

Chapter 438

Tests for the Ratio of Two Negative Binomial Rates

Introduction

Count data arise from counting the number of events of a particular type that occur during a specified time interval. Examples include 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 the night. Clinical trial often result in count data, with examples including the number of patients who are admitted to the hospital or the number of patients who respond favorably (or unfavorably) to a particular treatment.

Traditionally, the Poisson distribution (e.g. Poisson regression) has been used to model count data. The Poisson model assumes that the mean and variance are equal, but in many clinical trials the variance is observed to be greater than the mean in a condition called *overdispersion*. When overdispersion occurs, the Poisson model provides a poor fit to the data. As an alternative, the negative binomial model is increasingly being used to model overdispersed count data. While the Poisson distribution is characterized by a single parameter which represents both the mean and the variance, the negative binomial distribution includes two parameters, allowing for greater flexibility in modeling the mean-variance relationship that is observed in overdispersed, heterogeneous count data.

This procedure is based on the formulas and results outlined in Zhu and Lakkis (2014) and calculates the power and sample size for testing whether the ratio of two negative binomial event rates is different from one. The test is often performed using the Wald (or likelihood ratio) test statistic in the context of generalized linear models. Such an analysis is available within SAS Proc GENMOD. These asymptotic tests are appropriate when the sample size is greater than 50 per group. When the sample size is less than 50 per group, the results from this procedure can be used to obtain a rough estimate of the power (see Zhu and Lakkis (2014), page 381).

Technical Details

The Negative Binomial Model

As in Zhu and Lakkis (2014), define y_{ij} as the number of events during time t_{ij} for subject i ($i = 1$ to n_j) in group j ($j = 1, 2$). Usually, group 1 is considered the control or reference group and group 2 is considered the treatment group. If y_{ij} follows a negative binomial distribution with mean μ_{ij} and dispersion parameter κ , the probability function for y_{ij} is

$$P(y_{ij}) = \frac{\Gamma(\kappa^{-1} + y_{ij})}{\Gamma(\kappa^{-1})y_{ij}!} \left(\frac{\kappa\mu_{ij}}{1 + \kappa\mu_{ij}} \right)^{y_{ij}} \left(\frac{1}{1 + \kappa\mu_{ij}} \right)^{1/\kappa}$$

Tests for the Ratio of Two Negative Binomial Rates

where $\Gamma(\cdot)$ is the gamma function. Using negative binomial regression, we can model μ_{ij} as

$$\log(\mu_{ij}) = \log(t_{ij}) + \beta_0 + \beta_1 x_{ij}$$

such that

$$\log\left(\frac{\mu_{ij}}{t_{ij}}\right) = \beta_0 + \beta_1 x_{ij}$$

where $x_{ij} = 0$ if the i^{th} subject is in group 1 and $x_{ij} = 1$ if the i^{th} subject is in group 2.

Further define λ_1 and λ_2 as the mean event rates per time unit for groups 1 and 2, respectively, and $RR = \lambda_2/\lambda_1$ as the ratio of event rates. Using the negative binomial model, it follows then that

$$\begin{aligned}\lambda_1 &= e^{\beta_0} \\ \lambda_2 &= e^{\beta_0 + \beta_1} \\ RR &= \frac{\lambda_2}{\lambda_1} = e^{\beta_1}\end{aligned}$$

If we define $\hat{\beta}_1$ as the asymptotic maximum likelihood estimate of β_1 , then the variance of $\hat{\beta}_1$ can be written as

$$\text{Var}(\hat{\beta}_1) = \frac{1}{n_1} \left[\frac{1}{\mu_t} \left(\frac{1}{\lambda_1} + \frac{1}{R\lambda_2} \right) + \frac{(1+R)\kappa}{R} \right]$$

where n_1 and n_2 are the sample sizes and λ_1 and λ_2 are the event rates from groups 1 and 2, respectively, $R = n_2/n_1$ is the sample allocation ratio, κ is the negative binomial dispersion parameter (assumed to be constant for power calculations), and μ_t is the average exposure time across all subjects (i.e. $t_{ij} = \mu_t$ for all i, j).

Hypothesis Test

The two-sided null and alternative hypotheses for testing equality of the two event rates can be written as

$$H_0: \beta_1 = 0 \quad \text{vs.} \quad H_A: \beta_1 \neq 0$$

or equivalently in terms of $RR = \lambda_2/\lambda_1$ as

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

The upper and lower one-sided tests, respectively, are

$$H_0: \beta_1 \leq 0 \quad \text{vs.} \quad H_A: \beta_1 > 0$$

$$H_0: \beta_1 \geq 0 \quad \text{vs.} \quad H_A: \beta_1 < 0$$

or equivalently in terms of $RR = \lambda_2/\lambda_1$ as

$$H_0: RR \leq 1 \quad \text{vs.} \quad H_A: RR > 1$$

$$H_0: RR \geq 1 \quad \text{vs.} \quad H_A: RR < 1$$

These hypotheses are most commonly tested using the Wald test statistic within generalized linear models. The likelihood ratio test statistic is also used. Such an analysis can be performed for the negative binomial distribution using SAS Proc GENMOD with a logarithmic link function and an indicator variable for group (1 or 2) as the single independent variable. For more information see Zhu and Lakkis (2014) or the SAS help manual.

