

Chapter 170

Tests for the Odds Ratio of Two Proportions 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 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).

The design can be analyzed using a simple z -test if we ignore period and sequence effects or using a more complex random effects logistic regression model that adjusts for period and sequence effects. The sample size calculations herein ignore period and sequence effects. Julious (2010) suggests on page 175 that the bias due to ignoring period effects if a period-adjusted analysis is planned is not great and that sample size calculations that ignore period effects are adequate.

Cross-Over Design

The discussions that follow summarize the results in Lui (2016). Consider a 2×2 cross-over design and let $x_{ij}^{(g)}$ represent the binary response (0 or 1) from the j^{th} subject, $j = 1, \dots, n_g$, in the i^{th} period ($i = 1, 2$), in sequence g ($g = 1, 2$). Let $n_{rc}^{(g)}$ represent the number of subjects among n_g subjects in sequence g with the response vector ($x_{1j} = r, x_{2j} = c$). We can then summarize the results in terms of counts from a cross-over design with the following table for sequences 1 and 2 as

SEQUENCE 1 (Control → Treatment)					SEQUENCE 2 (Treatment → Control)				
		Period 2 (Treatment)					Period 2 (Control)		
		Yes	No	Total			Yes	No	Total
Period 1 (Control)	Yes	$n_{11}^{(1)}$	$n_{10}^{(1)}$	$n_{1\cdot}^{(1)}$	Period 1 (Treatment)	Yes	$n_{11}^{(2)}$	$n_{10}^{(2)}$	$n_{1\cdot}^{(2)}$
	No	$n_{01}^{(1)}$	$n_{00}^{(1)}$	$n_{0\cdot}^{(1)}$		No	$n_{01}^{(2)}$	$n_{00}^{(2)}$	$n_{0\cdot}^{(2)}$
Total		$n_{\cdot 1}^{(1)}$	$n_{\cdot 0}^{(1)}$	n_1	Total		$n_{\cdot 1}^{(2)}$	$n_{\cdot 0}^{(2)}$	n_2

Tests for the Odds Ratio of Two Proportions in a 2x2 Cross-Over Design

In terms of proportions, the 2x2 cross-over design tables can be summarized as

SEQUENCE 1 (Control → Treatment)					SEQUENCE 2 (Treatment → Control)				
		Period 2 (Treatment)					Period 2 (Control)		
		Yes	No	Total			Yes	No	Total
Period 1 (Control)	Yes	$P_{11}^{(1)}$	$P_{10}^{(1)}$	$P_{1\cdot}^{(1)}$	Period 1 (Treatment)	Yes	$P_{11}^{(2)}$	$P_{10}^{(2)}$	$P_{1\cdot}^{(2)}$
	No	$P_{01}^{(1)}$	$P_{00}^{(1)}$	$P_{0\cdot}^{(1)}$		No	$P_{01}^{(2)}$	$P_{00}^{(2)}$	$P_{0\cdot}^{(2)}$
Total		$P_{\cdot 1}^{(1)}$	$P_{\cdot 0}^{(1)}$	1	Total		$P_{\cdot 1}^{(2)}$	$P_{\cdot 0}^{(2)}$	1

with the individual proportions estimated as

$$\hat{P}_{rc}^{(g)} = \frac{n_{rc}^{(g)}}{n_g}$$

Lui (2016) indicates on pages 32-42 that the odds ratio for the treatment versus the control (O_T/O_C) is defined for a 2x2 cross-over design as

$$OR = \sqrt{\frac{P_{01}^{(1)} P_{10}^{(2)}}{P_{10}^{(1)} P_{01}^{(2)}}}$$

with estimate

$$\widehat{OR} = \sqrt{\frac{\hat{P}_{01}^{(1)} \hat{P}_{10}^{(2)}}{\hat{P}_{10}^{(1)} \hat{P}_{01}^{(2)}}}$$

