Chapter 873

Logrank Tests for Paired Survival Data

Introduction

This module computes the sample size and power of logrank tests for paired subjects (such as twins or eyes). The design parameters include uniform accrual for a fixed time followed by a fixed period of follow-up. It includes a loss to follow-up parameter.

It is used to compare the survival of two groups, usually treatment and control.

The procedure is documented in Jung (2008).

Technical Details

Paired Logrank Test Statistic

The following details follow closely the results in Jung (2008).

Suppose *N* pairs of subjects are enrolled in a study during the accrual period of length t_a and then observed during a follow-up period of length t_f. Let T_{1i} and T_{2i} denote the bivariate survival-time variables with marginal cumulative hazard functions $\Lambda_1(t)$ for $\Lambda_2(t)$. The null hypotheses to be tested is $H_0: \Lambda_1(t) = \Lambda_2(t)$ versus the alternative hypothesis $H_a: \Lambda_1(t) \neq \Lambda_2(t)$ for some $t \geq 0$. Let C_i denote the censoring time of the i^{th} pair. It is assumed that the C_i are independent of T_{1i} and T_{2i}. Let $X_{ki} = \min\,(T_{ki}, C_i)$ and $\Delta_{ki} = \mathrm{I}(T_{ki}, C_i)$. Furthermore, let $Y_{ki}(t) = \mathrm{I}\,(X_{ki} > t)$ and $Y_k(t) = \sum_{i=1}^{N} Y_{ki}(t)$.

The logrank statistic is defined as

$$W = \sqrt{N}\int_0^\infty H(t)\{\mathrm{d}\widehat{\Lambda}_1(t) - \mathrm{d}\widehat{\Lambda}_2(t)\}$$

where

$$H(t) = \frac{Y_1(t)Y_2(t)}{Y_1(t) + Y_2(t)}$$

$$\widehat{\Lambda}_k(t) = \int_0^t Y_k(s)^{-1}\mathrm{d}N_k(s)$$

$$N_k(t) = \sum_{i=1}^{N} N_{ki}(t)$$

$$N_{ki}(t) = \Delta_{ki}\mathrm{I}(X_{ki} \leq t)$$

Jung (1999) shows that for paired data, under the null hypothesis, W is asymptotically normal with mean 0 and variance σ^2.

Power Calculation

The sample size calculations below make the following assumptions

1. The paired subjects are accrued and censored at the same time points.
2. Pairs are uniformly accrued and then followed for a common period of t. The accrual period is t_a and the following period is t_f.
3. The lost to follow-up is assumed to be exponentially distributed with hazard rate is *v*.
4. The paired logrank test will be used for the data analysis.
5. The dependency in survival times within each pair is modelled by the positive stable frailty model based on the frailty coefficient θ which is functionally related to the correlation between survival times T_{1i} and T_{2i}.
6. Assume that the marginal survival distributions are exponential as defined below.

Based on these assumptions, we define the following relationships

$$S(t_1, t_2) = P(T_{1i} \geq t_1, T_{2i} \geq t_2) = \exp\left(\left[\{-\log S_1(t_1)\}^{\frac{1}{\theta}} + \{-\log S_2(t_2)\}^{\frac{1}{\theta}}\right]^{\theta}\right)$$

$$S_k(t) = P(T_{1k} \geq t_k) = \exp(-\lambda_k t) \text{ for } k = 1,2$$

This reduces to

$$S(t_1, t_2) = \exp\left(-\left[(\lambda_1 t_1)^{\frac{1}{\theta}} + (\lambda_2 t_2)^{\frac{1}{\theta}}\right]^{\theta}\right)$$

The censoring time distribution is

$$G(t) = \begin{cases} \exp(-vt) & \text{if } t < t_f \\ \exp(-vt)\left[1 - \dfrac{t - t_f}{t_a}\right] & \text{if } t_f < t < \mathrm{t_a} + t_f \\ 0 & \text{otherwise} \end{cases}$$

The correlation between survival times T_{1i} and T_{2i} is based only on the frailty coefficient as shown in the following definition

$$\rho = \frac{\int_0^\infty \int_0^\infty \{S_1(t_1, t_2) - S_1(t_1)S_2(t_2)\}\, \mathrm{d}t_1 \mathrm{d}t_2}{\sqrt{\operatorname{var}(T_{1i})\operatorname{var}(T_{2i})}} = \int_0^1 \int_0^1 \frac{S_u(u_1, u_2)}{u_1 u_2} \mathrm{d}u_1 \mathrm{d}u_2 - 2$$

where a change of variables is made from t_k to $u_k = \exp(-t_k)$. Note that the denominator of the second integral was inadvertently left out of Jung (2008) and was supplied by the author.