Estimating the Variance under the Null and Alternative Hypotheses

Asymptotically, the variance of $\hat{\beta}_1$ is

$$\text{Var}(\hat{\beta}_1) = \frac{1}{n_1} \left[\frac{1}{\mu_t} \left(\frac{1}{\lambda_1} + \frac{1}{R\lambda_2} \right) + \frac{(1+R)\kappa}{R} \right]$$

If we define V_A under the alternative hypothesis using the true rates λ_1 and λ_2 as

$$V_A = \frac{1}{\mu_t} \left(\frac{1}{\lambda_1} + \frac{1}{R\lambda_2} \right) + \frac{(1+R)\kappa}{R}$$

then the variance of $\hat{\beta}_1$ under the alternative hypothesis can be written as

$$\begin{aligned} \text{Var}_A(\hat{\beta}_1) &= \frac{1}{n_1} \left[\frac{1}{\mu_t} \left(\frac{1}{\lambda_1} + \frac{1}{R\lambda_2} \right) + \frac{(1+R)\kappa}{R} \right] \\ &= \frac{1}{n_1} [V_A] \end{aligned}$$

If we define V_0 using the rates $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ estimated under the null hypothesis as

$$V_0 = \frac{1}{\mu_t} \left(\frac{1}{\tilde{\lambda}_1} + \frac{1}{R\tilde{\lambda}_2} \right) + \frac{(1+R)\kappa}{R}$$

then the variance of $\hat{\beta}_1$ under the null hypothesis is

$$\begin{aligned} \text{Var}_0(\hat{\beta}_1) &= \frac{1}{n_1} \left[\frac{1}{\mu_t} \left(\frac{1}{\tilde{\lambda}_1} + \frac{1}{R\tilde{\lambda}_2} \right) + \frac{(1+R)\kappa}{R} \right] \\ &= \frac{1}{n_1} [V_0] \end{aligned}$$

The portion V_0 (and therefore $\text{Var}_0(\hat{\beta}_1)$) can be estimated in three different ways. Define $V_{0|M}$ as the estimate of V_0 given the chosen method M . It follows then that

$$\text{Var}_{0|M}(\hat{\beta}_1) = \frac{1}{n_1} [V_{0|M}]$$

where $M = 1, 2, 3$ indicates the method used to estimate V_0 .

The three methods for estimating V_0 are as follows:

- **Method 1: Use the Event Rate from Group 1 (λ_1)**

Under $H_0: RR = 1$, the event rates are equal (i.e. $\lambda_2 = \lambda_1$), and the null variance is estimated using

$$V_{0|1} = \frac{1+R}{\mu_t R \lambda_1} + \frac{(1+R)\kappa}{R}$$

- **Method 2: Use the True Event Rates (λ_1 and λ_2)**

The true rates λ_1 and λ_2 are used, and the null variance is estimated using

$$V_{0|2} = \frac{1}{\mu_t} \left(\frac{1}{\lambda_1} + \frac{1}{R\lambda_2} \right) + \frac{(1+R)\kappa}{R}$$

This is equivalent to the estimation of V_A under the alternative hypothesis, i.e. $V_{0|2} = V_A$.

Tests for the Ratio of Two Negative Binomial Rates

- **Method 3: Use Maximum Likelihood Estimation**

The maximum likelihood estimate of λ under $H_0: RR = 1$, a weighted average of λ_1 and λ_2 , is used, and the null variance is estimated using

$$V_{0|3} = \frac{(1+R)^2}{\mu_t R(\lambda_1 + R\lambda_2)} + \frac{(1+R)\kappa}{R}$$

Simulation studies suggest that sample sizes calculated using methods 2 and 3 are more accurate than those calculated using method 1 (see Zhu and Lakkis (2014), page 385).

Computing Sample Size

From Zhu and Lakkis (2014), page 378, the sample size in group 1 required to achieve power of $1 - \beta$ for the two-sided Wald or likelihood ratio test at significance level α can be calculated using calculation method M for V_0 as

$$n_1 \geq \frac{(z_{\alpha/2}\sqrt{V_{0|M}} + z_{\beta}\sqrt{V_A})^2}{(\log(RR))^2}$$

with RR , V_A , and $V_{0|M}$ as defined earlier. The power of the one-sided Wald or likelihood ratio test at significance level α using calculation method M for V_0 is

$$n_1 \geq \frac{(z_{\alpha}\sqrt{V_{0|M}} + z_{\beta}\sqrt{V_A})^2}{(\log(RR))^2}$$

Computing Power

From Zhu and Lakkis (2014), page 379, the power of the two-sided Wald or likelihood ratio test can be calculated at significance level α using calculation method M for V_0 as

$$\text{Power}_{2\text{-sided}} = 1 - \beta = \Phi\left(\frac{\sqrt{n_1}|\log(RR)| - z_{\alpha/2}\sqrt{V_{0|M}}}{\sqrt{V_A}}\right)$$

with RR , V_A , and $V_{0|M}$ as defined earlier. The power of the one-sided Wald or likelihood ratio test at significance level α using calculation method M for V_0 is

$$\text{Power}_{1\text{-sided}} = 1 - \beta = \Phi\left(\frac{\sqrt{n_1}|\log(RR)| - z_{\alpha}\sqrt{V_{0|M}}}{\sqrt{V_A}}\right)$$

The power calculations are accurate for the Wald and likelihood ratio tests when the group sample sizes are greater than 50. When the sample size is less than 50 per group the validity of the Wald and likelihood ratio tests is questionable, and these formulas should be used only to obtain a rough estimate of the power (see Zhu and Lakkis (2014), page 381).

