

Chapter 569

Two-Sample Equivalence Tests for Survival Data using Cox Regression

Introduction

This procedure uses Cox (proportional hazards) regression analysis, which models the relationship between a set of one or more covariates and the hazard rate, for making inference about the equivalence of a treatment hazard rate compared to a control hazard rate using survival data with independent groups. The question of interest is whether the treatment is equivalent to the control, that is, differ at most by a small margin.

Additional discrete or continuous covariates may be included for adjustment before comparison of the two groups. Inference about the equivalence of the treatment versus the control is based on the hazard ratio (or risk ratio), HR, where

$$HR = \frac{\text{Hazard}(\text{Treatment Group})}{\text{Hazard}(\text{Control Group})}$$

The details regarding Cox Regression are available in the Cox Regression chapter of the documentation and will not be repeated here. Only the details related specifically to equivalence testing will be covered in this chapter.

The Cox Regression Model

Survival analysis refers to the analysis of elapsed time. The response variable is the time between a *time origin* and an *end point*. The end point is either the occurrence of the event of interest, referred to as a *death* or *failure*, or the end of the subject's participation in the study. These elapsed times have two properties that invalidate standard statistical techniques, such as t-tests, analysis of variance, and multiple regression. First of all, the time values are often positively skewed. Standard statistical techniques require that the data be normally distributed. Although this skewness could be corrected with a transformation, it is easier to adopt a more realistic data distribution.

The second problem with survival data is that part of the data are *censored*. An observation is censored when the end point has not been reached when the subject is removed from study. This may be because the study ended before the subject's response occurred, or because the subject withdrew from active participation. This may be because the subject died for another reason, because the subject moved, or because the subject quit following the study protocol. All that is known is that the response of interest did not occur while the subject was being studied.

Two-Sample Equivalence Tests for Survival Data using Cox Regression

When analyzing survival data, two functions are of fundamental interest—the *survivor function* and the *hazard function*. Let T be the survival time. That is, T is the elapsed time from the beginning point, such as diagnosis of cancer, and death due to that disease. The values of T can be thought of as having a *probability distribution*. Suppose the *probability density function* of the random variable T is given by $f(T)$. The *probability distribution function* of T is then given by

$$\begin{aligned} F(T) &= \Pr(t < T) \\ &= \int_0^T f(t) dt \end{aligned}$$

The *survivor function*, $S(T)$, is the probability that an individual survives past T . This leads to

$$\begin{aligned} S(T) &= \Pr(T \geq t) \\ &= 1 - F(T) \end{aligned}$$

The *hazard function* is the probability that a subject experiences the event of interest (death, relapse, etc.) during a small time interval given that the individual has survived up to the beginning of that interval. The mathematical expression for the hazard function is

$$\begin{aligned} h(T) &= \lim_{\Delta T \rightarrow 0} \frac{\Pr(T \leq t < (T + \Delta T) | T \leq t)}{F(T + \Delta T) - F(T)} \\ &= \lim_{\Delta T \rightarrow 0} \frac{f(T)}{\Delta T} \\ &= \frac{f(T)}{S(T)} \end{aligned}$$

The cumulative hazard function $H(T)$ is the sum of the individual hazard rates from time zero to time T . The formula for the cumulative hazard function is

$$H(T) = \int_0^T h(u) du$$

Thus, the hazard function is the derivative, or slope, of the cumulative hazard function. The cumulative hazard function is related to the cumulative survival function by the expression

$$S(T) = e^{-H(T)}$$

or

$$H(T) = -\ln(S(T))$$

We see that the distribution function, the hazard function, and the survival function are mathematically related. As a matter of convenience and practicality, the hazard function is used in the basic Cox regression model.

