usually an inequality test is two-sided. The value of alpha will be divided by 2 and the result will be put evenly into both tails.

One-Sided

The power and sample size will be determined for an upper-, or a lower-, tailed test.

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.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

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.

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 is the combined sample size of all sequences in the cross-over design. It is assumed that the subjects are divided as evenly as possible among the sequences.

You can enter a single value such as *20* or a series of values such as *10 20 30* or *10 to 100 by 10*.

Effect Size – Mean Ratio

R1 (μT / μR)

This is the value of the ratio of the two means (Treatment Mean / Reference Mean) under the alternative hypothesis, Ha. It is the value at which the power is to be calculated.

Positive values greater than zero are possible, except for the value one. Usually, the range will be between 0.25 and 4.0.

If the two means were equal, the ratio would be 1.0. This is the value of R1 under the null hypothesis, H0. So 1.0 should not be used for R1.

Values of R1 < 1 imply that $\mu T < \mu R$.

Values of R1 > 1 imply that $\mu T > \mu R$.

You can enter a single value or a range of values.

Effect Size – Coefficient of Variation

COV (Coefficient of Variation)

A response variable Y has mean μ_Y and standard deviation σ_Y. The coefficient of variation of Y, σ_Y/μ_Y, is used to specify the variation in Y.

Suppose that Y is skewed, so a logarithmic transformation will be applied. That is, the variable $X = Ln(Y)$ will be analyzed using an appropriate ANOVA procedure. The analysis assumes that X is normally distributed with mean μ_X and standard deviation σ_X. So Y is lognormal and X is normal. Power and sample size calculations are made on X. However, interpretation may be easier in the Y scale.

If Y is truly lognormal, there is a direct relationship between μ_Y and σ_Y and μ_X and σ_X. Using this relationship, it can be shown that

$$COV_Y = \sqrt{Exp(\sigma_X^2) - 1}$$

Solving for σ_X, we obtain

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

There are two ways to obtain a value for COV_Y.

1. Calculate COV_Y directly from the mean and SD of a set of Y values obtained from a previous (or pilot) study.

2. Obtain σw_X as the square root of the within mean square error in an ANOVA table of X. Then use the fact that $\sigma_X = \sigma w_X$.

We have found that method 1 **does not** provide a reliable estimate of σ_X, **so we recommend method 2**.

Example 1 – Finding Power

Researchers want to calculate the power of a two-sided, inequality test using data from a three-period, two-sequence, dual cross-over design. They want to investigate sample sizes 5, 10, 15, 20, 35, and 50. They want a sample large enough to detect a mean ratio of 1.25 or 1.5. The significance level is 0.05. Past experience leads the researchers to set the COV to 0.40.

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	**Power**
Design Type...	**3x2 (Three-Period, Two-Sequence Dual: ABB|BAA)**
Alternative Hypothesis............................	**Two-Sided**
Alpha..	**0.05**
N (Total Sample Size).............................	**5 10 15 20 35 50**
R1 (µT / µR)	**1.25 1.5**
COV (Coefficient of Variation)	**0.4**

Annotated Output

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

Numeric Results

Numeric Results for an Inequality Test of the Mean Ratio in a Three-Period, Two-Sequence Dual Design
Treatment Sequences: ABB | BAA
Alternative Hypothesis: Two-Sided

Power	Total Sample Size N	Mean Ratio R1 (µT / µR)	Coefficient of Variation COV	Alpha
0.1070	5	1.250	0.4000	0.050
0.2706	10	1.250	0.4000	0.050
0.4123	15	1.250	0.4000	0.050
0.5344	20	1.250	0.4000	0.050
0.7871	35	1.250	0.4000	0.050
0.9114	50	1.250	0.4000	0.050
0.3091	5	1.500	0.4000	0.050
0.7207	10	1.500	0.4000	0.050
0.8928	15	1.500	0.4000	0.050
0.9611	20	1.500	0.4000	0.050
0.9985	35	1.500	0.4000	0.050
1.0000	50	1.500	0.4000	0.050

References
Chow, S.C. and Liu, J.P. 1999. Design and Analysis of Bioavailability and Bioequivalence Studies. Marcel Dekker. New York.
Chow, S.C.; Shao, J.; Wang, H. 2003. Sample Size Calculations in Clinical Research. Marcel Dekker. New York.
Chen, K.W.; Chow, S.C.; and Li, G. 1997. 'A Note on Sample Size Determination for Bioequivalence Studies with Higher-Order Crossover Designs.' Journal of Pharmacokinetics and Biopharmaceutics, Volume 25, No. 6, pages 753-765.

Report Definitions
Power is the probability of rejecting H0 when H0 is false.
N is the total number of subjects. They are divided evenly among all sequences.
µT is the treatment mean. It is usually associated with the letter 'A' in the design.
µR is the reference mean. It is usually associated with the letter 'B' in the design.
R1 is the ratio of the means at which the power is computed.
COV is the coefficient of variation on the original scale.
Alpha is the probability of falsely rejecting H0 (falsely concluding inequality).

Summary Statements
In an inequality test on data drawn from a three-period, two-sequence dual cross-over design, a
total sample size of 5 achieves 11% power at a 5% significance level when the true ratio of the
means is 1.250 and the coefficient of variation is 0.4000.

