Chapter 555

Kaplan-Meier Curves (Logrank Tests)

Introduction

This procedure computes the nonparametric Kaplan-Meier and Nelson-Aalen estimates of survival and associated hazard rates. It can fit complete, right censored, left censored, interval censored (readout), and grouped data values. It outputs various statistics and graphs that are useful in reliability and survival analysis.

It also performs several logrank tests and provides both the parametric and randomization test significance levels. This procedure also computes restricted mean survival time (RMST) and restricted mean time lost (RMTL) statistics and associated between-group comparisons.

Overview of Survival Analysis

We will give a brief introduction to the subject in this section. For a complete account of survival analysis, we suggest the book by Klein and Moeschberger (2003).

Survival analysis is the study of the distribution of life times. That is, it is the study of the elapsed time between an initiating event (birth, start of treatment, diagnosis, or start of operation) and a terminal event (death, relapse, cure, or machine failure). The data values are a mixture of complete (terminal event occurred) and censored (terminal event has not occurred) observations. From the data values, the survival analyst makes statements about the survival distribution of the failure times. This distribution allows questions about such quantities as survivability, expected life time, and mean time to failure to be answered.

Let T be the elapsed time until the occurrence of a specified event. The event may be death, occurrence of a disease, disappearance of a disease, appearance of a tumor, etc. The probability distribution of T may be specified using one of the following basic functions. Once one of these functions has been specified, the others may be derived using the mathematical relationships presented.

1. Probability density function, $f(t)$. This is the probability that an event occurs at time t.
2. Cumulative distribution function, $F(t)$. This is the probability that an individual survives until time t.

$$F(t) = \int_0^t f(x)dx$$

3. Survival function, *S(T)*. This is the probability that an individual survives beyond time *T*. This is usually the primary quantity of interest. It is estimated using the nonparametric Kaplan-Meier curve.

$$S(T) = \int_T^\infty f(x)dx$$
$$= 1 - F(T)$$

$$S(T) = \exp\left[-\int_0^T h(x)dx\right]$$
$$= \exp[-H(T)]$$

4. Hazard rate, *h(T)*. This is the probability that an individual at time *T* experiences the event in the next instant. It is a fundamental quantity in survival analysis. It is also known as the conditional failure rate in reliability, the force of mortality in demography, the intensity function in stochastic processes, the age-specific failure rate in epidemiology, and the inverse of Mill's ratio in economics. The empirical hazard rate may be used to identify the appropriate probability distribution of a particular mechanism, since each distribution has a different hazard rate function. Some distributions have a hazard rate that decreases with time, others have a hazard rate that increases with time, some are constant, and some exhibit all three behaviors at different points in time.

$$h(T) = \frac{f(T)}{S(T)}$$

5. Cumulative hazard function, *H*(*T*). This is integral of *h*(*T*) from 0 to *T*.

$$H(T) = \int_0^T h(x)dx$$
$$= -\ln[S(T)]$$

Nonparametric Estimators of Hazard and Survival

All of the following results are from Klein and Moeschberger (2003).

The recommended nonparametric estimator of the survival distribution, *S*(*T*), is the Kaplan-Meier product-limit estimator. The recommended nonparametric estimator of the cumulative hazard function, *H*(*T*), is the Nelson-Aalen estimator. Although each of these estimators could be used to estimate the other quantity using the relationship

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

or

$$S(T) = \exp[-H(T)]$$

this is not recommended.

The following notation will be used to define both of these estimators. Let $t = 1, \cdots, M$ index the *M* unique termination (failure or death) times $T_1, T_2, ..., T_M$. Note that *M* does not include duplicate times or times at which only censored observations occur. Associated with each of these failure times is an entry time E_t at which the subject began to be observed. Usually, these entry times are taken to be zero. If positive entry times are specified, the data are said to have been *left truncated*. When data are left truncated, it is often necessary to define a minimum time, *A*, below which failures are not considered. When a positive *A* is used, the unconditional survival function *S*(*T*) is changed to a conditional survival function *S*(*T*|*T*>*A*).

The set of all failures (deaths) that occur at time T_t is referred to as D_t and the number in this set is given by d_t. The *risk set* at *t*, R_t, is the set of all individuals that are at risk immediately before time T_t. This set includes all individuals whose entry and termination times include T_t. That is, R_t is made up of all individuals with times such that $E_j < T_t \le T_j$ and $A \le T_t$. The number of individuals in the risk set is given by r_t.

Kaplan-Meier Product-Limit Estimator

Using the above notation, the Kaplan-Meier product-limit estimator is defined as follows in the range of time values for which there are data.

$$\hat{S}(T) = \begin{cases} 1 & \text{if } T_{\min} > T \\ \prod_{A \le T_i \le T} \left[1 - \frac{d_i}{r_i}\right] & \text{if } T_{\min} \le T \end{cases}$$

The variance of *S(T)* is estimated by Greenwood's formula

$$\hat{V}\left[\hat{S}(T)\right] = \hat{S}(T)^2 \sum_{A \le T_i \le T} \frac{d_i}{r_i(r_i - d_i)}$$

Pointwise Confidence Intervals of Survival

A pointwise confidence interval for the survival probability at a specific time T_0 of $S(T_0)$ is represented by two confidence limits which have been constructed so that the probability that the true survival probability lies between them is $1-\alpha$. Note that these limits are constructed for a single time point. Several of them cannot be used together to form a confidence band such that the entire survival function lies within the band. When these are plotted with the survival curve, these limits must be interpreted on an individual, point by point, basis.

Three difference confidence intervals are available. All three confidence intervals perform about the same in large samples. The linear (Greenwood) interval is the most commonly used. However, the log-transformed and the arcsine-square intervals behave better in small to moderate samples, so they are recommended.

Linear (Greenwood) Pointwise Confidence Interval for S(T)

This estimator may be used to create a confidence interval at a specific time point T_0 of $S(T_0)$ using the formula

$$\hat{S}(T_0) \pm z_{1-\alpha/2}\sigma_S(T_0)$$

where

$$\sigma_S^2(T_0) = \frac{\hat{V}\left[\hat{S}(T_0)\right]}{\hat{S}^2(T_0)}$$

and *z* is the appropriate value from the standard normal distribution.

Log-Transformed Pointwise Confidence Interval for S(T)

Better confidence limits may be calculated using the logarithmic transformation of the hazard functions. These limits are

$$\hat{S}(T_0)^{1/\theta}, \hat{S}(T_0)^{\theta}$$

where

$$\theta = \exp\left\{ \frac{z_{1-\alpha/2}\sigma_S(T_0)}{\log[\hat{S}(T_0)]} \right\}$$

ArcSine-Square Root Pointwise Confidence Interval for S(T)

Another set of confidence limits using an improving transformation is given by the formula

$$\sin^2\left\{ \max\left[0, \arcsin\left\{\hat{S}(T_0)\right\}^{1/2} - 0.5 z_{1-\alpha/2}\sigma_S(T_0)\left(\frac{\hat{S}(T_0)}{1-\hat{S}(T_0)}\right)^{1/2} \right] \right\}$$
$$\leq S(T_0) \leq$$
$$\sin^2\left\{ \min\left[\frac{\pi}{2}, \arcsin\left\{\hat{S}(T_0)\right\}^{1/2} + 0.5 z_{1-\alpha/2}\sigma_S(T_0)\left(\frac{\hat{S}(T_0)}{1-\hat{S}(T_0)}\right)^{1/2} \right] \right\}$$

Nelson-Aalen Hazard Estimator

The Nelson-Aalen estimator is recommended as the best estimator of the cumulative hazard function, $H(T)$. This estimator is give as

$$\tilde{H}(T) = \begin{cases} 0 & if\ T_{\min} > T \\ \sum_{A \leq T_i \leq T} \frac{d_i}{r_i} & if\ T_{\min} \leq T \end{cases}$$

Three estimators of the variance of this estimate are mentioned on page 34 of Therneau and Grambsch (2000). These estimators differ in the way they model tied event times. When there are no event time ties, they give almost identical results.

1. Simple (Poisson) Variance Estimate

This estimate assumes that event time ties occur because of rounding and a lack of measurement precision. This estimate is the largest of the three, so it gives the widest, most conservative, confidence limits. The formula for this estimator, derived assuming a Poisson model for the number of deaths, is

$$\sigma^2_{\tilde{H}1}(T) = \sum_{A \leq T_i \leq T} \frac{d_i}{r_i^2}$$

2. Plug-in Variance Estimate

This estimate also assumes that event time ties occur because of rounding and a lack of measurement precision. The formula for this estimator, derived by substituting sample quantities in the theoretical variance formula, is

$$\sigma^2_{\tilde{H}2}(T) = \sum_{A \leq T_i \leq T} \frac{d_i(r_i - d_i)}{r_i^3}$$

Note that when $r_i = 1$, a '1' is substituted for $(r_i - d_i)/r_i$ in this formula.

3. Binomial Variance Estimate

This estimate assumes that event time ties occur because the process is fundamentally discrete rather than due to lack of precision and/or rounding. The formula for this estimator, derived assuming a binomial model for the number of events, is

$$\sigma^2_{\tilde{H}3}(T) = \sum_{A \le T_i \le T} \frac{d_i\left(r_i - d_i\right)}{r_i^2\left(r_i - 1\right)}$$

Note that when $r_i = 1$, a '1' is substituted for $\left(r_i - d_i\right) / \left(r_i - 1\right)$ in this formula.

Which Variance Estimate to Use

Therneau and Grambsch (2000) indicate that, as of the writing of their book, there is no clear-cut champion. The simple estimate is often suggested because it is always largest and thus gives the widest, most conservative confidence, confidence limits. In practice, there is little difference between them and the choice of which to use will make little difference in the final interpretation of the data. We have included all three since each occurs alone in various treatises on survival analysis.

Pointwise Confidence Intervals of Cumulative Hazard

A pointwise confidence interval for the cumulative hazard at a specific time T_0 of $H\left(T_0\right)$ is represented by two confidence limits which have been constructed so that the probability that the true hazard lies between them is $1-\alpha$. Note that these limits are constructed for a single time point. Several of them cannot be used together to form a confidence band such that the entire hazard function lies within the band. When these are plotted with the hazard curve, these limits must be interpreted on an individual, point by point, basis.

Three difference confidence intervals are available. All three confidence intervals perform about the same in large samples. The linear (Greenwood) interval is the most commonly used. However, the log-transformed and the arcsine-square intervals behave better in small to moderate samples, so they are recommended.

Linear Pointwise Confidence Interval for H(T)

This estimator may be used to create a confidence interval at a specific time point T_0 of $H\left(T_0\right)$ using the formula

$$\tilde{H}(T_0) \pm z_{1-\alpha/2}\sigma_{\tilde{H}}(T_0)$$

where *z* is the appropriate value from the standard normal distribution.

