Chapter 733

Equivalence Tests for Two Survival Curves using Cox's Proportional Hazards Model

Introduction

A clinical trial may be employed to test the equivalence of two survival distributions. Because survival times are not normally distributed and because some survival times are censored, Cox proportional-hazards regression is often used to analyze the data. The formulation for testing the significance of a Cox regression coefficient is identical to the standard logrank test. Thus, the power and sample size formulas for one analysis also work for the other.

The Cox Regression model has the added benefit over the exponential model that it does not assume that the hazard rates are constant, but only that they are proportional. That is, that the hazard ratio remains constant throughout the experiment, even if the hazard rates vary.

This procedure is documented in Chow, Shao, and Wang (2008) which summarizes the work of Schoenfeld (1981, 1983). Note that there was an error in Chow, Shao, and Wang (2008) page 179 which caused the sample size to be doubled. This error has been corrected in this edition.

Technical Details

Cox's Proportional Hazards Regression

Cox's proportional hazards regression is widely used for survival data. The regression model is

$$h(t|z) = h(t|0)\exp(bz)$$

where

b is the regression coefficient which is equal to $\log[h(t|1)/h(t|0)] = \log(HR)$

z is a binary indicator variable of treatment group

t is elapsed time

h(*t*|*z*) is the hazard rate at time t, given covariate z

HR is the hazard ratio $h(t|1)/h(t|0)$

Equivalence Hypothesis

The equivalence of two hazard rates is established by concluding that their ratio is within a clinically insignificant margin from one. The statistical hypotheses that yields this conclusion when the null hypothesis is rejected is

$$H_0: |\log(HR)| \geq \log(HR_0) \quad \text{vs.} \quad H_a: |\log(HR)| < \log(HR_0), \quad \text{assuming } HR_0 > 1$$

Test Statistic

It can be shown that the test of *b* based on the partial likelihood method of Cox (1972) coincides with the common logrank test statistic. When testing equivalence, you can use *b* from a Cox regression or calculate the modified logrank statistic as follows

Logrank Test

In this case, the test statistic uses two one-sided logrank tests. Define

$$L_U = \frac{\sum_{k=1}^{K}\left(I_k - \frac{Y_{1i}HR_0}{Y_{1i}HR_0 + Y_{2i}}\right)}{\left[\sum_{k=1}^{K}\left(\frac{Y_{1i}Y_{2i}HR_0}{(Y_{1i}HR_0 + Y_{2i})^2}\right)\right]^{-\frac{1}{2}}}$$

$$L_L = \frac{\sum_{k=1}^{K}\left(I_k - \frac{Y_{1i}(1/HR_0)}{Y_{1i}(1/HR_0) + Y_{2i}}\right)}{\left[\sum_{k=1}^{K}\left(\frac{Y_{1i}Y_{2i}(1/HR_0)}{(Y_{1i}(1/HR_0) + Y_{2i})^2}\right)\right]^{-\frac{1}{2}}}$$

where *K* is the number of deaths, Y_{ij} is the number of subjects at risk just prior to the j^{th} observed event in the i^{th} group, and I_k is a binary variable indicating whether the k^{th} event is from group 1 or not.

The null hypothesis is rejected if $L_U < z_\alpha$ and $L_L > z_{1-\alpha}$.

The distribution of L_U is approximately normal with mean $\left(\log(HR) - \log(HR_0)\right)\sqrt{P_1P_2dN}$ and unit variance and the distribution of L_L is approximately normal with mean $\left(\log(HR) - \log(1/HR_0)\right)\sqrt{P_1P_2dN}$ and unit variance, where

P_1 is the proportion of *N* that is in the control group

P_2 is the proportion of *N* that is in the treatment group

N is the total sample size

N_1 is the sample size from the control group, $N_1 = N(P_1)$

N_2 is the sample size from the treatment group, $N_2 = N(P_2)$

Pev_1 is probability of the event of interest in the control group

Pev_2 is probability of the event of interest in the treatment group

d is the overall probability of an event, $d = Pev_1P_1 + Pev_2P_2$

HR is the observed hazard ratio

HR_0 is the equivalence boundary (limit) of the hazard ratio

Power Calculations

The power of this test is given by

$$\Phi\left(\left(\log(HR_0) - \log(HR_1)\right)\sqrt{P_1P_2dN} - z_{1-\alpha}\right) + \Phi\left(\left(\log(HR_0) + \log(HR_1)\right)\sqrt{P_1P_2dN} - z_{1-\alpha}\right) - 1$$

where HR_1 is the actual value of the hazard ratio under the alternative hypothesis.