The numerator function becomes

$$S_u(u_1,u_2) = \exp\left[-\left\{(-\log u_1)^{\frac{1}{\theta}} + (-\log u_2)^{\frac{1}{\theta}}\right\}^{\theta}\right]$$

which only depends on θ. This gives a functional relationship between the frailty and the correlation which can be solved numerically.

Jung (2008) shows that the mean and variance of the normally distributed test statistic W are given by

$$\mu = (\lambda_1 - \lambda_2)\left\{\int_0^{t_a+t_f}\left[\frac{\mathrm{e}^{-(\lambda_1+\lambda_2+v)t}}{\mathrm{e}^{-\lambda_1 t}+\mathrm{e}^{-\lambda_2 t}}\right]\mathrm{d}t - \frac{1}{t_a}\int_{t_f}^{t_a+t_f}\left[\frac{(t-t_f)\mathrm{e}^{-(\lambda_1+\lambda_2+v)t}}{\mathrm{e}^{-\lambda_1 t}+\mathrm{e}^{-\lambda_2 t}}\right]\mathrm{d}t\right\}$$

$$\sigma^2 = \sigma_k^2 + \sigma_k^2 - 2\sigma_{12}$$

where

$$\sigma_k^2 = \lambda_k\left\{\int_0^{t_a+t_f}\left[\frac{\mathrm{e}^{-(\lambda_k+2\lambda_{k'}+v)t}}{(\mathrm{e}^{-\lambda_1 t}+\mathrm{e}^{-\lambda_2 t})^2}\right]\mathrm{d}t - \frac{1}{t_a}\int_{t_f}^{t_a+t_f}\left[\frac{(t-t_f)\mathrm{e}^{-(\lambda_k+2\lambda_{k'}+v)t}}{(\mathrm{e}^{-\lambda_1 t}+\mathrm{e}^{-\lambda_2 t})^2}\right]\mathrm{d}t\right\} \text{ for } k \neq k' \in \{1,2\}$$

$$\sigma_{12} = \int_0^{t_a+t_f}\int_0^{t_a+t_f}\frac{S(t_1,t_2)G(t_1 \vee t_2)\mathrm{e}^{-\lambda_2 t_1-\lambda_1 t_2}}{(\mathrm{e}^{-\lambda_1 t_1}+\mathrm{e}^{-\lambda_2 t_1})(\mathrm{e}^{-\lambda_1 t_2}+\mathrm{e}^{-\lambda_2 t_2})}\mathrm{d}A(t_1,t_2)$$

$$\mathrm{d}A(t_1,t_2) = \{\lambda(t_1,t_2) - \lambda_2[\lambda_{1|2}(t_1|t_2)] - \lambda_1[\lambda_{2|1}(t_2|t_1)] + \lambda_1\lambda_2\}\mathrm{d}t_1\mathrm{d}t_2$$

$$\lambda(t_1,t_2) = \lambda_1\lambda_2(\lambda_1\lambda_2 t_1 t_2)^{\frac{1}{\theta}-1}\left\{(\lambda_1 t_1)^{\frac{1}{\theta}} + (\lambda_2 t_2)^{\frac{1}{\theta}}\right\}^{\theta-2}\left[\left\{(\lambda_1 t_1)^{\frac{1}{\theta}} + (\lambda_2 t_2)^{\frac{1}{\theta}}\right\}^{\theta} + \frac{1-\theta}{\theta}\right]$$

$$\lambda_{k|k'}(t_k|t_{k'}) = \lambda_k(\lambda_k t_k)^{\frac{1}{\theta}-1}\left\{(\lambda_1 t_1)^{\frac{1}{\theta}} + (\lambda_2 t_2)^{\frac{1}{\theta}}\right\}^{\theta-1} \text{ for } k \neq k' \in \{1,2\}$$

Note that $(X \vee Y) = \max(X,Y)$.

