

Chapter 245

Non-Inferiority Tests for the Difference Between Two Poisson Rates in a Cluster-Randomized Design

Introduction

This procedure calculates power and sample size for non-inferiority tests of two rates in a cluster-randomized design in which the outcome variable is a count. It uses the work of Wang, Zhang, and Ahn (2018) which give the power for the case of varying cluster sizes. The analysis uses a simple z-test comparing the two rates.

Cluster-randomized designs are those in which whole clusters of subjects (classes, hospitals, communities, etc.) are put into the treatment group or the control group. Generally speaking, the larger the cluster sizes and the higher the correlation among subjects within the same cluster, the larger will be the overall sample size necessary to detect an effect with the same power.

The Statistical Hypotheses

Non-inferiority tests are examples of directional (one-sided) tests. This program module provides the input and output in formats that are convenient for these types of tests. This section will review the specifics of non-inferiority testing.

Remember that in the usual test setting, the null (H_0) and alternative (H_1) hypotheses for one-sided tests are defined as follows, assuming that $\delta = \mu_1 - \mu_2$ is to be tested.

$$H_0: \delta \leq 0 \quad \text{versus} \quad H_1: \delta > 0$$

Rejecting this test implies that the mean difference is larger than the value δ . This test is called an *upper-tailed test* because it is rejected in samples in which the difference between the sample means is larger than D .

Following is an example of a *lower-tailed test*.

$$H_0: \delta \geq 0 \quad \text{versus} \quad H_1: \delta < 0$$

Non-inferiority tests are special cases of the above directional tests. It will be convenient to adopt the following specialized notation for the discussion of these tests.

Non-Inferiority Tests for the Difference Between Two Poisson Rates in a Cluster-Randomized Design

<u>Parameter</u>	<u>PASS Input/Output</u>	<u>Interpretation</u>
λ_1	λ_1	<i>Incidence rate</i> of population 1. Population 1 is assumed to consist of those who have received the new treatment.
λ_2	λ_2	<i>Incidence rate</i> of population 2. Population 2 is assumed to consist of those who have received the control (standard or reference) treatment.
ε	D0	<i>Margin of non-inferiority</i> . This is a tolerance value that defines the magnitude of the amount that is not of practical importance. This may be thought of as the largest change from the baseline that is considered to be trivial. The absolute value is shown to emphasize that this is a magnitude. The sign of the value will be determined by the specific design that is being used.
δ	D1	<i>True difference</i> . This is the value of $\lambda_1 - \lambda_2$, the difference between the rates.

Non-Inferiority Tests

A *non-inferiority test* tests that the treatment rate is not worse than the control rate by more than the non-inferiority margin. The actual direction of the hypothesis depends on the nature of the response variable being studied.

Case 1: High Values are Better

In this case, higher values are better. The hypotheses are arranged so that rejecting the null hypothesis implies that the treatment rate is no less than a small amount below the control rate. The value of δ is often set to zero. The following are equivalent sets of hypotheses.

$$\begin{aligned}
 H_0: \delta \leq -\varepsilon & \text{ versus } H_1: \delta > -\varepsilon, \varepsilon > 0 \\
 H_0: \lambda_1 - \lambda_2 \leq -\varepsilon & \text{ versus } H_1: \lambda_1 - \lambda_2 > -\varepsilon, \varepsilon > 0 \\
 H_0: \lambda_1 \leq \lambda_2 - \varepsilon & \text{ versus } H_1: \lambda_1 > \lambda_2 - \varepsilon, \varepsilon > 0
 \end{aligned}$$

Case 2: High Values are Worse

In this case, lower values are better. The hypotheses are arranged so that rejecting the null hypothesis implies that the treatment rate is no more than a small amount above the control rate. The value of δ is often set to zero. The following are equivalent sets of hypotheses.

$$\begin{aligned}
 H_0: \delta \geq \varepsilon & \text{ versus } H_1: \delta < \varepsilon, \varepsilon > 0 \\
 H_0: \lambda_1 - \lambda_2 \geq \varepsilon & \text{ versus } H_1: \lambda_1 - \lambda_2 < \varepsilon, \varepsilon > 0 \\
 H_0: \lambda_1 \geq \lambda_2 + \varepsilon & \text{ versus } H_1: \lambda_1 < \lambda_2 + \varepsilon, \varepsilon > 0
 \end{aligned}$$

Technical Details

The following discussion summarizes the results in Wang, Zhang, and Ahn (2018).