Example 1 – Finding the Sample Size

A researcher is planning an equivalence trial using a parallel, two-group, equal sample allocation design to compare the survivability of a new treatment with that of the current treatment. The equivalence hazard ratio is 1.35. The desired power is 0.90 and significance level is 0.05. HR1 will be between 0.9 and 1.1. The probability of observing an event is 0.70 in the control group and 0.40 in the treatment group.

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**
Power	**0.90**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
Pev1 (Event Probability in Group 1)	**0.70**
Pev2 (Event Probability in Group 2)	**0.40**
HR1 (Actual Hazard Ratio)	**0.9 0.95 1.0 1.05 1.1**
HR0 (Equivalence Hazard Ratio)	**1.35**

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: HR ≤ 1/HR0 or HR ≥ HR0 vs. Ha: 1/HR0 < HR < HR0

	Sample Size			Percent Group 1	Number of Events			Hazard Ratio		Event Probability		
Power	N	N1	N2	%N1	E	E1	E2	Equivalence HR0	Actual HR1	Pev1	Pev2	Alpha
0.9001	1643	821	822	50	903.5	574.7	328.8	1.35	0.90	0.7	0.4	0.05
0.9000	1029	514	515	50	565.8	359.8	206.0	1.35	0.95	0.7	0.4	0.05
0.9000	874	437	437	50	480.7	305.9	174.8	1.35	1.00	0.7	0.4	0.05
0.9004	1014	507	507	50	557.7	354.9	202.8	1.35	1.05	0.7	0.4	0.05
0.9001	1486	743	743	50	817.3	520.1	297.2	1.35	1.10	0.7	0.4	0.05

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 groups 1 and 2, respectively.
%N1	The percent of the total sample that is in group 1, the control group.
E, E1, and E2	The number of events required in both groups and groups 1 and 2, respectively.
HR	The hazard ratio. HR = h2/h1
HR0	The equivalence hazard ratio.
HR1	The actual hazard ratio at which power is calculated.
Pev1 and Pev2	The probabilities of an event in the control and the treatment groups, respectively.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A parallel, two-group design will be used to test whether the Group 1 (control) hazard rate is equivalent to the Group 2 (treatment) hazard rate, with an equivalence bound of 1.35 (H0: HR ≤ 1 / 1.35 or HR ≥ 1.35 versus Ha: 1 / 1.35 < HR < 1.35, HR = h2 / h1). The comparison will be made using two one-sided, two-sample Cox proportional hazards regression tests (or equivalent equivalence logrank tests) with an overall Type I error rate (α) of 0.05. It is anticipated that the probability of observing an event during the course of the study is 0.7 for each member of the control group and 0.4 for each member of the treatment group. The calculations are based on the assumption that the hazard ratio is constant throughout the study. To detect a hazard ratio (h2 / h1) of 0.9 with 90% power, the number of needed subjects will be 821 in Group 1 and 822 in Group 2 (totaling 1643 subjects). The corresponding required number of events is 574.7 in Group 1 and 328.8 in Group 2 (totaling 903.5 events).

