Chapter 437

Tests for Two Poisson Means

Introduction

The Poisson probability law gives the probability distribution of the number of events occurring in a specified interval of time or space. The Poisson distribution is often used to fit count data, such as the number of defects on an item, the number of accidents at an intersection during a year, the number of calls to a call center during an hour, or the number of meteors seen in the evening sky during an hour.

The Poisson distribution is characterized by a single parameter which is the mean number of occurrences during the specified interval.

The procedure documented in this chapter calculates the power or sample size for testing whether the ratio of two Poisson means is different from a specified value (usually one). The test procedure is described in Gu et al. (2008).

Test Procedure

Assume that all subjects in each group are observed for a fixed time period and the number of events (outcomes or defects) is recorded. The following table presents the various terms that are used.

Group	1	2
Fixed time interval	t_1	t_2
Sample Size	N_1	N_2
Number of events	X_1	X_2
Individual event rates	λ_1	$\lambda_2 = \rho\lambda_1$
Distribution of X	$\text{Poisson}(\lambda_1 t_1)$	$\text{Poisson}(\lambda_2 t_2)$

Gu (2008) considered several test statistics that can be used to test one-sided hypotheses about the ratio.

$$H_0 : \lambda_2 / \lambda_1 = \rho \text{ versus } H_a : \lambda_2 / \lambda_1 > \rho .$$

Two test statistics are available in this case. The first is based on unconstrained maximum likelihood estimates

$$W_1 = \frac{X_2 - X_1\left(\dfrac{\rho}{d}\right)}{\sqrt{X_2 + X_1\left(\dfrac{\rho}{d}\right)^2}}$$

The second test is based on constrained maximum likelihood estimates

$$W_2 = \frac{X_2 - X_1\left(\dfrac{\rho}{d}\right)}{\sqrt{(X_2 + X_1)\left(\dfrac{\rho}{d}\right)}}$$

An equivalent pair of test statistics are available if logarithms are used. The statistical hypothesis is

$$H_0 : \ln(\lambda_2 / \lambda_1) - \ln(\rho) = 0 \text{ versus } H_a : \ln(\lambda_2 / \lambda_1) - \ln(\rho) > 0$$

Two test statistics are available in this case as well. The first is based on unconstrained maximum likelihood estimates

$$W_3 = \frac{\ln\left(\dfrac{X_2}{X_1}\right) - \ln\left(\dfrac{\rho}{d}\right)}{\sqrt{\dfrac{1}{X_2} + \dfrac{1}{X_1}}}$$

The second test is based on constrained maximum likelihood estimates

$$W_4 = \frac{\ln\left(\dfrac{X_2}{X_1}\right) - \ln\left(\dfrac{\rho}{d}\right)}{\sqrt{\dfrac{\left(2 + \dfrac{d}{\rho} + \dfrac{\rho}{d}\right)}{X_1 + X_2}}}$$

After extensive simulation, they recommend the following extension of the variance-stabilized test proposed by Huffman (1984) for the case when $\rho / d > 1$.

$$W_5 = \frac{2\left[\sqrt{X_2 + k} - \sqrt{\dfrac{\rho}{d}(X_1 + k)}\right]}{\sqrt{1 + \dfrac{\rho}{d}}}$$

where $d = t_1 N_1 / t_2 N_2$ and $k = 3/8$.

Gu et al. (2008) show that all of these test statistics are approximately distributed as a standard normal and thus use the normal distribution as the basis of significance testing and power analysis.

Assumptions

The assumptions of the two-sample *Poisson* test are:

1. The data in each group are counts (discrete) that follow the Poisson distribution.

2. Each sample is a simple random sample from its population. Unlike most designs, in this design the sample size involves a fixed time parameter. That is, instead of specifying the number of people in a study, the number of man-hours is specified. Hence, a sample size of 10 hours could be achieved by ten people being observed for one hour or two people being observed for five hours.

    Also, since time is continuous, the sample size can be fractional.