Cox (1972) expressed the relationship between the hazard rate and a set of covariates using the model

$$\ln[h(T)] = \ln[h_0(T)] + \sum_{i=1}^p x_i \beta_i$$

or

$$h(T) = h_0(T) e^{\sum_{i=1}^p x_i \beta_i}$$

Two-Sample Equivalence Tests for Survival Data using Cox Regression

where x_1, x_2, \dots, x_p are covariates, $\beta_1, \beta_2, \dots, \beta_p$ are regression coefficients to be estimated, T is the elapsed time, and $h_0(T)$ is the baseline hazard rate when all covariates are equal to zero. Thus the linear form of the regression model is

$$\ln \left[\frac{h(T)}{h_0(T)} \right] = \sum_{i=1}^p x_i \beta_i$$

Taking the exponential of both sides of the above equation, we see that this is the ratio between the actual hazard rate and the baseline hazard rate (i.e. the hazard ratio), and is sometimes called the *relative risk*. This can be rearranged to give the model

$$\begin{aligned} \frac{h(T)}{h_0(T)} &= \exp \left(\sum_{i=1}^p x_i \beta_i \right) \\ &= e^{x_1 \beta_1} e^{x_2 \beta_2} \dots e^{x_p \beta_p} \end{aligned}$$

The regression coefficients can thus be interpreted as the relative risk when the value of the covariate is increased by one unit.

Note that unlike most regression models, this model does not include an intercept term. This is because if an intercept term were included, it would become part of $h_0(T)$.

Also note that the above model does not include T on the right-hand side. That is, the relative risk is constant for all time values. This is why the method is called *proportional hazards*.

An interesting attribute of this model is that you only need to use the ranks of the failure times to estimate the regression coefficients. The actual failure times are not used except to generate the ranks. Thus, you will achieve the same regression coefficient estimates regardless of whether you enter the time values in days, months, or years.

TOST Equivalence Test (Two One-Sided Tests) using the Cox Proportional Hazard Regression Model

Schuurmann's (1987) two one-sided tests (TOST) approach is used to test equivalence. This equivalence test essentially reverses the roles of the usual null and alternative hypothesis. Using $HR = \frac{h_T(T)}{h_C(T)}$ (where $h_T(T)$ is the hazard rate of the treatment group at time T and $h_C(T)$ is the hazard rate of the reference or control group at time T) to represent the hazard ratio of the treatment and the control, L to represent the lower equivalence bound, and U to represent the upper equivalence bound, the null and alternative hypotheses are

$$\begin{aligned} H_0: HR < L \text{ or } HR > U \\ H_1: L < HR < U \end{aligned}$$

The null hypothesis is made up of two simple one-sided hypotheses:

$$\begin{aligned} H_{01}: HR < L \\ H_{02}: HR > U \end{aligned}$$

The null hypotheses can be tested at significance level α by constructing a $100(1 - 2\alpha)\%$ confidence interval for HR (not the usual $100(1 - \alpha)\%$ like you might suspect). If the upper bound of the $100(1 - 2\alpha)\%$ confidence interval is less than U , and the lower bound is greater than L , then you can conclude that the treatment and the control are equivalent. This is the same as conducting two one-sided tests of the null hypotheses using the Wald test. If both of these one-sided tests are rejected, we conclude H1 that the two groups are equivalent (their difference is confined within a small margin). Schuurmann showed that if we want the alpha level of the equivalence test to be α , then each of the one-sided tests should be α as well (not $\alpha/2$ as you might expect). The probability level (p-value) of the equivalence test is equal to the maximum of the probability levels of the two one-sided tests.

Two-Sample Equivalence Tests for Survival Data using Cox Regression

Hazard Ratio (or Risk Ratio)

The hazard ratio for the group variable is computed from the Cox regression coefficient for the grouping variable as

$$HR = \exp(b_{Grp})$$

100(1 – 2α)% Confidence Interval for HR

The 100(1 – 2α)% confidence interval for the hazard ratio computed by NCSS is based on the Wald statistic, which is valid for large samples. The formula for the limits of a 100(1 – 2α)% two-sided confidence interval for HR is

$$\exp\left(b_{Grp} \pm z_{1-\alpha} s_{b_{Grp}}\right),$$

where b_{Grp} is the Cox regression coefficient corresponding to the Group variable, and $s_{b_{Grp}}$ is the standard error of b_{Grp} .

Wald Test

The formula for the Wald test statistics are

$$Z_L = \frac{b_{Grp} - \ln(L)}{s_{b_{Grp}}} \quad \text{and} \quad Z_U = \frac{b_{Grp} - \ln(U)}{s_{b_{Grp}}}$$

where b_{Grp} is the Cox regression coefficient corresponding to the Group variable, $s_{b_{Grp}}$ is the standard error of b_{Grp} , and L and U are the lower and upper equivalence bounds, respectively. With large sample sizes, the distribution of Z is closely approximated by the normal distribution. With small and moderate sample sizes, the normal approximation is described as “adequate.”