Suppose you are interested in comparing the incidence rates of two groups (treatment and control) with a non-inferiority test. Further suppose that the response is known to be related to other covariates (such as age, race, or gender) and so their impact needs to be adjusted for. This may be accomplished by stratifying on the covariates and forming hypotheses about a common mean difference across all clusters and strata. Often, the stratification is based on cluster size, but this is not required.

Let Y_{jki} be the count outcome of the i th ($i = 1, \dots, M_{jk}$) subject in the k th ($k = 1, \dots, K_j$) cluster of the j th ($j = 1, 2$) group. Assuming that Y_{jki} follows a Poisson model with

$$\lambda_j = E(Y_{jki}) = \text{Var}(Y_{jki})$$

and a common intraclass correlation coefficient (ICC) $\rho = \text{corr}(Y_{jki}, Y_{jki'})$ for $i \neq i'$.

Test Statistic

An unbiased estimator of λ_j is

$$\hat{\lambda}_j = \frac{\sum_{k=1}^{K_j} \sum_{i=1}^{M_{jk}} Y_{jki}}{\sum_{k=1}^{K_j} M_{jk}}$$

with estimated variance

$$s_j^2 = \frac{\hat{\lambda}_j \sum_{k=1}^{K_j} M_{jk} [1 + (M_{jk} - 1)\hat{\rho}]}{\left(\sum_{k=1}^{K_j} M_{jk}\right)^2}$$

Using a one-sided hypothesis $H_0: \lambda_1 = \lambda_2$ versus $H_1: \lambda_1 < \lambda_2$, H_0 is rejected if

$$\frac{\hat{\lambda}_1 - \hat{\lambda}_2}{\sqrt{s_1^2 + s_2^2}} > z_{1-\alpha}$$

Sample Size and Power

Wang, Zhang, and Ahn (2018) provide the following formula for estimating K_2 (the number of clusters in the control group) when there are an unequal number of clusters in each group.

$$K_2 = \frac{\left(z_{1-\frac{\alpha}{2}} + z_{1-\beta}\right)^2 \left(\frac{\lambda_1}{R} + \lambda_2\right)}{(\lambda_1 - \lambda_2)^2} \left[(1 - \rho) \frac{1}{\theta} + \rho + \rho\gamma^2 \right]$$

where $R = K_1/K_2$, $M = E(M_{jk})$, $\tau^2 = \text{Var}(M_{jk})$, and $\gamma = \tau/M$. Here M is the average cluster size of all clusters in the study and γ is the coefficient of variation of the cluster sizes. Note that $z_x = \Phi(x)$ is the standard normal distribution function.

This equation can easily be rearranged to provide a formula for power $(1 - \beta)$.

Example 1 – Finding the Number of Clusters

A non-inferiority study is being planned to investigate whether a new intervention will decrease the incidence rate of a certain disease over the rate achieved by the current intervention. This response is a count. For a number of reasons, the researchers decide to administer the intervention to whole clusters (clinics) rather than randomize the treatment to individuals within a cluster. The number of clinics receiving each treatment will be balanced.

The average number of subjects per clinic is 21. The coefficient of variation of the cluster sizes is 0.42.

Prior studies have obtained an incidence rate of 0.35 for the current treatment and an ICC of 0.07. The non-inferiority margin is set to 0.05. The researchers want to compare the necessary sample size when the actual incidence rate of the new treatment is 0.3, 0.25, 0.2, and 0.15.

