Chapter 247

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

Introduction

This procedure calculates power and sample size for superiority by a margin tests of the ratio 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.

Hypotheses

When higher rates are better, the superiority by a margin test hypotheses are

$$H_0: \frac{\lambda_1}{\lambda_2} \leq R_0 \quad \text{vs.} \quad H_1: \frac{\lambda_1}{\lambda_2} > R_0$$

where $R_0 > 1$.

When higher rates are worse, the superiority by a margin test hypotheses are

$$H_0: \frac{\lambda_1}{\lambda_2} \geq R_0 \quad \text{vs.} \quad H_1: \frac{\lambda_1}{\lambda_2} < R_0$$

where $R_0 < 1$.

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

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

and a common intracluster correlation coefficient (ICC) $\rho = \mathrm{corr}(Y_{jki}, Y_{jki'})$ 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

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

Using a one-sided hypothesis $H0: \lambda_1 = \lambda_2$ versus $H1: \lambda_1 < \lambda_2$, *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 $R = K_1/K_2$, $M = \mathrm{E}(M_{jk})$, $\tau^2 = \mathrm{Var}(M_{jk})$, 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 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. The 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 ratio is set to 0.86. The researchers want to compare the necessary sample size when the actual ratio is 0.43, 0.57, and 0.71.

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.

Design Tab	
Solve For	**K1 (Number of Clusters in Group 1)**
Higher Rates Are	**Worse (H1: λ1 / λ2 < R0, where R0 < 1)**
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	**Ratios (R0 and R1)**
R0 (Superiority Ratio)	**0.86**
R1 (Actual Ratio)	**0.43 0.57 0.71**
λ2 (Incidence Rate of Control Group)	**0.35**
ρ (Intracluster Correlation, ICC)	**0.07**

Output

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

Numeric Reports

Numeric Results

Solve For:	K1 (Number of Clusters in Group 1)
Groups:	1 = Treatment, 2 = Control
Higher Rates Are:	Worse
Hypotheses:	H0: λ1 / λ2 ≥ R0 vs. H1: λ1 / λ2 < R0

	Number of Clusters			Cluster Size		Total Sample Size	Incidence Rate: Treatment		Incidence Rate: Control	Ratio		Intracluster Correlation	
Power	Treatment K1	Control K2	Total K	Average M	CV	N	Superiority λ1.0	Actual λ1.1	λ2	Superiority R0	Actual R1	ρ	Alpha
0.80020	187	187	374	21	0.42	7854	0.301	0.1505	0.35	0.86	0.43	0.07	0.025
0.80153	208	208	416	21	0.42	8736	0.301	0.1995	0.35	0.86	0.57	0.07	0.025
0.80090	228	228	456	21	0.42	9576	0.301	0.2485	0.35	0.86	0.71	0.07	0.025

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
K1, K2, and K	The number of clusters in groups 1, 2, and both, respectively.
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.
N	The total sample size, i.e., the total number of subjects from all clusters.
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.
R0	The incidence rate ratio assumed by the null hypothesis. This is often called the superiority boundary. R0 = λ1.0 / λ2.
R1	The incidence rate ratio assumed by the alternative hypothesis. This is the ratio at which the power is computed. R1 = 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

