Chapter 516

Non-Inferiority Tests for the Ratio of Two Poisson Rates in a 2x2 Cross-Over Design

Introduction

Senn (2002) defines a *cross-over* design as one in which each subject receives all treatments and the objective 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 the other. It is assumed that there is a *washout* period between treatments during which the response returns back to its baseline value. If this does not occur, there is said to be a *carry-over* effect.

A 2×2 cross-over design contains to two *sequences* (treatment orderings) and two time periods (occasions). One sequence receives treatment A followed by treatment B. The other sequence receives B and then A. The design includes a washout period between responses to make certain that the effects of the first drug do not carry over to the second. Thus, the groups in this design are defined by the sequence in which the drugs are administered, not by the treatments they receive. Indeed, higher-order cross-over designs have been used in which the same treatment is used at both occasions.

Cross-over designs are employed because, if the no-carryover assumption is met, treatment differences are measured within a subject rather than between subjects—making a more precise measurement. Examples of the situations that might use a cross-over design are the comparison of anti-inflammatory drugs in arthritis and the comparison of hypotensive agents in essential hypertension. In both cases, symptoms are expected to return to their usual baseline level shortly after the treatment is stopped.

The sample size calculations in the procedure are based on the formulas presented in Lui (2016).

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 into the study may be easier because each patient will receive both treatments. Finally, it is often more difficult to obtain a subject than to obtain a measurement.

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. It may be difficult to separate the treatment effect from the period effect, the carry-over effect of the previous treatment, and the interaction between period and treatment.

The design cannot be used when the treatment (or the measurement of the response) alters the subject permanently. Hence, it should not be used to compare treatments that are intended to provide a cure.

Because subjects must be measured at least twice, it is often more difficult to keep patients enrolled in the study. It is arguably simpler to measure a subject once than to obtain their measurement twice. This is particularly true when the measurement process is painful, uncomfortable, embarrassing, or time consuming.

Technical Details

The 2×2 crossover design may be described as follows. Randomly assign the subjects to one of two sequence groups so that there are n_1 subjects in sequence one and n_2 subjects in sequence two. In order to achieve design balance, the sample sizes n_1 and n_2 are assumed to be equal so that $n_1 = n_2 = n = N/2$.

Sequence one is given the control (A) followed by the treatment (B). Sequence two is given the treatment (B) followed by the control (A).

Cross-Over Design

The discussions that follow summarize the results in Lui (2016) on pages 75-88. Consider a 2×2 cross-over design and let $Y_{ij}^{(g)}$ represent the frequency of event occurrences for the j^{th} subject, j = 1, …, n_g, in the i^{th} period (i = 1, 2), in sequence g (g = 1, 2). Let $X_{ij}^{(g)}$ represent the treatment-received covariate for the j^{th} subject, j = 1, …, n_g, in the i^{th} period (i = 1, 2), in sequence g (g = 1, 2) such that $X_{ij}^{(g)} = 1$ for a subject receiving the experimental treatment and $X_{ij}^{(g)} = 0$ for a subject receiving the control or standard treatment. Let $Z_{ij}^{(g)}$ represent the period covariate for the j^{th} subject, j = 1, …, n_g, in the i^{th} period (i = 1, 2), in sequence g (g = 1, 2) such that $Z_{ij}^{(g)} = 1$ for period 2 and $Z_{ij}^{(g)} = 0$ for period 1. Finally, assume that the $Y_{ij}^{(g)}$ follow a Poisson distribution with mean

$$E\left(Y_{ij}^{(g)}\right) = \mu_j^{(g)} \exp\left(\eta X_{ij}^{(g)} + \gamma Z_{ij}^{(g)}\right)$$

where $\mu_j^{(g)}$ represents the random effect of the j^{th} subject assigned to sequence g and has overall mean μ, η is the relative effect of the treatment to the control, and γ is the relative effect of period 2 to period 1. For a fixed period, the ratio, R, of mean event rates for the treatment versus the control is

