Chapter 506

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

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

Test Statistic

For a two-sided test of the hypotheses

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

or equivalently,

$$H_0: \eta = 0 \quad \text{vs} \quad H_A: \eta \neq 0$$

since $\eta = \log(R)$. The power and sample size calculations are based on the test statistic

$$Z = \frac{\log\left(\hat{R}\right)}{\sqrt{\widehat{Var}_{H_0}\left(\log\left(\hat{R}\right)\right)}}$$

which is asymptotically distributed as standard normal under the null hypothesis. The event rate ratio estimate, $\hat{R}$, and the null variance estimate, $\widehat{Var}_{H_0}\left(\log\left(\hat{R}\right)\right)$, 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

$$|Z| > Z_{1-\alpha/2}$$

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

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

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

Power Calculation

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

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

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

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

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

Under $H_0: R = 1$ (or $H_0: \eta = 0$), $V(\mu, \eta, \gamma)$ can be further approximated by

$$V_{H_0}(\mu, \eta, \gamma) = \frac{1}{4\bar{p}(1-\bar{p})}\left(\frac{1}{\mu(1+e^{\eta+\gamma})} + \frac{1}{\mu(e^{\eta}+e^{\gamma})}\right)$$

where the pooled proportion $\bar{p}$ is calculated as

$$\bar{p} = \frac{e^{\eta+\gamma}+e^{\gamma}}{1+e^{\eta+\gamma}+e^{\eta}+e^{\gamma}}.$$

Derived from the sample size formula given in Lui (2016) on page 87, the power for the two-sided test of $H_0: R = 1$ versus $H_A: R \neq 1$ is

$$\Phi\left(\frac{\sqrt{n}|\log(R_1)| - Z_{1-\alpha/2}\sqrt{V_{H_0}(\mu,\eta,\gamma)}}{\sqrt{V(\mu,\eta,\gamma)}}\right)$$

where $\Phi()$ is the standard normal distribution function, R_1 is a value of the event rate 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

$$n = Ceiling\left\{\left(\frac{\left(Z_{1-\alpha/2}\sqrt{V_{H_0}(\mu,\eta,\gamma)} + Z_{1-\beta}\sqrt{V(\mu,\eta,\gamma)}\right)}{\log(R_1)}\right)^2\right\}.$$

The power for a one-sided test is

$$\Phi\left(\frac{\sqrt{n}|\log(R_1)| - Z_{1-\alpha}\sqrt{V_{H_0}(\mu,\eta,\gamma)}}{\sqrt{V(\mu,\eta,\gamma)}}\right)$$

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

$$n = Ceiling\left\{\left(\frac{\left(Z_{1-\alpha}\sqrt{V_{H_0}(\mu,\eta,\gamma)} + Z_{1-\beta}\sqrt{V(\mu,\eta,\gamma)}\right)}{\log(R_1)}\right)^2\right\}.$$

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 (R1)* 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*.

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

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

- **Event Rates**

  Enter λt (the event rate for the treatment or experimental group) and λc (the event rate for the control, reference, or standard group). The event rate ratio is calculated from these values as $R1 = \lambda t/\lambda c$.

- **Ratios**

  Enter the event rate ratio under the alternative hypothesis ($R1 = \lambda t/\lambda c$).

R1 (Ratio|H1 = λt/λc))

Enter a value for the event rate ratio under the alternative hypothesis, H1. The power calculations assume that this is the actual value of the event rate ratio. 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 R1 > 0, R1 ≠ 1.

λt (Treatment Event Rate|H1)

Enter a value for the treatment or experimental group event rate under the alternative hypothesis, H1. The power calculations assume that this is the actual value of the event rate. The event rate ratio, R1, is calculated from λt and λc as R1 = λt/λc. 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 λt > 0, λt ≠ λc.

λc (Control Event Rate)

Enter a value for the control, reference, or standard event rate. The event rate ratio, R1, is calculated from λt and λc as R1 = λt/λc. 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 λc > 0, λc ≠ λt.

