Chapter 708

Non-Inferiority Tests for the Difference of Two Hazard Rates Assuming an Exponential Model

Introduction

A clinical trial is often employed to test the equality of survival distributions of two treatment groups. The two-sample t-test is not appropriate for two reasons. First, the data are not normally distributed. Second, some survival times are *censored*. For these reasons, special test statistics such as the logrank test have been developed. This module computes the sample size and power of the logrank test assuming survival times follow exponential distributions. Accrual time and follow-up time are included among the input parameters.

This procedure is based on the *unconditional* method of Chow, Shao, and Wang (2008) which, in turn, is based on the *conditional* methods of Lachin and Foulkes (1986). The conditional procedure does not extend to this case (see Chow, Shao, and Wang (2008) page 173).

Technical Details

This section presents the *unconditional* method of Chow, Shao, and Wang (2008).

Basic Model

Suppose a clinical trial consists of two independent groups labeled “1” and “2” (where group 1 is the control group and group 2 is the treatment group). The total sample size is N and the sizes of the two groups are N_1 and N_2. Usually, you would plan to have $N_1 = N_2$.

Non-Inferiority Hypothesis

Assuming that lower hazard rates are better, non-inferiority is established by concluding that the treatment hazard rate is at most, only higher than the control hazard rate by the margin Δ. The statistical hypotheses that yields this conclusion when the null hypothesis is rejected is

$$H_0\colon\ (h_2 - h_1) \geq \Delta \quad \text{versus} \quad H_a\colon (h_2 - h_1) < \Delta$$

or

$$H_0\colon\ h_2 \geq h_1 + \Delta \quad \text{versus} \quad H_a\colon h_2 < h_1 + \Delta$$

If, however, higher hazard rates are better, non-inferiority is established by concluding that the treatment hazard rate is at most, only slightly lower than the control hazard rate. The statistical hypotheses that yields this conclusion when the null hypothesis is rejected is

$$H_0\colon\ (h_2 - h_1) \le -\Delta \quad \text{versus} \quad H_a\colon (h_2 - h_1) > -\Delta$$

or

$$H_0\colon\ h_2 \le h_1 - \Delta \quad \text{versus} \quad H_a\colon h_2 > h_1 - \Delta.$$

Test Statistic

The power and sample size formulas presented below are for the difference of two exponential hazard rates. Simulation studies have shown that they also approximate the power of the logrank test. It is anticipated that the actual test statistic is the regression coefficient from a Cox regression.

Test Comparing Hazard Rates

The original test statistic is the difference of the hazard rates estimated by maximum likelihood divided by their standard error. The maximum likelihood estimate of an exponential hazard rate for a particular group is

$$\hat{h} = \frac{\text{number of events}}{\text{sum of study time of all subjects}}$$

Chow, Shao, and Wang (2008) indicate that the test statistic

$$Z = \frac{\left(\hat{h}_2 - \hat{h}_1\right) - \Delta}{\sqrt{\frac{\sigma^2\left(\hat{h}_1\right)}{N_1} + \frac{\sigma^2\left(\hat{h}_2\right)}{N_2}}}$$

where

$$\sigma^2(h) = \frac{h^2}{1 + \frac{e^{-hT}(1 - e^{hR})}{hR}}$$

This Z statistic is approximately normally distributed.

Power Calculations

Assuming an exponential model with hazard rates h_1 and h_2 for the two groups, Chow et al. (2008) give the following equation relating *N* and power of a non-inferiority test.

$$\frac{(h_2 - h_1) + \Delta}{\sqrt{\frac{\sigma^2(h_1, \omega_1, A)}{N_1} + \frac{\sigma^2(h_2, \omega_2, A)}{N_2}}} - z_{1-\alpha} = z_{1-\beta}$$

where

$$\sigma^2(h_i, \omega_i, A) = \frac{h_i^2}{E(d_i|h_i, \omega_i, A)}$$

$$E(d_i|h_i, \omega_i, A) = \left(\frac{h_i}{h_i + \omega_i}\right)\left(1 + \frac{A\exp\{-(h_i + \omega_i)T\}[1 - \exp\{(h_i + \omega_i - A)R\}]}{(h_i + \omega_i - A)[1 - \exp\{-AR\}]}\right)$$

$$E(d_i|h_i, \omega_i, 0) = \left(\frac{h_i}{h_i + \omega_i}\right)\left(1 + \frac{\exp\{-(h_i + \omega_i)T\}[1 - \exp\{(h_i + \omega_i)R\}]}{(h_i + \omega_i)R}\right)$$

These parameters are interpreted as follows.