Log-Transformed Pointwise Confidence Interval for H(T)

Better confidence limits may be calculated using the logarithmic transformation of the hazard functions. These limits are

$$\tilde{H}(T_0)/\phi,\ \tilde{H}(T_0)\phi$$

where

$$\phi = \exp\left\{\frac{z_{1-\alpha/2}\sigma_{\tilde{H}}(T_0)}{\tilde{H}\left(T_0\right)}\right\}$$

ArcSine-Square Root Pointwise Confidence Interval for H(T)

Another set of confidence limits using an improving transformation is given by the formula

$$-2\ln\left\{\sin\left[\min\left(\frac{\pi}{2},\arcsin\left[\exp\left\{-\frac{\tilde{H}(T_0)}{2}\right\}\right]+\frac{z_{1-\alpha/2}\sigma_{\tilde{H}}(T_0)}{2\sqrt{\exp\left\{\tilde{H}(T_0)\right\}-1}}\right)\right]\right\}$$
$$\leq H(T_0)\leq$$
$$-2\ln\left\{\sin\left[\max\left(0,\arcsin\left[\exp\left\{-\frac{\tilde{H}(T_0)}{2}\right\}\right]-\frac{z_{1-\alpha/2}\sigma_{\tilde{H}}(T_0)}{2\sqrt{\exp\left\{\tilde{H}(T_0)\right\}-1}}\right)\right]\right\}$$

Survival Quantiles

The median survival time is an example of a quantile of the survival distribution. It is the smallest value of T such that $\hat{S}(T)=0.50$. In fact, more general results are available for any quantile p. The pth quantile is estimated by

$$T_p=\inf\left\{T:\hat{S}(T)\leq 1-p\right\}$$

In words, T_p is smallest time at which $\hat{S}(T)$ is less than or equal to 1 - p.

A $100(1-\alpha)\%$ confidence interval for T_p can be generated using each of the three estimation methods. These are given next.

Linear Pointwise Confidence Interval for Tp

This confidence interval is given by the set of all times such that

$$-z_{1-\alpha/2}\leq\frac{\hat{S}(T)-(1-p)}{\sqrt{\hat{V}\left[\hat{S}(T)\right]}}\leq z_{1-\alpha/2}$$

where z is the appropriate value from the standard normal distribution.

Log-Transformed Pointwise Confidence Interval for Tp

This confidence interval is given by the set of all times such that

$$-z_{1-\alpha/2}\leq\frac{\left[\ln\left\{-\ln\left[\hat{S}(T)\right]\right\}-\ln\left\{-\ln[1-p]\right\}\right]\left[\hat{S}(T)\ln\left[\hat{S}(T)\right]\right]}{\sqrt{\hat{V}\left[\hat{S}(T)\right]}}\leq z_{1-\alpha/2}$$

where z is the appropriate value from the standard normal distribution.

ArcSine-Square Root Pointwise Confidence Interval for Tp

This confidence interval is given by the set of all times such that

$$-z_{1-\alpha/2}\leq\frac{2\left\{\arcsin\left[\sqrt{\hat{S}(T)}\right]-\arcsin\left[\sqrt{1-p}\right]\right\}\sqrt{\hat{S}(T)\left[1-\hat{S}(T)\right]}}{\sqrt{\hat{V}\left[\hat{S}(T)\right]}}\leq z_{1-\alpha/2}$$

where z is the appropriate value from the standard normal distribution.

Hazard Rate Estimation

The characteristics of the failure process are best understood by studying the hazard rate, *h*(*T*), which is the derivative (slope) of the cumulative hazard function *H*(*T*). The hazard rate is estimated using kernel smoothing of the Nelson-Aalen estimator as given in Klein and Moeschberger (2003). The formulas for the estimated hazard rate and its variance are given by

$$\hat{h}(T) = \frac{1}{b} \sum_{A \le T_i \le T} K\left(\frac{T - T_i}{b}\right) \Delta\tilde{H}(T_i)$$

$$\sigma^2\left[\hat{h}(T)\right] = \frac{1}{b^2} \sum_{A \le T_i \le T} K\left(\frac{T - T_i}{b}\right)^2 \Delta\hat{V}\left[\tilde{H}(T_i)\right]$$

where *b* is the bandwidth about *T* and

$$\Delta\tilde{H}(T_k) = \tilde{H}(T_k) - \tilde{H}(T_{k-1})$$

$$\Delta\hat{V}\left[\tilde{H}(T_k)\right] = \hat{V}\left[\tilde{H}(T_k)\right] - \hat{V}\left[\tilde{H}(T_{k-1})\right]$$

Three choices are available for the kernel function *K*(x) in the above formulation. These are defined differently for various values of *T*. Note that the T_i's are for failed items only and that T_{Max} is the maximum failure time. For the *uniform kernel* the formulas for the various values of *T* are

$$K(x) = \frac{1}{2} \quad for \quad T - b \le T \le T + b$$

$$K_L(x) = \frac{4\left(1 + q^3\right)}{(1 + q)^4} + \frac{6(1 - q)}{(1 + q)^3} x \quad for \quad T < b$$

$$K_R(x) = \frac{4\left(1 + r^3\right)}{(1 + r)^4} - \frac{6(1 - r)}{(1 + r)^3} x \quad for \quad T_{Max} - b < T < T_{Max}$$

where

$$q = \frac{T}{b}$$

and

$$r = \frac{T_{Max} - T}{b}$$

For the *Epanechnikov kernel* the formulas for the various values of *T* are

$$K(x) = \frac{3}{4}\left(1 - x^2\right) \quad for \quad T - b \le T \le T + b$$

$$K_L(x) = K(x)(A + Bx) \quad for \quad T < b$$

$$K_R(x) = K(-x)(A - Bx) \quad for \quad T_{Max} - b < T < T_{Max}$$

where

$$A = \frac{64\left(2 - 4q + 6q^2 - 3q^3\right)}{\left(1+q\right)^4\left(19 - 18q + 3q^2\right)}$$

$$B = \frac{240\left(1-q\right)^2}{\left(1+q\right)^4\left(19 - 18q + 3q^2\right)}$$

$$q = \frac{T}{b}$$

$$r = \frac{T_{Max} - T}{b}$$

For the *biweight kernel* the formulas for the various values of T are

$$K(x) = \frac{15}{16}\left(1 - x^2\right)^2 \quad for \quad T - b \le T \le T + b$$

$$K_L(x) = K(x)(A + Bx) \quad for \quad T < b$$

$$K_R(x) = K(-x)(A - Bx) \quad for \quad T_{Max} - b < T < T_{Max}$$

where

$$A = \frac{64\left(8 - 24q + 48q^2 - 45q^3 + 15q^4\right)}{\left(1+q\right)^5\left(81 - 168q + 126q^2 - 40q^3 + 5q^4\right)}$$

$$B = \frac{1120\left(1-q\right)^3}{\left(1+q\right)^5\left(81 - 168q + 126q^2 - 40q^3 + 5q^4\right)}$$

$$q = \frac{T}{b}$$

$$r = \frac{T_{Max} - T}{b}$$

Confidence intervals for $h(T)$ are given by

$$\hat{h}(T)\exp\left[\pm \frac{z_{1-\alpha/2}\,\sigma\left[\hat{h}(T)\right]}{\hat{h}(T)}\right]$$

Care must be taken when using this kernel-smoothed estimator since it is actually estimating a smoothed version of the hazard rate, not the hazard rate itself. Thus, it may be biased. Also, it is greatly influenced by the choice of the bandwidth b. We have found that you must experiment with b to find an appropriate value for each dataset.

Hazard Ratio

Often, it will be useful to compare the hazard rates of two groups. This is most often accomplished by creating the *hazard ratio* (*HR*). The hazard ratio is discussed in depth in Parmar and Machin (1995) and we refer you to this reference for details which we summarize here. The Cox-Mantel estimate of *HR* for two groups A and B is given by

$$HR_{CM} = \frac{H_A}{H_B} = \frac{O_A / E_A}{O_B / E_B}$$

where O_i is the observed number of events (deaths) in group i, E_i is the expected number of events (deaths) in group *i*, and H_i is the overall hazard rate for the ith group. The calculation of the E_i is explained in Parmar and Machin (1995).

A confidence interval for *HR* is found by first transforming to the log scale which is better approximated by the normal distribution, calculating the limits, and then transforming back to the original scale. The calculation is made using

$$\ln\left(HR_{CM}\right) \pm z_{1-\alpha/2}\left(SE_{\ln HR_{CM}}\right)$$

where

$$SE_{\ln HR_{CM}} = \sqrt{\frac{1}{E_A} + \frac{1}{E_B}}$$

which results in the limits

$$\exp\left[\ln\left(HR_{CM}\right) - z_{1-\alpha/2}\left(SE_{\ln HR_{CM}}\right)\right]$$

and

$$\exp\left[\ln\left(HR_{CM}\right) + z_{1-\alpha/2}\left(SE_{\ln HR_{CM}}\right)\right]$$

An alternative estimate of *HR* that is sometimes used is the Mantel-Haenszel estimator which is calculated using

$$HR_{MH} = \exp\left(\frac{O_A - E_A}{V}\right)$$

where *V* is the hypergeometric variance. For further details, see Parmar and Machin (1995). A confidence interval for *HR* is found by first transforming to the log scale which is better approximated by the normal distribution, calculating the limits, and then transforming back to the original scale. The calculation is made using

$$\ln\left(HR_{MH}\right) \pm z_{1-\alpha/2}\left(SE_{\ln HR_{MH}}\right)$$

where

$$SE_{\ln HR_{MH}} = \sqrt{\frac{1}{V}}$$

which results in the limits

$$\exp\left[\ln\left(HR_{MH}\right) - z_{1-\alpha/2}\left(SE_{\ln HR_{MH}}\right)\right]$$

and

$$\exp\left[\ln\left(HR_{MH}\right) + z_{1-\alpha/2}\left(SE_{\ln HR_{MH}}\right)\right]$$

Restricted Mean Survival Time (RMST)

The mean survival time, μ, is computed from the survival function, $S(t)$, as

$$\mu = \int_0^\infty S(t)dt.$$

This corresponds to the area under the survival curve and is only appropriate when the largest observation corresponds to a death. If we restrict the mean estimate to a time interval $[0, \tau]$, where τ represents either the largest observed time (as in the case of Efron's tail correction) or a preassigned interval maximum time, then the mean survival time can be estimated using the nonparametric Kaplan-Meier curve as

$$\hat{\mu}_\tau = \int_0^\tau \hat{S}(t)dt.$$

This is called the Restricted Mean Survival Time (RMST) and corresponds to the area under the Kaplan-Meier survival curve up to time τ. That is,

$$RMST_\tau = \hat{\mu}_\tau.$$

The estimated variance of $\hat{\mu}_\tau$ is

$$\hat{V}(\hat{\mu}_\tau) = \sum_{i=1}^{D} \left[\int_{t_i}^{\tau} \hat{S}(t)dt\right]^2 \frac{d_i}{Y_i(Y_i - d_i)},$$

where D is the total number of events, Y_i is the number of subjects at risk at time t_i, and d_i is the number of subjects failing at time t_i. The standard error is

$$\widehat{SE}(\hat{\mu}_\tau) = \sqrt{\hat{V}(\hat{\mu}_\tau)}.$$

A 100(1 – α)% confidence interval is given by

$$\hat{\mu}_\tau \pm \widehat{SE}(\hat{\mu}_\tau) Z_{1-\alpha/2}.$$

Restricted Mean Time Lost (RMTL)

As described in Uno et al (2014), the Restricted Mean Time Lost (RMTL) corresponds to the area above the Kaplan-Meier survival curve up to time τ and is computed as

$$RMTL_\tau = \tau - \hat{\mu}_\tau.$$

The variance and standard error estimates are the same as for $\hat{\mu}_\tau$. A 100(1 – α)% confidence interval for $RMTL_\tau$ is given by

$$\tau - \hat{\mu}_\tau \pm \widehat{SE}(\hat{\mu}_\tau) Z_{1-\alpha/2}.$$

Hypothesis Tests

This section presents methods for testing that the survival curves, and thus the hazard rates, of two or more populations are equal. The specific hypothesis set that is being tested is

$$H_0: h_1(T) = h_2(T) = \cdots = h_K(T), \quad \text{for all } T \le \tau.$$
$$H_A: h_i(T) \ne h_j(T) \text{ for at least one value of } i, j, \text{ and } T \le \tau.$$

Here τ is taken to be the largest observed time in the study.