Data Structure

Survival data sets require up to three components for the survival time: the ending survival time, the beginning survival time during which the subject was not observed, and an indicator of whether the observation was censored or failed.

Based on these three components, various types of data may be analyzed. *Right censored* data are specified using only the ending time variable and the censor variable. *Left truncated* and *Interval* data are entered using all three variables.

The table below shows survival data ready for analysis. These data are in the CoxRegSub dataset. The variables are

Time	months of survival
Status	censor indicator
Treatment	treatment category (C = Control, T1 = Treatment)
Age	age of patient in years
Gender	gender of patient
Count	frequency (or count) variable

Two-Sample Equivalence Tests for Survival Data using Cox Regression

CoxRegSub dataset (subset)

Time	Censor	Treatment	Age	Gender	Count
14.29	1	C	20	F	1
16.68	0	C	20	F	5
3.38	1	C	20	F	1
3.97	1	C	20	F	1
10.43	1	C	20	M	1
11.66	0	C	20	M	4
6.51	1	C	20	M	1
11.86	1	C	20	M	1
5.71	1	T1	20	F	1
6.35	0	T1	20	F	6
11.36	1	T1	20	F	1
8.43	1	T1	20	F	1
.
.
.

Procedure Options

This section describes the options available in this procedure that are basic to equivalence testing. All other options are described in the Cox Regression chapter.

Variables, Model Tab

This panel lets you designate which variables and model are used in the analysis.

Variables

Time Variable

This variable contains the length of time that an individual was observed. This may represent a failure time or a censor time. Whether the subject actually died is specified by the Censor variable. Since the values are elapsed times, they must be positive. Zeroes and negative values are treated as missing values.

During the maximum likelihood calculations, a risk set is defined for each individual. The risk set is defined to be those subjects who were being observed at this subject's failure and who lived as long or longer. It may take several rows of data to specify a subject's history.

This variable and the Entry Time variable define a period during which the individual was at risk of failing. If the Entry Time is not specified, its value is assumed to be zero.

Several types of data may be entered. These will be explained next.

- **Failure**

This type of data occurs when a subject is followed from their entrance into the study until their death. The failure time is entered in this variable and the Censor variable is set to the failed code, which is often a one.

The Entry Time Variable is not necessary. If an Entry Time variable is used, its value should be zero for this type of observation.

Two-Sample Equivalence Tests for Survival Data using Cox Regression

- **Interval Failure**

This type of data occurs when a subject is known to have died during a certain interval. The subject may, or may not, have been observed during other intervals. If they were, they are treated as Interval Censored data. An individual may require several rows on the database to record their complete follow-up history.

For example, suppose the condition of the subjects is only available at the end of each month. If a subject fails during the fifth month, two rows of data would be required. One row, representing the failure, would have a Time of 5.0 and an Entry Time of 4.0. The Censor variable would contain the failure code. A second row, representing the prior periods, would have a Time of 4.0 and an Entry Time of 0.0. The Censor variable would contain the censor code.

- **Censored**

This type of data occurs when a subject has not failed up to the specified time. For example, suppose that a subject enters the study and does not die until after the study ends 12 months later. The subject's time (365 days) is entered here. The Censor variable contains the censor code.

- **Interval Censored**

This type of data occurs when a subject is known not to have died during a certain interval. The subject may, or may not, have been observed during other intervals. An individual may require several rows on the database to record their complete follow-up history.

For example, suppose the condition of the subjects is only available at the end of each month. If a subject fails during the fifth month, two rows of data would be required. One row, representing the failure, would have a Time of 5.0 and an Entry Time of 4.0. The Censor variable would contain the failure code. A second row, representing the prior periods, would have a Time of 4.0 and an Entry Time of 0.0. The Censor variable would contain the censor code.

Ties Method

The basic Cox regression model assumes that all failure times are unique. When ties exist among the failure times, one of two approximation methods is used to deal with the ties. When no ties are present, both of these methods result in the same estimates.

