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 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 λ_{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 λ_1 and λ_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 α to detect an event rate difference is given by

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

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

Solve For

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are *Power*, *Ki*, *M*, and *λ2*.

Under most situations, you will select either *Power* to calculate power or *Ki* 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 statistical test is two-sided or one-sided.

- **Two-Sided**

  This option tests whether the two event rates are different (H1: λ1 ≠ λ2). This is the option that is usually selected.

- **One-Sided**

  When this option is used and the value of λ1 is less than λ2, rejecting the null hypothesis results in the conclusion that the group 1 (control) event rate (λ1) is LESS THAN the group 2 (treatment) event rate (λ2).

  Otherwise, rejecting the null hypothesis results in the conclusion that the group 1 (control) event rate (λ1) is GREATER THAN the group 2 (treatment) event rate (λ2).

Appropriate Alpha

When you use a one-sided test, you should usually divide your alpha level by two to keep your results standard. For example, you should use 0.025 rather than 0.05.

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 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Sample Size – Number of Clusters & Cluster Size

Ki (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 to 20 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 *100 150 200*.

Effect Size

λ1 (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 λ2 by the statistical test. The difference in the rates, λ2 - λ1, 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 λ1

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

$\lambda1 = 40/(200*1) = 0.2$ 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

$\lambda1 = 40/(200*2) = 0.1$ per patient-year

Event Rate Difference

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

$$Diff = \lambda 2 - \lambda 1$$

such that

$$\lambda 1 = \lambda 2 - Diff$$

The range of acceptable values is λ1 > 0. You can enter a single value such as *1* or a series of values such as *1 to 2 by 0.5*.

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

λ2 = 40/(200*1) = 0.2 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

λ2 = 40/(200*2) = 0.1 per patient-year

Event Rate Difference

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

$$Diff = \lambda 2 - \lambda 1$$

such that

$$\lambda 2 = \lambda 1 + Diff$$

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 λ1 = 0.50; λ2=0.6; CV1 = CV2 = 0.25; M = 20, 40, 60, or 80; Ki = 20, 40, 60, 80, or 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 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**.

Option	Value
Design Tab	
Solve For	**Power**
Alternative Hypothesis	**Two-Sided**
Alpha	**0.05**
Ki (Number of Clusters per Group)	**20 40 60 80 100**
M (Person-Years per Cluster)	**20 40 60 80**
λ1 (Event Rate of Group 1)	**0.5**
Enter λ2, Diff, or Ratio for Group 2	**λ2 (Event Rate of Group 2)**
λ2 (Event Rate of Group 2)	**0.6**
CV1 (COV of Rates in Group 1)	**0.25**
CV2 (COV of Rates in Group 2)	**CV1**

Annotated Output

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

Numeric Results

Numeric Results for a Test of Event-Rate Difference
Group 1 = Control, Group 2 = Treatment
Alternative Hypothesis: Two-Sided

Power	Total Pers Years N	Total Clus Cnt K	Pers Years per Grp Ni	Clus Cnt per Grp Ki	Pers Years per Clus M	Event Rate Gr 1 λ1	Event Rate Gr 2 λ2	Event Rate Diff λ2-λ1	Event Rate Ratio λ2/λ1	COV Gr 1 CV1	COV Gr 2 CV2	Alpha
0.2975	800	40	400	20	20	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.3980	1600	40	800	20	40	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.4501	2400	40	1200	20	60	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.4816	3200	40	1600	20	80	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.5345	1600	80	800	40	20	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.6836	3200	80	1600	40	40	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.7480	4800	80	2400	40	60	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.7829	6400	80	3200	40	80	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.7113	2400	120	1200	60	20	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.8505	4800	120	2400	60	40	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.8984	7200	120	3600	60	60	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.9211	9600	120	4800	60	80	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.8296	3200	160	1600	80	20	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.9344	6400	160	3200	80	40	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050
0.9625	9600	160	4800	80	60	0.5000	0.6000	0.1000	1.20	0.250	0.250	0.050

(Report Continues)

References

Hayes, R.J. and Bennett, S. 1999. 'Simple sample size calculation for cluster-randomized trials'. International Journal of Epidemiology. Vol 28, pages 319-326.
Hayes, R.J. and Moulton, L.H. 2009. Cluster Randomised Trials. CRC Press. New York.
Campbell, M.J. and Walters, S.J. 2014. How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research. Wiley. New York.