In words, the null hypothesis is that the hazard rates of all populations are equal at all times less than the maximum observed time and the alternative hypothesis is that at least two of the hazard rates are different at some time less than the observed maximum time.

In the remainder of this section, we will present a general formulation that includes many of the most popular tests. We use the same notation as before, except that now we add an additional subscript, k, that represents one of the K populations. The test is formed by making a comparison of the actual versus the expected hazard rates. The various hazard rates may be weighted differently. These different weights result in different tests with different properties.

The test is based on the K-1 statistics

$$Z_k(\tau) = \sum_{A \le T_i \le T} W(T_i) r_{ik} \left(\frac{d_{ik}}{r_{ik}} - \frac{d_i}{r_i} \right), k = 1, 2, \cdots, K-1$$

where

$$d_i = \sum_{k=1}^{K} d_{ik}$$

$$r_i = \sum_{k=1}^{K} r_{ik}$$

The Z's have a covariance matrix Σ with elements

$$\sigma_{kg} = \sum_{A \le T_i \le T} W(T_i)^2 \left(\frac{r_{ik}}{r_i} \right) \left(\delta_{kg} - \frac{r_{ig}}{r_i} \right) \left(\frac{r_i - d_i}{r_i - 1} \right) d_i$$

where

$$\delta_{kg} = \begin{cases} 1 \text{ if } k = g \\ 0 \text{ if } k \ne g \end{cases}$$

If we let Z represent the vector of K-1 statistics and Σ represent the covariance matrix, the test statistic is given by

$$Q = Z' \Sigma^{-1} Z$$

In large samples, Q is approximately distributed as a chi-squared random variable with K-1 degrees of freedom. Details of the above formulas can be found in Klein and Moeschberger (2003), pages 205-216 and Andersen, Borgan, Gill, and Keiding (1992), pages 345-356.

Ten different choices for the weight function, $W(T)$, result in the ten different tests that are available in **NCSS**. The most commonly used test is the logrank test which has equal weighting. The other nine tests shift the heaviest weighting to the beginning or end of the trial. This may be appropriate in some studies, but the use of one of these other weighting schemes should be designated before the data have been seen. Also, even though ten tests are displayed, you should only use one of them. Because of the different weighting patterns, they will often give quite different results. It is bad science to look at all the tests and pick the one that matches your own conclusions. That is why you must designate the test you will use before you have seen the data.

The following table describes each of these tests.

Test	Weight	Comments
Logrank	1	This is the most commonly used test and the one we recommend. Equal weights across all times. This test has optimum power when the hazard rates of the *K* populations are proportional to each other.
Gehan	r_i	Places very heavy weight on hazards at the beginning of the study.
Tarone-Ware	$\sqrt{r_i}$	Places heavy weight on hazards at the beginning of the study.
Peto-Peto	$\tilde{S}(T_i)$	Places a little more weight on hazards at the beginning of the study.
Modified Peto-Peto	$\tilde{S}(T_i)r_i/(r_i+1)$	Places a little more weight on hazards at the beginning of the study.
Fleming-Harrington (0,0)	$1-\hat{S}(T_{i-1})$	Places heavy weight on hazards at the end of the study.
Fleming-Harrington (1,0)	$\hat{S}(T_{i-1})$	Places almost equal weight at all times.
Fleming-Harrington (1,1)	$\hat{S}(T_{i-1})(1-\hat{S}(T_{i-1}))$	Places heavy weight on hazards at the end of the study.
Test	**Weight**	**Comments**
Fleming-Harrington (0.5,0.5)	$\sqrt{\hat{S}(T_{i-1})(1-\hat{S}(T_{i-1}))}$	Places a little more weight on hazards at the end of the study.
Fleming-Harrington (0.5,2)	$(1-\hat{S}(T_{i-1}))^2\sqrt{\hat{S}(T_{i-1})}$	Places very heavy weight on hazards at the end of the study.

This table uses the following definitions.

$$\hat{S}(T)=\prod_{T_i\leq T}\left(1-\frac{d_i}{r_i}\right)$$

$$\tilde{S}(T)=\prod_{T_i\leq T}\left(1-\frac{d_i}{r_i+1}\right)$$

Logrank Tests

The logrank test is perhaps the most popular test for testing equality of hazard functions. This test uses $W(T) = 1$, that is, equal weighting. This test has optimum power when the hazard rates of the K populations are proportional to each other.

Note that this version of the logrank test is different from the version used in **NCSS**'s Logrank procedure. That procedure uses the permutation covariance matrix of Lee and Desu which is only valid for equal censoring. The covariance matrix used here is valid for any random censoring pattern, so it is much less restrictive.

Cox-Mantel and Mantel-Haenszel Logrank Tests

When there are only two groups, two versions of the logrank test are commonly used. These tests test the hypothesis that the hazard ratio (*HR*) is one; that is, that the two hazard rates being compared are zero. Note that these tests are equivalent except in small samples.

Cox-Mantel Logrank Test

Using the notation given above in the section on the hazard ratios, the *Cox-Mantel logrank test* statistic is computed using

$$\chi^2_{CM} = \frac{\left(O_A - E_A\right)^2}{E_A} + \frac{\left(O_B - E_B\right)^2}{E_B}$$

This test statistic is approximately distributed as a chi-square random variable with one degree of freedom.

Mantel-Haenszel Logrank Test

The *Mantel-Haenszel logrank test* statistic is computed using

$$\chi^2_{CM} = \frac{\left(O_A - E_A\right)^2}{V}$$

This test statistic is also approximately distributed as a chi-square random variable with one degree of freedom.

Randomization Probability Levels

Because of assumptions that must be made when using this procedure, **NCSS** also includes a randomization test as outlined by Edgington (1987). Randomization tests are becoming more and more popular as the speed of computers allows them to be computed in seconds rather than hours.

A randomization test is conducted by forming a Monte Carlo sampling of all possible permutations of the sample data, calculating the test statistic for each sampled permutation, and counting the number of permutations that result in a chi-square value greater than or equal to the actual chi-square value. Dividing this count by the number of permutations sampled gives the significance level of the test. Edgington suggests that at least 1,000 permutations be selected.

Between-Group Mean Survival Comparisons

Restricted Mean Survival Time (RMST) and Restricted Mean Survival Time (RMTL) comparisons between groups *i* and *j* that follow are described in Royston and Parmar (2013) and Uno et al. (2014) and summarized below.

Restricted Mean Survival Time (RMST) Difference

The RMST difference between groups *i* and *j* is calculated as

$$\widehat{D}_{RMST,ij} = \hat{\mu}_{\tau,i} - \hat{\mu}_{\tau,j},$$

and has estimated standard deviation

$$\widehat{SE}\left(\hat{\mu}_{\tau,i} - \hat{\mu}_{\tau,j}\right) = \sqrt{\hat{V}\left(\hat{\mu}_{\tau,i}\right) + \hat{V}\left(\hat{\mu}_{\tau,j}\right)}.$$

A 100(1 – α)% confidence interval for the difference is given by

$$\widehat{D}_{RMST,ij} \pm \widehat{SE}\left(\hat{\mu}_{\tau,i} - \hat{\mu}_{\tau,j}\right) Z_{1-\alpha/2}.$$

A level α test of the null hypothesis

$$H_0: \mu_{\tau,i} - \mu_{\tau,j} = 0$$

versus the alternative

$$H_A: \mu_{\tau,i} - \mu_{\tau,j} \neq 0$$

can be conducted using the test statistic

$$Z_{ij} = \frac{\hat{\mu}_{\tau,i} - \hat{\mu}_{\tau,j}}{\widehat{SE}\left(\hat{\mu}_{\tau,i} - \hat{\mu}_{\tau,j}\right)},$$

which follows a standard normal distribution.

Restricted Mean Survival Time (RMST) Ratio

The RMST ratio of groups *i* and *j* is calculated as

$$\hat{R}_{RMST,ij} = \frac{\hat{\mu}_{\tau,i}}{\hat{\mu}_{\tau,j}}.$$

The standard error of the ratio is estimated by first taking the log of $\hat{R}_{RMST,ij}$ and computing the standard error of the difference of logged means as

$$\widehat{SE}\left(log\left(\hat{\mu}_{\tau,i}\right) - log\left(\hat{\mu}_{\tau,j}\right)\right) = \sqrt{\frac{\hat{V}\left(\hat{\mu}_{\tau,i}\right)}{\hat{\mu}_{\tau,i}^{\ 2}} + \frac{\hat{V}\left(\hat{\mu}_{\tau,j}\right)}{\hat{\mu}_{\tau,j}^{\ 2}}}.$$

A 100(1 – α)% confidence interval for the ratio is given by

$$\hat{R}_{RMST,ij} \pm e^{\widehat{SE}\left(log\left(\hat{\mu}_{\tau,i}\right) - log\left(\hat{\mu}_{\tau,j}\right)\right) Z_{1-\alpha/2}}.$$

A level α test of the null hypothesis

$$H_0: \frac{\mu_{\tau,i}}{\mu_{\tau,j}} = 1$$

versus the alternative

$$H_A: \frac{\mu_{\tau,i}}{\mu_{\tau,j}} \neq 1$$

can be conducted using the test statistic

$$Z_{ij} = \frac{log\left(\hat{\mu}_{\tau,i}\right) - log\left(\hat{\mu}_{\tau,j}\right)}{\widehat{SE}\left(log\left(\hat{\mu}_{\tau,i}\right) - log\left(\hat{\mu}_{\tau,j}\right)\right)},$$

which follows a standard normal distribution.

Restricted Mean Time Lost (RMTL) Ratio

The RMTL ratio of groups *i* and *j* is calculated as

$$\hat{R}_{RMTL,ij} = \frac{\tau - \hat{\mu}_{\tau,i}}{\tau - \hat{\mu}_{\tau,j}}.$$

The standard error of the ratio is estimated by first taking the log of $\hat{R}_{RMTL,ij}$ and computing the standard error of the difference of logged means as

$$\widehat{SE}\left(log(\tau - \hat{\mu}_{\tau,i}) - log(\tau - \hat{\mu}_{\tau,j})\right) = \widehat{SE}\left(log(\hat{\mu}_{\tau,i}) - log(\hat{\mu}_{\tau,j})\right)$$

$$= \sqrt{\frac{\hat{V}(\hat{\mu}_{\tau,i})}{\hat{\mu}_{\tau,i}^{\ 2}} + \frac{\hat{V}(\hat{\mu}_{\tau,j})}{\hat{\mu}_{\tau,j}^{\ 2}}}.$$

A 100(1 – α)% confidence interval for the ratio is given by

$$\hat{R}_{RMTL,ij} \pm e^{\widehat{SE}\left(log(\hat{\mu}_{\tau,i}) - log(\hat{\mu}_{\tau,j})\right) Z_{1-\alpha/2}}.$$

A level α test of the null hypothesis

$$H_0: \frac{\tau - \mu_{\tau,i}}{\tau - \mu_{\tau,j}} = 1$$

versus the alternative

$$H_A: \frac{\tau - \mu_{\tau,i}}{\tau - \mu_{\tau,j}} \neq 1$$

can be conducted using the test statistic

$$Z_{ij} = \frac{log(\tau - \hat{\mu}_{\tau,i}) - log(\tau - \hat{\mu}_{\tau,j})}{\widehat{SE}\left(log(\hat{\mu}_{\tau,i}) - log(\hat{\mu}_{\tau,j})\right)},$$

which follows a standard normal distribution.