- **Breslow**

This method was suggested first and is the default in many programs. However, the Efron method has been shown to be more accurate in most cases. The Breslow method is only used when you want to match the results of some other (older) Cox regression package.

- **Efron**

This method has been shown to be more accurate, but requires slightly more time to calculate. This is the recommended method.

Group Variable

Specify the column that contains the treatment group for each row. This variable must be categorical and have exactly 2 unique values.

Reference Group

This is the value to be used as the reference or control group to which the treatment group will be compared. If this option is left blank, the first value in the Group Variable after sorting is used as the reference group.

Entry Time Variable

This optional variable contains the elapsed time before an individual entered the study. Usually, this value is zero. However, in cases such as *left truncation* and *interval censoring*, this value defines a time period before which the individual was not observed.

Negative entry times are treated as missing values. It is possible for the entry time to be zero.

Two-Sample Equivalence Tests for Survival Data using Cox Regression

Censor Variable

The values in this variable indicate whether the value of the Time Variable represents a censored time or a failure time. These values may be text or numeric. The interpretation of these codes is specified by the Failed and Censored options to the right of this option.

Only two values are used, the Failure code and the Censor code. The Unknown Type option specifies what is to be done with values that do not match either the Failure code or the Censor code.

Rows with missing values (blanks) in this variable are omitted from the estimation phase, but results are shown in any reports that output predicted values.

Failure

This value identifies those values of the Censor Variable that indicate that the Time Variable gives a failure time. The value may be a number or a letter.

We suggest the letter 'F' or the number '1' when you are in doubt as to what to use.

A failed observation is one in which the time until the event of interest was measured exactly; for example, the subject died of the disease being studied. The exact failure time is known.

Left Censoring

When the exact failure time is not known, but instead only an upper bound on the failure time is known, the time value is said to have been *left censored*. In this case, the time value is treated as if it were the true failure time, not just an upper bound. So left censored observations should be coded as failed observations.

Censor

This value identifies those values of the Censor Variable that indicate that the individual recorded on this row was censored. That is, the actual failure time occurs sometime after the value of the Time Variable.

We suggest the letter 'C' or the number '0' when you are in doubt as to what to use.

A censored observation is one in which the time until the event of interest is not known because the individual withdrew from the study, the study ended before the individual failed, or for some similar reason.

Note that it does not matter whether the censoring was Right or Interval. All you need to indicate here is that they were censored.

Other Censor

This option specifies what the program is to assume about observations whose censor value is not equal to either the *Failure* code or the *Censor* code. Note that observations with missing censor values are always treated as missing.

- **Censored**
Observations with unknown censor values are assumed to have been censored.
- **Failed**
Observations with unknown censor values are assumed to have failed.
- **Missing**
Observations with unknown censor values are assumed to be missing and they are removed from the analysis.

Frequencies

This is an optional variable containing the frequency (observation count) for each row. Usually, you would leave this option blank and let each row receive the default frequency of one.

If your data have already been summarized, this option lets you specify how many actual rows each physical row represents.

Equivalence Test Options

Hazard Ratio Upper (Lower) Equivalence Bound

Enter the desired value of the upper (lower) equivalence bound for the hazard ratio. If the upper limit of the $100(1-2\alpha)\%$ confidence interval for HR is less than the Upper Equivalence Bound and the lower limit is greater than the Lower Equivalence Bound, then the treatment will be considered to be equivalent to the control or reference.

Additional Covariates

All options in this section not included below are described in detail in the Cox Regression chapter and will not be repeated here.

Adjust for Covariates

Because this method utilizes Cox Regression, it is possible to adjust the results of the two-sample test for both numeric and categorical covariates. Check this option to include additional variables in the model. The two-sample test will be conducted based on hazard ratio estimates that are adjusted for these covariates.

Iteration Tab

This panel lets you control the maximum likelihood estimation algorithm.

Iteration Options

These options control the number of iterations used while the algorithm is searching for the maximum likelihood solution.

Maximum Iterations

This option specifies the maximum number of iterations used while finding a solution. If this number is reached, the procedure is terminated prematurely. This is used to prevent an infinite loop and to reduce the running time of lengthy variable selection runs.

Usually, no more than 20 iterations are needed. In fact, most runs converge in about 7 or 8 iterations.