Tests for the Ratio of Two Negative Binomial Rates

Estimating the Negative Binomial Dispersion Parameter, κ , from a Previous Study that was Analyzed using Poisson Regression

Admittedly, the hardest value to determine among those required for these sample size and power calculations is the value for κ , the negative binomial dispersion parameter. If a suitable value for κ is not known, then you can estimate κ from a similar study that was analyzed using Poisson regression.

Given a Poisson mean event rate λ and an overdispersion factor ϕ , estimated from a similar Poisson regression study, the relationship between λ , ϕ , and κ is

$$\phi = 1 + \kappa\lambda$$

such that an estimate of κ can be calculated using overall estimates of λ and ϕ as

$$\hat{\kappa} = \frac{\hat{\phi} - 1}{\hat{\lambda}}$$

If $\hat{\phi}$ (the estimate for the overdispersion factor) is not directly reported in the previous study, Zhu and Lakkis (2014) suggest on page 384 that the Poisson overdispersion factor can be estimated from the total number of events, Y , the total exposure time, T , and the standard error, $SE(\log \hat{\lambda})$, as

$$\hat{\phi} = \frac{\text{Var}(Y)}{\hat{\lambda}T} \approx T\hat{\lambda} \left(SE(\log \hat{\lambda}) \right)^2$$

since for overdispersed Poisson

$$\text{Var}(Y) = \phi\lambda T$$

The maximum likelihood estimator of the event rate of the Poisson distribution, λ , is

$$\hat{\lambda} = \frac{Y}{T}$$

The standard error, $SE(\log \hat{\lambda})$, can be determined by back-calculating from reported $100(1 - \alpha)\%$ confidence interval endpoints for λ from Poisson regression as

$$SE(\log \hat{\lambda}) = \frac{\left(\frac{\log(\text{Upper Bound}) - \log(\text{Lower Bound})}{2} \right)}{z_{1-\alpha/2}}$$

Finally, from a previous two-group Poisson regression study, κ can be calculated from the estimated Poisson event rates $\hat{\lambda}_1$ and $\hat{\lambda}_2$ and the estimated Poisson overdispersion factors $\hat{\phi}_1$ and $\hat{\phi}_2$ from each group as

$$\hat{\kappa} = \frac{\left(\frac{\hat{\phi}_1 + \hat{\phi}_2}{2} \right) - 1}{\left(\frac{\hat{\lambda}_1 + \hat{\lambda}_2}{2} \right)}$$

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.

Tests for the Ratio of Two Negative Binomial Rates

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters.

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

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

Select $\mu(t)$ (*Average Exposure Time*) to calculate the required exposure time for a set number of subjects to achieve a desired power.

Select *RR (Ratio of Event Rates)* to find the detectable ratio for a given power and sample size.

Test

Alternative Hypothesis

Specify whether the alternative hypothesis of the test is one-sided or two-sided. If a one-sided test is chosen, the hypothesis test direction is chosen based on whether the event rate ratio (*RR*) is greater than or less than one.

The options are

- **Two-Sided**

The null and alternative hypotheses are

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

- **One-Sided**

The upper and lower null and alternative hypotheses are

$$\text{Upper: } H_0: RR \leq 1 \quad \text{vs.} \quad H_A: RR > 1$$

$$\text{Lower: } H_0: RR \geq 1 \quad \text{vs.} \quad H_A: RR < 1$$

Null Variance Calculation Method

Choose how the variance under the null hypothesis will be calculated using the values for the event rates λ_1 and λ_2 . This choice affects the overall power calculations.

Three different methods are available:

- **Method 1: Use Event Rate from Group 1 (λ_1)**
- **Method 2: Use True Event Rates (λ_1 and λ_2)**
- **Method 3: Use Maximum Likelihood Estimation**

See the section “Estimating the Variance under the Null and Alternative Hypotheses” above for a description of each of these choices.

Recommendation

Simulation studies suggest that methods 2 and 3 perform better than method 1 (see Zhu and Lakkis (2014), page 385). **PASS** uses method 3 by default because it is the method used in the sample size calculation example given on page 382 of Zhu and Lakkis (2014).

Tests for the Ratio of Two Negative Binomial Rates

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

$\mu(t)$ (Average Subject Exposure Time)

Enter a value (or range of values) for the average time of exposure for all subjects in the study. The unit of time measurement for this value should be consistent with the unit of time used to enter λ_1 and λ_2 . The range of acceptable values is $\mu(t) > 0$. You can enter a single value such as *1* or a series of values such as *0.8 to 1.0 by 0.1*.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on N_1 or N_2 or both.