Data Structure

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

Sample Dataset

Most survival data sets require at least two variables: the failure time variable and a censor variable that indicates whether time is a failure or was censored. Optional variables include a count variable which gives the number of items occurring at that time and a group variable that identifies which group this observation belongs to. If the censor variable is omitted, all time values represent failed items. If the count variable is omitted, all counts are assumed to be one.

The table below shows a dataset reporting on a two-group time-to-tumor study. In this data set, time-to-tumor (in days) is given for twelve mice. The twelve mice were randomly divided into two groups. The first group served as a control group, while the second group received a dose of a certain chemical. These data are contained in the Survival dataset.

Survival dataset

Tumor6	Censor6	Trtmnt6
8	1	1
8	1	1
10	1	1
12	1	1
12	1	1
13	1	1
9	1	2
12	1	2
15	1	2
20	1	2
30	0	2
30	0	2

Example 1 – Kaplan-Meier Survival Analysis

This section presents an example of how to analyze a typical set of survival data. In this study, thirty subjects were watched to see how long until a certain event happened after the subject received a certain treatment. The study was terminated at 152.7 hours. At this time, the event had not occurred in eighteen of the subjects. The data used are recorded in the Weibull dataset.

Setup

To run this example, complete the following steps:

1 Open the Weibull example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Weibull** and click **OK**.

2 Specify the Kaplan-Meier Curves (Logrank Tests) procedure options

- Find and open the **Kaplan-Meier Curves (Logrank Tests)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
(Elapsed) Time Variable	**Time**
Censor Variable	**Censor**
Frequency Variable	**Count**
Reports Tab	
Times	**10:150(10)**
Plots Tab	
Kaplan-Meier Survival/Reliability Plot	**Checked**
Cumulative Hazard Function Plot	**Checked**
Hazard Rate Plot	**Checked**
Kaplan-Meier Survival/Reliability Plot Format *(Click the Button)*	
Kaplan-Meier Survival Line	**Checked**
Confidence Limits	**Checked**
Cumulative Hazard Function Plot Format *(Click the Button)*	
Hazard Function Line	**Checked**
Confidence Limits	**Checked**
Hazard Rate Plot Format *(Click the Button)*	
Hazard Rate Line	**Checked**
Confidence Limits	**Checked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Data Summary

Data Summary

Type	Rows	Count	Percent (%)	Minimum	Maximum
Failed	12	12	40.00%	12.5	152.7
Censored	1	18	60.00%	152.7	152.7
Total	13	30	100.00%	12.5	152.7

This report displays a summary of the amount of data that were analyzed. Scan this report to determine if there were any obvious data errors by double checking the counts and the minimum and maximum times.

Median and Mean Survival Estimates

Median and Mean Survival Estimates

Type	Failed	Total	Estimate	Standard Error	Lower 95% C.L.	Upper 95% C.L.
Median Survival Time	12	30			123.2	152.7
Restricted Mean Survival Time*	12	30	126.2	7.508	111.5	140.9
Restricted Mean Time Lost*	12	30	26.5	7.508	11.8	41.2

* Estimates are based on an interval upper limit of τ = 152.7, the maximum observed time.

This report displays point estimates for the median and restricted means, along with lower and upper confidence limits.

Type

This is the statistic being reported on this line.

Failed

This is the number of failed observations.

Total

This is the total sample size.

Estimate

This is the estimate of the corresponding statistic.

Standard Error

This is the standard error of the estimate.

Lower and Upper Confidence Limits

These are the lower and upper confidence limits for the estimate.

Kaplan-Meier Product-Limit Survival Analysis

Kaplan-Meier Product-Limit Survival Analysis

Event Time T	Cumulative Survival S(T)	Standard Error of S(T)	Lower 95% C.L. for S(T)	Upper 95% C.L. for S(T)	At Risk	Count	Total Events
12.5	0.9667	0.0328	0.9024	1.0000	30	1	1
24.4	0.9333	0.0455	0.8441	1.0000	29	1	2
58.2	0.9000	0.0548	0.7926	1.0000	28	1	3
68.0	0.8667	0.0621	0.7450	0.9883	27	1	4
69.1	0.8333	0.0680	0.7000	0.9667	26	1	5
95.5	0.8000	0.0730	0.6569	0.9431	25	1	6
96.6	0.7667	0.0772	0.6153	0.9180	24	1	7
97.0	0.7333	0.0807	0.5751	0.8916	23	1	8
114.2	0.7000	0.0837	0.5360	0.8640	22	1	9
123.2	0.6667	0.0861	0.4980	0.8354	21	1	10
125.6	0.6333	0.0880	0.4609	0.8058	20	1	11
152.7	0.6000	0.0894	0.4247	0.7753	19	1	12
152.7+					18	18	12

This report displays the Kaplan-Meier product-limit survival distribution along with confidence limits. The formulas used were presented earlier.

Also note that the sample size is given for each time period. As time progresses, participants are removed from the study, reducing the sample size. Hence, the survival results near the end of the study are based on only a few participants and are therefore less precise. This shows up as a widening of the confidence limits.

Event Time (T)

This is the time point being reported on this line. The time values are specific event times that occurred in the data.

Note that censored observations are marked with a plus sign on their time value. The survival functions are not calculated for censored observations.

Cumulative Survival S(T)

This is the probability that a subject does not have the event until after the event time given on this line. This probability is estimated using the Kaplan-Meier product limit method. The estimate is given by the formula

$$\hat{S}(T) = \begin{cases} 1 & if\ T_{\min} > T \\ \prod_{A \le T_i \le T} \left[1 - \frac{d_i}{r_i}\right] & if\ T_{\min} \le T \end{cases}$$

Standard Error of S(T)

This is the estimated standard error of the Kaplan-Meier survival probability. The variance of *S(T)* is estimated by Greenwood's formula

$$\hat{V}\left[\hat{S}(T)\right] = \hat{S}(T)^2 \sum_{A \le T_i \le T} \frac{d_i}{r_i(r_i - d_i)}$$

The standard error is the square root of this variance.

Lower and Upper Confidence Limits for S(T)

The lower and upper confidence limits provide a pointwise confidence interval for the survival probability at each time point. These limits are constructed so that the probability that the true survival probability lies between them is $1-\alpha$. Note that these limits are constructed for a single time point. Several of them cannot be used together to form a confidence band such that the entire survival function lies within the band.

Three difference confidence intervals are available. All three confidence intervals perform about the same in large samples. The linear (Greenwood) interval is the most commonly used. However, the log-transformed and the arcsine-square intervals behave better in small to moderate samples, so they are recommended. The formulas for these limits were given at the beginning of the chapter and are not repeated here.

At Risk

This value is the number of individuals at risk. The number at risk is all those under study who died or were censored at a time later than the current time. As the number of individuals at risk is decreased, the estimates become less reliable.

Count

This is the number of individuals having the event (failing) at this time point.

Total Events

This is the cumulative number of individuals having the event (failing) up to and including this time value.

Kaplan-Meier Product-Limit Survival Analysis at Specific Times

Kaplan-Meier Product-Limit Survival Analysis at Specific Times

Time T	Cumulative Survival S(T)	Standard Error of S(T)	Lower 95% C.L. for S(T)	Upper 95% C.L. for S(T)	At Risk
10.0	1.0000	0.0000	1.0000	1.0000	30
20.0	0.9667	0.0328	0.9024	1.0000	29
30.0	0.9333	0.0455	0.8441	1.0000	28
40.0	0.9333	0.0455	0.8441	1.0000	28
50.0	0.9333	0.0455	0.8441	1.0000	28
60.0	0.9000	0.0548	0.7926	1.0000	27
70.0	0.8333	0.0680	0.7000	0.9667	25
80.0	0.8333	0.0680	0.7000	0.9667	25
90.0	0.8333	0.0680	0.7000	0.9667	25
100.0	0.7333	0.0807	0.5751	0.8916	22
110.0	0.7333	0.0807	0.5751	0.8916	22
120.0	0.7000	0.0837	0.5360	0.8640	21
130.0	0.6333	0.0880	0.4609	0.8058	19
140.0	0.6333	0.0880	0.4609	0.8058	19
150.0	0.6333	0.0880	0.4609	0.8058	19

This report displays the Kaplan-Meier product-limit survival probabilities at the specified time points. The formulas used were presented earlier.

Event Time (T)

This is the time point being reported on this line. The time values were specified in the Times box under the Report tab.

Cumulative Survival S(T)

This is the probability that a subject does not have the event until after the event time given on this line. This probability is estimated using the Kaplan-Meier product limit method. The estimate is given by the formula

$$\hat{S}(T) = \begin{cases} 1 & \text{if } T_{\min} > T \\ \prod_{A \le T_i \le T} \left[1 - \frac{d_i}{r_i} \right] & \text{if } T_{\min} \le T \end{cases}$$

Standard Error of S(T)

This is the estimated standard error of the Kaplan-Meier survival probability. The variance of *S(T)* is estimated by Greenwood's formula

$$\hat{V}\left[\hat{S}(T)\right] = \hat{S}(T)^2 \sum_{A \le T_i \le T} \frac{d_i}{r_i(r_i - d_i)}$$

The standard error is the square root of this variance.

Lower and Upper Confidence Limits for S(T)

The lower and upper confidence limits provide a pointwise confidence interval for the survival probability at each time point. These limits are constructed so that the probability that the true survival probability lies between them is $1-\alpha$. Note that these limits are constructed for a single time point. Several of them cannot be used together to form a confidence band such that the entire survival function lies within the band.

Three difference confidence intervals are available. All three confidence intervals perform about the same in large samples. The linear (Greenwood) interval is the most commonly used. However, the log-transformed and the arcsine-square intervals behave better in small to moderate samples, so they are recommended. The formulas for these limits were given at the beginning of the chapter and are not repeated here.

At Risk

This value is the number of individuals at risk. The number at risk is all those under study who died or were censored at a time later than the current time. As the number of individuals at risk is decreased, the estimates become less reliable.