This report shows the power for the indicated scenarios.

Plots Section
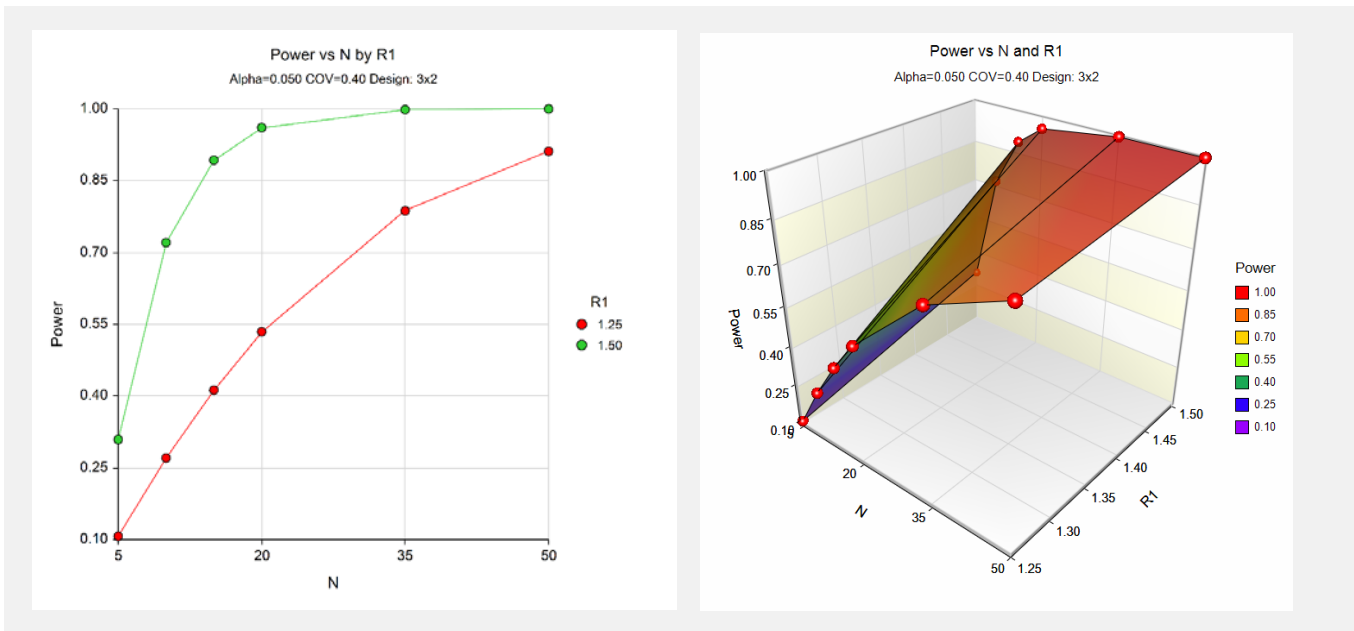
These plots show the power versus the sample size.

Example 2 – Finding Sample Size

Continuing with Example 1, the researchers want to find the exact sample size needed to achieve 90% power.

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	**Sample Size (Exact)**
Design Type...	**3x2 (Three-Period, Two-Sequence Dual: ABB|BAA)**
Alternative Hypothesis	**Two-Sided**
Power..	**0.90**
Alpha...	**0.05**
R1 (μT / μR)...	**1.25 1.5**
COV (Coefficient of Variation)	**0.4**

Output

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

Numeric Results

Numeric Results for an Inequality Test of the Mean Ratio in a Three-Period, Two-Sequence Dual Design
Treatment Sequences: ABB | BAA
Alternative Hypothesis: Two-Sided

Power	Total Sample Size N	Mean Ratio R1 (μT / μR)	Coefficient of Variation COV	Alpha
0.9058	49	1.250	0.4000	0.050
0.9121	16	1.500	0.4000	0.050

This report shows the sample sizes.

Example 3 – Validation

We will validate this procedure by comparing its results to the *Superiority by a Margin Tests for the Ratio of Two Means in a Higher-Order Cross-Over Design (Log-Normal Data)* procedure which has already been validated. The two routines should achieve the same power when SM is set to zero in that routine.

In the superiority routine, if we set the Design Type to 3x2, Higher Means to Better, Alpha to 0.05, N to 10, SM to 0, Difference to 1.5, and COV to 0.50, the power is computed to be 0.6778.

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 3** by going to the **File** menu and choosing **Open Example Template**.

Option   **Value**

Design Tab
Solve For .... **Power**
Design Type... **3x2 (Three-Period, Two-Sequence Dual)**
Alternative Hypothesis .... **One-Sided**
Alpha... **0.05**
N (Total Sample Size)............................... **10**
R1 (μT / μR)... **1.5**
COV (Coefficient of Variation) **0.5**

Output

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

Numeric Results

Numeric Results for an Inequality Test of the Mean Ratio in a Three-Period, Two-Sequence Dual Design
Treatment Sequences: ABB | BAA
Alternative Hypotheses: One-Sided

Power	Total Sample Size N	Mean Ratio R1 (μT / μR)	Coefficient of Variation COV	Alpha
0.6778	10	1.500	0.5000	0.050

This procedure has also calculated the power to be 0.6778.