$$R = \frac{\lambda_T}{\lambda_C} = e^{\eta}.$$

Similarly, the ratio of mean event rates for period 2 versus period 1 is

$$R_p = \frac{\lambda_2}{\lambda_1} = e^{\gamma}.$$

Non-Inferiority Test Statistics

Higher Event Rates Better

When higher event rates are better, the null and alternative hypotheses for a one-sided non-inferiority test are

$$H_0: R \leq R_0 \quad \text{vs} \quad H_A: R > R_0$$

or equivalently,

$$H_0: \eta \leq \eta_0 \quad \text{vs} \quad H_A: \eta > \eta_0$$

since $\eta = \log(R)$. R_0 is the lower non-inferiority bound (i.e. the smallest event rate ratio (λ_T/λ_C) for which the treatment will still be considered non-inferior to the standard or control). When higher event rates are better, R_0 should be less than one.

The power and sample size calculations are based on the test statistic

$$Z = \frac{\log(\hat{R}) - \log(R_0)}{\sqrt{\widehat{Var}\left(\log(\hat{R})\right)}}$$

which is asymptotically distributed as standard normal under the null hypothesis. The event rate ratio estimate, $\hat{R}$, and the variance estimate, $\widehat{Var}\left(\log(\hat{R})\right)$, are calculated as described in Lui (2016) on pages 77 through 79.

The null hypothesis is rejected in favor of the alternative at level α if

$$Z > Z_{1-\alpha}$$

where $Z_{1-\alpha}$ is the upper $1-\alpha$ percentile of the standard normal distribution.

Higher Event Rates Worse

When higher event rates are worse, the null and alternative hypotheses for a one-sided non-inferiority test are

$$H_0: R \geq R_0 \quad \text{vs} \quad H_A: R < R_0$$

or equivalently,

$$H_0: \eta \geq \eta_0 \quad \text{vs} \quad H_A: \eta < \eta_0$$

since $\eta = \log(R)$. R_0 is the upper non-inferiority bound (i.e. the largest event rate ratio (λ_T/λ_C) for which the treatment will still be considered non-inferior to the standard or control). When higher event rates are worse, R_0 should be greater than one.

The power and sample size calculations are based on the test statistic

$$Z = \frac{\log(\hat{R}) - \log(R_0)}{\sqrt{\widehat{Var}\left(\log(\hat{R})\right)}}$$

which is asymptotically distributed as standard normal under the null hypothesis. The event rate ratio estimate, $\hat{R}$, and the variance estimate, $\widehat{Var}\left(\log(\hat{R})\right)$, are calculated as described in Lui (2016) on pages 77 through 79.

The null hypothesis is rejected in favor of the alternative at level α if

$$Z < Z_{\alpha}$$

where Z_{α} is the lower α percentile of the standard normal distribution.

Power Calculation

If $\hat{R}$ is the estimate of the event rate ratio, then $\hat{\eta} = \log(\hat{R})$ has asymptotic variance

$$Var\left(\log(\hat{R})\right) = \frac{V(\mu,\eta,\gamma)}{n}$$

where

$$V(\mu,\eta,\gamma) = \frac{1}{4}\left(\frac{1}{\mu(1+e^{\eta+\gamma})p_1(1-p_1)} + \frac{1}{\mu(e^{\eta}+e^{\gamma})p_2(1-p_2)}\right)$$

with

$$p_1 = \frac{e^{\eta+\gamma}}{1+e^{\eta+\gamma}}$$

$$p_2 = \frac{e^{\gamma}}{e^{\eta}+e^{\gamma}}.$$

Higher Event Rates Better

Derived from the sample size formula given in Lui (2016) on page 87, the power for the one-sided non-inferiority test of $H_0: R \le R_0$ versus $H_A: R > R_0$ is

$$\Phi\left(\frac{\sqrt{n}\left(\log(R_1)-\log(R_0)\right)}{\sqrt{V(\mu,\eta,\gamma)}} - Z_{1-\alpha}\right)$$

where $\Phi()$ is the standard normal distribution function, R_1 is the actual value of the event rate ratio under the alternative hypothesis, and $Z_{1-\alpha}$ is the upper $1-\alpha$ percentile of the standard normal distribution. The sample size calculation formula is

$$n = Ceiling\left\{\left(\frac{\left(Z_{1-\alpha}+Z_{1-\beta}\right)\sqrt{V(\mu,\eta,\gamma)}}{\log(R_1)-\log(R_0)}\right)^2\right\}.$$