References

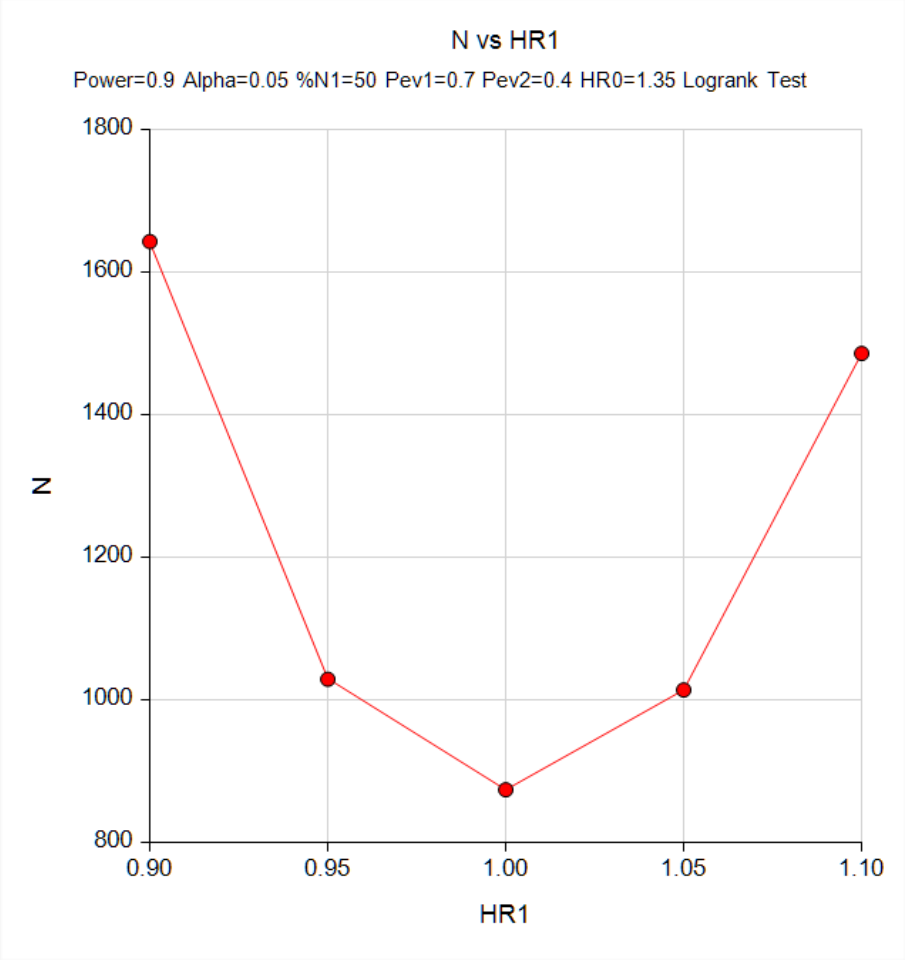
Chow, S.C., Shao, J., Wang, H. 2008. Sample Size Calculations in Clinical Research, 2nd Edition. Chapman & Hall/CRC.

Schoenfeld, David A. 1983. 'Sample Size Formula for the Proportional-Hazards Regression Model', Biometrics, Volume 39, Pages 499-503.

These reports show the values of each of the parameters, one scenario per row.

Plots Section

Plots



This plot shows the relationship between HR1 and N.

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

Chow et al. (2008) page 179 presents an example that we will use for validation. In their example, HR1 = 1, log(HR0) = 0.5, Pev1 = Pev2 = 0.8, P1 = 0.5, alpha = 0.05, and power = 0.8. They should have obtained a total sample size of about 172 (they incorrectly double their result).

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**
Power	**0.80**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
Pev1 (Event Probability in Group 1)	**0.8**
Pev2 (Event Probability in Group 2)	**Pev1**
HR1 (Actual Hazard Ratio)	**1.0**
HR0 (Equivalence Hazard Ratio)	**1.6487 (which is exp(0.5))**

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: HR ≤ 1/HR0 or HR ≥ HR0 vs. Ha: 1/HR0 < HR < HR0

	Sample Size			Percent	Number of Events			Hazard Ratio		Event Probability		
Power	N	N1	N2	Group 1 %N1	E	E1	E2	Equivalence HR0	Actual HR1	Pev1	Pev2	Alpha
0.8021	172	86	86	50	137.6	68.8	68.8	1.649	1	0.8	0.8	0.05

PASS also calculates the value of N = 172.