The options are

- **Equal ($N_1 = N_2$)**
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 N_1 , solve for N_2**
Select this option when you wish to fix N_1 at some value (or values), and then solve only for N_2 . Please note that for some values of N_1 , there may not be a value of N_2 that is large enough to obtain the desired power.
- **Enter N_2 , solve for N_1**
Select this option when you wish to fix N_2 at some value (or values), and then solve only for N_1 . Please note that for some values of N_2 , there may not be a value of N_1 that is large enough to obtain the desired power.
- **Enter $R = N_2/N_1$, solve for N_1 and N_2**
For this choice, you set a value for the ratio of N_2 to N_1 , and then PASS determines the needed N_1 and N_2 , with this ratio, to obtain the desired power. An equivalent representation of the ratio, R , is

$$N_2 = R * N_1.$$
- **Enter percentage in Group 1, solve for N_1 and N_2**
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 N_1 and N_2 with this percentage to obtain the desired power.

Tests for the Ratio of Two Negative Binomial Rates

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 $\geq Y$.

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 Not 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 $N2$.

Tests for the Ratio of Two Negative Binomial Rates

- **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:

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

where the value $[Y]$ is the next integer $\geq Y$.

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.

Tests for the Ratio of Two Negative Binomial Rates

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 – Event (Incidence) Rates

λ_1 (Event Rate of Group 1)

Enter a value (or range of values) for the mean event rate per time unit in group 1 (control).

Example of Estimating λ_1

If 200 patients were exposed for 1 year (i.e. $\mu(t) = 1$ year) and 40 experienced the event of interest, then the mean event rate would be

$$\lambda_1 = 40/(200*1) = 0.2 \text{ per patient-year}$$

If 200 patients were exposed for 2 years (i.e. $\mu(t) = 2$ years) and 40 experienced the event of interest, then the mean event rate would be

$$\lambda_1 = 40/(200*2) = 0.1 \text{ per patient-year}$$

Event Rate Ratio

λ_1 is used with λ_2 to calculate the event rate ratio as

$$RR = \lambda_2/\lambda_1$$

such that

$$\lambda_1 = \lambda_2/RR$$

The range of acceptable values is $\lambda_1 > 0$. You can enter a single value such as 1 or a series of values such as 1 to 2 by 0.5.

Enter λ_2 or Ratio for Group 2

Indicate whether to enter the group 2 event rate (λ_2) or the event rate ratio (RR) to specify the effect size.

The event rate ratio is calculated from λ_2 and λ_1 as

$$RR = \lambda_2/\lambda_1$$

RR (Ratio of Event Rates)

This option is displayed only if Enter λ_2 or Ratio for Group 2 = "RR (Ratio of Event Rates)."

This is the value of the ratio of the two event rates, λ_1 and λ_2 , at which the power is to be calculated.

The event rate ratio is calculated from λ_1 and λ_2 as

$$RR = \lambda_2/\lambda_1$$

The range of acceptable values is $RR > 0$ and $RR \neq 1$.

You can enter a single value such as 1.15 or a series of values such as 1.05 to 1.20 by 0.05.

RR (Ratio of Event Rates) [when solving for RR (Ratio of Event Rates)]

Specify whether to solve for event rate ratios that are less than one or greater than one.

Tests for the Ratio of Two Negative Binomial Rates

λ_2 (Event Rate of Group 2)

This option is displayed only if Enter λ_2 or Ratio for Group 2 = " λ_2 (Event Rate of Group 2)."

Enter a value (or range of values) for the mean event rate per time unit in group 2 (treatment).

Example of Estimating λ_2

If 200 patients were exposed for 1 year (i.e. $\mu(t) = 1$ year) and 40 experienced the event of interest, then the mean event rate would be

$$\lambda_2 = 40/(200*1) = 0.2 \text{ per patient-year}$$

If 200 patients were exposed for 2 years (i.e. $\mu(t) = 2$ years) and 40 experienced the event of interest, then the mean event rate would be

$$\lambda_2 = 40/(200*2) = 0.1 \text{ per patient-year}$$

Event Rate Ratio

λ_2 is used with λ_1 to calculate the event rate ratio as

$$RR = \lambda_2/\lambda_1$$

such that

$$\lambda_2 = RR*\lambda_1$$

The range of acceptable values is $\lambda_2 > 0$. You can enter a single value such as 1 or a series of values such as 1 to 2 by 0.5 .

Effect Size – Negative Binomial Distribution

κ (Negative Binomial Dispersion)

Enter a value (or range of values) for the dispersion parameter of the negative binomial distribution. This value is difficult to obtain in some situations. In these cases, your best option is to estimate κ from the Poisson overdispersion factor, ϕ , calculated from a similar study that has already been completed.

The range of acceptable values is $\kappa \geq 0$. You can enter a single value such as 1 or a series of values such as 0.8 to 1 by 0.1 .

Calculating κ from the Poisson Overdispersion Factor, ϕ

A suitable value for κ can be estimated from the Poisson overdispersion factor, ϕ , and the average event rate across groups, λ , from a previous study as

$$\kappa = (\phi - 1)/\lambda$$

For additional details and suggestions about calculating κ from a previous study, see the section "Estimating the Negative Binomial Dispersion Parameter, κ , from a Previous Study that was Analyzed using Poisson Regression" above.