Higher Event Rates Worse

Derived from the sample size formula given in Lui (2016) on page 87, the power for the one-sided non-inferiority test of $H_0: R \ge R_0$ versus $H_A: R < R_0$ is

$$\Phi\left(\frac{\sqrt{n}\left(\log(R_0)-\log(R_1)\right)}{\sqrt{V(\mu,\eta,\gamma)}} + Z_{\alpha}\right)$$

where $\Phi()$ is the standard normal distribution function, R_1 is the actual value of the event rate ratio under the alternative hypothesis, and Z_{α} is the lower α percentile of the standard normal distribution. The sample size calculation formula is

$$n = Ceiling\left\{\left(\frac{\left(Z_{1-\alpha}+Z_{1-\beta}\right)\sqrt{V(\mu,\eta,\gamma)}}{\log(R_1)-\log(R_0)}\right)^2\right\}.$$

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 most of the parameters and options that you will be concerned with.

Solve For

Solve For

This option specifies the parameter to be calculated from the values of the other parameters. Under most conditions, you would select either *Power* or *Sample Size*.

Select *Sample Size* when you want to determine the sample size needed to achieve a given power and alpha level.

Select *Power* when you want to calculate the power of an experiment that has already been run.

Select *Effect Size (R1)* when you want to calculate the minimum effect size that can be detected for a particular design.

Test

Higher Event Rates Are

Use this option to specify the direction of the test.

If Higher Event Rates are “Better”, the alternative hypothesis is H1: $\lambda t/\lambda c > R0$.

If Higher Event Rates are “Worse”, the alternative hypothesis is H1: $\lambda t/\lambda c < R0$.

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 equal means when in fact the means are different.

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. In this procedure, a type-I error occurs when you reject the null hypothesis of equal means when in fact the means are equal.

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 (Sample Size per Sequence)

This is the sample size of each sequence or group (AB and BA) in the cross-over design. The individual sequence sample sizes are assumed to be equal such that the total sample size is equal to $N = 2n$.

You can enter a single value such as *50* or a list of values using the syntax *50 100 150 200 250* or *50 to 250 by 50*.

Effect Size – Event Rate Ratio

R0 (Non-Inferiority Ratio)

Specify the non-inferiority event rate ratio.

When higher event rates are "Better", the non-inferiority event rate ratio is the smallest ratio ($\lambda t/\lambda c$) for which the treatment will still be considered non-inferior to the standard or control.

When higher event rates are "Worse", the non-inferiority event rate ratio is the largest ratio ($\lambda t/\lambda c$) for which the treatment will still be considered non-inferior to the standard or control.

You can enter a single value such as *0.8* or a series of values such as *0.8 0.85 0.9* or *0.8 to 0.9 by 0.05* in the range R0 > 0, R0 ≠ R1. When higher event rates are "Better", R0 should be less than one. When higher event rates are "Worse", R0 should be greater than one.

R1 (Actual Ratio)

Specify the actual event rate ratio at which power is calculated.

You can enter a single value such as *1* or a series of values such as *1 1.5 2* or *1 to 2 by 0.5* in the range R1 > 0, R1 ≠ R0. When higher event rates are "Better", R1 should be greater than R0. When higher event rates are "Worse", R1 should be less than R0.

Effect Size – Additional Parameters

Fixed Mean Rate (μ)

Enter a value for the fixed mean rate of underlying random effects for the two treatments. This is usually estimated from a previous study. You can enter a single value such as *1* or a series of values such as *1 1.1 1.2* or *1 to 1.2 by 0.1* in the range μ > 0.

Period Rate Ratio (Rp)

Enter a value for the rate ratio for period 2 vs. period 1 on a given subject, given a fixed treatment. This is usually estimated from a previous study. You can enter a single value such as *1* or a series of values such as *1 1.1 1.2* or *1 to 1.2 by 0.1* in the range Rp > 0.