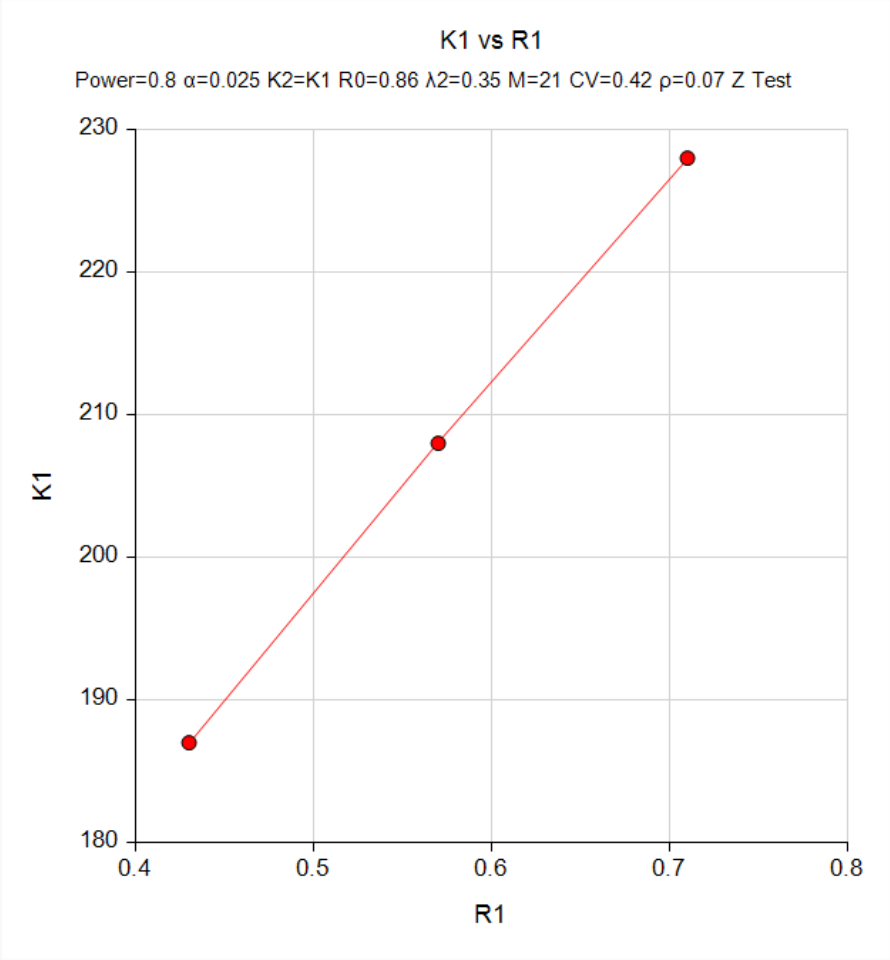
A parallel two-group cluster-randomized design will be used to test whether the Group 1 (treatment) incidence rate (λ1) is superior to the Group 2 (control) incidence rate (λ2) by a margin, with an incidence rate superiority ratio of 0.86 (H0: λ1 / λ2 ≥ 0.86 versus H1: λ1 / λ2 < 0.86). The comparison will be made using a one-sided incidence rate difference Z-test with a Type I error rate (α) of 0.025. The intracluster correlation coefficient is assumed to be 0.07. The control group incidence rate is assumed to be 0.35. 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. To detect an incidence rate ratio (λ1 / λ2) of 0.43 (or treatment group incidence rate of 0.1505) with 80% power, with an average of 21 subjects per cluster, the number of clusters needed will be 187 in Group 1 (treatment) and 187 in Group 2 (control) (for an overall total of 7854 subjects).

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.

Plots Section

Plots



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.

Design Tab	
Solve For	**K1 (Number of Clusters in Group 1)**
Higher Rates Are	**Better (H1: λ1 / λ2 > R0, where R0 > 1)**
Power	**0.9**
Alpha	**0.025**
K2 (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 (λ1.0 and λ1.1)**
λ1.0 (Superiority Incidence Rate)	**0.6**
λ1.1 (Actual Incidence Rate)	**0.5**
λ2 (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

Solve For: K1 (Number of Clusters in Group 1)
Groups: 1 = Treatment, 2 = Control
Higher Rates Are: Better
Hypotheses: H0: λ1 / λ2 ≤ R0 vs. H1: λ1 / λ2 > R0

	Number of Clusters			Cluster Size		Total Sample Size	Incidence Rate			Ratio		Intracluster Correlation	
							Treatment						
Power	Treatment K1	Control K2	Total K	Average M	CV	N	Superiority λ1.0	Actual λ1.1	Control λ2	Superiority R0	Actual R1	ρ	Alpha
0.90572	26	26	52	50	0.2	2600	0.6	0.5	0.5	1.2	1	0.002	0.025

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