Parameter	**Interpretation**
$\sigma^2(h, \omega, A)$	Variance of $\hat{h}$
$E(d_i\|h_i, \omega_i, A)$	Expected proportion of events (deaths) in group *i*
d_i	Indicates a person does $(d_i = 1)$ or does not $(d_i = 0)$ die in group *i*
h_i	Hazard rate of group *i* (see below)
ω_i	Loss to follow-up hazard rate of group *i* (see below)
A	Patient entry parameter (see below)
R	Accrual time
T	Total time
$T - R$	Follow-up time

Exponential Distribution

The hazard rate from the exponential distribution, h, is usually estimated using maximum likelihood techniques. In the planning stages, you have to obtain an estimate of this parameter. To see how to accomplish this, let's briefly review the exponential distribution. The density function of the exponential is defined as

$$f(t) = h\exp\{-ht\}, \quad t \geq 0, h > 0.$$

The cumulative survival distribution function is

$$\mathrm{S}(t) = \exp\{-ht\}, \quad \mathrm{t} \geq 0.$$

Solving this for h yields

$$h = -\frac{\log\{S(t)\}}{t}$$

Note that *S*(*t*) gives the probability of surviving *t* years. To obtain a planning estimate of h, you need only know the proportion surviving during a particular time period. You can then use the above equation to calculate h.

Patient Entry

Patients are enrolled during the accrual period. **PASS** lets you specify the pattern in which subjects are enrolled. Suppose patient entry times are distributed as *g*(*t*) where t_i is the entry time of the i^{th} individual and $0 \leq t_i \leq R$. Let *g*(*t*) follow the truncated exponential distribution with parameter *A*, which has the density

$$g(t) = \begin{cases} \dfrac{A\exp\{-At\}}{1 - A\exp\{-AR\}} & \text{if } 0 \leq t \leq R, \ A \neq 0 \\ 1 & \text{otherwise} \end{cases}$$

where

R is accrual time.

A is interpreted as follows:

$A > 0$ results in a convex (faster than expected) entry distribution.

$A < 0$ results in a concave (slower than expected) entry distribution.

$A = 0$ results in the uniform entry distribution in which *g*(*t*) =1/*R*.

Rather than specify *A* directly, **PASS** has you enter the percentage of the accrual time that will be needed to enroll 50% of the subjects. Using an iterative search, the value of *A* corresponding to this percentage is calculated and used in the calculations.

Losses to Follow-Up

The staggered patient entry over the accrual period results in censoring times ranging from *T* - *R* to *T* years during the follow-up period. This is often referred to as administrative censoring, since it is caused by the conclusion of the study rather than by some random factor working on an individual. To model the losses to follow-up in each group which come from other causes, we use the exponential distribution again, this time with hazard rates ω_1 and ω_2. You can obtain appropriate loss-to-follow-up hazard rates using the following formula or by using the Survival Parameter Conversion Tool available from the Tools menu or by pressing the small button to the rate of the loss-to-follow-up hazard rate box.

$$\omega = -\frac{\log\{1 - P_{loss}(R)\}}{R}$$

Example 1 – Finding the Sample Size

Suppose the hazard rate when using the current treatment of a disease is 2. A new treatment for the disease is cheaper and has fewer side effects. The company that developed this new treatment wants to establish that its hazard rate is no worse than 25% of the current treatment. How large of a sample is needed if the recruitment period is one-year after which the study will continue for an additional two-years? It is assumed that patients will enter the study uniformly over the recruitment period. The researcher estimates the loss-to-follow rate to be 0.165 in both the current and the groups. The company would like to compare sample sizes when the power is 0.80 and 0.90 and when D is between -1 and 0. The researcher will test at the 0.05 significance level.