Technical Details

Computing Power

The power analysis for testing the hypothesis

$$H_0 : \lambda_2 / \lambda_1 = \rho_0 \text{ versus } H_a : \lambda_2 / \lambda_1 = \rho_a > \rho_0$$

using the test statistics defined above is completed as follows.

1. **Find the critical value**. Choose the critical value $z_{1-\alpha}$ using the standard normal distribution so that the probability of rejecting H_0 when it is true is α.

2. **Compute the power**. Compute the power for each test as follows.

    For W_1, W_3, and W_4, the power is given by

    $$Power(W_i) = 1 - \Phi\left(\frac{z_{1-\alpha}\sigma_i - \mu_i}{\sigma_i}\right)$$

    where

    $$\Phi(z) = \int_{-\infty}^{z} Normal(0,1)$$

    $$\mu_1 = \left(\frac{\rho_a}{d} - \frac{\rho_0}{d}\right)t_1 N_1 \lambda_1, \qquad \sigma_1^2 = \left(\frac{d\rho_a + \rho_0^2}{d^2}\right)t_1 N_1 \lambda_1$$

    $$\mu_3 = \ln\left(\frac{\rho_a}{\rho_0}\right), \qquad \sigma_3^2 = \frac{d + \rho_a}{t_1 N_1 \lambda_1 \rho_a}$$

    $$\mu_4 = \ln\left(\frac{\rho_a}{\rho_0}\right), \qquad \sigma_4^2 = \frac{\left(2 + \dfrac{d}{\rho_0} + \dfrac{\rho_0}{d}\right)}{t_1 N_1 \lambda_1 \left(1 + \dfrac{\rho_a}{d}\right)}$$

For W_2 the power is computed using

$$Power(W_2) = 1 - \Phi\left(\frac{Ez_{1-\alpha} - F}{G}\right)$$

where

$$E = \sqrt{\left(\frac{\rho_0}{\rho_a}\right)^2 + \frac{\rho_0^2}{\rho_a d}}, \qquad F = \left(1 - \frac{\rho_0}{\rho_a}\right)\sqrt{\frac{\lambda_1 t_1 N_1 \rho_0}{d}},$$

$$G = \sqrt{\frac{\rho_0}{\rho_a}\left(1 + \frac{\rho_0^2}{d\rho_a}\right)}$$

For W_2 the power is computed using

$$Power(W_5) = 1 - \Phi\left(\frac{|A|\sqrt{B} - z_{1-\alpha}C}{D}\right)$$

where

$$A = 2\left(1 - \sqrt{\frac{\rho_0}{\rho_a}}\right), \qquad B = \lambda_1 t_1 N_1 + k,$$

$$C = \sqrt{\frac{\rho_0 + d}{\rho_a}}, \qquad D = \sqrt{\frac{\rho_a + d}{\rho_a}}$$

Computing Sample Size

The sample size is found using the formula

$$N_1 = \frac{\left(\dfrac{z_{1-\alpha}C + z_{Power}D}{A}\right)^2 - k}{\lambda_1 t_1}.$$

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 solved for from the other parameters. Under most situations, you will select either *Power* or *Sample Size*.

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

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

Test

Test Statistic

This option allows you to select which test statistic will be used. Possible choices are

- **W1 = MLE**

  This is the unconstrained maximum likelihood test in which $\lambda 1$ and $\lambda 2$ are estimated separately.

- **W2 = CMLE**

  This is the constrained maximum likelihood test in which $\lambda 1$ and $\lambda 2$ are estimated jointly.

- **W3 = Ln(MLE)**

  This is the logarithmic version of the unconstrained maximum likelihood test in which $\lambda 1$ and $\lambda 2$ are estimated separately.

- **W4 = Ln(CMLE)**

  This is the logarithmic version of the constrained maximum likelihood test $\lambda 1$ and $\lambda 2$ are estimated jointly.

- **W5 = Variance Stabilized**

  This is the variance stabilized version of the unconstrained maximum likelihood test.