Kaplan-Meier Quantiles of Survival Time

Kaplan-Meier Quantiles of Survival Time

Proportion Surviving	Proportion Failing	Survival Time	Lower 95% C.L. Survival Time	Upper 95% C.L. Survival Time
0.9500	0.0500	24.4	12.5	69.1
0.9000	0.1000	58.2	24.4	96.6
0.8500	0.1500	69.1	24.4	114.2
0.8000	0.2000	95.5	58.2	125.6
0.7500	0.2500	97.0	68.0	152.7
0.7000	0.3000	114.2	69.1	152.7
0.6500	0.3500	125.6	96.6	152.7
0.6000	0.4000	152.7	97.0	152.7
0.5500	0.4500		114.2	152.7
0.5000	0.5000		123.2	152.7
0.4500	0.5500		152.7	152.7
0.4000	0.6000			152.7
0.3500	0.6500			152.7
0.3000	0.7000			152.7
0.2500	0.7500			152.7
0.2000	0.8000			152.7
0.1500	0.8500			152.7
0.1000	0.9000			152.7
0.0500	0.9500			152.7

This report displays the estimated survival times for various survival proportions. For example, it gives the median survival time if it can be estimated.

Proportion Surviving

This is the proportion surviving that is reported on this line. The proportion values were specified in the Percentiles box under the Report tab.

Proportion Failing

This is the proportion failing. The proportion is equal to one minus the proportion surviving.

Survival Time

This is the time value corresponding to the proportion surviving. The *p*th quantile is estimated by

$$T_p = \inf\left\{T : \hat{S}(T) \le 1 - p\right\}$$

In words, T_p is smallest time at which $\hat{S}(T)$ is less than or equal to 1 - *p*.

For example, this table estimates that 95% of the subjects will survive longer than 24.4 hours.

Lower and Upper Confidence Limits on Survival Time

These values provide a pointwise $100(1-\alpha)\%$ confidence interval for T_p. For example, if *p* is 0.50, this provides a confidence interval for the median survival time.

Three methods are available for calculating these confidence limits. The method is designated under the Variables tab in the Confidence Limits box. The formulas for these confidence limits were given in the Survival Quantiles section. Note that because of censoring, estimates and confidence limits are not available for all survival proportions.

Nelson-Aalen Cumulative Hazard

Nelson-Aalen Cumulative Hazard

Event Time T	Cumulative Hazard H(T)	Standard Error of H(T)	Lower 95% C.L. for H(T)	Upper 95% C.L. for H(T)	At Risk	Count	Total Events
12.5	0.0333	0.0333	0.0000	0.0987	30	1	1
24.4	0.0678	0.0480	0.0000	0.1618	29	1	2
58.2	0.1035	0.0598	0.0000	0.2207	28	1	3
68.0	0.1406	0.0703	0.0027	0.2784	27	1	4
69.1	0.1790	0.0802	0.0219	0.3362	26	1	5
95.5	0.2190	0.0896	0.0434	0.3946	25	1	6
96.6	0.2607	0.0988	0.0670	0.4544	24	1	7
97.0	0.3042	0.1079	0.0926	0.5158	23	1	8
114.2	0.3496	0.1171	0.1201	0.5792	22	1	9
123.2	0.3972	0.1264	0.1494	0.6451	21	1	10
125.6	0.4472	0.1360	0.1808	0.7137	20	1	11
152.7	0.4999	0.1458	0.2141	0.7856	19	1	12
152.7+					18	18	12

This report displays estimates of the cumulative hazard function at the time points encountered in the dataset. The formulas used were presented earlier.

Event Time (T)

This is the time point being reported on this line. The time values are specific event times that occurred in the data.

Note that censored observations are marked with a plus sign on their time value. The survival functions are not calculated for censored observations.

Cumulative Hazard H(T)

This is the Nelson-Aalen estimator of the cumulative hazard function, $H(T)$.

Standard Error of H(T)

This is the estimated standard error of the above cumulative hazard function. The formula used was specified under the Variables tab in the Variance box. These formulas were given above in the section discussing the Nelson-Aalen estimator.

The standard error is the square root of this variance.

Lower and Upper Confidence Limits for H(T)

The lower and upper confidence limits provide a pointwise confidence interval for the cumulative hazard at each time point. These limits are constructed so that the probability that the true cumulative hazard lies between them is $1-\alpha$. Note that these limits are constructed for a single time point. Several of them cannot be used together to form a confidence band such that the entire cumulative hazard function lies within the band.

Three difference confidence intervals are available. All three confidence intervals perform about the same in large samples. The linear (Greenwood) interval is the most commonly used. However, the log-transformed and the arcsine-square intervals behave better in small to moderate samples, so they are recommended. The formulas for these limits were given at the beginning of the chapter and are not repeated here.

At Risk

This value is the number of individuals at risk. The number at risk is all those under study who died or were censored at a time later than the current time. As the number of individuals at risk is decreased, the estimates become less reliable.

Count

This is the number of individuals having the event (failing) at this time point.

Total Events

This is the cumulative number of individuals having the event (failing) up to and including this time value.

Nelson-Aalen Cumulative Hazard at Specific Times

Nelson-Aalen Cumulative Hazard at Specific Times

Time T	Cumulative Hazard H(T)	Standard Error of H(T)	Lower 95% C.L. for H(T)	Upper 95% C.L. for H(T)	At Risk
10.0	0.0000	0.0000	0.0000	0.0000	30
20.0	0.0333	0.0333	0.0000	0.0987	29
30.0	0.0678	0.0480	0.0000	0.1618	28
40.0	0.0678	0.0480	0.0000	0.1618	28
50.0	0.0678	0.0480	0.0000	0.1618	28
60.0	0.1035	0.0598	0.0000	0.2207	27
70.0	0.1790	0.0802	0.0219	0.3362	25
80.0	0.1790	0.0802	0.0219	0.3362	25
90.0	0.1790	0.0802	0.0219	0.3362	25
100.0	0.3042	0.1079	0.0926	0.5158	22
110.0	0.3042	0.1079	0.0926	0.5158	22
120.0	0.3496	0.1171	0.1201	0.5792	21
130.0	0.4472	0.1360	0.1808	0.7137	19
140.0	0.4472	0.1360	0.1808	0.7137	19
150.0	0.4472	0.1360	0.1808	0.7137	19

This report displays estimates of the cumulative hazard function at the specified time points. The formulas used were presented earlier.

Event Time (T)

This is the time point being reported on this line. The time values were specified in the Times box under the Report tab.

Cumulative Hazard H(T)

This is the Nelson-Aalen estimator of the cumulative hazard function, $H(T)$. This estimator is give by

$$\tilde{H}(T) = \begin{cases} 0 & if\ T_{\min} > T \\ \sum_{A \le T_i \le T} \frac{d_i}{r_i} & if\ T_{\min} \le T \end{cases}$$

Standard Error of H(T)

This is the estimated standard error of the above cumulative hazard function. The formula used was specified under the Variables tab in the Variance box. These formulas were given above in the section discussing the Nelson-Aalen estimator.

The standard error is the square root of this variance.

Lower and Upper Confidence Limits for H(T)

The lower and upper confidence limits provide a pointwise confidence interval for the cumulative hazard at each time point. These limits are constructed so that the probability that the true cumulative hazard lies between them is $1-\alpha$. Note that these limits are constructed for a single time point. Several of them cannot be used together to form a confidence band such that the entire cumulative hazard function lies within the band.

Three difference confidence intervals are available. All three confidence intervals perform about the same in large samples. The linear (Greenwood) interval is the most commonly used. However, the log-transformed and the arcsine-square intervals behave better in small to moderate samples, so they are recommended. The formulas for these limits were given at the beginning of the chapter and are not repeated here.

At Risk

This value is the number of individuals at risk. The number at risk is all those under study who died or were censored at a time later than the current time. As the number of individuals at risk is decreased, the estimates become less reliable.

Hazard Rates Section

Hazard Rates

Failure Time	Nonparametric Hazard Rate	Std Error of Hazard Rate	95% Lower Conf. Limit of Hazard Rate	95% Upper Conf. Limit of Hazard Rate
10.0	0.0018	0.0014	0.0004	0.0085
20.0	0.0018	0.0013	0.0005	0.0073
30.0	0.0015	0.0010	0.0004	0.0055
40.0	0.0016	0.0009	0.0005	0.0047
50.0	0.0022	0.0012	0.0008	0.0063
60.0	0.0026	0.0015	0.0008	0.0080
70.0	0.0034	0.0016	0.0014	0.0084
80.0	0.0042	0.0017	0.0019	0.0095
90.0	0.0043	0.0019	0.0018	0.0101
100.0	0.0048	0.0021	0.0020	0.0111
110.0	0.0054	0.0022	0.0024	0.0121
120.0	0.0047	0.0021	0.0019	0.0113
130.0	0.0038	0.0020	0.0014	0.0105
140.0	0.0036	0.0025	0.0009	0.0143
150.0	0.0066	0.0066	0.0009	0.0468

This report displays estimates of the hazard rates at the specified time points. The formulas used were presented earlier.

Failure Time

This is the time point being reported on this line. The time values were specified in the Times box under the Report tab.

Nonparametric Hazard Rate

The characteristics of the failure process are best understood by studying the hazard rate, $h(T)$, which is the derivative (slope) of the cumulative hazard function $H(T)$. The hazard rate is estimated using kernel smoothing of the Nelson-Aalen estimator as given in Klein and Moeschberger (2003). The formulas used were given earlier and are not repeated here.

Care must be taken when using this kernel-smoothed estimator since it is actually estimating a smoothed version of the hazard rate, not the hazard rate itself. Thus, it may be biased. Also, it is greatly influenced by the choice of the bandwidth b. We have found that you must experiment with b to find an appropriate value for each dataset.

The values of the smoothing parameters are specified under the Hazards tab.

Standard Error of Hazard Rate

This is the estimated standard error of the above hazard rate. The formula used was specified under the Variables tab in the Variance box. These formulas were given above in the section discussing the Nelson-Aalen estimator.

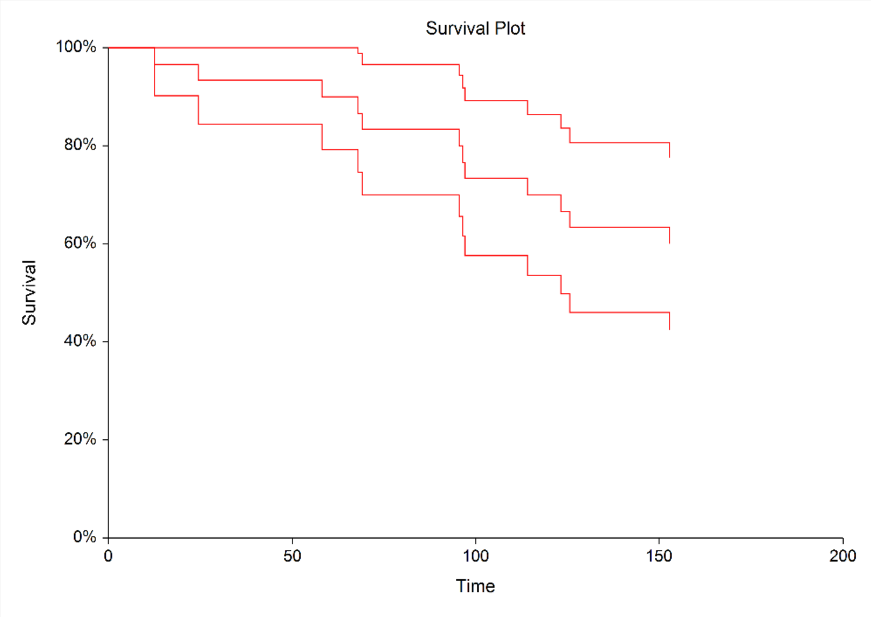
The standard error is the square root of this variance.