During a variable selection run, it may be advisable to reset this value to 4 or 5 to speed up the variable selection. Usually, the last few iterations make little difference in the estimated values of the regression coefficients.

Convergence Zero

This option specifies the convergence target for the maximum likelihood estimation procedure. The algorithm finds the maximum relative change of the regression coefficients. If this amount is less than the value set here, the maximum likelihood procedure is terminated.

For large datasets, you might want to increase this value to about 0.0001 so that fewer iterations are used, thus decreasing the running time of the procedure.

Regression Coefficient Starting Values

These options control the starting regression coefficient values (the B's).

Start B's at

Select a starting value (or enter a list of individual starting values) for the regression coefficients (the B's).

Although the B's can be any numeric value, the typical range is between -1 and 1. So starting values in this range usually allow the algorithm to converge to a useful solution.

The Cox regression algorithm solves for the maximum likelihood estimates of the regression coefficients by the iterative Newton-Raphson algorithm. This algorithm begins at a set of starting values for the regression

Two-Sample Equivalence Tests for Survival Data using Cox Regression

coefficients and, at each iteration, modifies the B's in a way that leads to a local maximum of the likelihood function. Sometimes the algorithm does not converge to the global maximum, so a different set of starting values must be tried.

Typically, a 'good' solution is one with all B's less than 50 in absolute value. If your solution has one or more B's that are large (over 10,000), you should rerun the algorithm with a different set of starting values.

List of Starting B's (If Start B's at = "List")

Enter a list of starting values. If not enough values are entered, the last value will be used over and over.

Although the B's can be any numeric value, the typical range is between -1 and 1. So starting values in this range usually allow the algorithm to converge to a useful solution.

Reports Tab

The following options control which reports are displayed.

Alpha

Alpha Level

Alpha is the significance level used in the hypothesis tests. One minus alpha is the confidence level of the confidence intervals. A value of 0.05 is most commonly used. This corresponds to a chance of error of 1 in 20. You should not be afraid to use other values since 0.05 became popular in pre-computer days when it was the only value available.

Typical values range from 0.001 to 0.20.

Select Reports – Summary

Run Summary

Indicate whether to display this summary report.

Select Reports – Test

100(1 – 2 α)% Confidence Interval Test for Equivalence (Wald Test)

Indicate whether to display the equivalence test report.

Select Reports – Estimation

Regression Coefficients ... C.L. of Regression Coefficients

Indicate whether to display these estimation reports.

Select Reports – Goodness-of-Fit

Analysis of Deviance ... Baseline Hazard and Survival

Indicate whether to display these model goodness-of-fit reports.

Select Reports – Row-by-Row Lists

Residuals ... Predicted Values

Indicate whether to display these list reports. Note that since these reports provide results for each row, they may be too long for normal use when requested on large databases.

Two-Sample Equivalence Tests for Survival Data using Cox Regression

Order of Row Reports

This option specifies the order of the observations displayed on reports that display a separate value for each row. The rows can be displayed in the original order of the database or sorted by the time value, from lowest to highest.

Report Options Tab

The options on this panel control the label and decimal options of the report.

Variable Labels

Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

Stagger label and output if label length is \geq

The names of the indicator variables can be too long to fit in the space provided. If the name contains more characters than the number specified here, only the name is shown on the first line of the report and the rest of the output is placed on the next line.

Enter *1* when you want the each variable's results printed on two lines.

Enter *100* when you want each variable's results printed on a single line.

Decimal Places

Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

Time ... Z Decimals

These options specify the number of decimal places shown on the reports for the indicated values.

Plots Tab

These options control the attributes of the various plots.

Select Plots

Null Martingale Resid vs X Plot ... Deviance Resid vs Time Plot

Indicate whether to display these plots. Click the plot format button to change the plot settings.

Storage Tab

These options let you specify if, and where on the dataset, various statistics are stored.

Warning: Any data already in these columns are replaced by the new data. Be careful not to specify columns that contain important data.

Data Storage Options

Storage Option

This option controls whether the values indicated below are stored on the dataset when the procedure is run.