Example 1 – Power Analysis

Suppose you want to consider the power of non-inferiority test based on the ratio in a balanced cross-over design with a Poisson count endpoint for sequence sample sizes between 50 and 300. The non-inferiority ratio is 0.8 and the actual ratio is 1.0, the fixed mean rate is estimated to be 1, and the period rate ratio is estimated to be between 0.9 and 1.1. The significance level is 0.05.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Non-Inferiority Tests for the Ratio of Two Poisson Rates in a 2x2 Cross-Over Design** procedure window by expanding **Rates and Counts**, then **Cross-Over (2x2) Design**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Ratio of Two Poisson Rates 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 Event Rates Are	**Better**
Alpha	**0.05**
n (Sample Size per Sequence)	**50 to 300 by 50**
R0 (Non-Inferiority Ratio)	**0.8**
R1 (Actual Ratio)	**1**
Fixed Mean Rate (μ)	**1**
Period Rate Ratio (Rp)	**0.9 1.0 1.1**

Annotated Output

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

Numeric Results

Numeric Results for a Non-Inferiority Test
H0: R ≤ R0 vs. H1: R > R0

Power	Sequence Sample Size n	Total Sample Size N	Non-Inf. Ratio R0	Actual Ratio R1	Fixed Mean Rate μ	Period Rate Ratio Rp	Alpha
0.45657	50	100	0.800	1.000	1.000	0.900	0.050
0.47329	50	100	0.800	1.000	1.000	1.000	0.050
0.48809	50	100	0.800	1.000	1.000	1.100	0.050
0.70093	100	200	0.800	1.000	1.000	0.900	0.050
0.72126	100	200	0.800	1.000	1.000	1.000	0.050
0.73862	100	200	0.800	1.000	1.000	1.100	0.050
0.84499	150	300	0.800	1.000	1.000	0.900	0.050
0.86172	150	300	0.800	1.000	1.000	1.000	0.050
0.87542	150	300	0.800	1.000	1.000	1.100	0.050
0.92317	200	400	0.800	1.000	1.000	0.900	0.050
0.93459	200	400	0.800	1.000	1.000	1.000	0.050
0.94353	200	400	0.800	1.000	1.000	1.100	0.050
0.96321	250	500	0.800	1.000	1.000	0.900	0.050
0.97017	250	500	0.800	1.000	1.000	1.000	0.050
0.97537	250	500	0.800	1.000	1.000	1.100	0.050
0.98287	300	600	0.800	1.000	1.000	0.900	0.050
0.98679	300	600	0.800	1.000	1.000	1.000	0.050
0.98959	300	600	0.800	1.000	1.000	1.100	0.050

Non-Inferiority Tests for the Ratio of Two Poisson Rates in a 2x2 Cross-Over Design

References
Lui, Kung-Jong. 2016. Crossover Designs: Testing, Estimation, and Sample Size. John Wiley & Sons Ltd. Chichester, West Sussex, England.
Lui, Kung-Jong. 2013. Sample size determination for testing equality in Poisson frequency data under an AB/BA crossover trial. Pharmaceutical Statistics. Volume 12, pages 74-81.

Report Definitions
Power is the probability of rejecting a false null hypothesis. It should be close to one.
n is the sample size in each sequence (or group).
N is the total sample size from both sequences. The sample is divided equally among sequences.
R0 is the non-inferiority ratio used to specify the hypothesis test.
R1 is the actual event rate ratio (λt/λc) at which power is calculated.
μ is the fixed mean rate of underlying random effects for the two treatments.
Rp is the rate ratio for period 2 vs. period 1 on a given subject, given a fixed treatment.
Alpha is the probability of rejecting a true null hypothesis. It should be small.

Summary Statements

For a 2x2 cross-over design, a sample size of 50 in each sequence for a total of 100 achieves 45.657% power to detect an event rate ratio of 1.000 using a one-sided non-inferiority test against a bound of 0.800 with a significance level of 0.050 when the fixed mean rate of underlying random effects for the two treatments is 1.000 and the rate ratio for period 2 vs. period 1 on a given patient, given a fixed treatment, is 0.900.

