Chapter 439

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

Introduction

Cluster-randomized designs are those in which whole clusters of subjects (classes, hospitals, communities, etc.) are sampled, rather than individual subjects. The difference between the event rates of two groups, each consisting of of \( K_i \) clusters of \( M_{ij} \) individuals each, is tested using a two-sample t-test.

The formulas used here are based on Hayes and Bennett (1999) as quoted by Campbell and Walters (2014). These results are also available in Hayes and Moulton (2009).

Technical Details

Our formulation comes from Hayes and Bennett (1999). Let \( K1 \) and \( K2 \) represent the number of clusters in groups 1 (control) and 2 (treatment), respectively. Assume that \( K1 = K2 = Ki \). Let \( M \) represent the number of person-years of observation in each cluster. Let \( \lambda_{ij} \) represent the true event rate in the \( j \)th cluster of the \( i \)th group and \( r_{ij} \) represent the corresponding observed rate. Let \( \bar{r_i} \) represent the means of the two cluster rates. Assume that \( E(r_{ij}) = \lambda_i \) and \( V(r_{ij}) = \sigma_B^2 \). Let the coefficient of variation in the \( i \)th group be \( CV_i = \sigma_{Bi}/\lambda_i \). Let \( s_i^2 \) be the sample variances computed from the \( r_{ij} \).

The inequality of the \( \lambda_1 \) and \( \lambda_2 \) can be tested by the following two-sample t-test

\[
t_{K1+K2-2} = \frac{\bar{r}_2 - \bar{r}_1}{\sqrt{\frac{s_1^2}{K_1} + \frac{s_2^2}{K_2}}}
\]
The formula for the power, given by Hayes and Moulton (2009) for a two-sided significance test of level $\alpha$ to detect an event rate difference is given by

$$
Power = \Phi \left[ \frac{(K_1 - 1)(\lambda_2 - \lambda_1)^2}{(\lambda_1 + \lambda_2) + (CV_1 \lambda_1)^2 + (CV_2 \lambda_2)^2} - z_{1-\alpha/2} \right]
$$

where $z_x = \Phi(x)$ is the standard normal distribution function.

### 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

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are Power, $K_i$, $M$, and $\lambda_2$.

Under most situations, you will select either Power to calculate power or $K_i$ to calculate the number of clusters. Occasionally, you may want to fix the number of clusters and find the necessary cluster size.

Note that the value selected here always appears as the vertical axis on the charts.

The program is set up to calculate power directly. To find appropriate values of the other parameters, a binary search is made using an iterative procedure until an appropriate value is found.

Note that when searching for $M$, some scenarios with small $K_i$'s are not feasible.

#### Test

**Alternative Hypothesis**

Specify whether the test is one-sided or two-sided. The one-sided option specifies a one-tailed 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 commonly used.

A single value may be entered 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. Usually, the value of 0.05 is used for two-sided tests and 0.025 is used for one-sided tests.

You may enter a range of values such as $0.01 \text{ to } 0.10$ or $0.01 \text{ to } 0.10 \text{ by } 0.01$.

Sample Size – Number of Clusters & Cluster Size

$K_i$ (Number of Clusters per Group)
Enter a value (or range of values) for the number of clusters in each group.

You may enter a range of values such as $10 \text{ to } 20 \text{ by } 2$.

$M$ (Person-Years per Cluster)
This is the average number of person-year per cluster in both groups. This value must be a positive number that is at least one. You can use a list of values such as $100 \text{ to } 200$.

Effect Size

$\lambda_1$ (Event Rate of Group 1)
Enter a value (or range of values) for the mean event rate per unit time in group 1 (control group). The value must be greater than zero. This value is compared to $\lambda_2$ by the statistical test. The difference in the rates, $\lambda_2 - \lambda_1$, is the amount that this design can detect.

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

Example of Estimating $\lambda_1$
If 200 patients were exposed for 1 year and 40 experienced the event of interest, then the mean event rate would be

$$\lambda_1 = \frac{40}{200 \times 1} = 0.2 \text{ per patient-year}$$

If 200 patients were exposed for 2 years and 40 experienced the event of interest, then the mean event rate would be

$$\lambda_1 = \frac{40}{200 \times 2} = 0.1 \text{ per patient-year}$$

Event Rate Difference
$\lambda_1$ is used with $\lambda_2$ to calculate the event rate difference as

$$\text{Diff} = \lambda_2 - \lambda_1$$

such that

$$\lambda_1 = \lambda_2 - \text{Diff}$$

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 $1 \text{ to } 2 \text{ by } 0.5$. 

© NCSS, LLC. All Rights Reserved.
Enter λ2, Diff, or Ratio for Group 2

This option lets you indicate how you want to enter λ2. The options are

- **λ2 (Event Rate of Group 2)**
  Enter the value of λ2 directly.

- **Diff (Difference Between Event Rates)**
  Enter values for the difference between the event rates (Diff = λ2 - λ1). The value of λ2 is equal to λ1 + Diff.