The estimated log odds ratio, $\log(\widehat{OR})$, has asymptotic variance σ^2/n with

$$\sigma^2 = \frac{1}{4} \left(\frac{1}{P_{01}^{(1)}} + \frac{1}{P_{10}^{(1)}} + \frac{1}{P_{01}^{(2)}} + \frac{1}{P_{10}^{(2)}} \right)$$

which can be estimated as

$$\hat{\sigma}^2 = \frac{1}{4} \left(\frac{1}{\hat{P}_{01}^{(1)}} + \frac{1}{\hat{P}_{10}^{(1)}} + \frac{1}{\hat{P}_{01}^{(2)}} + \frac{1}{\hat{P}_{10}^{(2)}} \right)$$

The standard deviation, then, is

$$SD = \sigma = \sqrt{\sigma^2}$$

with estimate

$$\widehat{SD} = \hat{\sigma} = \sqrt{\hat{\sigma}^2}$$

Test Statistic

For a two-sided test of the hypotheses

$$H_0: OR = 1 \text{ vs } H_A: OR \neq 1$$

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

$$Z = \frac{\log(\widehat{OR})}{\frac{SD}{\sqrt{n}}}$$

which is asymptotically distributed as standard normal under the null hypothesis. The null hypothesis is rejected in favor of the alternative at level α if

$$\frac{|\log(\widehat{OR})|}{\frac{SD}{\sqrt{n}}} > Z_{1-\alpha/2}$$

where $Z_{1-\alpha/2}$ is the upper $1 - \alpha/2$ percentile of the standard normal distribution. One-sided tests reject the null hypothesis at level α if

$$\frac{|\log(\widehat{OR})|}{\frac{SD}{\sqrt{n}}} > Z_{1-\alpha}$$

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

Power Calculation

Derived from the sample size formula given in Lui (2016) on page 42, the power for the two-sided test of $H_0: OR = 1$ versus $H_A: OR \neq 1$ is

$$\Phi\left(\frac{|\log(OR_1)|}{\frac{SD}{\sqrt{n}}} - Z_{1-\alpha/2}\right)$$

where $\Phi()$ is the standard normal distribution function, OR_1 is a value of the odds ratio under the alternative hypothesis, and $Z_{1-\alpha/2}$ is the upper $1 - \alpha/2$ percentile of the standard normal distribution. The sample size calculation formula for a two-sided test is

$$n = \text{Ceiling}\left\{\left(\frac{(Z_{1-\alpha/2} + Z_{1-\beta})SD}{\log(OR_1)}\right)^2\right\}.$$

The power for a one-sided test is

$$\Phi\left(\frac{|\log(OR_1)|}{\frac{SD}{\sqrt{n}}} - Z_{1-\alpha}\right)$$

where $Z_{1-\alpha}$ is the upper $1 - \alpha$ percentile of the standard normal distribution. The sample size calculation formula for a one-sided test is

$$n = \text{Ceiling}\left\{\left(\frac{(Z_{1-\alpha} + Z_{1-\beta})SD}{\log(OR_1)}\right)^2\right\}.$$

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

Test

Alternative Hypothesis

Specify whether the test is one-sided or two-sided. When a two-sided hypothesis is selected, the value of alpha is split in half. Everything else remains the same.

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*.

Tests for the Odds Ratio of Two Proportions in a 2x2 Cross-Over Design

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 – Odds Ratio

Input Type

Indicate what type of values to enter to specify the effect size. Regardless of the entry type chosen, the calculations are the same. This option is simply given for convenience in specifying the effect size.

The choices are

- **Proportions**

Enter P_t (the proportion for the treatment or experimental group) and P_c (the proportion for the control, reference, or standard group). The odds ratio is calculated from these values as $[P_t/(1-P_t)] / [P_c/(1-P_c)]$.

- **Odds Ratios**

Enter the odds ratio $[P_t/(1-P_t)] / [P_c/(1-P_c)]$ under the alternative hypothesis.

OR1 (Odds Ratio|H1 = Ot/Oc)

The odds ratio is the ratio of the odds of the treatment group showing a response to the odds of the control group showing a response.

$$OR1 = O_t/O_c = [P_t/(1-P_t)] / [P_c/(1-P_c)] \text{ or } \sqrt{[P_{01}(1)P_{10}(2)] / P_{10}(1)P_{01}(2)}$$

For example, an odds ratio of 2 means that the odds of a response in the treatment group is twice the odds of a positive response in the control group.