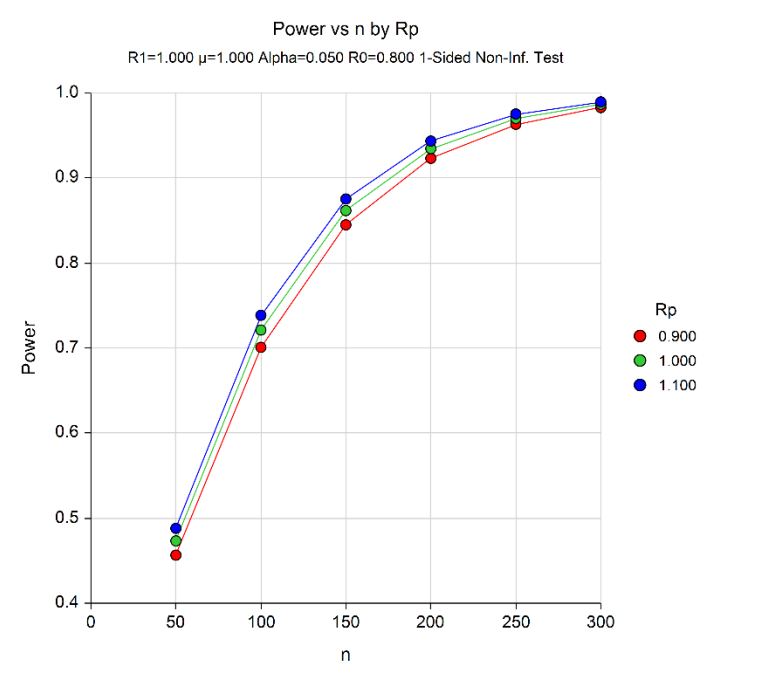
Dropout-Inflated Sample Size

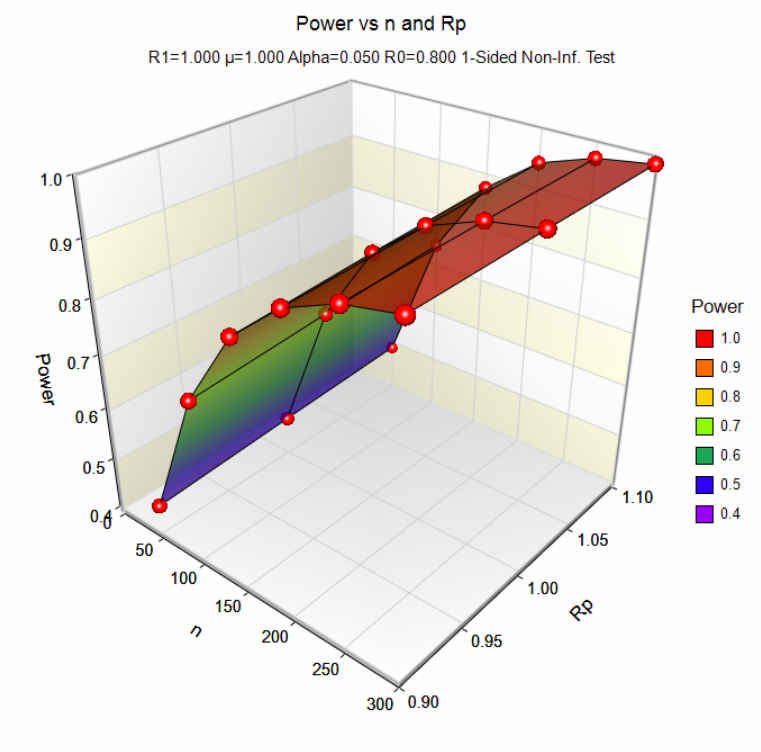
	Sample Size		Dropout-Inflated Enrollment Sample Size		Expected Number of Dropouts	
Dropout Rate	n	N	n'	N'	d	D
20%	50	100	63	126	13	26
20%	100	200	125	250	25	50
20%	150	300	188	376	38	76
20%	200	400	250	500	50	100
20%	250	500	313	626	63	126
20%	300	600	375	750	75	150