- **Do not store data**
No data are stored even if they are checked.
- **Store in empty columns only**
The values are stored in empty columns only. Columns containing data are not used for data storage, so no data can be lost.
- **Store in designated columns**
Beginning at the *Store First Item In*, the values are stored in this column and those to the right. If a column contains data, the data are replaced by the storage values. Care must be used with this option because it cannot be undone.

Store First Variable In

The first item is stored in this column. Each additional item that is checked is stored in the columns immediately to the right of this column.

Leave this value blank if you want the data storage to begin in the first blank column on the right-hand side of the data.

Warning: any existing data in these columns is automatically replaced, so be careful.

Data Storage Options – Select Items to Store

Expanded X Values ... Covariance Matrix

Indicate whether to store these row-by-row values, beginning at the column indicated by the *Store First Item In* option. Note that several of these values include a different value for each covariate and so they require several columns when they are stored.

Expanded X Values

This option refers to the experimental design matrix. They include all binary and interaction variables generated.

Example 1 – Two-Sample TOST Equivalence Test for Survival Data

This section presents an example of how to perform a two-sample equivalence test with survival data using Cox regression. Researchers want to determine if a new treatment (T1) is equivalent to the current standard treatment (C). The upper and lower hazard ratio bounds of equivalence are set at 1.25 and 0.8, respectively. The data for this study are contained in the **CoxRegSub** dataset.

Note: This example only shows the output directly related to the equivalence test. For description about the other reports available, see the Cox Regression chapter.

You may follow along here by making the appropriate entries or load the completed template **Example 1** by clicking on Open Example Template from the File menu of the Two-Sample Equivalence Tests for Survival Data using Cox Regression window.

1 Open the CoxRegSub dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **CoxRegSub.NCSS**.
- Click **Open**.

2 Open the Two-Sample Equivalence Tests for Survival Data using Cox Regression window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Two-Sample Equivalence Tests for Survival Data using Cox Regression** procedure.
- On the menus, select **File**, then **New Template**. This will load the default template.

3 Specify the variables.

- Select the **Variables, Model tab**.
- Enter **Time** in the **Time** box.
- Enter **Treatment** in the **Group** box.
- For **Reference Group**, enter **C**.
- Enter **Censor** in the **Censor Variable** box.
- Enter **Count** in the **Frequency Variable** box.

4 Specify the equivalence test options.

- Set **Hazard Ratio Upper Equivalence Bound** to **1.25**.
- Set **Hazard Ratio Lower Equivalence Bound** to **0.80**.

5 Specify the reports.

- Select the **Reports tab**.
- Leave only the default reports checked.

6 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

The following reports and charts will be displayed in the Output window.

Two-Sample Equivalence Tests for Survival Data using Cox Regression

Run Summary

Parameter	Value	Parameter	Value
Rows Read	49	Time Variable	Time
Rows Filtered Out	0	Censor Variable	Censor
Rows Missing X's	0	Frequency Variable	Count
Rows Processed	49		
Rows Prediction Only	0	Ind. Var's Available	1
Rows Failed	36	No. of X's in Model	1
Rows Censored	13	Iterations	3
Sum of Frequencies	93	Maximum Iterations	20
Sum Censored Freqs	57	Convergence Criterion	1E-09
Sum Failed Freqs	36	Achieved Convergence	4.853289E-10
Final Log Likelihood	-142.3606	Completion Status	Normal completion
Starting B's	0		

This report summarizes the characteristics of the dataset and provides useful information about the analysis. It should be studied to make sure that the data were read in properly and that the estimation algorithm terminated normally. We will only discuss those parameters that need special explanation.

Rows Read

This is the number of rows processed during the run. Check this count to make certain it agrees with what you anticipated.

Iterations

This is the number of iterations used by the maximum likelihood procedure. This value should be compared against the value of the Maximum Iterations option to see if the iterative procedure terminated early.

Achieved Convergence

This is the maximum of the relative changes in the regression coefficients on the last iteration. If this value is less than the Convergence Criterion, the procedure converged normally. Otherwise, the specified convergence precision was not achieved.

Final Log Likelihood

This is the log likelihood of the model.