Example 1 – Finding the Sample Size (Validation 1 using Zhu and Lakkis (2014))

Zhu and Lakkis (2014) conducted numerous simulation studies to investigate the performance of the three null variance calculation methods. We'll use their example to demonstrate how to calculate sample size for various effect sizes, group 1 event rates, event rate ratios, and negative binomial dispersion values.

The settings for this example are similar to those that might be encountered when studying exacerbation events in COPD studies. A balanced, parallel study is designed to detect a 10-20% rate reduction from the control to the treatment group when the reference rate is about one exacerbation event per patient per year. They studied control rates of 0.8, 1.0, 1.2, and 1.4 events per patient-year and a rate ratio of 0.85, representing a 15% reduction in the treatment event rate relative to the control. They also included a rate ratio of 1.15 to represent a 15% increase in the treatment rate relative to the control for completeness. They assumed an average subject exposure time of 0.75 years since similar studies usually see discontinued participation by some patients for various reasons, resulting in an average exposure time that is less than the usual designed study length of 1 year. They studied dispersion parameter values of 0.4, 0.7, 1.0, and 1.5. They calculate the sample sizes required to achieve 80% power at a significance level of 0.05. They investigated all three null variance calculation methods, so this example will be presented in 3 parts, one for each method.

Their sample size calculation results are given in Table I on page 381. By running this example, you'll see that **PASS** matches their sample size calculation results exactly for all parameter combinations and for all three variance calculation methods.

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 Negative Binomial Rates** procedure window. You may then make the appropriate entries as listed below, or open **Example 1a** 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
Null Variance Calculation Method	Use Maximum Likelihood Estimation
Power	0.8
Alpha	0.05
$\mu(t)$ (Average Subject Exposure Time) ...	0.75
Group Allocation	Equal (N1 = N2)
λ_1 (Event Rate of Group 1)	0.8 1.0 1.2 1.4
Enter λ_2 or Ratio for Group 2	RR (Ratio of Event Rates)
RR (Ratio of Event Rates)	0.85 1.15
κ (Negative Binomial Dispersion)	0.4 0.7 1.0 1.5

Tests for the Ratio of Two Negative Binomial Rates

Annotated Output

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

Numeric Results

Numeric Results for Testing the Ratio of Two Negative Binomial Rates using the Wald or LR Test

Alternative Hypothesis: Two-Sided (H0: RR = 1 vs. Ha: RR ≠ 1)

Null Variance Calculation Method: Use Maximum Likelihood Estimation

Power	N1	N2	N	Ave Expos Time $\mu(t)$	Grp 1 Event Rate λ_1	Grp 2 Event Rate λ_2	Event Rate Ratio RR	Neg Binom Disp κ	Alpha
0.80008	1311	1311	2622	0.75	0.80	0.68	0.8500	0.4	0.050
0.80025	1490	1490	2980	0.75	0.80	0.68	0.8500	0.7	0.050
0.80016	1668	1668	3336	0.75	0.80	0.68	0.8500	1.0	0.050
0.80010	1965	1965	3930	0.75	0.80	0.68	0.8500	1.5	0.050
0.80019	1570	1570	3140	0.75	0.80	0.92	1.1500	0.4	0.050
0.80015	1811	1811	3622	0.75	0.80	0.92	1.1500	0.7	0.050
0.80011	2052	2052	4104	0.75	0.80	0.92	1.1500	1.0	0.050
0.80012	2454	2454	4908	0.75	0.80	0.92	1.1500	1.5	0.050
0.80031	1097	1097	2194	0.75	1.00	0.85	0.8500	0.4	0.050
0.80017	1275	1275	2550	0.75	1.00	0.85	0.8500	0.7	0.050
0.80007	1453	1453	2906	0.75	1.00	0.85	0.8500	1.0	0.050
0.80002	1750	1750	3500	0.75	1.00	0.85	0.8500	1.5	0.050
0.80010	1320	1320	2640	0.75	1.00	1.15	1.1500	0.4	0.050
0.80006	1561	1561	3122	0.75	1.00	1.15	1.1500	0.7	0.050
0.80003	1802	1802	3604	0.75	1.00	1.15	1.1500	1.0	0.050
0.80006	2204	2204	4408	0.75	1.00	1.15	1.1500	1.5	0.050
0.80038	954	954	1908	0.75	1.20	1.02	0.8500	0.4	0.050
0.80022	1132	1132	2264	0.75	1.20	1.02	0.8500	0.7	0.050
0.80010	1310	1310	2620	0.75	1.20	1.02	0.8500	1.0	0.050
0.80004	1607	1607	3214	0.75	1.20	1.02	0.8500	1.5	0.050
0.80024	1154	1154	2308	0.75	1.20	1.38	1.1500	0.4	0.050
0.80017	1395	1395	2790	0.75	1.20	1.38	1.1500	0.7	0.050
0.80012	1636	1636	3272	0.75	1.20	1.38	1.1500	1.0	0.050
0.80013	2038	2038	4076	0.75	1.20	1.38	1.1500	1.5	0.050
0.80006	851	851	1702	0.75	1.40	1.19	0.8500	0.4	0.050
0.80031	1030	1030	2060	0.75	1.40	1.19	0.8500	0.7	0.050
0.80017	1208	1208	2416	0.75	1.40	1.19	0.8500	1.0	0.050
0.80009	1505	1505	3010	0.75	1.40	1.19	0.8500	1.5	0.050
0.80020	1035	1035	2070	0.75	1.40	1.61	1.1500	0.4	0.050
0.80013	1276	1276	2552	0.75	1.40	1.61	1.1500	0.7	0.050
0.80009	1517	1517	3034	0.75	1.40	1.61	1.1500	1.0	0.050
0.80011	1919	1919	3838	0.75	1.40	1.61	1.1500	1.5	0.050

References

Zhu, H. and Lakkis, H. 2014. 'Sample Size Calculation for Comparing Two Negative Binomial Rates.' Statistics in Medicine, Volume 33, Pages 376-387.