Definitions
Dropout Rate (DR) is 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”).
n and N are the evaluable group and total sample sizes, respectively, at which power is computed (as entered by the user). If n subjects from each group are evaluated out of the n' subjects that are enrolled in the study, the design will achieve the stated power. N = 2n.
n' and N' are the number of subjects that should be enrolled in the study in order to end up with n and N evaluable subjects, based on the assumed dropout rate. 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., and Wang, H. (2008) pages 39-40.). N' = 2n'.
d and D are the expected number of group and total dropouts, respectively. d = n' - n and D = 2d.

This report shows the values of each of the parameters, one scenario per row.

Chart Section





These plots show the relationship between sample size, Rp, and power. We see that sample sizes of between 150 and 200 per sequence are required to detect an actual event rate ratio of 1 with 90% power when the lower non-inferiority bound ratio is 0.8.

Example 2 – Calculating Sample Size (Validation using Hand Calculations)

In this example, we'll demonstrate how to compute sample size for a non-inferiority test of two Poisson rates from a 2x2 cross-over design. This example will also serve as validation for this procedure. We couldn't find any published examples of this test, so we'll validate the procedure by hand. Let's find the sample size required to detect and actual event rate ratio of 1 with 80% power at a significance level of 0.05 when the non-inferiority ratio is 0.8 and both the fixed mean rate and period rate ratio are equal to 1. These values are similar to those used in Table II on page 78 of Lui (2013) for a test of inequality.

First, we need to compute the variance component with

$$p_1 = \frac{e^{\eta+\gamma}}{1+e^{\eta+\gamma}} = \frac{e^{\eta}e^{\gamma}}{1+e^{\eta}e^{\gamma}} = \frac{1}{1+1} = 0.5$$

$$p_2 = \frac{e^{\gamma}}{e^{\eta}+e^{\gamma}} = \frac{1}{1+1} = 0.5$$

$$V(\mu,\eta,\gamma) = \frac{1}{4}\left(\frac{1}{\mu(1+e^{\eta+\gamma})p_1(1-p_1)} + \frac{1}{\mu(e^{\eta}+e^{\gamma})p_2(1-p_2)}\right)$$

$$= \frac{1}{4}\left(\frac{1}{1(1+1)0.5(1-0.5)} + \frac{1}{1(1+1)0.5(1-0.5)}\right)$$

$$= \frac{1}{4}\left(\frac{1}{0.5} + \frac{1}{0.5}\right) = 1.0$$

The formula for sample size given in Lui (2016) on page 87 is

$$n = Ceiling\left\{\left(\frac{(Z_{1-\alpha}+Z_{1-\beta})\sqrt{V(\mu,\eta,\gamma)}}{\log(R_1)-\log(R_0)}\right)^2\right\}$$

$$= Ceiling\left\{\left(\frac{(1.6449+0.8416)\sqrt{1.0}}{\log(1)-\log(0.8)}\right)^2\right\}$$

$$= Ceiling\{124.1651\} = 125.$$

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Non-Inferiority Tests for the Ratio of Two Poisson Rates in a 2x2 Cross-Over Design** procedure window by expanding **Rates and Counts**, then **Cross-Over (2x2) Design**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Ratio of Two Poisson Rates in a 2x2 Cross-Over Design**. 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**
Higher Event Rates Are	**Better**
Power	**0.80**
Alpha	**0.05**
R0 (Non-Inferiority Ratio)	**0.8**
R1 (Actual Ratio)	**1**
Fixed Mean Rate (μ)	**1**
Period Rate Ratio (Rp)	**1**

Output

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

Numeric Results

Numeric Results for a Non-Inferiority Test
H0: R ≤ R0 vs. H1: R > R0

Power	Sequence Sample Size n	Total Sample Size N	Non-Inf. Ratio R0	Actual Ratio R1	Fixed Mean Rate μ	Period Rate Ratio Rp	Alpha
0.80233	**125**	250	0.800	1.000	1.000	1.000	0.050

The sample size of 125 per sequence calculated by **PASS** matches our hand calculations.