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 a balanced cross-over design with a Poisson count endpoint where the test is computed based on the ratio for sequence sample sizes between 50 and 300. The ratio to detect is 1.2, 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 **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 **Test (Inequality)**, and then clicking on **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**
Alternative Hypothesis	**Two-Sided**
Alpha	**0.05**
n (Sample Size per Sequence)	**50 to 300 by 50**
Input Type	**Ratios**
R1 (Ratio|H1 = λt/λc)	**1.2**
Fixed Mean Rate (μ)	**1**
Period Rate Ratio (Rp)	**0.9 1 1.1**

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: R = 1 vs. H1: R ≠ 1

Power	Sequence Sample Size n	Total Sample Size N	Event Rate Ratio R1	Fixed Mean Rate μ	Period Rate Ratio Rp	Alpha
0.26068	50	100	1.200	1.000	0.900	0.050
0.27249	50	100	1.200	1.000	1.000	0.050
0.28310	50	100	1.200	1.000	1.100	0.050
0.46082	100	200	1.200	1.000	0.900	0.050
0.48103	100	200	1.200	1.000	1.000	0.050
0.49890	100	200	1.200	1.000	1.100	0.050
0.62483	150	300	1.200	1.000	0.900	0.050
0.64818	150	300	1.200	1.000	1.000	0.050
0.66832	150	300	1.200	1.000	1.100	0.050
0.74837	200	400	1.200	1.000	0.900	0.050
0.77072	200	400	1.200	1.000	1.000	0.050
0.78947	200	400	1.200	1.000	1.100	0.050
0.83615	250	500	1.200	1.000	0.900	0.050
0.85522	250	500	1.200	1.000	1.000	0.050
0.87075	250	500	1.200	1.000	1.100	0.050
0.89589	300	600	1.200	1.000	0.900	0.050
0.91092	300	600	1.200	1.000	1.000	0.050
0.92279	300	600	1.200	1.000	1.100	0.050

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.
R1 is the event rate ratio (λt/λc) under the alternative hypothesis.
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 26.068% power to detect an event rate ratio of 1.200 using a two-sided test 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.

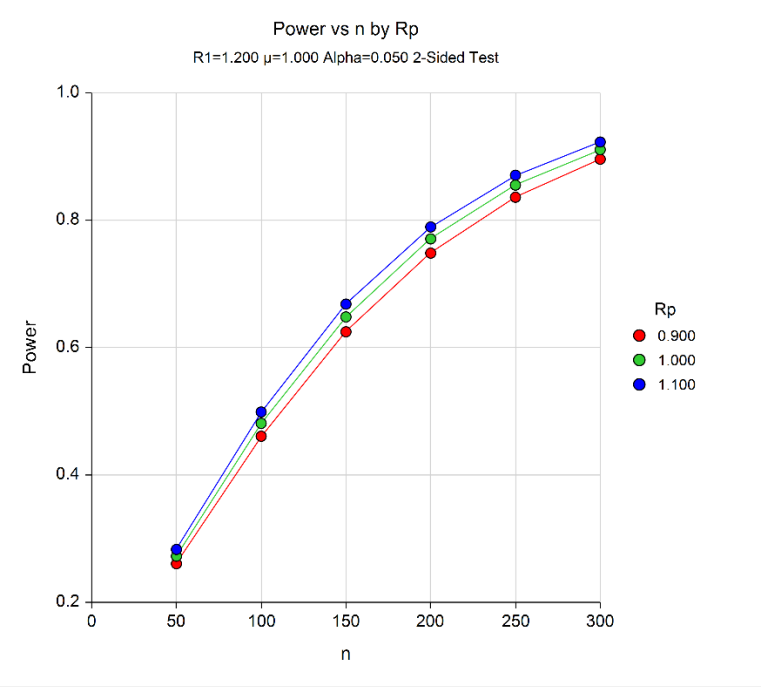
Dropout-Inflated Sample Size

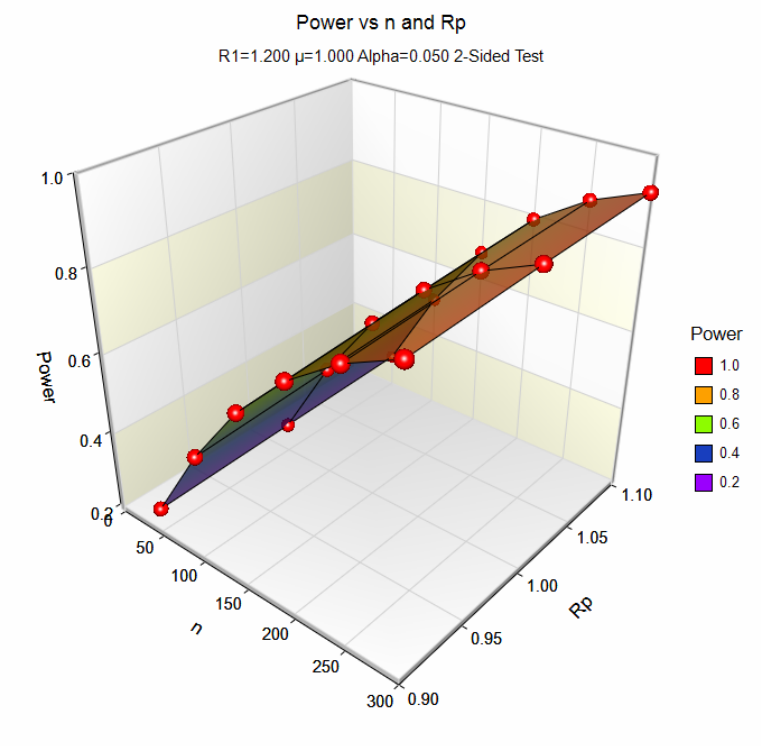
	Sample Size		Dropout-Inflated Enrollment Sample Size		Expected Number of Dropouts	
Dropout Rate	n	N	n'	N'	d	D
20%	50	100	63	126	13	26
20%	100	200	125	250	25	50
20%	150	300	188	376	38	76
20%	200	400	250	500	50	100
20%	250	500	313	626	63	126
20%	300	600	375	750	75	150