Report Definitions

Power is the probability of rejecting the null hypothesis when it is false. It should be close to one.

N1 and N2 are the number of subjects in groups 1 and 2, respectively.

N is the total sample size. $N = N1 + N2$.

$\mu(t)$ is the average exposure time across all subjects.

λ_1 is the mean event rate per time unit in group 1 (control). This is often the baseline event rate.

λ_2 is the mean event rate per time unit in group 2 (treatment).

RR is the ratio of the two event rates. $RR = \lambda_2/\lambda_1$.

κ is the negative binomial dispersion parameter.

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

Summary Statements

For a two-sided Wald or likelihood ratio test of the null hypothesis H0: RR = 1 vs. the alternative Ha: RR ≠ 1, samples of 1311 subjects in group 1 and 1311 subjects in group 2 achieve 80.008% power to detect an event rate ratio (RR) of 0.8500 when the event rate in group 1 (λ_1) is 0.80, the average exposure time in both groups ($\mu(t)$) is 0.75, the negative binomial dispersion parameter (κ) is 0.4, the significance level (alpha) is 0.050, and the null variance is calculated using maximum likelihood estimation.

Tests for the Ratio of Two Negative Binomial Rates

This report shows the sample size for each combination of the input parameters, one scenario per row. If you look at the results in the column labeled M3 under “Calculated N/group” in Table I on page 381 of Zhu and Lakkis (2014), you’ll see that the group sample sizes calculated by **PASS** match those exactly in all cases (though they are presented in a different order).

Power

Power is the probability of rejecting the null hypothesis when it is false. It should be close to one.

N1 and N2

N1 and N2 are the number of subjects in groups 1 and 2, respectively. To calculate the sample size in terms of person-time, you would multiply the sample size by the average exposure time, $\mu(t)$. For example, the group sample size in the first row calculated in person-years is $1311 \times 0.75 = 983.25$ person-years.

N

N is the total sample size. $N = N1 + N2$.

$\mu(t)$

$\mu(t)$ is the average exposure time across all subjects.

λ_1

λ_1 is the mean event rate per time unit in group 1 (control). This is often the baseline event rate.

λ_2

λ_2 is the mean event rate per time unit in group 2 (treatment).

RR

RR is the ratio of the two event rates. $RR = \lambda_2/\lambda_1$.

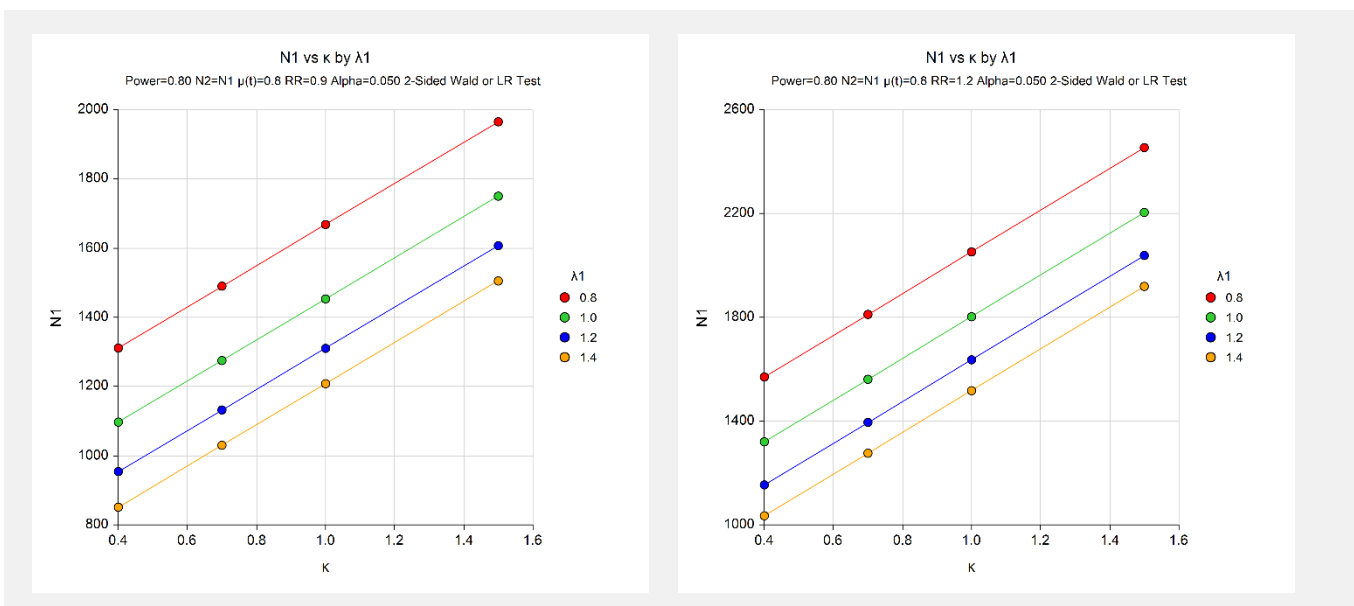
κ

κ is the negative binomial dispersion parameter.