You can enter a single value such as 2 or a series of values such as 2 2.5 3 or 2 to 3 by 0.5 in the range $OR1 > 0$, $OR1 \neq 1$.

Pt (Treatment Proportion|H1)

Enter a value for the treatment or experimental group proportion under the alternative hypothesis, H1. The power calculations assume that this is the actual value of the proportion.

The odds ratio, OR1, is calculated from P_t and P_c as

$$OR1 = [P_t/(1-P_t)] / [P_c/(1-P_c)]$$

You can enter a single value such as 0.1 or a series of values such as 0.1 0.2 0.3 or 0.1 to 0.9 by 0.1 in the range $0 < P_t < 1$, $P_t \neq P_c$.

Pc (Control Proportion)

Enter a value for the control, reference, or standard group proportion.

The odds ratio, OR1, is calculated from P_t and P_c as

$$OR1 = [P_t/(1-P_t)] / [P_c/(1-P_c)]$$

You can enter a single value such as 0.1 or a series of values such as 0.1 0.2 0.3 or 0.1 to 0.9 by 0.1 in the range $0 < P_c < 1$, $P_c \neq P_t$.

Effect Size – Standard Deviation of Log(Odds Ratio)

Standard Deviation Estimation Method

Choose how the standard deviation of the log odds ratio (SD) will be estimated.

The options are

- **Enter SD Directly**

Input log odds ratio SD values directly. This option allows you to input any value for SD, regardless of how it was estimated.

- **Use Estimated Discordant Cell Proportions**

Input discordant cell proportions from 2 tables from a previous 2x2 cross-over study. This is the method described in Lui (2016). These estimates are used only to calculate the standard deviation and are not necessarily related to the value used for OR1.

Standard Deviation (SD)

Enter a value for the standard deviation of the log odds ratio, SD.

Estimating SD using Previously-Estimated Discordant Cell Proportions

The standard deviation may be estimated using discordant cell proportions from a previous study using the method of Lui (2016) on page 42, where

$$SD = \sqrt{[(1/4)*(1/p_{01(1)} + 1/p_{10(1)} + 1/p_{01(2)} + 1/p_{10(2)})]}.$$

PASS will calculate the SD for you using this method if you set Estimation Method to “Use Estimated Discordant Cell Proportions” and enter discordant cell proportions from 2 tables from a previous 2x2 cross-over study.

Discordant Cell Proportions

Enter the corresponding discordant cell proportion from a previous 2x2 cross-over study.

The estimated proportions are defined as follows:

SEQUENCE 1 (C→T)

$p_{01(1)}$: Control = “No”, Treatment = “Yes”

$p_{10(1)}$: Control = “Yes”, Treatment = “No”

SEQUENCE 2 (T→C)

$p_{01(2)}$: Control = “Yes”, Treatment = “No”

$p_{10(2)}$: Control = “No”, Treatment = “Yes”

The standard deviation is calculated by PASS as described in Lui (2016) on page 42 as

$$SD = \sqrt{[(1/4)*(1/p_{01(1)} + 1/p_{10(1)} + 1/p_{01(2)} + 1/p_{10(2)})]}.$$

These estimates are used only to calculate the standard deviation and are not necessarily related to the value used for OR1.

Range: $0 < p_{01(g)} < 1$, $0 < p_{10(g)} < 1$, $p_{01(g)} + p_{10(g)} \leq 1$

Example 1 – Power Analysis

Suppose you want to consider the power of a balanced cross-over design with a binary endpoint where the test is computed based on the odds ratio for sequence sample sizes between 50 and 200. The odds ratio to detect is 2 and the estimated standard deviation of the log odds ratio is 2.5. 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 **Tests for the Odds Ratio of Two Proportions in a 2x2 Cross-Over Design** procedure window by expanding **Proportions**, then **Cross-Over (2x2) Design**, then clicking on **Test (Inequality)**, and then clicking on **Tests for the Odds Ratio of Two Proportions 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**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Alternative Hypothesis	Two-Sided
Alpha.....	0.05
n (Sample Size per Sequence).....	50 to 200 by 25
Input Type.....	Odds Ratios
OR1 (Odds Ratio H1 = Ot/Oc).....	2
Estimation Method.....	Enter SD Directly
Standard Deviation (SD).....	2.5