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**
Higher Hazards Are	**Worse (Ha: h2 < h1 + Δ)**
Power	**0.8 0.9**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
ω1 (Loss Hazard Rate of Control Group)	**0.165**
ω2 (Loss Hazard Rate of Treatment Group)	**ω1**
R (Accrual, or Recruitment, Time)	**1**
Percent of R Until 50% are Accrued	**50**
T-R (Follow-Up Time)	**2**
Specify Hazard Parameters Using	**Differences**
h1 (Hazard Rate of Control Group)	**2**
D (Hazard Rate Difference = h2-h1)	**-1 to 0 by 0.2**
Δ (Non-Inferiority Margin)	**0.5**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size
Groups: 1 = Control, 2 = Treatment
Hypotheses: H0: h2 ≥ h1 + Δ vs. Ha: h2 < h1 + Δ
Accrual: Uniform

	Sample Size			Hazard Rate		Hazard Rate Difference	Non-Inferiority		Loss Hazard Rate		Time			
Power	N	N1	N2	h1	h2	D	Margin Δ	Boundary B	ω1	ω2	Accrual R	Follow-Up T - R	Alpha	Report Row
0.8141	32	16	16	2	1.0	-1.0	0.5	2.5	0.165	0.165	1	2	0.05	1
0.8021	45	22	23	2	1.2	-0.8	0.5	2.5	0.165	0.165	1	2	0.05	2
0.8032	68	34	34	2	1.4	-0.6	0.5	2.5	0.165	0.165	1	2	0.05	3
0.8019	111	55	56	2	1.6	-0.4	0.5	2.5	0.165	0.165	1	2	0.05	4
0.8002	200	100	100	2	1.8	-0.2	0.5	2.5	0.165	0.165	1	2	0.05	5
0.8003	431	215	216	2	2.0	0.0	0.5	2.5	0.165	0.165	1	2	0.05	6
0.9084	44	22	22	2	1.0	-1.0	0.5	2.5	0.165	0.165	1	2	0.05	7
0.9028	62	31	31	2	1.2	-0.8	0.5	2.5	0.165	0.165	1	2	0.05	8
0.9018	94	47	47	2	1.4	-0.6	0.5	2.5	0.165	0.165	1	2	0.05	9
0.9003	153	76	77	2	1.6	-0.4	0.5	2.5	0.165	0.165	1	2	0.05	10
0.9000	277	138	139	2	1.8	-0.2	0.5	2.5	0.165	0.165	1	2	0.05	11
0.9002	597	298	299	2	2.0	0.0	0.5	2.5	0.165	0.165	1	2	0.05	12

	Number of Events			Percent Group 1	Hazard Ratio	Variance		Report
Power	E	E1	E2	%N1	HR	σ²(h1)	σ²(h2)	Row
0.8141	27.6	14.7	12.9	50	0.5	4.353	1.236	1
0.8021	39.7	20.2	19.5	50	0.6	4.353	1.698	2
0.8032	61.0	31.2	29.7	50	0.7	4.353	2.241	3
0.8019	100.6	50.5	50.1	50	0.8	4.353	2.863	4
0.8002	182.7	91.9	90.8	50	0.9	4.353	3.568	5
0.8003	396.0	197.5	198.5	50	1.0	4.353	4.353	6
0.9084	38.0	20.2	17.8	50	0.5	4.353	1.236	7
0.9028	54.8	28.5	26.3	50	0.6	4.353	1.698	8
0.9018	84.3	43.2	41.1	50	0.7	4.353	2.241	9
0.9003	138.7	69.8	68.8	50	0.8	4.353	2.863	10
0.9000	253.0	126.8	126.2	50	0.9	4.353	3.568	11
0.9002	548.5	273.8	274.7	50	1.0	4.353	4.353	12

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N	The total sample size.
N1 and N2	The sample sizes of the control and treatment groups, respectively.
h1 and h2	The hazard rates in the control and treatment groups, respectively.
D	The difference in hazard rates. D = h2 - h1.
Δ	The non-inferiority margin.
B	The non-inferiority boundary for h2. B = h1 + Δ.
ω1 and ω2	The rates at which subjects in groups 1 and 2 are lost to follow up, respectively.
R	The accrual (recruitment) time.
T - R	The follow-up time. Hence, T is the total time of the study.
Alpha	The probability of rejecting a true null hypothesis.
E	The total number of events required.
E1 and E2	The number of events required in the control and treatment groups.
%N1	The percent of the total sample that is in group 1, the control group.
HR	The hazard ratio. HR = h2 / h1.
σ²(h1) and σ²(h2)	The variances of the estimates of h1 and h2, respectively.