Report Definitions

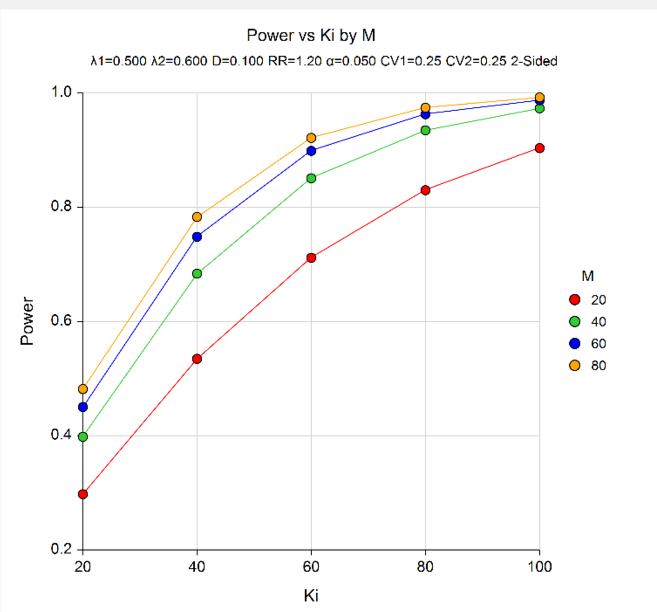
Power is the probability of rejecting a false null hypothesis. It should be close to one.
N is the total number of person-years in the design. N = N1 + N2.
K is the total number of clusters in the design. K = K1 + K2.
Ni represents N1 and N2, the number of person-years in each group. This formulation 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.

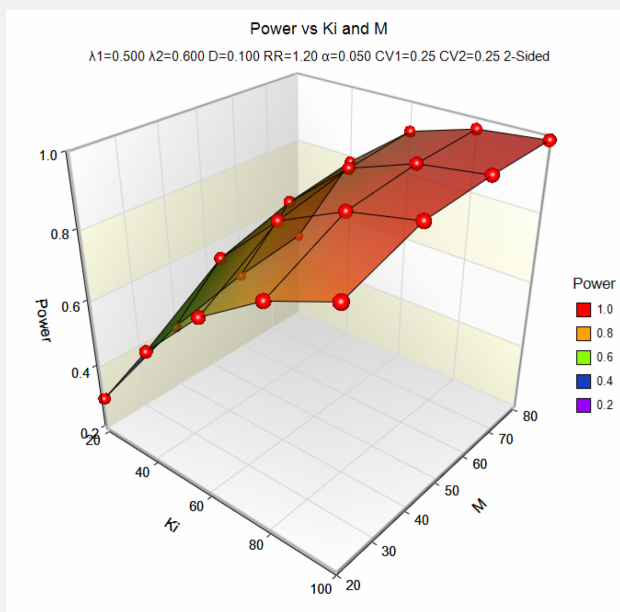
Summary Statements

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; 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 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**.

Option	Value
Design Tab	
Solve For	**Power**
Alternative Hypothesis	**Two-Sided**
Alpha	**0.05**
Ki (Number of Clusters per Group)	**28**
M (Person-Years per Cluster)	**424**
λ1 (Event Rate of Group 1)	**0.0148**
Enter λ2, Diff, or Ratio for Group 2	**λ2 (Event Rate of Group 2)**
λ2 (Event Rate of Group 2)	**0.0104**
CV1 (COV of Rates in Group 1)	**0.29**
CV2 (COV of Rates in Group 2)	**CV1**

Output

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

Numeric Results

Numeric Results for a Test of Event-Rate Difference
Group 1 = Control, Group 2 = Treatment
Alternative Hypothesis: Two-Sided

Power	Total Pers Years N	Total Clus Cnt K	Pers Years per Grp Ni	Clus Cnt per Grp Ki	Pers Years per Clus M	Event Rate Gr 1 λ1	Event Rate Gr 2 λ2	Event Rate Diff λ2-λ1	Event Rate Ratio λ2/λ1	COV Gr 1 CV1	COV Gr 2 CV2	Alpha
0.6886	23744	56	11872	28	424	0.0148	0.0104	-0.0044	0.70	0.290	0.290	0.050

PASS calculates the same power.