Using these results, the power of a two-sided test for a particular N (number of pairs) is given by.

$$Power = \Phi\left(z_{1-\alpha/2} - \frac{\sqrt{N}|\mu|}{\sigma}\right)$$

The required number of events is

$$E = N(d_1 + d_2)$$

where

$$d_k = 1 - (1 - \exp(-\lambda_k t_a)\left(\frac{\exp(-\lambda_k t_f)}{\lambda_k t_a}\right)$$

This power formula can be rearranged to give the following formula for the number of pairs.

$$N = \frac{\sigma^2 \left(z_{1-\frac{\alpha}{2}} + z_{1-\beta}\right)^2}{\mu^2}$$

To make certain that the required power is attained, the computed *N* is rounded up to the next integer and then the power is recomputed with this final value of *N*.

Alternative Hazard Rate Input Types

The hazard rates λ_1 and λ_2 can be given in terms of the hazard ratio *HR*, the median survival times M_1 and M_2, or the cumulative survival proportions S_1 and S_2 at time t_0. These various parameters are transformed to hazard rates using

$$HR = \frac{\lambda_1}{\lambda_2}$$

$$\lambda_k = \frac{\log 2}{M_k} = \frac{-\log S_k(t_0)}{t_0}$$

Example 1 – Finding Sample Size

A researcher is planning a trial to compare the ability to stop blindness of a new treatment versus the current treatment in subjects that have been recently diagnosed with diabetic retinopathy in both eyes. Each subject will receive both treatments, randomized to the eyes.

A prior study has been used to set the hazard rate of the treatment group to 0.012 and of the control group to 0.021. The power is set to 0.90 and the two-sided significance level to 0.05. Given the annual rates of enrollment, the accrual time is set to 0.85.

The researcher would like to compare the sample requirements if the follow-up period is 1, 2, or 3 years and if the lost to follow-up hazard rate is 0, 0.05, or 0.10.

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	**Sample Size (Number of Pairs)**
Alternative Hypothesis	**Two-Sided**
Power	**0.90**
Alpha	**0.05**
Ta (Accrual Time)	**0.85**
Tf (Follow-Up Time)	**1 2 3**
v (Lost to Follow-Up Hazard Rate)	**0 0.05 0.10**
Hazard Rates Input Type	**λ1, λ2 (Hazard Rates)**
λ1 (Hazard Rate - Group 1)	**0.012**
λ2 (Hazard Rate - Group 2)	**0.021**
Frailty Input Type	**Θ (Frailty Coefficient)**
Θ (Frailty Coefficient)	**0.3**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size (Number of Pairs)
Groups: 1 = Treatment, 2 = Control
Hypothesis Type: Two-Sided
Test Statistic: Logrank Test
Data Distribution: Exponential
Accrual Type: Uniform

	Sample Size		Lost to Follow-Up	Time			Hazard Rate				
Power	Number of Pairs N	Number of Events E	Hazard Rate v	Accrual Ta	Follow-Up Tf	Hazard Ratio HR	Treatment λ1	Control λ2	Frailty Coefficient ϴ	Within-Pair Correlation ρ	Alpha
0.90023	1002	46.5	0.00	0.85	1	0.57143	0.012	0.021	0.3	0.8029	0.05
0.90023	1039	48.2	0.05	0.85	1	0.57143	0.012	0.021	0.3	0.8029	0.05
0.90002	1076	49.9	0.10	0.85	1	0.57143	0.012	0.021	0.3	0.8029	0.05
0.90019	594	46.5	0.00	0.85	2	0.57143	0.012	0.021	0.3	0.8029	0.05
0.90028	631	49.4	0.05	0.85	2	0.57143	0.012	0.021	0.3	0.8029	0.05
0.90018	669	52.4	0.10	0.85	2	0.57143	0.012	0.021	0.3	0.8029	0.05
0.90062	425	46.6	0.00	0.85	3	0.57143	0.012	0.021	0.3	0.8029	0.05
0.90051	462	50.7	0.05	0.85	3	0.57143	0.012	0.021	0.3	0.8029	0.05
0.90040	501	54.9	0.10	0.85	3	0.57143	0.012	0.021	0.3	0.8029	0.05