Annotated Output

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

Numeric Results

Numeric Results for a Two-Sided Test

H0: OR = 1 vs. H1: OR ≠ 1

	Sequence Sample Size n	Total Sample Size N	Odds Ratio OR1	Standard Deviation SD	Alpha
Power	50	100	2.000	2.500	0.050
	75	150	2.000	2.500	0.050
	100	200	2.000	2.500	0.050
	125	250	2.000	2.500	0.050
	150	300	2.000	2.500	0.050
	175	350	2.000	2.500	0.050
	200	400	2.000	2.500	0.050

References

Lui, Kung-Jong. 2016. Crossover Designs: Testing, Estimation, and Sample Size. John Wiley & Sons Ltd. Chichester, West Sussex, England.

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 combined sample size from both sequences.

OR1 is the odds ratio under the alternative hypothesis.

SD is the user-entered standard deviation. This is estimated from a previous study.

Alpha is the probability of rejecting a true null hypothesis. It should be small.

Tests for the Odds Ratio of Two Proportions in a 2x2 Cross-Over Design

Summary Statements

For a 2x2 cross-over design, a sample size of 50 in each sequence for a total of 100 achieves 50.022% power to detect an odds ratio of 2.000 using a two-sided test with a significance level of 0.050 when the standard deviation is 2.500.

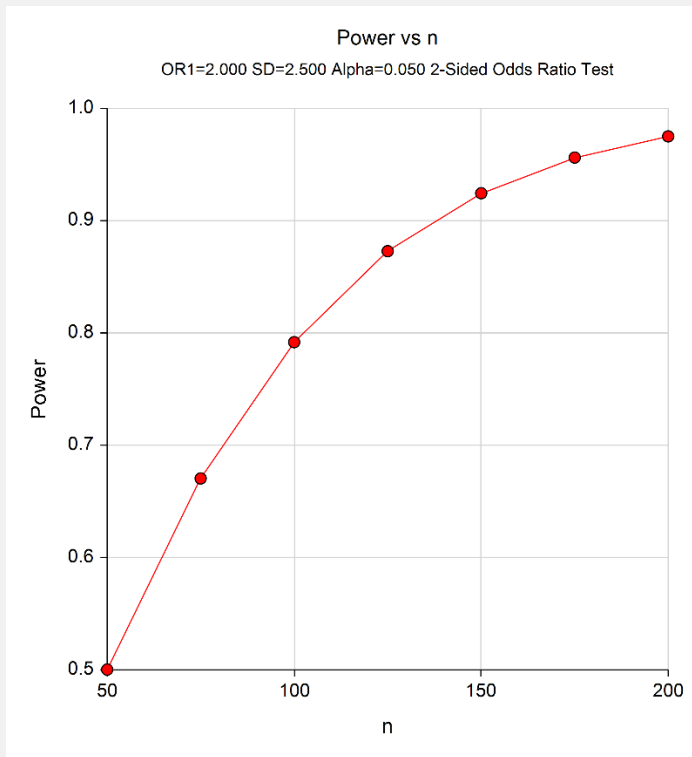
Dropout-Inflated Sample Size

Dropout Rate	Sample Size		Dropout-Inflated Enrollment Sample Size		Expected Number of Dropouts	
	n	N	n'	N'	d	D
20%	50	100	63	126	13	26
20%	75	150	94	188	19	38
20%	100	200	125	250	25	50
20%	125	250	157	314	32	64
20%	150	300	188	376	38	76
20%	175	350	219	438	44	88
20%	200	400	250	500	50	100

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 $D = 2d$.

Charts Section



This report shows the values of each of the parameters, one scenario per row. This plot shows the relationship between sample size and power. We see that a sample size of just over 100 per sequence is required to detect an odds ratio of 2 with 80% power.