- **RR (Ratio of Event Rates)**
  Enter values for the ratio of the event rates (RR = λ2/λ1). The value of λ2 is equal to λ1 x Ratio. Note that the hypothesis still concerns the difference. This is just a convenient way of specifying a value.

**λ2 (Event Rate of Group 2)**

*This option is displayed only if Enter λ2, Diff, or Ratio for Group 2 = “λ2 (Event Rate of Group 2).”*

Enter a value (or range of values) for the mean event rate per time unit in group 2 (treatment group). The value must be greater than zero and different from λ1. This value is compared to λ1 by the statistical test. The difference in the rates, λ2 - λ1, is the amount that this design can detect.

**Example of Estimating λ2**

If 200 patients were exposed for 1 year and 40 experienced the event of interest, then the mean event rate would be

\[ \lambda_2 = \frac{40}{200 \times 1} = 0.2 \text{ per patient-year} \]

If 200 patients were exposed for 2 years and 40 experienced the event of interest, then the mean event rate would be

\[ \lambda_2 = \frac{40}{200 \times 2} = 0.1 \text{ per patient-year} \]

**Event Rate Difference**

λ2 is used with λ1 to calculate the event rate difference as

\[ \text{Diff} = \lambda_2 - \lambda_1 \]

such that

\[ \lambda_2 = \lambda_1 + \text{Diff} \]

The range of acceptable values is λ2 > 0. You can enter a single value such as 1 or a series of values such as 1 to 2 by 0.5.

**CV1 (COV of Rates in Group 1)**

Enter values for CV1. Each cluster in group 1 has an event rate. This is the coefficient of variation of those cluster event rates. The coefficient of variation is equal to the standard deviation of the cluster event rates in group 1 divided by the average event rate, λ1.

If prior information is not available, Hayes and Bennett (1999) suggest that CV1 is usually less than 0.25 and seldom greater than 0.50.

**CV2 (COV of Rates in Group 2)**

Enter values for CV2. Each cluster in the treatment group has an event rate. This is the coefficient of variation of those cluster event rates. The coefficient of variation is equal to the standard deviation of the cluster event rates in group 2 divided by the average event rate, λ2.
If prior information is not available, Hayes and Bennett (1999) suggest that CV2 is usually less than 0.25 and seldom greater than 0.50.

Use CV1

If you enter “CV1”, the value of CV2 will be set to that of CV1.

Example 1 – Calculating Power

Suppose that a cluster randomized study is to be conducted in which $\lambda_1 = 0.50$; $\lambda_2 = 0.6$; CV1 = CV2 = 0.25; $M = 20, 40, 60, 80$; $K_i = 20, 40, 60, 80, 100$; and $\alpha = 0.05$. The power is to be calculated for a two-sided test.

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 the Difference Between Two Poisson Rates in a Cluster-Randomized Design procedure window. 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.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Tab</td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>Power</td>
</tr>
<tr>
<td>Alternative Hypothesis</td>
<td>Two-Sided</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
<tr>
<td>Ki (Number of Clusters per Group)</td>
<td>20 40 60 80 100</td>
</tr>
<tr>
<td>M (Person-Years per Cluster)</td>
<td>20 40 60 80</td>
</tr>
<tr>
<td>$\lambda_1$ (Event Rate of Group 1)</td>
<td>0.5</td>
</tr>
<tr>
<td>Enter $\lambda_2$, Diff, or Ratio for Group 2</td>
<td>$\lambda_2$ (Event Rate of Group 2)</td>
</tr>
<tr>
<td>$\lambda_2$ (Event Rate of Group 2)</td>
<td>0.6</td>
</tr>
<tr>
<td>CV1 (COV of Rates in Group 1)</td>
<td>0.25</td>
</tr>
<tr>
<td>CV2 (COV of Rates in Group 2)</td>
<td>CV1</td>
</tr>
</tbody>
</table>
Annotated Output