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N The number of pairs, assuming no subject lost to dropout or follow-up during the study.
E The total number of events (failures) in both groups expected during the study.
v The lost to follow-up hazard rate. This can be converted to the proportion lost to follow-up during a single time period.
Ta The length of the accrual time during which subjects are added to the study.
Tf The length of the follow-up time after the last subject is added to the study.
HR The hazard ratio is the treatment group's hazard rate divided by control group's hazard rate. HR = λ1 / λ2.
λ1 The hazard rate of the treatment group.
λ2 The hazard rate of the control group.
ϴ The frailty coefficient is an index of the frailty (propensity to die) in the next instant. This value varies from 0 to 1. Subjects with higher frailty coefficients tend to die sooner than those with a lower frailty.
ρ The within-pair correlation coefficient is the correlation between the two items of a pair. It is related to the frailty through a complicated, functional relationship.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A paired design will be used to test whether the treatment group hazard rate is different from the control group hazard rate. The comparison will be made using a two-sided logrank test, with a Type I error rate (α) of 0.05. It is assumed that the survival time distribution is approximated reasonably well by the exponential distribution. The lost to follow-up hazard rate is assumed to be 0. The accrual time will be 0.85 and the follow-up time (time after complete accrual) will be 1. The frailty coefficient will be 0.3 and the within-pair correlation coefficient will be 0.8029. To detect a hazard rate of 0.012 in the treatment group when the hazard rate of the control group is 0.021, with 90% power, the number of needed pairs will be 1002. The expected number of events from both groups during the study is 46.5.