Example 2 – Calculating Sample Size when Estimating the Standard Deviation from a Previous Study (Validation using Hand Calculations)

This example demonstrates how to calculate the sample size when estimating the standard deviation of the log odds ratio from data in a previous study using the method in Lui (2016) on page 42. In this example we'll find the sample size required to detect an odds ratio of 2 with 80% power at a significance level of 0.05, with the standard deviation estimated from discordant proportions from a previous study.

Table 3.2 of Lui (2016) on page 36 presents the following results from 279 subjects a simple 2x2 cross-over trial comparing two inhalation devices, A and B.

SEQUENCE 1 (Control (A) → Treatment (B))					SEQUENCE 2 (Treatment (B) → Control (A))				
		Period 2 (B)					Period 2 (A)		
		Yes	No	Total			Yes	No	Total
Period 1 (A)	Yes	26	41	67	Period 1 (B)	Yes	38	16	54
	No	15	57	72		No	32	54	86
Total		41	98	139	Total		70	70	140

The discordant proportions are estimated as

$$\hat{p}_{01}^{(1)} = \frac{15}{139} = 0.1079$$

$$\hat{p}_{10}^{(1)} = \frac{41}{139} = 0.2950$$

$$\hat{p}_{01}^{(2)} = \frac{32}{140} = 0.2286$$

$$\hat{p}_{10}^{(2)} = \frac{16}{140} = 0.1143$$

Using the method of Lui (2016) on page 42, SD is estimated as

$$\begin{aligned} \widehat{SD} &= \sqrt{\hat{\sigma}^2} \\ &= \sqrt{\frac{1}{4} \left(\frac{1}{\hat{p}_{01}^{(1)}} + \frac{1}{\hat{p}_{10}^{(1)}} + \frac{1}{\hat{p}_{01}^{(2)}} + \frac{1}{\hat{p}_{10}^{(2)}} \right)} \\ &= \sqrt{\frac{1}{4} \left(\frac{1}{0.1079} + \frac{1}{0.2950} + \frac{1}{0.2286} + \frac{1}{0.1143} \right)} \\ &= \mathbf{2.5388} \end{aligned}$$

PASS will calculate this SD value for you automatically when you input the discordant cell proportions.

Tests for the Odds Ratio of Two Proportions in a 2x2 Cross-Over Design

Since there is no example given for this calculation in the book, we'll validate this procedure using hand calculations. The sample size calculated by hand using the equation on Lui (2016) page 42 is

$$n = \text{Ceiling} \left\{ \left(\frac{(Z_{1-\alpha/2} + Z_{1-\beta})SD}{\log(OR_1)} \right)^2 \right\} = \text{Ceiling} \left\{ \left(\frac{(1.96 + 0.84)2.5388}{\log(2)} \right)^2 \right\}$$

$$= \text{Ceiling}\{105.177\}$$

$$= 106$$

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for the Odds Ratio of Two Proportions in a 2x2 Cross-Over Design** procedure window by expanding **Proportions**, then **Cross-Over (2x2) Design**, then clicking on **Test (Inequality)**, and then clicking on **Tests for the Odds Ratio of Two Proportions 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**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided
Power	0.80
Alpha	0.05
Input Type	Odds Ratios
OR1 (Odds Ratio H1 = Ot/Oc)	2
Estimation Method	Use Estimated Discordant Cell Proportions
p01(1)	0.1079
p10(1)	0.2950
p01(2)	0.2286
p10(2)	0.1143

Output

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

Numeric Results

Numeric Results for a Two-Sided Test

H0: OR = 1 vs. H1: OR ≠ 1

	Sequence Sample Size n	Total Sample Size N	Odds Ratio OR1	Standard Deviation SD*	Alpha
Power	106	212	2.000	2.539	0.050

* SD Estimated using Previously-Estimated Discordant Cell Proportions:
p01(1) = 0.1079, p10(1) = 0.295, p01(2) = 0.2286, p10(2) = 0.1143

This report indicates that the estimated standard deviation using the method of Lui (2016) is 2.539 and the required sample size is 106 per sequence. The discordant cell proportions are also listed. These values for SD and n computed by **PASS** match our hand calculations above.