Chapter 246

Superiority by a Margin Tests for the Difference Between Two Poisson Rates in a Cluster-Randomized Design

Introduction

This procedure calculates power and sample size for superiority by a margin 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

Superiority by a margin 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 superiority by a margin testing.

Remember that in the usual test setting, the null (H₀) and alternative (H₁) hypotheses for one-sided tests are defined as follows, assuming that \( \delta = \mu_1 - \mu_2 \) is to be tested.

\[
H₀: \delta \leq 0 \quad \text{versus} \quad H₁: \delta > 0
\]

Rejecting this test implies that the mean difference is larger than the value \( \delta \). 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.

\[
H₀: \delta \geq 0 \quad \text{versus} \quad H₁: \delta < 0
\]

Superiority by a margin 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.
PASS Sample Size Software

Superiority by a Margin Tests for the Difference of Two Poisson Rates in a Cluster-Randomized Design

<table>
<thead>
<tr>
<th>Parameter</th>
<th>PASS Input/Output</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>$\lambda_1$</td>
<td>Incidence rate of population 1. Population 1 is assumed to consist of those who have received the new treatment.</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>$\lambda_2$</td>
<td>Incidence rate of population 2. Population 2 is assumed to consist of those who have received the control (standard or reference) treatment.</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>$D0$</td>
<td>Margin of Superiority. 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.</td>
</tr>
<tr>
<td>$\delta$</td>
<td>$D1$</td>
<td>True difference. This is the value of $\lambda_1 - \lambda_2$, the difference between the rates.</td>
</tr>
</tbody>
</table>

Superiority by a Margin Tests

A superiority by a margin test tests that the treatment rate is better than the control rate by more than the superiority 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 some amount above the control rate. The following are equivalent sets of hypotheses.

\[
\begin{align*}
H_0: \delta & \leq \epsilon \quad \text{versus} \quad H_1: \delta > \epsilon, \ \epsilon > 0 \\
H_0: \lambda_1 - \lambda_2 & \leq \epsilon \quad \text{versus} \quad H_1: \lambda_1 - \lambda_2 > \epsilon, \ \epsilon > 0 \\
H_0: \lambda_1 & \leq \lambda_2 + \epsilon \quad \text{versus} \quad H_1: \lambda_1 > \lambda_2 + \epsilon, \ \epsilon > 0
\end{align*}
\]

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 some amount below the control rate. The following are equivalent sets of hypotheses.

\[
\begin{align*}
H_0: \delta & \geq -\epsilon \quad \text{versus} \quad H_1: \delta < -\epsilon, \ \epsilon > 0 \\
H_0: \lambda_1 - \lambda_2 & \geq -\epsilon \quad \text{versus} \quad H_1: \lambda_1 - \lambda_2 < -\epsilon, \ \epsilon > 0 \\
H_0: \lambda_1 & \geq \lambda_2 - \epsilon \quad \text{versus} \quad H_1: \lambda_1 < \lambda_2 - \epsilon, \ \epsilon > 0
\end{align*}
\]
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 superiority by a margin 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 (\( i = 1, ..., M_{jk} \)) subject in the \( k \)th (\( k = 1, ..., K_j \)) 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 intracluster correlation coefficient (ICC) \( \rho = \text{corr}(Y_{jki}, Y_{jki'}) \) for \( i \neq i' \).

Test Statistic

An unbiased estimator of \( \lambda_j \) 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

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

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.

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

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 \( \gamma \) 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 superiority 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 superiority 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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the Example 1 settings file. To load these settings to the procedure window, click Open Example Settings File in the Help Center or File menu.
Output