Alternative Hypothesis

This option specifies the alternative hypothesis. This implicitly specifies the direction of the hypothesis test. The null hypothesis is $H_0 : \rho_a = \rho_0$. Possible selections for the alternative hypothesis are:

1. $H_a : \rho_0 < \rho_a$. This option yields a *one-tailed t test*.

2. $H_a : \rho_0 > \rho_a$. This option yields a *one-tailed t 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. For one-sided tests such as this, the value of 0.025 is recommended for alpha. You may enter a range of values such as *0.025 0.05 0.10* or *0.025 to 0.05 by 0.005*.

Sample Size

t1 (Observation Time Group 1)

This is the fixed observation time for group 1. Each subject is observed for this amount of time. If the observation times are variable, this is the average observed time per subject.

t2 (Observation Time Group 2)

This is the fixed observation time for group 2. Each subject is observed for this amount of time. If the observation times are variable, this is the average observed time per subject.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on *N1* or *N2* or both.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter N1, solve for N2**

  Select this option when you wish to fix *N1* at some value (or values), and then solve only for *N2*. Please note that for some values of *N1*, there may not be a value of *N2* that is large enough to obtain the desired power.

- **Enter N2, solve for N1**

  Select this option when you wish to fix *N2* at some value (or values), and then solve only for *N1*. Please note that for some values of *N2*, there may not be a value of *N1* that is large enough to obtain the desired power.

- **Enter R = N2/N1, solve for N1 and N2**

  For this choice, you set a value for the ratio of *N2* to *N1*, and then PASS determines the needed *N1* and *N2*, with this ratio, to obtain the desired power. An equivalent representation of the ratio, *R*, is

$$N2 = R * N1.$$

- **Enter percentage in Group 1, solve for N1 and N2**

  For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed *N1* and *N2* with this percentage to obtain the desired power.

N1 (Sample Size, Group 1)

This option is displayed if Group Allocation = "Enter N1, solve for N2"

N1 is the number of items or individuals sampled from the Group 1 population.

N1 must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed if Group Allocation = "Enter N2, solve for N1"

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter R = N2/N1, solve for N1 and N2."

R is the ratio of *N2* to *N1*. That is,

$$R = N2 / N1.$$

Use this value to fix the ratio of *N2* to *N1* while solving for *N1* and *N2*. Only sample size combinations with this ratio are considered.

N2 is related to *N1* by the formula:

$$N2 = [R \times N1],$$

where the value *[Y]* is the next integer ≥ *Y*.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1 (e.g., *N1* = 10 and *N2* = 20, or *N1* = 50 and *N2* = 100).

R must be greater than 0. If *R* < 1, then *N2* will be less than *N1*; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for N1 and N2."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for *N1* and *N2*. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When <u>Not</u> Solving for Sample Size)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter N1 and N2 individually**

  This choice permits you to enter different values for *N1* and *N2*.

- **Enter N1 and R, where N2 = R * N1**

  Choose this option to specify a value (or values) for *N1*, and obtain *N2* as a ratio (multiple) of *N1*.

- **Enter total sample size and percentage in Group 1**

  Choose this option to specify a value (or values) for the total sample size (*N*), obtain *N1* as a percentage of *N*, and then *N2* as *N - N1*.

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal (N1 = N2)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for *N1*, and also the value for *N2*.

The Sample Size Per Group must be ≥ 2. You can enter a single value or a series of values.

N1 (Sample Size, Group 1)

This option is displayed if Group Allocation = "Enter N1 and N2 individually" or "Enter N1 and R, where N2 = R * N1."

N1 is the number of items or individuals sampled from the Group 1 population.

N1 must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter N1 and N2 individually."

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter N1 and R, where N2 = R * N1."

R is the ratio of *N2* to *N1*. That is,

$$R = N2/N1$$

Use this value to obtain *N2* as a multiple (or proportion) of *N1*.