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

### Numeric Results

#### Numeric Results for a Two-Sided Test of Event-Rate Difference

<table>
<thead>
<tr>
<th>Power</th>
<th>Total Pers Years</th>
<th>Total Clus Cnt</th>
<th>Pers Cnt per Grp</th>
<th>Clus Cnt per Grp</th>
<th>Pers Years per Clus</th>
<th>Event Rate Gr 1</th>
<th>Event Rate Gr 2</th>
<th>Event Rate Diff</th>
<th>Event Rate Ratio</th>
<th>COV Gr 1</th>
<th>COV Gr 2</th>
<th>Two-Sided Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2975</td>
<td>800</td>
<td>40</td>
<td>400</td>
<td>20</td>
<td>20</td>
<td>0.5000</td>
<td>0.6000</td>
<td>0.1000</td>
<td>1.20</td>
<td>0.250</td>
<td>0.250</td>
<td>0.050</td>
</tr>
<tr>
<td>0.3980</td>
<td>1600</td>
<td>40</td>
<td>800</td>
<td>20</td>
<td>40</td>
<td>0.5000</td>
<td>0.6000</td>
<td>0.1000</td>
<td>1.20</td>
<td>0.250</td>
<td>0.250</td>
<td>0.050</td>
</tr>
<tr>
<td>0.4501</td>
<td>2400</td>
<td>40</td>
<td>1200</td>
<td>20</td>
<td>60</td>
<td>0.5000</td>
<td>0.6000</td>
<td>0.1000</td>
<td>1.20</td>
<td>0.250</td>
<td>0.250</td>
<td>0.050</td>
</tr>
<tr>
<td>0.4816</td>
<td>3200</td>
<td>40</td>
<td>1600</td>
<td>20</td>
<td>80</td>
<td>0.5000</td>
<td>0.6000</td>
<td>0.1000</td>
<td>1.20</td>
<td>0.250</td>
<td>0.250</td>
<td>0.050</td>
</tr>
</tbody>
</table>

#### References


#### Report Definitions

- **Power** is the probability of rejecting a false null hypothesis. It should be close to one.
- **N** is the total number of person-years in the design. N = N1 + N2.
- **K** is the total number of clusters in the design. K = K1 + K2.
- **Ni** represents N1 and N2, the number of person-years in each group. This formulation assumes N1 = N2.
- **Ki** represents K1 and K2, the number of clusters in each group. This formulation assumes K1 = K2.
- **M** is the average number of person-years per cluster in all clusters.
- **λ1** is the event (or incidence) rate of the control group. This is the baseline rate.
- **λ2** is the event (or incidence) rate of the treatment group.
- **λ2 - λ1** is the difference between the treatment event rate and the control event rate.
- **λ2 / λ1** is the ratio of the treatment event rate and the control event rate.
- **CV1** is the coefficient of variation of the cluster event rates in the control group.
- **CV2** is the coefficient of variation of the cluster event rates in the treatment group.
- **Alpha** is the probability of rejecting a true null hypothesis, that is, rejecting when the event rates are actually equal.

#### Summary Statements

A total sample size of 800 person-years, which are obtained by sampling 40 clusters (20 in each group or arm) with an average of 20 person-years per cluster, achieve 30% power to detect a difference of 0.1000 between the treatment event rate 0.6000 and the control event rate 0.5000.

The between-cluster coefficient of variation in the control group was 0.250 and in the treatment group was 0.250. A two-sided t-test of the event-rate difference was used with a significance level of 0.050.

This report shows the power for each of the scenarios.
Plots Section

These plots show the power versus the cluster size for the two alpha values.
Example 2 – Validation using Hayes and Moulton (2009)

Hayes and Moulton (2009) on page 109 present a power calculation for this test. For the values $\lambda_1 = 0.0148; \lambda_2 = 0.0104; CV1 = CV2 = 0.29; M = 424; \alpha = 0.05; \text{ and } K1 = K2 = 28$. The resulting power value is 0.69.

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 the Difference Between Two Poisson Rates in a Cluster-Randomized Design 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.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Tab</td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>Power</td>
</tr>
<tr>
<td>Alternative Hypothesis</td>
<td>Two-Sided</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
<tr>
<td>Ki (Number of Clusters per Group)</td>
<td>28</td>
</tr>
<tr>
<td>M (Person-Years per Cluster)</td>
<td>424</td>
</tr>
<tr>
<td>$\lambda_1$ (Event Rate of Group 1)</td>
<td>0.0148</td>
</tr>
<tr>
<td>Enter $\lambda_2$, Diff, or Ratio for Group 2</td>
<td>$\lambda_2$ (Event Rate of Group 2)</td>
</tr>
<tr>
<td>$\lambda_2$ (Event Rate of Group 2)</td>
<td>0.0104</td>
</tr>
<tr>
<td>CV1 (COV of Rates in Group 1)</td>
<td>0.29</td>
</tr>
<tr>
<td>CV2 (COV of Rates in Group 2)</td>
<td>CV1</td>
</tr>
</tbody>
</table>

Output

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

**Numeric Results**

<table>
<thead>
<tr>
<th>Power</th>
<th>Total Pers Years</th>
<th>Total Clus</th>
<th>Pers Cnt per Grp</th>
<th>Clus Pers Years per Grp</th>
<th>Pers Cnt per Clus</th>
<th>Event Rate Gr 1</th>
<th>Event Rate Gr 2</th>
<th>Event Rate Diff</th>
<th>Event Rate Ratio</th>
<th>COV Gr 1</th>
<th>COV Gr 2</th>
<th>Two-Sided</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6886</td>
<td>23744</td>
<td>56</td>
<td>11872</td>
<td>28</td>
<td>424</td>
<td>0.0148</td>
<td>0.0104</td>
<td>-0.0044</td>
<td>0.70</td>
<td>0.290</td>
<td>0.290</td>
<td>0.050</td>
</tr>
</tbody>
</table>

PASS calculates the same power.