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 size of between 200 and 250 per sequence is required to detect an event rate ratio of 1.2 with 80% power.

Example 2 – Calculating Sample Size (Validation using Lui (2013))

Table II on page 78 of Lui (2013) (summarized below) presents sample size calculations to achieve 80% power at alpha = 0.05 for a two-sided test and various combinations of *R1* (0.5, 1.2, 1.5), μ (0.5, 1.0, 3.0), and R_p (0.90, 1.0, 1.1). This example will replicate all of the sample size calculations of that table.

		R_p		
		0.90	**1.0**	**1.1**
μ	*R1*	***n***	***n***	***n***
0.5	0.5	48	46	44
	1.2	455	431	411
	1.5	82	78	74
1.0	0.5	24	23	22
	1.2	228	216	206
	1.5	41	39	37
3.0	0.5	8	8	8
	1.2	76	72	69
	1.5	14	13	13

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 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 **Test (Inequality)**, and then clicking on **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**
Alternative Hypothesis	**Two-Sided**
Power	**0.80**
Alpha	**0.05**
Input Type	**Ratios**
R1 (Ratio|H1 = λt/λc)	**0.5 1.2 1.5**
Fixed Mean Rate (μ)	**0.5 1.0 3.0**
Period Rate Ratio (Rp)	**0.90 1.0 1.1**

Output

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

Numeric Results

Numeric Results for a Two-Sided Test
H0: R = 1 vs. H1: R ≠ 1

Power	Sequence Sample Size n	Total Sample Size N	Event Rate Ratio R1	Fixed Mean Rate μ	Period Rate Ratio Rp	Alpha
0.80247	48	96	0.500	0.500	0.900	0.050
0.80685	46	92	0.500	0.500	1.000	0.050
0.80755	44	88	0.500	0.500	1.100	0.050
0.80247	24	48	0.500	1.000	0.900	0.050
0.80685	23	46	0.500	1.000	1.000	0.050
0.80755	22	44	0.500	1.000	1.100	0.050
0.80247	8	16	0.500	3.000	0.900	0.050
0.82244	8	16	0.500	3.000	1.000	0.050
0.83885	8	16	0.500	3.000	1.100	0.050
0.80060	455	910	1.200	0.500	0.900	0.050
0.80056	431	862	1.200	0.500	1.000	0.050
0.80017	411	822	1.200	0.500	1.100	0.050
0.80146	228	456	1.200	1.000	0.900	0.050
0.80147	216	432	1.200	1.000	1.000	0.050
0.80112	206	412	1.200	1.000	1.100	0.050
0.80146	76	152	1.200	3.000	0.900	0.050
0.80147	72	144	1.200	3.000	1.000	0.050
0.80300	69	138	1.200	3.000	1.100	0.050
0.80170	82	164	1.500	0.500	0.900	0.050
0.80329	78	156	1.500	0.500	1.000	0.050
0.80091	74	148	1.500	0.500	1.100	0.050
0.80170	41	82	1.500	1.000	0.900	0.050
0.80329	39	78	1.500	1.000	1.000	0.050
0.80091	37	74	1.500	1.000	1.100	0.050
0.81091	14	28	1.500	3.000	0.900	0.050
0.80329	13	26	1.500	3.000	1.000	0.050
0.82086	13	26	1.500	3.000	1.100	0.050

The results from **PASS** match the sample sizes calculated in Lui (2013) exactly.