N2 is calculated from *N1* using the formula:

$$N2=[R \; x \; N1],$$

where the value *[Y]* is the next integer ≥ *Y*.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If *R* < 1, then *N2* will be less than N1; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines *N1* and *N2*.

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.

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

Percent in Group 1

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Effect Size – Means

λ1 (Group 1 Event Rate)

This option specifies one or more values of the mean occurrence rate in group 1 assuming the null hypothesis. This value must be greater than zero. This mean is compared to λ2 by the statistical test.

The difference in the ratios of $\lambda2$ and $\lambda1$ under the null and alternative hypotheses is the amount that this design can detect. You can enter a range of values such as "1 1.2 1.4" or "1 to 5 by 1."

ρ0 (λ_2 / λ_1 Assuming H0)

This is the value of the ratio of the two mean event rates assumed by the null hypothesis, H0. Usually, $\rho0 = 1.0$ which implies that the two rates are equal. However, you may test other values of $\rho0$ as well.

Strictly speaking, any positive number is valid. However, usually 1.0 is used.

Warning: you cannot use the same value for both $\rho0$ and ρa.

ρa (λ_2 / λ_1 Assuming Ha)

This is the value of the ratio of the two mean event rates assumed by the alternative hypothesis, Ha. The magnitude of the difference between this value and $\rho0$ is the amount that the design can detect.

Warning: you cannot use the same value for both $\rho0$ and ρa.

Example 1 – Finding the Sample Size

We will use the example of Gu (2008) in which epidemiologist wish to examine the relationship of post-menopausal hormone use and coronary heart disease (CHD). The incidence rate for those not using the hormone is 0.0005 ($\lambda_1 = 0.0005$). How large of a sample is needed to detect a change in the incidence ratio from $\rho_0 = 1$ to $\rho_a = 2, 3, 4, 5,$ or 6. Assume that 90% power is required and $\alpha = 0.05$. Assume that each subject will be observed for two years and that the design calls for an equal number of subjects in both groups.

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 Two Poisson Means** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Non-Normal Data**, and then clicking on **Tests for Two Poisson Means**. 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	**Sample Size**
Test Statistic	**W5 = Variance Stabilized**
Ha (Alternative Hypothesis)....................	**Ha: ρ0 < ρa**
Power..	**0.90**
Alpha...	**0.05**
t1 (Observation Time Group 1)..............	**2**
t2 (Observation Time Group 2)..............	**2**
Group Allocation	**Equal (N1 = N2)**
λ1 (Group 1 Event Rate)........................	**0.0005**
ρ0 (λ2 / λ1 Assuming H0)	**1**
ρa (λ2 / λ1 Assuming Ha)	**2 3 4 5 6**

Annotated Output

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

Numeric Results

Numeric Results for Two-Sample Poisson Test
Null Hypothesis: λ2 / λ1 = ρ0   Alternative Hypothesis: λ2 / λ1 = ρa > ρ0
Fixed Observation Time Length: Group 1 = 2, Group 2 = 2

Power	N1	N2	N	Mean Rate 1 λ1	Mean Rate 2 λ2 | ρa	H0 Ratio ρ0	Ha Ratio ρa	Alpha
0.9000	29736.2	29736.2	59472.5	0.00	0.00	1.0000	2.0000	0.050
0.9000	10776.9	10776.9	21553.7	0.00	0.00	1.0000	3.0000	0.050
0.9000	6363.7	6363.7	12727.4	0.00	0.00	1.0000	4.0000	0.050
0.9000	4512.5	4512.5	9025.0	0.00	0.00	1.0000	5.0000	0.050
0.9000	3513.9	3513.9	7027.7	0.00	0.00	1.0000	6.0000	0.050

References
Gu, K., Ng, H.K.T., Tang, M.L., and Schucany, W. 2008. 'Testing the Ratio of Two Poisson Rates.' Biometrical
    Journal, 50, 2, 283-298.
Huffman, Michael. 1984. 'An Improved Approximate Two-Sample Poisson Test.' Applied Statistics, 33, 2, 224-226.