Lower and Upper Confidence Limits of Hazard Rate

The lower and upper confidence limits provide a pointwise confidence interval for the smoothed hazard rate at each time point. These limits are constructed so that the probability that the true hazard rate lies between them is $1-\alpha$. Note that these limits are constructed for a single time point. Several of them cannot be used together to form a confidence band such that the entire hazard rate function lies within the band.

Kaplan-Meier Survival Curve(s)

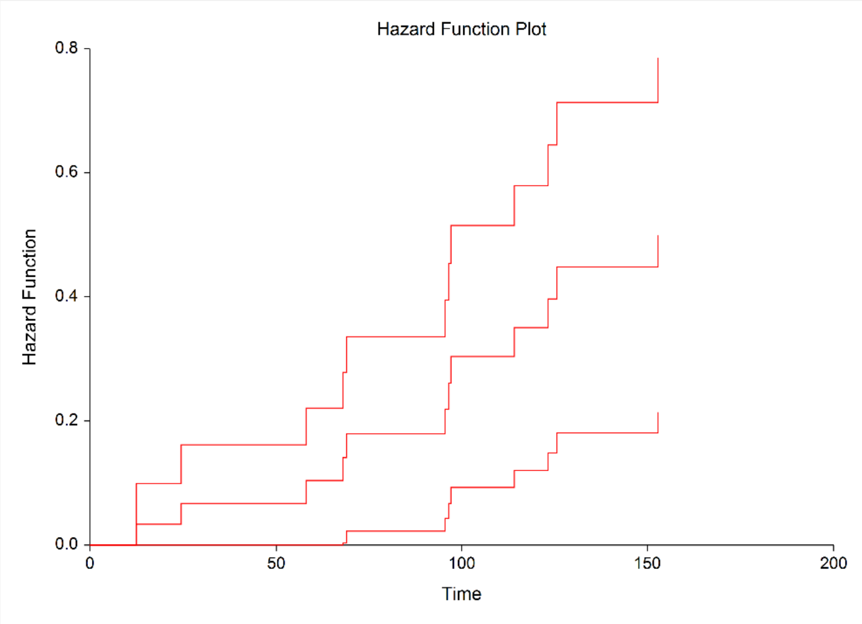
Kaplan-Meier Survival Curve(s)



This plot shows the product-limit survivorship function as well as the pointwise confidence intervals. If there are several groups, a separate line is drawn for each group.

Cumulative Hazard Function Plot(s)

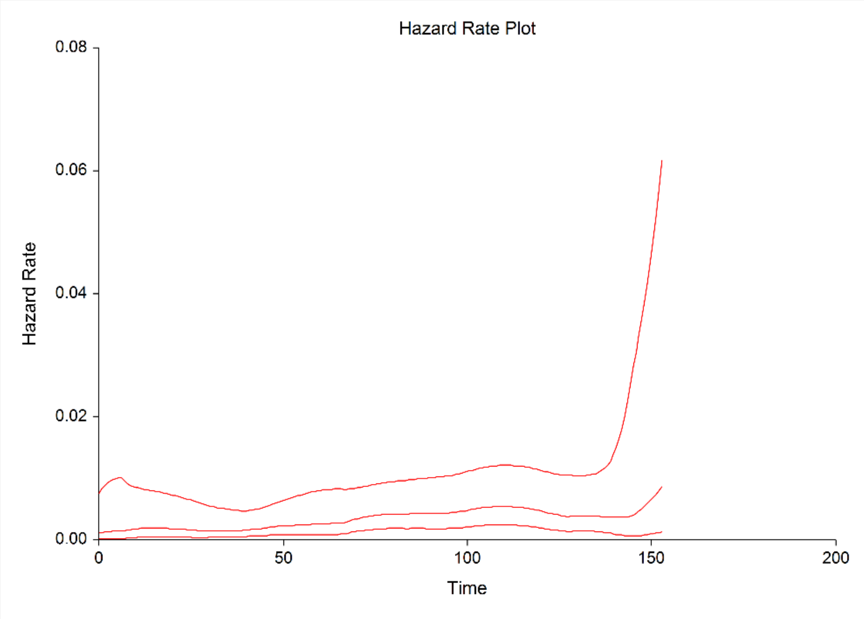
Cumulative Hazard Function Plot(s)



This plot shows the Nelson-Aalen cumulative hazard function for the data analyzed. Confidence limits are also given. If you have several groups, then a separate line is drawn for each group.

Hazard Rate Plot(s)

Hazard Rate Plot(s)



This plot shows the hazard rate with associated confidence limits.

Example 2 – Logrank Tests and Restricted Mean Survival Time (RMST) and Restricted Mean Time Lost (RMTL) Comparisons

This section presents an example of how to use a logrank test to compare the hazard rates of two or more groups. The data used are recorded in the variables Tumor6, Censor6, and Trtmnt6 of the Survival dataset.

Setup

To run this example, complete the following steps:

1 Open the Survival example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Survival** and click **OK**.

2 Specify the Kaplan-Meier Curves (Logrank Tests) procedure options

- Find and open the **Kaplan-Meier Curves (Logrank Tests)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 2** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
(Elapsed) Time Variable	**Tumor6**
Censor Variable	**Censor6**
Group Variable	**Trtmnt6**
Reports Tab	
Data Summary	**Checked**
Logrank Tests Summary	**Checked**
Logrank Tests Details	**Checked**
Hazard Ratio Details	**Checked**
Hazard Ratio Confidence Intervals	**Checked**
Hazard Ratio Logrank Tests	**Checked**
Median and Mean Survival Estimates	**Checked**
Between-Group Mean Survival Comparisons	**Checked**
Run Logrank Randomization Tests	**Checked**
Monte Carlo Samples	**1000**
Random Seed	**5946705** (for reproducibility)
All Other Reports	**Unchecked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Data Summary

Data Summary

Group	Type	Rows	Count	Percent (%)	Minimum	Maximum
1	Failed	6	6	100.00%	8	13
	Censored	0	0	0.00%		
	Total	6	6	50.00%*	8	13
2	Failed	4	4	66.67%	9	20
	Censored	2	2	33.33%	30	30
	Total	6	6	50.00%*	9	30
Combined	Failed	10	10	83.33%	8	20
	Censored	2	2	16.67%	30	30
	Total	12	12	100.00%	8	30

* Percent (%) of Combined Total

This report displays a summary of the amount of data that were analyzed. Scan this report to determine if there were any obvious data errors by double checking the counts and the minimum and maximum times.

Logrank Tests Summary

Logrank Tests Summary

Test Name	Chi-Square	DF	Prob Level*	Randomization Test Prob Level*†	Weighting of Hazard Comparisons Across Time
Logrank	4.996	1	0.0254	0.0410	Equal
Gehan-Wilcoxon	3.956	1	0.0467	0.0510	High++ to Low++
Tarone-Ware	4.437	1	0.0352	0.0480	High to Low
Peto-Peto	3.729	1	0.0535	0.0650	High+ to Low+
Mod. Peto-Peto	3.618	1	0.0572	0.0740	High+ to Low+
F-H (1, 0)	3.956	1	0.0467	0.0510	Almost Equal
F-H (.5, .5)	3.507	1	0.0611	0.1010	Low+ to High+
F-H (1, 1)	4.024	1	0.0449	0.0790	Low to High
F-H (0, 1)	5.212	1	0.0224	0.0670	Low to High
F-H (.5, 2)	5.942	1	0.0148	0.0490	Low++ to High++

* These probability levels are only valid when a single test was selected before this analysis is seen. You cannot select a test to report after viewing this table without adding bias to the results. Unless you have good reason to do otherwise, you should use the equal weighting (logrank) test.
† Randomization Test results are based on 1000 Monte Carlo samples with a random seed of 5946705.

This report gives the results of the ten logrank type tests that are provided by this procedure. We strongly suggest that you select the test that will be used before viewing this report. Unless you have a good reason for doing so, we recommend that you use the first (Logrank) test.

Chi-Square

This is the chi-square value of the test. Each of these tests is approximately distributed as a chi-square in large samples.

DF

This is the degrees of freedom of the chi-square distribution. It is one less than the number of groups.

Prob Level

This is the significance level of the test. If this value is less than your chosen significance level (often 0.05), the test is significant, and the hazard rates of the groups are not identical at all time values.

Randomization Test Prob Level

This is the significance level of the corresponding randomization test. This significance level is exact if the assumption that any censoring is independent of which group the subject was in.

In this example, several of the tests that were just significant at the 0.05 level are not significant using corresponding the randomization test. In cases like this, the randomization test should be considered more accurate than the chi-square test.

Weighting of Hazard Comparisons Across Time

The type of weighting pattern that is used by this test is given here.

Logrank Tests Details

Logrank Tests Details

Logrank Test
Chi-Square = 4.996, DF = 1, Probability Level = 0.0254
Randomization Test Probability Level = 0.0410 (1000 Monte Carlo Samples, Random Seed = 5946705)
Weighting of Hazard Comparisons Across Time = Equal

Group	Z-Value	Standard Error	Standardized Z-Value
1	2.831	1.266	2.235
2	-2.831	1.266	-2.235

Gehan-Wilcoxon Test
Chi-Square = 3.956, DF = 1, Probability Level = 0.0467
Randomization Test Probability Level = 0.0510 (1000 Monte Carlo Samples, Random Seed = 5946705)
Weighting of Hazard Comparisons Across Time = High++ to Low++

Group	Z-Value	Standard Error	Standardized Z-Value
1	24.000	12.066	1.989
2	-24.000	12.066	-1.989

.
.
.

report continues for all ten tests.

This report gives the details of each of the ten logrank tests that are provided by this procedure. We strongly suggest that you select the test that will be used before viewing this report. Unless you have a good reason for doing so, we recommend that you use the first (Logrank) test.

Group

This is the group reported about on this line.

Z-Value

This is a weighted average of the difference between the observed hazard rates of this group and the expected hazard rates under the null hypothesis of hazard rate equality. The expected hazard rates are found by computing new hazard rates based on all that data as if they all came from a single group.

By considering the magnitudes of these values, you can determine which group (or groups) are different from the rest.

Standard Error

This is the standard error of the above z-value. It is used to standardize the z-values.

Standardized Z-Value

The standardized z-value is created by dividing the z-value by its standard error. This provides an index number that will usually very between -3 and 3. Larger values represent groups that quite different from the typical group, at least at some time values.

Hazard Ratio Details

Hazard Ratio Details

Group Ratio	Sample Size n_A | n_B	Observed Events O_A | O_B	Expected Events E_A | E_B	Hazard Rates HR_A | HR_B	Cox-Mantel Hazard Ratio HR	Hyper-geometric Variance V
1/2	6 | 6	6 | 4	3.17 | 6.83	1.89 | 0.59	3.23	1.60
2/1	6 | 6	4 | 6	6.83 | 3.17	0.59 | 1.89	0.31	1.60

This report gives the details of the hazard ratio calculation. One line of the report is devoted to each pair of groups.

Groups

These are the two groups reported about on this line, separated by a slash.

Sample Size

These are the sample sizes of the two groups.

Observed Events

These are the number of events (deaths) observed in the two groups.