Summary Statements

A parallel, two-group design (where higher hazard rates are considered worse) will be used to test whether the Group 2 (treatment) hazard rate is non-inferior to the Group 1 (control) hazard rate, with a non-inferiority margin of 0.5 (H0: h2 - h1 ≥ 0.5 versus Ha: h2 - h1 < 0.5).The comparison will be made using a one-sided, two-sample maximum likelihood estimation Z test with a Type I error rate (α) of 0.05. Patients will enter the study during an accrual period of 1 time period. 50% of the enrollment will be complete when 50% of the accrual time has passed (uniform accrual). A follow-up period of 2 time periods will have a 0.165 loss to follow-up hazard rate in the control group and a 0.165 loss to follow-up hazard rate in the treatment group. The calculations are based on the assumption that the survival times are exponentially distributed. To detect a hazard rate difference of -1 (h1 = 2, h2 = 1) with 80% power, the number of needed subjects will be 16 in Group 1 and 16 in Group 2 (totaling 32 subjects). The corresponding required number of events is 14.7 in Group 1 and 12.9 in Group 2 (totaling 27.6 events).

References

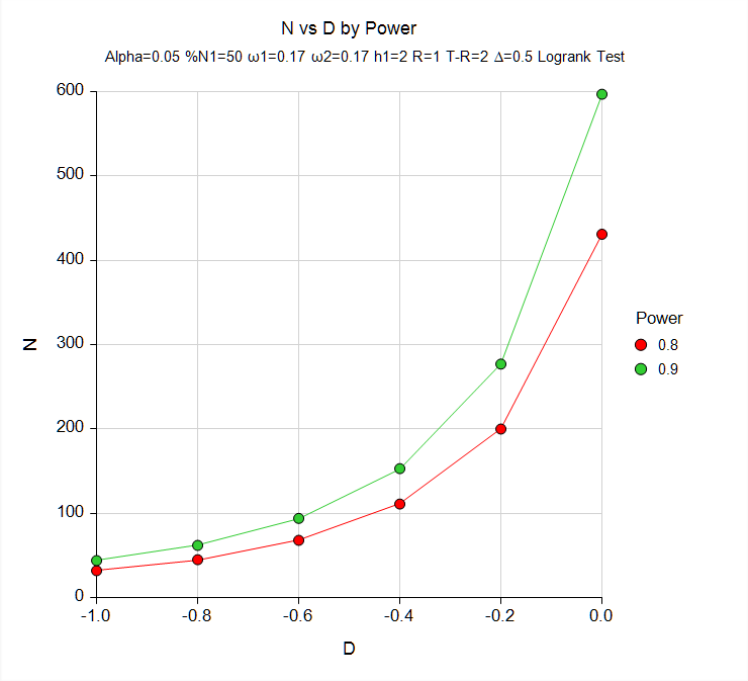
Chow, S.C., Shao, J., Wang, H. 2008. Sample Size Calculations in Clinical Research, 2nd Edition. Chapman & Hall/CRC.