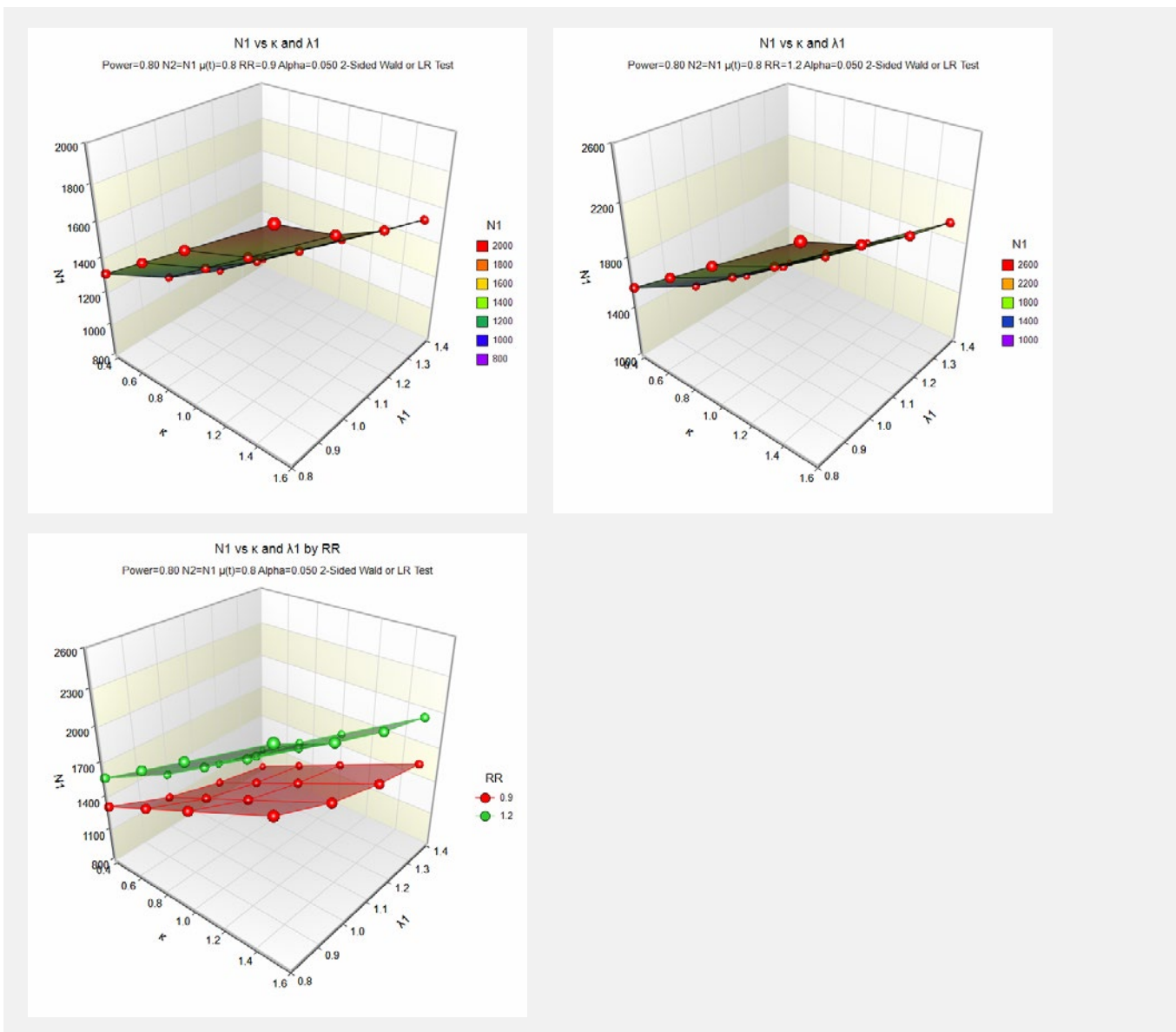
Alpha

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

Plots Section



Tests for the Ratio of Two Negative Binomial Rates



These plots show the relationship between sample size, κ , λ_1 , and the event rate ratio, RR.

To match the results for Method 1 (M1) in Table I of Zhu and Lakkis (2014) on page 381, change Null Variance Calculation Method to “Use Event Rate from Group 1 (λ_1)” in **PASS** and then re-calculate the sample size (or load the example template **Example 1b**). The results are not displayed here.

To match the results for Method 2 (M2) in Table I of Zhu and Lakkis (2014) on page 381, change Null Variance Calculation Method to “Use True Event Rates (λ_1 and λ_2)” in **PASS** and then re-calculate the sample size (or load the example template **Example 1c**). The results are not displayed here.