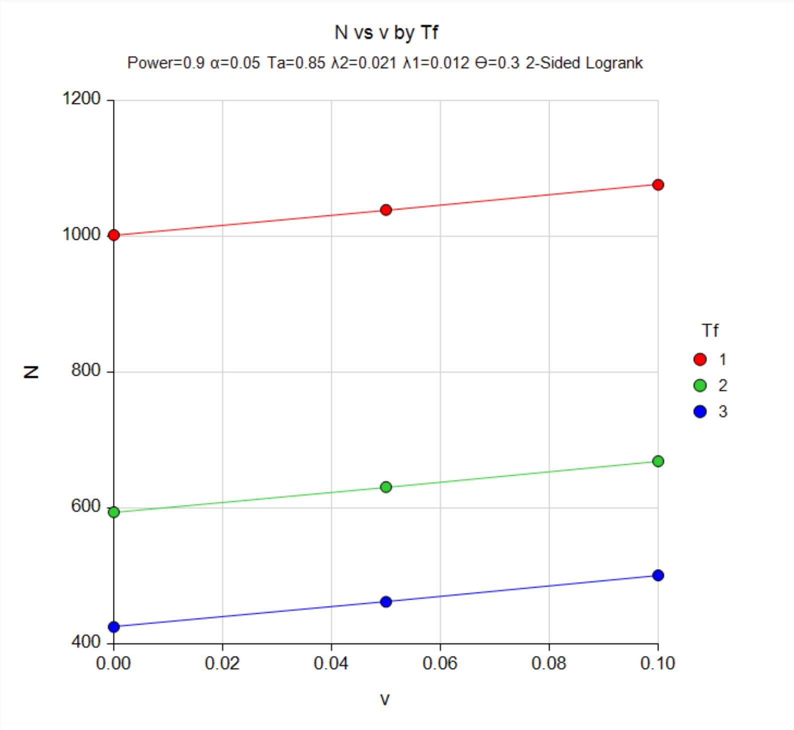
References

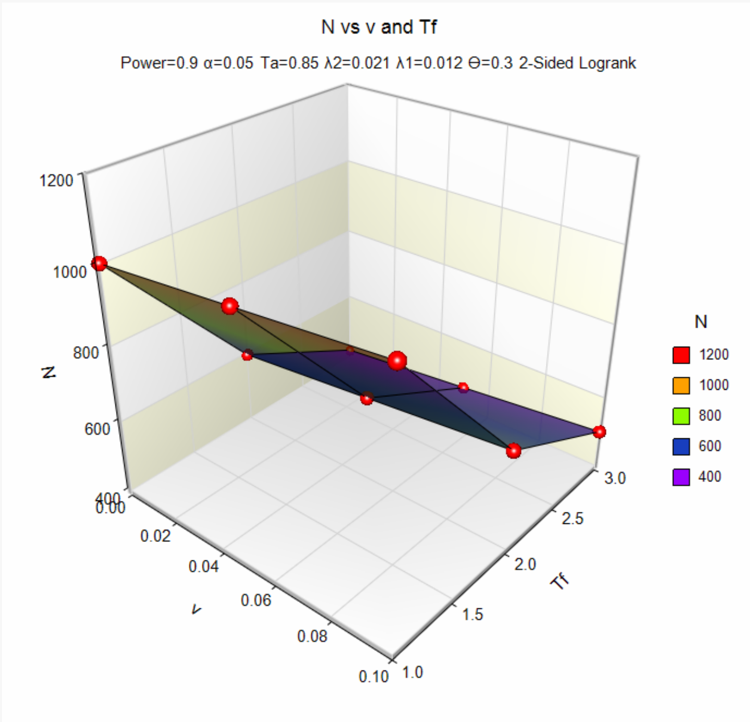
Jung, Sin-Ho. 1999. 'Rank tests for matched survival data', Lifetime Data Analysis, Vol. 5, pages 67-79.
Jung, Sin-Ho. 2008. 'Sample size calculation for the weighted rank statistics with paired survival data', Statistics in Medicine, Vol. 27, pages 3350-3365.
Su, P-F, Li, C-I, Shyr, Y. 2014. 'Sample size determination for paired right-censored data based on the difference of Kaplan-Meier estimates', Computational Statistics and Data Analysis, Vol. 74, pages 39-51.
Hougaard, Philip. 1986. 'A Class of Multivariate Failure Time Distributions', Biometrika, Vol. 73, No. 3, pages 671-678.

This report presents the calculated sample sizes for each scenario as well as the values of the other parameters.

Plots Section

Plots





These plots show the relationship between sample size, follow-up time, and v.

Example 2 – Validation using Jung (2008)

Jung (2008) page 3360 gives an example of a two-sided test in which power = 0.90, alpha = 0.05, Ta = 3, Tf = 2, $\lambda_1 = 0.3$, $\lambda_2 = 0.5$, *v* = 0, 0.1, and $\theta = 0.9$. Jung calculates N to be 106 and 121.

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	**Sample Size (Number of Pairs)**
Alternative Hypothesis	**Two-Sided**
Power	**0.90**
Alpha	**0.05**
Ta (Accrual Time)	**3**
Tf (Follow-Up Time)	**2**
v (Lost to Follow-Up Hazard Rate)	**0 0.10**
Hazard Rates Input Type	**λ1, λ2 (Hazard Rates)**
λ1 (Hazard Rate - Group 1)	**0.3**
λ2 (Hazard Rate - Group 2)	**0.5**
Frailty Input Type	**ϴ (Frailty Coefficient)**
ϴ (Frailty Coefficient)	**0.9**

Output

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

Numeric Results

Solve For: Sample Size (Number of Pairs)
Groups: 1 = Treatment, 2 = Control
Hypothesis Type: Two-Sided
Test Statistic: Logrank Test
Data Distribution: Exponential
Accrual Type: Uniform

	Sample Size		Lost to Follow-Up Hazard	Time			Hazard Rate				
Power	Number of Pairs N	Number of Events E	Rate v	Accrual Ta	Follow-Up Tf	Hazard Ratio HR	Treatment λ1	Control λ2	Frailty Coefficient ϴ	Within-Pair Correlation ρ	Alpha
0.90056	107	154.9	0.0	3	2	0.6	0.3	0.5	0.9	0.10349	0.05
0.90218	122	176.6	0.1	3	2	0.6	0.3	0.5	0.9	0.10349	0.05

PASS has calculated N to be 107 and 122. Note that both values are one more than in Jung's article. This is because **PASS** guarantees at least the target power of 0.9. The powers for 106 and 121 were slightly less than 0.90.