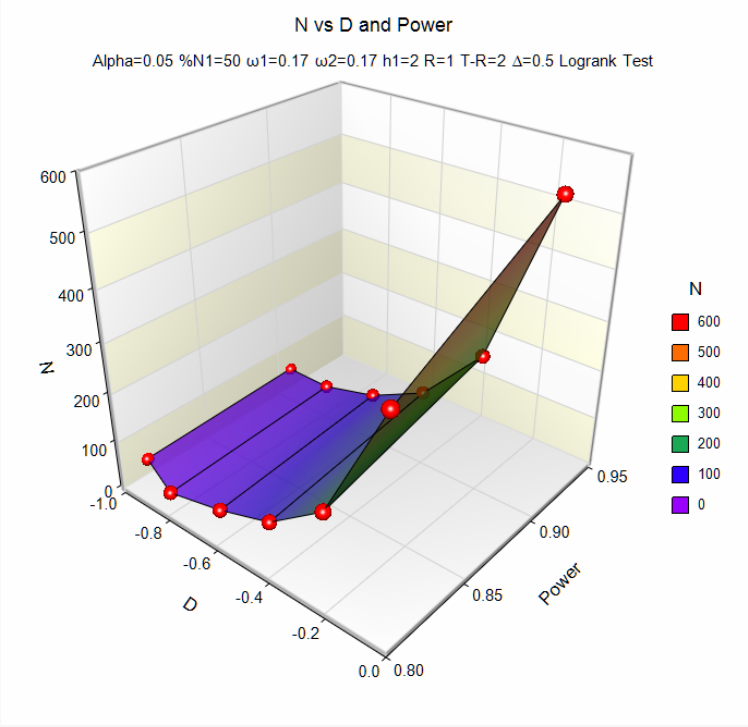
Lachin, John M. and Foulkes, Mary A. 1986. 'Evaluation of Sample Size and L.Power for Analyses of Survival with Allowance for Nonuniform Patient Entry, Losses to Follow-up, Noncompliance, and Stratification', Biometrics, Volume 42, September, pages 507-516.

These reports show the values of each of the parameters, one scenario per row. The second report presents information about the number of events that are necessary.

Plots Section

Plots





These plots show the relationship between power and sample size.

Example 2 – Validation using Chow et al. (2008)

Chow et al. (2008) page 172 presents an example of a two-group, equal sample allocation superiority design to compare the hazard rates of a new treatment with that of the current treatment using a logrank test. The sample size is to be large enough to detect non-inferiority when h1 = 2, h2 = 1, and Δ = 0.2. A 3-year study is contemplated with a 1-year, uniform accrual. There is no loss-to-follow up. Alpha is set to 0.05 and power is 0.80. Since this example is for a superiority test, but the validation is for a non-inferiority test, we have to substitute Δ = -0.2 in their calculations.

Chow et al. (2008) carried out their calculations to only two decimal places. Their values, with the substitution described above, yields

$$\begin{aligned} N1 &= \left(\frac{1.64 + 0.84}{2 - 1 + 0.2}\right)^2 (.97 + 3.94) \\ &= 20.9712 \\ &\approx 21 \end{aligned}$$

If they had kept four significant digits, they would have obtained

$$\begin{aligned} N1 &= \left(\frac{1.6449 + 0.8416}{2 - 1 + 0.2}\right)^2 (1.094 + 4.032) \\ &= 22.0086 \\ &\approx 22 \end{aligned}$$

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**
Higher Hazards Are	**Worse (Ha: h2 < h1 + Δ)**
Power	**0.80**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
ω1 (Loss Hazard Rate of Control Group)	**0**
ω2 (Loss Hazard Rate of Treatment Group)	**ω1**
R (Accrual, or Recruitment, Time)	**1**
Percent of R Until 50% are Accrued	**50**
T-R (Follow-Up Time)	**2**
Specify Hazard Parameters Using	**Differences**
h1 (Hazard Rate of Control Group)	**2**
D (Hazard Rate Difference = h2-h1)	**-1**
Δ (Non-Inferiority Margin)	**0.2**

Output

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

Numeric Results

Solve For: Sample Size
Groups: 1 = Control, 2 = Treatment
Hypotheses: H0: h2 ≥ h1 + Δ vs. Ha: h2 < h1 + Δ
Accrual: Uniform

	Sample Size			Hazard Rate		Hazard Rate Difference	Non-Inferiority		Loss Hazard Rate		Time		
Power	N	N1	N2	h1	h2	D	Margin Δ	Boundary B	ω1	ω2	Accrual R	Follow-Up T - R	Alpha
0.8031	45	22	23	2	1	-1	0.2	2.2	0	0	1	2	0.05

	Number of Events			Percent Group 1	Hazard Ratio	Variance	
Power	E	E1	E2	%N1	HR	$\sigma^2(h1)$	$\sigma^2(h2)$
0.8031	42.9	21.8	21	50	0.5	4.032	1.094

The value of N1 = 22 is identical to our hand calculations above and very close to Chow's. Note that since the hand calculated value was slightly greater than 22, **PASS** has added one to N2 so that the total N is 45.