Example 2 – Estimating the Negative Binomial Dispersion Parameter from a Previous Study (Validation 2 using Zhu and Lakkis (2014))

Zhu and Lakkis (2014) suggests that the hardest parameter to obtain in these sample size and power calculations is usually the negative binomial dispersion parameter, κ . When a suitable value for κ is not known, then you can estimate κ from a similar study that was analyzed using Poisson regression. Given a Poisson mean event rate λ and an overdispersion factor ϕ , estimated from a similar Poisson regression study, κ can be estimated as

$$\hat{\kappa} = \frac{\hat{\phi} - 1}{\hat{\lambda}}$$

Zhu and Lakkis (2014) go even further and demonstrate how to obtain the Poisson overdispersion factor estimate when it is not reported directly. (See the section “Estimating the Negative Binomial Dispersion Parameter, κ , from a Previous Study that was Analyzed using Poisson Regression” above for details).

In their example (pages 382 through 384), a one-year study is being designed for a drug that is intended to reduce asthma exacerbations in a particular set of patients. They found a similar study to that being designed that would provide a basis for the sample size calculations of the present study. The previous study was analyzed using Poisson regression, but did not report the estimated overdispersion factor directly. It did report total exposure times, estimated exacerbation rates, and asymmetric 95% confidence limits for both groups:

Group 1 (Placebo): Exposure Time = 397 patient-years, $\hat{\lambda}_1 = 0.663$, 95% Confidence Interval = (0.573, 0.768)

Group 2 (Tiotropium): Exposure Time = 399 patient-years, $\hat{\lambda}_2 = 0.530$, 95% Confidence Interval = (0.450, 0.625)

Using this information and the equations presented above, they back-calculate $SE(\log \hat{\lambda}_1)$ and $SE(\log \hat{\lambda}_2)$ from the 95% confidence limits as

$$SE(\log \hat{\lambda}_1) = \frac{\left(\frac{\log(\text{Upper Bound}) - \log(\text{Lower Bound})}{2} \right)}{z_{1-\alpha/2}} = \frac{\left(\frac{\log(0.768) - \log(0.573)}{2} \right)}{1.96} = 0.0747$$

$$SE(\log \hat{\lambda}_2) = \frac{\left(\frac{\log(\text{Upper Bound}) - \log(\text{Lower Bound})}{2} \right)}{z_{1-\alpha/2}} = \frac{\left(\frac{\log(0.625) - \log(0.450)}{2} \right)}{1.96} = 0.0838$$

From this they calculate $\hat{\phi}_1$ and $\hat{\phi}_2$ as

$$\hat{\phi}_1 = T_1 \hat{\lambda}_1 \left(SE(\log \hat{\lambda}_1) \right)^2 = 397 \times 0.663 \times (0.0747)^2 = 1.47$$

$$\hat{\phi}_2 = T_2 \hat{\lambda}_2 \left(SE(\log \hat{\lambda}_2) \right)^2 = 399 \times 0.530 \times (0.0838)^2 = 1.49$$

Finally, an estimate for the negative binomial dispersion κ is calculated as

$$\hat{\kappa} = \frac{\left(\frac{\hat{\phi}_1 + \hat{\phi}_2}{2} \right) - 1}{\left(\frac{\hat{\lambda}_1 + \hat{\lambda}_2}{2} \right)} = \frac{\left(\frac{1.47 + 1.49}{2} \right) - 1}{\left(\frac{0.663 + 0.530}{2} \right)} = 0.8$$

Using this estimated negative binomial dispersion of 0.8, the placebo rate of 0.66 events per patient-year, an event rate ratio of 0.8 (representing a 20% reduction in exacerbations with the new drug), an average exposure time of 0.9 years (since not everybody is expected to be followed for the full year), and using method 3 to estimate the null variance, they calculate a per-group sample size of 1131 for the new study to achieve 90% power at a significance level of 0.05. **PASS** matches this result exactly as follows.

Tests for the Ratio of Two Negative Binomial Rates

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 Negative Binomial Rates** 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**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided
Null Variance Calculation Method	Use Maximum Likelihood Estimation
Power	0.90
Alpha	0.05
$\mu(t)$ (Average Subject Exposure Time) ...	0.9
Group Allocation	Equal (N1 = N2)
λ_1 (Event Rate of Group 1)	0.66
Enter λ_2 or Ratio for Group 2	RR (Ratio of Event Rates)
RR (Ratio of Event Rates)	0.8
κ (Negative Binomial Dispersion)	0.8

Output

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

Numeric Results**Numeric Results for Testing the Ratio of Two Negative Binomial Rates using the Wald or LR Test**Alternative Hypothesis: Two-Sided (H0: RR = 1 vs. Ha: RR \neq 1)

Null Variance Calculation Method: Use Maximum Likelihood Estimation

				Ave Expos Time $\mu(t)$	Grp 1 Event Rate λ_1	Grp 2 Event Rate λ_2	Event Rate Ratio RR	Neg Binom Disp κ	Alpha
Power	N1	N2	N						
0.90000	1131	1131	2262	0.90	0.66	0.53	0.8000	0.8	0.050

The group sample size of 1131 calculated by **PASS** matches the result obtained by Zhu and Lakkis (2014) exactly.