The one-sided significance level is set to 0.025 and the power is set to 0.8.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load this procedure. 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	K1 (Number of Clusters in Group 1)
Higher Rates Are	Worse (H1: $\lambda_1 - \lambda_2 < D_0$, where $D_0 > 0$)
Power.....	0.80
Alpha.....	0.025
K2 (Clusters in Group 2).....	K1
M (Average of Cluster Sizes).....	21
CV (Coef of Variation of Cluster Sizes)	0.42
Incidence Rate Input Type.....	Differences (D0 and D1)
D0 (Non-Inferiority Difference).....	0.05
D1 (Actual Difference)	-0.2 -0.15 -0.1 -0.05
λ_2 (Incidence Rate of Control Group).....	0.35
ρ (Intracluster Correlation, ICC).....	0.07

Annotated Output

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

Numeric Results

Numeric Results													
Higher Rates Are:		Worse											
Hypotheses:		H0: $\lambda_1 - \lambda_2 \geq D_0$ vs. H1: $\lambda_1 - \lambda_2 < D_0$											
Power	Sample Size	Clusters		Cluster Size		Incidence Rates			Rate Differences			Alpha	
		Grp 1	Grp 2	Ave	COV	N.I.	Act	Cntl	N.I.	Act	ICC		
	N	K1	K2	M	CV	$\lambda_{1.0}$	$\lambda_{1.1}$	λ_2	D0	D1	ρ		
0.80060	9198	219	219	21	0.42	0.4	0.15	0.35	0.05	-0.20	0.07	0.025	
0.80075	10038	239	239	21	0.42	0.4	0.20	0.35	0.05	-0.15	0.07	0.025	
0.80087	10878	259	259	21	0.42	0.4	0.25	0.35	0.05	-0.10	0.07	0.025	
0.80098	11718	279	279	21	0.42	0.4	0.30	0.35	0.05	-0.05	0.07	0.025	

Non-Inferiority Tests for the Difference Between Two Poisson Rates in a Cluster-Randomized Design

References

Wang, J., Zhang, S., and Ahn, C. 2018. 'Sample size calculation for count outcomes in cluster randomization trials with varying cluster sizes.' Communications in Statistics--Theory and Methods, DOI: 10.1080/03610926.2018.1532004.

Report Definitions

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

N is the total sample size, i.e., the total number of subjects from all clusters.

K1 is the number of clusters assigned to group 1, the treatment group.

K2 is the number of clusters assigned to group 2, the control group.

M is the average cluster size. It is the anticipated average number of subjects per cluster.

CV is the coefficient of variation of the cluster sizes. It is the standard deviation of cluster sizes divided by the average cluster size.

$\lambda_{1.0}$ is the non-inferiority bound on incidence rate in the treatment group.

$\lambda_{1.1}$ is the actual incidence rate of the treatment group at which the power is computed.

λ_2 is the incidence rate in the control group.

D0 is the difference in incidence rates assumed by the null hypothesis. This is often called the non-inferiority margin. $D0 = \lambda_{1.0} - \lambda_2$.

D1 is the difference in incidence rates assumed by the alternative hypothesis. This is the difference at which the power is computed. $D1 = \lambda_{1.1} - \lambda_2$.

ρ is the intracluster correlation coefficient. This is the correlation between any two subjects within a particular cluster.

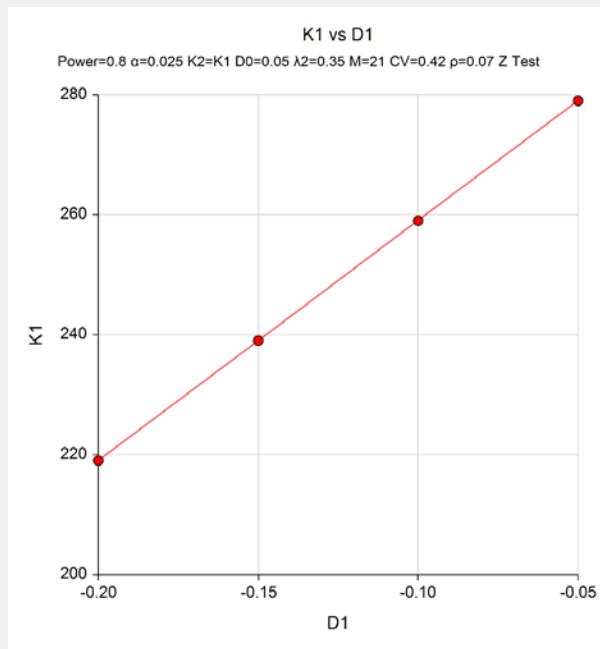
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements

A cluster-randomized, non-inferiority design consists of 219 treatment clusters and 219 control clusters. Each cluster is assumed to have an average of 21 subjects per cluster. The outcome of the study will be analyzed by comparing the difference in incidence rates of a treatment and a control group to a non-inferiority margin. Higher values of these incidence rates are assumed to be worse. The expected sample size of 9198 subjects achieves 80% power to reject inferiority when the actual difference is -0.2. The non-inferiority margin of the incidence rate difference is 0.05. The incidence rate of the control group 0.35. The incidence rate of the treatment group assumed by the alternative hypothesis is 0.15. The incidence rate of the treatment group which is the non-inferiority bound is 0.4. The intracluster correlation coefficient is 0.07. The significance level of the test is 0.025. The individual cluster sizes (the number of subjects per cluster) are assumed to vary according to a discrete distribution with mean 21 and coefficient of variation 0.42.

Chart Section

Chart Section



The values from the Numeric Results report are displayed in this plot.

Example 2 – Validation using a Previously Validated Procedure

We could not find a validation example in the literature so we will use a previously validated procedure in **PASS** to validate this procedure.

Suppose in a non-inferiority design, higher rates are worse, power = 0.9, alpha = 0.025, $K_2 = K_1$, $M = 50$, $CV = 0.2$, $\lambda_{1.0} = 0.6$, $\lambda_{1.1} = 0.5$, $\lambda_2 = 0.5$, and $\rho = 0.002$. Solve for K_1 .

This scenario can be solved using the procedure “Tests for the Difference Between Two Poisson Rates in a Cluster-Randomized Design” with the following settings.

Set the alternative hypothesis to “One-Sided ($H_1: \delta > 0$)”, power = 0.9, alpha = 0.025, $K_2 = K_1$, $M = 50$, $CV = 0.2$, $\lambda_1 = 0.6$, $\lambda_2 = 0.5$, and $\rho = 0.002$. The solution is $K_1 = K_2 = 26$.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the **PASS** Home window, load this procedure. 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	K1 (Number of Clusters in Group 1)
Higher Rates Are	Worse ($H_1: \lambda_1 - \lambda_2 < D_0$, where $D_0 > 0$)
Power.....	0.9
Alpha.....	0.025
K_2 (Clusters in Group 2).....	K1
M (Average of Cluster Sizes).....	50
CV (Coef of Variation of Cluster Sizes)	0.2
Incidence Rate Input Type.....	Incidence Rates ($\lambda_{1.0}$ and $\lambda_{1.1}$)
$\lambda_{1.0}$ (Non-Inferiority Incidence Rate).....	0.6
$\lambda_{1.1}$ (Actual Incidence Rate)	0.5
λ_2 (Incidence Rate of Control Group).....	0.5
ρ (Intracluster Correlation, ICC).....	0.002

Output

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

Numeric Results

Numeric Results													
Higher Rates Are: Worse													
Hypotheses: $H_0: \lambda_1 - \lambda_2 \geq D_0$ vs. $H_1: \lambda_1 - \lambda_2 < D_0$													
	Sample Size	Clusters		Cluster Size		Incidence Rates			Rate Differences			ICC	Alpha
		Grp 1	Grp 2	Ave	COV	N.I.	Act	Cntl	N.I.	Act	D0		
Power	N	K1	K2	M	CV	$\lambda_{1.0}$	$\lambda_{1.1}$	λ_2	D0	D1	ρ		
0.90572	2600	26	26	50	0.2	0.6	0.5	0.5	0.1	0	0.002		0.025

PASS has also obtained $K_1 = K_2 = 26$. Thus, the procedure is validated.