100(1 - 2 α)% Confidence Interval Test for Equivalence (Two One-Sided Wald Tests)

100(1 - 2 α)% Confidence Interval Test for Equivalence (Two One-Sided Wald Tests)

Hazard Ratio (HR) = Hazard(Treatment="T1") / Hazard(Treatment="C")

Equivalence Hypothesis: $0.8 < HR < 1.25$

Test	Alternative Hypothesis	Hazard Ratio HR*	Lower 90.0% C.L. of HR	Upper 90.0% C.L. of HR	Wald Z-Value	Wald Prob Level	Reject H0 at $\alpha = 0.05$?
Lower Boundary	HR > 0.8	0.8590	0.4958	1.4884	0.2131	0.4156	No
Upper Boundary	HR < 1.25	0.8590	0.4958	1.4884	-1.1225	0.1308	No
Equivalence	$0.8 < HR < 1.25$	0.8590	0.4958	1.4884		0.4156	No

* In Cox Regression, the Hazard Ratio (HR) is commonly referred to as the Risk Ratio and is equal to $\text{Exp}(B)$, where B is the estimated regression coefficient.

This report displays the results of the two one-sided tests for equivalence. The Wald P-value for the overall equivalence test is 0.4156, indicating that we cannot reject the null and cannot conclude equivalence at the 0.05 level. Notice that the upper limit of the 90% confidence interval for the hazard ratio (HR) is not less than the equivalence bound of 1.25 and the lower limit of the confidence interval is not greater than 0.80.

Analysis of Deviance Section

Term(s) Omitted	DF	-2 Log Likelihood	Increase From Model Deviance (Chi ²)	Prob Level
All Terms	1	284.9279	0.2066	0.6495
Treatment	1	284.9279	0.2066	0.6495
None(Model)	1	284.7213		

This report is the Cox regression analog of the analysis of variance table. It displays the results of a chi-square test used to test whether each of the individual terms are statistically significant after adjusting for all other terms in the model.

The DF (degrees of freedom) column indicates the number of binary variables needed to represent each term. The chi² test is used to test the significance of all binary variables associated with a particular term.

Log Likelihood & R² Section

Term(s) Omitted	DF	Log Likelihood	R ² Of Remaining Term(s)	Reduction From Model R ²
All Terms	1	-142.4639	0.0000	0.0022
Treatment	1	-142.4639	0.0000	0.0022
None(Model)	1	-142.3606	0.0022	0.0000

This report displays the Log Likelihood and R² that is achieved when each term is omitted from the regression model. The DF (degrees of freedom) column indicates the number of binary variables needed to represent each term. The chi² test is used to test the significance of all binary variables associated with a particular term.

Example 2 – Two-Sample TOST Equivalence Test for Survival Data with Adjustment for Additional Covariates

Continuing with Example 1, this example will demonstrate how to include additional covariates in the model in which the Group variable is tested for equivalence.

You may follow along here by making the appropriate entries or load the completed template **Example 2** by clicking on Open Example Template from the File menu of the Two-Sample Equivalence Tests for Survival Data using Cox Regression window.

1 Open the CoxRegSub dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **CoxRegSub.NCSS**.
- Click **Open**.

2 Open the Two-Sample Equivalence Tests for Survival Data using Cox Regression window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Two-Sample Equivalence Tests for Survival Data using Cox Regression** procedure.
- On the menus, select **File**, then **New Template**. This will load the default template.

3 Specify the variables.

- Select the **Variables, Model** tab.
- Enter **Time** in the **Time** box.
- Enter **Treatment** in the **Group** box.
- For **Reference Group**, enter **C**.
- Enter **Censor** in the **Censor Variable** box.
- Enter **Count** in the **Frequency Variable** box.

4 Specify the equivalence test options.

- Set **Hazard Ratio Upper Equivalence Bound** to **1.25**.
- Set **Hazard Ratio Lower Equivalence Bound** to **0.80**.

5 Specify the additional covariates.

- Check **Adjust for Covariates**.
- For **Categorical X's** enter **Age, Gender**.

6 Specify the reports.

- Select the **Reports** tab.
- Leave all the default reports checked.
- Check **Regression Coefficients** and **Confidence Limits of Regression Coefficients**.