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

Numeric Reports

<table>
<thead>
<tr>
<th>Power</th>
<th>Sample Size</th>
<th>Clusters</th>
<th>Cluster Size</th>
<th>Incidence Rates</th>
<th>Rate Diffs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Grp 1 K1</td>
<td>Ave M</td>
<td>Sup λ1.0</td>
<td>Sup D0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grp 2 K2</td>
<td>COV CV</td>
<td>Act λ1.1</td>
<td>Act D1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cntl λ2</td>
<td>ICC ρ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Alpha</td>
</tr>
<tr>
<td>0.80020</td>
<td>7518</td>
<td>179</td>
<td>21</td>
<td>0.3</td>
<td>-0.05</td>
</tr>
<tr>
<td>0.80042</td>
<td>8358</td>
<td>199</td>
<td>21</td>
<td>0.3</td>
<td>-0.05</td>
</tr>
<tr>
<td>0.80060</td>
<td>9198</td>
<td>219</td>
<td>21</td>
<td>0.3</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N The total sample size, i.e., the total number of subjects from all clusters.
K1 The number of clusters assigned to group 1, the treatment group.
K2 The number of clusters assigned to group 2, the control group.
M The average cluster size. It is the anticipated average number of subjects per cluster.
CV The coefficient of variation of the cluster sizes. It is the standard deviation of cluster sizes divided by the average cluster size.
λ1.0 The superiority bound on incidence rate in the treatment group.
λ1.1 The actual incidence rate of the treatment group at which the power is computed.
λ2 The incidence rate in the control group.
D0 The difference in incidence rates assumed by the null hypothesis. This is often called the superiority margin. D0 = λ1.0 - λ2.
D1 The difference in incidence rates assumed by the alternative hypothesis. This is the difference at which the power is computed. D1 = λ1.1 - λ2.
ρ The intracluster correlation coefficient. This is the correlation between any two subjects within a particular cluster.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A cluster-randomized, superiority by a margin design consists of 179 treatment clusters and 179 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 superiority margin. Higher values of these incidence rates are assumed to be worse. The expected sample size of 7518 subjects achieves 80% power to reject non-superiority when the actual difference is -0.2. The superiority 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 superiority bound is 0.3. 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.

References

Plots 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 superiority design, higher rates are better, power = 0.9, alpha = 0.025, K2 = K1, M = 50, CV = 0.2, λ1.0 = 0.6, λ1.1 = 0.5, λ2 = 0.5, and ρ = 0.002. Solve for K1.

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 (H1: δ > 0)”, power = 0.9, alpha = 0.025, K2 = K1, M = 50, CV = 0.2, λ1 = 0.6, λ2 = 0.5, and ρ = 0.002. The solution is K1 = K2 = 26.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

<table>
<thead>
<tr>
<th>Design Tab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solve For .......................................................(K_1) (Number of Clusters in Group 1)</td>
</tr>
<tr>
<td>Higher Rates Are ...........................................(H_1: \lambda_1 - \lambda_2 &gt; D_0, \text{ where } D_0 &gt; 0)</td>
</tr>
<tr>
<td>Power ...............................................................0.9</td>
</tr>
<tr>
<td>Alpha ...............................................................0.025</td>
</tr>
<tr>
<td>K2 (Clusters in Group 2) ................................ (K_1)</td>
</tr>
<tr>
<td>M (Average of Cluster Sizes) ............................50</td>
</tr>
<tr>
<td>CV (Coef of Variation of Cluster Sizes) ...........0.2</td>
</tr>
<tr>
<td>Incidence Rate Input Type ..................Incidence Rates ((\lambda_{1.0}) and (\lambda_{1.1}))</td>
</tr>
<tr>
<td>(\lambda_{1.0}) (Superiority Incidence Rate) ..............0.6</td>
</tr>
<tr>
<td>(\lambda_{1.1}) (Actual Incidence Rate) ...................0.5</td>
</tr>
<tr>
<td>(\lambda_{2}) (Incidence Rate of Control Group) ............0.5</td>
</tr>
<tr>
<td>(\rho) (Intracluster Correlation, ICC) ..................0.002</td>
</tr>
</tbody>
</table>
Output

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

<table>
<thead>
<tr>
<th>Numeric Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solve For: K1 (Number of Clusters in Group 1)</td>
</tr>
<tr>
<td>Higher Rates Are: Better</td>
</tr>
<tr>
<td>Hypotheses: H0: λ1 - λ2 ≤ D0 vs. H1: λ1 - λ2 &gt; D0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Power</th>
<th>Sample Size N</th>
<th>Clusters Grp 1 K1</th>
<th>Clusters Grp 2 K2</th>
<th>Cluster Size Ave M</th>
<th>COV CV</th>
<th>Incidence Rates Sup λ1.0</th>
<th>Act λ1.1</th>
<th>Cntl λ2</th>
<th>Rate Diffs Sup D0</th>
<th>Act D1</th>
<th>ICC ρ</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90572</td>
<td>2600</td>
<td>26</td>
<td>26</td>
<td>50</td>
<td>0.2</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.1</td>
<td>0</td>
<td>0.002</td>
<td>0.025</td>
</tr>
</tbody>
</table>

PASS has also obtained K1 = K2 = 26. Thus, the procedure is validated.