Expected Events

These are the number of events (deaths) expected in each group under the hypothesis that the two hazard rates are equal.

Hazard Rates

These are the hazard rates of the two groups. The hazard rate is computed as the ratio of the number of observed and expected events.

Cox-Mantel Hazard Ratio

This is the value of the Cox-Mantel hazard ratio. This is the ratio of the two hazard rates.

Hypergeometric Variance

This is the value of *V*, the hypergeometric variance. This value is used to compute the Mantel-Haenszel hazard ratio and confidence interval.

Hazard Ratio Confidence Intervals

Hazard Ratio Confidence Intervals

Group Ratio	Cox-Mantel Hazard Ratio HR	Log Hazard Ratio Value	Log Hazard Ratio SE	Lower 95% C.L. for HR	Upper 95% C.L. for HR
1/2	3.23	1.1733	0.6796	0.85	12.25
2/1	0.31	-1.1733	0.6796	0.08	1.17

This report gives the details of the Cox-Mantel confidence interval for the hazard ratio. The formulas for these quantities were given earlier in this chapter. One line of the report is devoted to each pair of groups.

Groups

These are the two groups reported about on this line, separated by a slash.

Cox-Mantel Hazard Ratio (HR)

This is the value of the Cox-Mantel hazard ratio.

Log Hazard Ratio Value

This is the natural logarithm of the hazard ratio. The logarithmic transformation is applied because the distribution is better approximated by the normal distribution.

Log Hazard Ratio S.E.

This is the standard deviation of the log hazard ratio.

Lower & Upper 95% C.L. for HR

These are the lower and upper confidence limits of the Cox-Mantel confidence interval of the hazard ratio.

Hazard Ratio Logrank Tests

Hazard Ratio Logrank Tests

	Cox-Mantel Logrank Test			Mantel-Haenszel Logrank Test		
Group Ratio	Hazard Ratio	Chi-Square	Prob Level	Hazard Ratio	Chi-Square	Prob Level
1/2	3.23	3.701	0.0544	5.84	4.996	0.0254
2/1	0.31	3.701	0.0544	0.17	4.996	0.0254

This report gives the two logrank tests which test the null hypothesis that the hazard ratio is one (that is, that the hazard rates are equal). The formulas for these quantities were given earlier in this chapter. One line of the report is devoted to each pair of groups.

Groups

These are the two groups reported about on this line, separated by a slash.

Cox-Mantel Hazard Ratio

This is the value of the Cox-Mantel hazard ratio.

Cox-Mantel Logrank Test Chi-Square

This is the test statistic for the Cox-Mantel logrank test. This value is approximately distributed as a chi-square with one degree of freedom.

Note that this test is more commonly used than the Mantel-Haenszel test.

Cox-Mantel Prob Level

This is the significance level of the Cox-Mantel logrank test. The hypothesis of hazard rate equality is rejected if this value is less than 0.05 (or 0.01).

Mantel-Haenszel Hazard Ratio

This is the value of the Mantel-Haenszel hazard ratio.

Mantel-Haenszel Logrank Test Chi-Square

This is the test statistic for the Mantel-Haenszel logrank test. This value is approximately distributed as a chi-square with one degree of freedom.

Mantel-Haenszel Prob Level

This is the significance level of the Mantel-Haenszel logrank test. The hypothesis of hazard rate equality is rejected if this value is less than 0.05 (or 0.01).

Median and Mean Survival Estimates

Median and Mean Survival Estimates

Median Survival Time

Group	Failed	Total	Median	Lower 95% C.L.	Upper 95% C.L.
1	6	6	10.0	8.0	10.0
2	4	6	15.0	12.0	20.0

Restricted Mean Survival Time (RMST)

Group	Failed	Total	τ	RMST(τ)*	Standard Error	Lower 95% C.L.	Upper 95% C.L.
1	6	6	13.0	10.5	0.808	8.9	12.1
2	4	6	13.0	12.2	0.597	11.0	13.3

* Estimates are based on an interval upper limit of τ = 13.0, the smallest maximum observed time among all groups.

Restricted Mean Time Lost (RMTL)

Group	Failed	Total	τ	RMTL(τ)*	Standard Error	Lower 95% C.L.	Upper 95% C.L.
1	6	6	13.0	2.5	0.808	0.9	4.1
2	4	6	13.0	0.8	0.597	-0.3	2.0

* Estimates are based on an interval upper limit of τ = 13.0, the smallest maximum observed time among all groups.

This report displays point estimates for the median and restricted means for each group, along with lower and upper confidence limits.

Between-Group Mean Survival Comparisons

Between-Group Mean Survival Comparisons

Restricted Mean Survival Time (RMST) Difference
Hypotheses: H0: $RMST_i(\tau) - RMST_j(\tau) = 0$ vs. H1: $RMST_i(\tau) - RMST_j(\tau) \neq 0$

Comparison	τ	RMST(τ) Difference*	Standard Error	Z-Value	Prob Level	Lower 95% C.L.	Upper 95% C.L.
1 - 2	13.0	-1.7	1.005	-1.659	0.0971	-3.6	0.3
2 - 1	13.0	1.7	1.005	1.659	0.0971	-0.3	3.6

* Comparison estimates are based on an interval upper limit of τ = 13.0, the smallest maximum observed time among all groups.

Restricted Mean Survival Time (RMST) Ratio
Hypotheses: H0: $RMST_i(\tau)/RMST_j(\tau) = 1$ vs. H1: $RMST_i(\tau)/RMST_j(\tau) \neq 1$

Comparison	τ	RMST(τ) Ratio*	Standard Error	Z-Value	Prob Level	Lower 95% C.L.	Upper 95% C.L.
1/2	13.0	0.86	0.091	-1.614	0.1065	0.72	1.03
2/1	13.0	1.16	0.091	1.614	0.1065	0.97	1.39

* Comparison estimates are based on an interval upper limit of τ = 13.0, the smallest maximum observed time among all groups.

Restricted Mean Time Lost (RMTL) Ratio
Hypotheses: H0: $RMTL_i(\tau)/RMTL_j(\tau) = 1$ vs. H1: $RMTL_i(\tau)/RMTL_j(\tau) \neq 1$

Comparison	τ	RMTL(τ) Ratio*	Standard Error	Z-Value	Prob Level	Lower 95% C.L.	Upper 95% C.L.
1/2	13.0	3.00	0.786	1.398	0.1622	0.64	14.00
2/1	13.0	0.33	0.786	-1.398	0.1622	0.07	1.56

* Comparison estimates are based on an interval upper limit of τ = 13.0, the smallest maximum observed time among all groups.

This report displays hypothesis tests for the RMST difference and ratio and RMTL ratio along with confidence limits. The three tests and confidence limits indicate that the two groups are not significantly different based on RMST difference and ratio and RMTL ratio. Note that the RMTL difference is not computed because it would be the same as the RMST difference.

Comparison

This is the comparison (difference or ratio) being reported on this line.

τ

This is the value of the upper limit used in calculations.

RMST(τ) Difference, RMST(τ) Ratio, RMTL(τ) Ratio

This is estimated difference or ratio.

Standard Error

This is the standard error of the comparison estimate.

Z-value

This is the computed Z-statistic for the hypothesis test.

Prob Level

This is the computed p-value for the hypothesis test.

Lower and Upper Confidence Limits

These are the lower and upper confidence limits for the comparison estimate.

Example 3 – Validation of Kaplan-Meier Product Limit Estimator using Collett (1994)

This section presents validation of the Kaplan-Meier product limit estimator and associated statistics. Collett (1994) presents an example on page 5 of the time to discontinuation of use of an IUD. The data are as follows: 10, 13+, 18+, 19, 23+, 30, 36, 38+, 54+, 56+, 59, 75, 93, 97, 104+, 107, 107+, 107+. These data are contained in the Collett dataset.

On page 26, Collett (1994) gives the product-limit estimator, its standard deviation, and 95% confidence interval. A partial list of these results is given here:

Time	S(T)	SE	95% C.I
10	0.9444	0.0540	(0.839, 1.000)
36	0.7459	0.1170	(0.529, 0.963)
93	0.4662	0.1452	(0.182, 0.751)
107	0.2486	0.1392	(0.000, 0.522)

We will now run these data through this procedure to see that **NCSS** gets these same results.

Setup

To run this example, complete the following steps:

1 Open the Collett5 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Collett5** and click **OK**.

2 Specify the Kaplan-Meier Curves (Logrank Tests) procedure options

- Find and open the **Kaplan-Meier Curves (Logrank Tests)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 3** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
(Elapsed) Time Variable	**Time**
Censor Variable	**Censor**
Reports Tab	
Kaplan-Meier Product-Limit Survival Analysis	**Checked**
All Other Reports	**Unchecked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Kaplan-Meier Product-Limit Survival Analysis Output

Kaplan-Meier Product-Limit Survival Analysis

Event Time T	Cumulative Survival S(T)	Standard Error of S(T)	Lower 95% C.L. for S(T)	Upper 95% C.L. for S(T)	At Risk	Count	Total Events
10.0	**0.9444**	**0.0540**	**0.8386**	**1.0000**	**18**	**1**	**1**
13.0+					17	1	1
18.0+					16	1	1
19.0	0.8815	0.0790	0.7267	1.0000	15	1	2
23.0+					14	1	2
30.0	0.8137	0.0978	0.6220	1.0000	13	1	3
36.0	**0.7459**	**0.1107**	**0.5290**	**0.9628**	**12**	**1**	**4**
38.0+					11	1	4
54.0+					10	1	4
56.0+					9	1	4
59.0	0.6526	0.1303	0.3972	0.9081	8	1	5
75.0	0.5594	0.1412	0.2827	0.8361	7	1	6
93.0	**0.4662**	**0.1452**	**0.1816**	**0.7508**	**6**	**1**	**7**
97.0	0.3729	0.1430	0.0927	0.6532	5	1	8
104.0+					4	1	8
107.0	**0.2486**	**0.1392**	**0.0000**	**0.5215**	**3**	**1**	**9**
107.0+					2	2	9

You can check this table to see that the results are the same as Collett’s.

Example 4 – Validation of Nelson-Aalen Estimator using Klein and Moeschberger (2003)

This section presents validation of the Nelson-Aalen estimator and associated statistics. Klein and Moeschberger (2003) present an example of output for the cumulative hazard function on page 97. The data are available on their website. These data are contained in the BMT dataset.

A partial list of these results for the ALL group (our group 1) is given here:

Time	H(T)	SE
1	0.0263	0.0263
332	0.5873	0.1449
662	1.0152	0.2185

We will now run these data through this procedure to see that **NCSS** gets these same results.