7 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

Two-Sample Equivalence Tests for Survival Data using Cox Regression

Output

Run Summary

Parameter	Value	Parameter	Value
Rows Read	49	Time Variable	Time
Rows Filtered Out	0	Censor Variable	Censor
Rows Missing X's	0	Frequency Variable	Count
Rows Processed	49		
Rows Prediction Only	0	Ind. Var's Available	3
Rows Failed	36	No. of X's in Model	4
Rows Censored	13	Iterations	5
Sum of Frequencies	93	Maximum Iterations	20
Sum Censored Freqs	57	Convergence Criterion	1E-09
Sum Failed Freqs	36	Achieved Convergence	1.996924E-16
Final Log Likelihood	-141.0197	Completion Status	Normal completion
Starting B's	0		

100(1 - 2 α)% Confidence Interval Test for Equivalence (Two One-Sided Wald Tests)

Hazard Ratio (HR) = Hazard(Treatment="T1") / Hazard(Treatment="C")

Additional Categorical X's: Age, Gender

Equivalence Hypothesis: 0.8 < HR < 1.25

Test	Alternative Hypothesis	Hazard Ratio HR*	Lower 90.0% C.L. of HR	Upper 90.0% C.L. of HR	Wald Z-Value	Wald Prob Level	Reject H0 at $\alpha = 0.05$?
Lower Boundary	HR > 0.8	0.8108	0.4599	1.4296	0.0390	0.4844	No
Upper Boundary	HR < 1.25	0.8108	0.4599	1.4296	-1.2555	0.1046	No
Equivalence	0.8 < HR < 1.25	0.8108	0.4599	1.4296		0.4844	No

* In Cox Regression, the Hazard Ratio (HR) is commonly referred to as the Risk Ratio and is equal to $\text{Exp}(B)$, where B is the estimated regression coefficient.

Regression Coefficients

Independent Variable	Regression Coefficient (B)	Standard Error of B	Risk Ratio Exp(B)	Mean	Wald Z-Value	Prob Level	Pseudo R ²
B1 (Treatment="T1")	-0.209688	0.344742	0.8108	0.4946	-0.6082	0.5430	0.0105
B2 (Age=40)	0.232735	0.413718	1.2620	0.3226	0.5625	0.5737	0.0090
B3 (Age=60)	0.695165	0.425375	2.0040	0.3333	1.6342	0.1022	0.0712
B4 (Gender="M")	0.084476	0.344212	1.0881	0.5269	0.2454	0.8061	0.0017

Confidence Limits

Independent Variable	Regression Coefficient (B)	Lower 95.0% Confidence Limit of B	Upper 95.0% Confidence Limit of B	Risk Ratio Exp(B)	Lower 95.0% C.L. of Exp(B)	Upper 95.0% C.L. of Exp(B)
B1 (Treatment="T1")	-0.209688	-0.885371	0.465994	0.8108	0.4126	1.5936
B2 (Age=40)	0.232735	-0.578138	1.043608	1.2620	0.5609	2.8394
B3 (Age=60)	0.695165	-0.138555	1.528885	2.0040	0.8706	4.6130
B4 (Gender="M")	0.084476	-0.590167	0.759119	1.0881	0.5542	2.1364

Analysis of Deviance

Term(s) Omitted	DF	-2 Log Likelihood	Increase From Model Deviance (Chi ²)	Prob Level
All Terms	4	284.9279	2.8884	0.5767
Treatment	1	282.4091	0.3697	0.5432
Age	2	284.7057	2.6662	0.2637
Gender	1	282.0997	0.0603	0.8061
None(Model)	4	282.0394		

The Prob Level is for testing the significance of that term after adjusting for all other terms.

Two-Sample Equivalence Tests for Survival Data using Cox Regression

Log Likelihood & R ²				
Term(s) Omitted	DF	Log Likelihood	R ² Of Remaining Term(s)	Reduction From Model R ²
All Terms	4	-142.4639	0.0000	0.0306
Treatment	1	-141.2046	0.0267	0.0039
Age	2	-142.3528	0.0024	0.0282
Gender	1	-141.0498	0.0300	0.0006
None(Model)	4	-141.0197	0.0306	0.0000

Adjusting for other covariates, it is still not possible to reject the null hypothesis and conclude equivalence.

The Regression Coefficients and Confidence Limits reports give the Cox regression parameter estimates for both the group variable and the other variables included in the model. The hypotheses tested by these reports are that the regression coefficients are equal to zero.