Report Definitions
Power is the probability of rejecting a false null hypothesis.
N1 and N2 are the number of subjects in the first and second groups.
N is the total sample size, N1 + N2.
λ1 is the mean response rate for a group 1 subject during the fixed observation time.
λ2 | ρa is the mean response rate for a group 2 subject during the fixed observation time assuming ρa.
ρ0 is the ratio λ2 / λ1 assuming the null hypothesis.
ρa is the ratio λ2 / λ1 assuming the alternative hypothesis.
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements
Samples of 29736.2 subjects in group 1 observed for 2.0 time periods and 29736.2 subjects in
group 2 observed for 2 time periods achieve 90% power to detect a change in the mean response
ratio (λ2 / λ1) from 1.0000 to 2.0000 when λ1 = 0.0005 and the significance level (alpha)
is 0.0500 using a one-sided test.

This report shows the values of each of the parameters, one scenario per row. The values of power and beta were calculated from the other parameters.

Power

Power is the probability of rejecting a false null hypothesis.

N1

N1 is the number of subjects in the first group. In the case of Poisson random variables, N1 does not have to be a whole number. A fractional subject simple refers to a subject that is observed for a fraction of the time period. Also, the effective sample size is the number of subjects multiplied by the fixed time period. For example, the sample size in the first row is 29736.2 x 2 = 59472.4 person-years.

N2

N2 is the number of subjects in the second group.

N

N is the total sample size, N1 + N2.

λ1

λ1 is the mean response rate for a group 1 subject during the fixed observation time.

λ2 | ρa

λ2 | ρa is the mean response rate for a group 2 subject during the fixed observation time assuming ρa.

ρ0

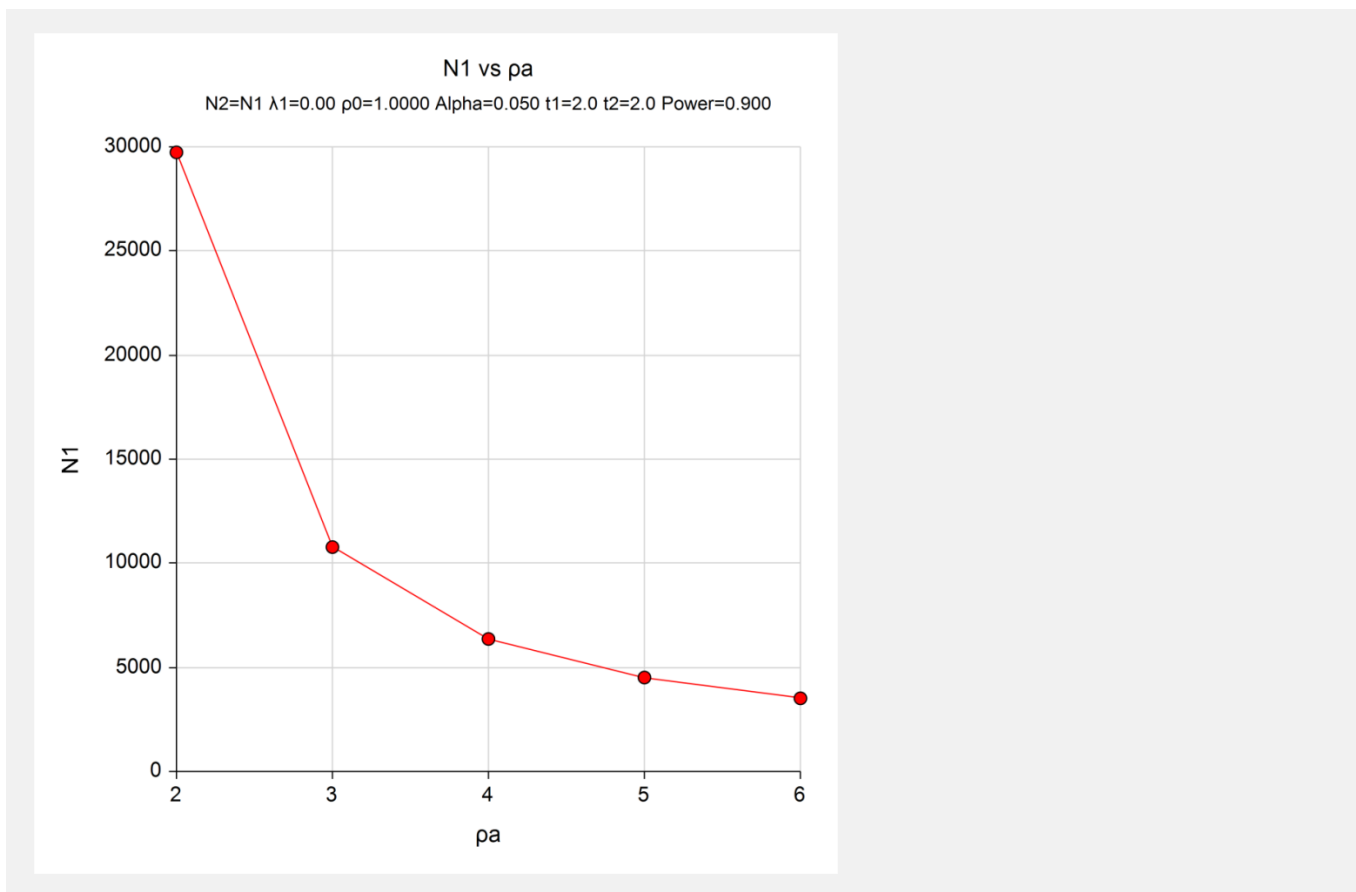ρ0 is the ratio λ2 / λ1 assuming the null hypothesis.

ρa

ρa is the ratio λ2 / λ1 assuming the alternative hypothesis.

Alpha

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

Plots Section

This plot shows the relationship between sample size and ρa.

Example 2 – Validation using Gu

Gu et al. (2008) present an example that we will use to validate this procedure. Using the scenario cited in Example 1 above, they give a sample size calculation on page 295. In this example, $\lambda_1 = 0.0005$, $\rho_0 = 1$, $\rho_a = 4$, $t_1 = t_2 = 2$, $\alpha = 0.05$, $R = 0.5$, and power = 0.9. In their Table 6, they list the sample size for $p_5^{(A)}$ in this scenario as 8627. However, this number is inaccurate because of the two-decimal place rounding that was done during their calculation. In a private communication, they agreed that the more accurate number is 8589.

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 Two Poisson Means** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Non-Normal Data**, and then clicking on **Tests for Two Poisson Means**. 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**
Test Statistic	**W5 = Variance Stabilized**
Ha (Alternative Hypothesis)	**Ha: ρ0 < ρa**
Power..	**0.90**
Alpha...	**0.05**
t1 (Observation Time Group 1)..............	**2**
t2 (Observation Time Group 2)..............	**2**
Group Allocation	**Enter R = N2/N1, solve for N1 and N2**
R	**0.5**
λ1 (Group 1 Event Rate)........................	**0.0005**
ρ0 (λ2 / λ1 Assuming H0)	**1**
ρa (λ2 / λ1 Assuming Ha)	**4**

Output

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

Numeric Results

Numeric Results for Two-Sample Poisson Test
Null Hypothesis: λ2 / λ1 = ρ0   Alternative Hypothesis: λ2 / λ1 = ρa > ρ0
Fixed Observation Time Length: Group 1 = 2, Group 2 = 2

Power	N1	N2	N	R	Mean Rate 1 λ1	Mean Rate 2 λ2 | ρa	H0 Ratio ρ0	Ha Ratio ρa	Alpha
0.9000	8589.4	4294.7	12884.1	0.50	0.0005	0.0020	1.0000	4.0000	0.050

These results match the more accurate value of 8589.