Setup

To run this example, complete the following steps:

1 Open the BMT example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **BMT** and click **OK**.

2 Specify the Kaplan-Meier Curves (Logrank Tests) procedure options

- Find and open the **Kaplan-Meier Curves (Logrank Tests)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 4** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
(Elapsed) Time Variable	**Time**
Censor Variable	**D3**
Group Variable	**Group**
Reports Tab	
Nelson-Aalen Cumulative Hazard	**Checked**
All Other Reports	**Unchecked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Nelson-Aalen Cumulative Hazard Output

Nelson-Aalen Cumulative Hazard for Group = 1

Event Time T	Cumulative Hazard H(T)	Standard Error of H(T)	Lower 95% C.L. for H(T)	Upper 95% C.L. for H(T)	At Risk	Count	Total Events
1.0	0.0263	0.0263	0.0000	0.0779	38	1	1
55.0	0.0533	0.0377	0.0000	0.1273	37	1	2
74.0	0.0811	0.0468	0.0000	0.1729	36	1	3
86.0	0.1097	0.0549	0.0021	0.2172	35	1	4
104.0	0.1391	0.0623	0.0171	0.2611	34	1	5
107.0	0.1694	0.0692	0.0337	0.3051	33	1	6
109.0	0.2007	0.0760	0.0518	0.3495	32	1	7
110.0	0.2329	0.0825	0.0712	0.3947	31	1	8
122.0	0.2996	0.0950	0.1133	0.4859	30	2	10
129.0	0.3353	0.1015	0.1363	0.5343	28	1	11
172.0	0.3723	0.1081	0.1605	0.5842	27	1	12
192.0	0.4108	0.1147	0.1860	0.6356	26	1	13
194.0	0.4508	0.1215	0.2127	0.6889	25	1	14
226.0+					24	1	14
230.0	0.4943	0.1290	0.2414	0.7472	23	1	15
276.0	0.5397	0.1368	0.2716	0.8079	22	1	16
332.0	0.5873	0.1449	0.3034	0.8713	21	1	17
383.0	0.6373	0.1532	0.3370	0.9377	20	1	18
418.0	0.6900	0.1620	0.3724	1.0076	19	1	19
466.0	0.7455	0.1713	0.4098	1.0813	18	1	20
487.0	0.8044	0.1811	0.4494	1.1593	17	1	21
526.0	0.8669	0.1916	0.4913	1.2424	16	1	22
530.0+					15	1	22
609.0	0.9383	0.2045	0.5375	1.3390	14	1	23
662.0	1.0152	0.2185	0.5870	1.4434	13	1	24
996.0+					12	1	24
1111.0+					11	1	24
1167.0+					10	1	24
1182.0+					9	1	24
1199.0+					8	1	24
1330.0+					7	1	24
1377.0+					6	1	24
1433.0+					5	1	24
1462.0+					4	1	24
1496.0+					3	1	24
1602.0+					2	1	24
2081.0+					1	1	24

You can check this table to see that the results are the same as Klein and Moeschberger's.

Example 5 – Validation of Logrank Tests using Klein and Moeschberger (2003)

This section presents validation of the logrank tests. Klein and Moeschberger (2003) present an example of output for the ten logrank tests on page 210. The data are available on their website. These data are contained in the Klein6 dataset.

A list of these results is given here:

Test	Chi-Square	P-Value
Logrank	2.53	0.112
Gehan	0.002	0.964
Tarone-Ware	0.40	0.526
Peto-Peto	1.40	0.237
Modified Peto-Peto	1.28	0.259
Fleming-Harrington(0,1)	9.67	0.002
Fleming-Harrington(1,0)	1.39	0.239
Fleming-Harrington(1,1)	9.83	0.002
Fleming-Harrington(0.5,0.5)	9.28	0.002
Fleming-Harrington(0.5,2)	8.18	0.004

We will now run these data through this procedure to see that **NCSS** gets these same results.

Setup

To run this example, complete the following steps:

1 Open the Klein6 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Klein6** and click **OK**.

2 Specify the Kaplan-Meier Curves (Logrank Tests) procedure options

- Find and open the **Kaplan-Meier Curves (Logrank Tests)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 5** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
(Elapsed) Time Variable	**Time**
Censor Variable	**Censor**
Group Variable	**Group**
Reports Tab	
Logrank Tests Summary	**Checked**
All Other Reports	**Unchecked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Logrank Tests Summary Output

Logrank Tests Summary

Test Name	Chi-Square	DF	Prob Level*	Weighting of Hazard Comparisons Across Time
Logrank	2.530	1	0.1117	Equal
Gehan-Wilcoxon	0.002	1	0.9636	High++ to Low++
Tarone-Ware	0.403	1	0.5257	High to Low
Peto-Peto	1.399	1	0.2369	High+ to Low+
Mod. Peto-Peto	1.276	1	0.2587	High+ to Low+
F-H (1, 0)	1.387	1	0.2390	Almost Equal
F-H (.5, .5)	9.285	1	0.0023	Low+ to High+
F-H (1, 1)	9.834	1	0.0017	Low to High
F-H (0, 1)	9.668	1	0.0019	Low to High
F-H (.5, 2)	8.179	1	0.0042	Low++ to High++

* These probability levels are only valid when a single test was selected before this analysis is seen. You cannot select a test to report after viewing this table without adding bias to the results. Unless you have good reason to do otherwise, you should use the equal weighting (logrank) test.

You can check this table to see that the results are the same as Klein and Moeschberger's. Note that the order of the tests is different.

Example 6 – Validation of Restricted Mean Survival Time (RMST) Calculations using Klein and Moeschberger (2003)

This section presents validation of the RMST estimates. Klein and Moeschberger (2003) present an example of RMST calculations on pages 119 and 120. The data are available on their website. These data are contained in the BMT dataset. Group 1 is ALL, Group 2 is AML low risk, and Group 3 is AML high risk.

Page 119

Group	τ	Mean	SE	95% Confidence Interval
1: ALL	2081	899.28	150.34*	(606.61, 1193.95)*
2: AML low risk	2569	1548.84	150.62	(1253.62, 1844.07)
3: AML high risk	2640	792.31	158.25	(482.15, 1102.5)

Page 120

Group	τ	Mean	SE	95% Confidence Interval
1: ALL	2081	899.3	150.3*	(606.6, 1193.9)*
2: AML low risk	2081	1315.2	118.8	(1082.4, 1548.0)
3: AML high risk	2081	655.67	122.9	(414.8, 896.5)

*The SE for the ALL group is incorrect as stated in the book. It should have been 146.13, which would have resulted in a confidence interval of (612.87, 1185.69). This can be verified by hand calculations.

We will now run these data through this procedure to see that **NCSS** gets these same results.

Setup

To run this example, complete the following steps:

1 Open the BMT example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **BMT** and click **OK**.

2 Specify the Kaplan-Meier Curves (Logrank Tests) procedure options

- Find and open the **Kaplan-Meier Curves (Logrank Tests)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 6a** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
(Elapsed) Time Variable	**Time**
Censor Variable	**D3**
Group Variable	**Group**
Interval Upper Limit (τ)	**Maximum Observed Time Within Each Group (Variable)**
Reports Tab	
Median and Mean Survival Estimates	**Checked**
All Other Reports	**Unchecked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Median and Mean Survival Estimates Output

Median and Mean Survival Estimates

Median Survival Time

Group	Failed	Total	Median	Lower 95% C.L.	Upper 95% C.L.
1	24	38	418.0	194.0	662.0
2	25	54	2204.0	704.0	2204.0
3	34	45	183.0	115.0	363.0

Restricted Mean Survival Time (RMST)

Group	Failed	Total	τ	RMST(τ)*	Standard Error	Lower 95% C.L.	Upper 95% C.L.
1	24	38	2081.0	899.2	146.131	612.8	1185.6
2	25	54	2569.0	1548.8	150.625	1253.6	1844.1
3	34	45	2640.0	792.3	158.246	482.2	1102.5

* Estimates are based on variable interval upper limits (τ) that equal the maximum observed time within each group.

Restricted Mean Time Lost (RMTL)

Group	Failed	Total	τ	RMTL(τ)*	Standard Error	Lower 95% C.L.	Upper 95% C.L.
1	24	38	2081.0	1181.8	146.131	895.4	1468.2
2	25	54	2569.0	1020.2	150.625	724.9	1315.4
3	34	45	2640.0	1847.7	158.246	1537.5	2157.8

* Estimates are based on variable interval upper limits (τ) that equal the maximum observed time within each group.

These results are the same as Klein and Moeschberger's on page 119 (except for group 1, which matches the correct result, not the incorrect result in the book).

To match the results on page 120, change **Interval Upper Limit (τ)** to **Smallest Maximum Observed Time Among All Groups** (Example 6b) or change **Interval Upper Limit (τ)** to **Custom (User-Entered)** and enter **2081** for **Custom τ** (Example 6c). Both give the same results since 2081 is the smallest maximum observed time.

Example 6b

1 Change Interval Upper Limit (τ) to Smallest Maximum Observed Time Among All Groups

- The settings for this example are listed below and are stored in the **Example 6b** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Interval Upper Limit (τ)	Smallest Maximum Observed Time Among All Groups

2 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Example 6c

1 Change Interval Upper Limit (τ) to Custom (User-Entered)

- The settings for this example are listed below and are stored in the **Example 6c** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Interval Upper Limit (τ)	**Custom (User-Entered)**
Custom τ	**2081**

2 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Example 6b and 6c Output

Restricted Mean Survival Time (RMST)

Group	Failed	Total	τ	RMST(τ)*	Standard Error	Lower 95% C.L.	Upper 95% C.L.
1	24	38	2081.0	899.2	146.131	612.8	1185.6
2	25	54	2081.0	1315.2	118.793	1082.3	1548.0
3	34	45	2081.0	655.7	122.895	414.8	896.5

* Estimates are based on the user-entered interval upper limit of τ = 2081.

Again, these results are the same as Klein and Moeschberger's on page 120 (except for group 1, which matches the correct result, not the incorrect result in the book).

Example 7 – Adding an At-Risk Table to a Survival Plot and a Cumulative Hazard Plot

This section demonstrates how to add a table containing the number of subjects at risk, the cumulative number of censored observations, and the cumulative number of events to the bottom of a survival plot and a cumulative hazard plot. The cumulative hazard plot will be created with default settings, but we'll modify the survival plot to highlight some of the available options. The data used in this example are contained in the BMT dataset.

Setup

To run this example, complete the following steps:

1 Open the BMT example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **BMT** and click **OK**.

2 Specify the Kaplan-Meier Curves (Logrank Tests) procedure options

- Find and open the **Kaplan-Meier Curves (Logrank Tests)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 7** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
(Elapsed) Time Variable	**Time**
Censor Variable	**D3**
Group Variable	**Group**
Reports Tab	
All Reports	**Unchecked**
Report Options Tab	
Value Labels	**Value Labels**
Plots Tab	
Kaplan-Meier Survival/Reliability Plot	**Checked**
Cumulative Hazard Function Plot	**Checked**
Hazard Rate Plot	**Unchecked**
Individual-Group Plots	**Unchecked**
Combined Plots	**Checked**
Kaplan-Meier Survival/Reliability Plot Format *(Click the Button)*	
Survival Tab	
Kaplan-Meier Survival Line	**Checked**
Tick Marks at Censor Points	**Checked**
At-Risk Table Tab	
General Sub-Tab	
Show At-Risk Table	**Checked**
Display	**Number At Risk (Number Censored) (Number of Events)**
Layout Sub-Tab	
Top Outside Margin	**0**

Groups Sub-Tab
Display (Group Headings)................ **Line + Symbol + Label**
Add a Label Prefix............................ **Unchecked**
Label Colors From............................ **Line Fill**
Value Colors From .... **Line Fill**

Value Layout (*Click the Button*)
Alignment..................................... **Center**

Cumulative Hazard Function Plot Format *(Click the Button)*

Haz. Function Tab
Hazard Function Line .... **Checked**

At-Risk Table Tab

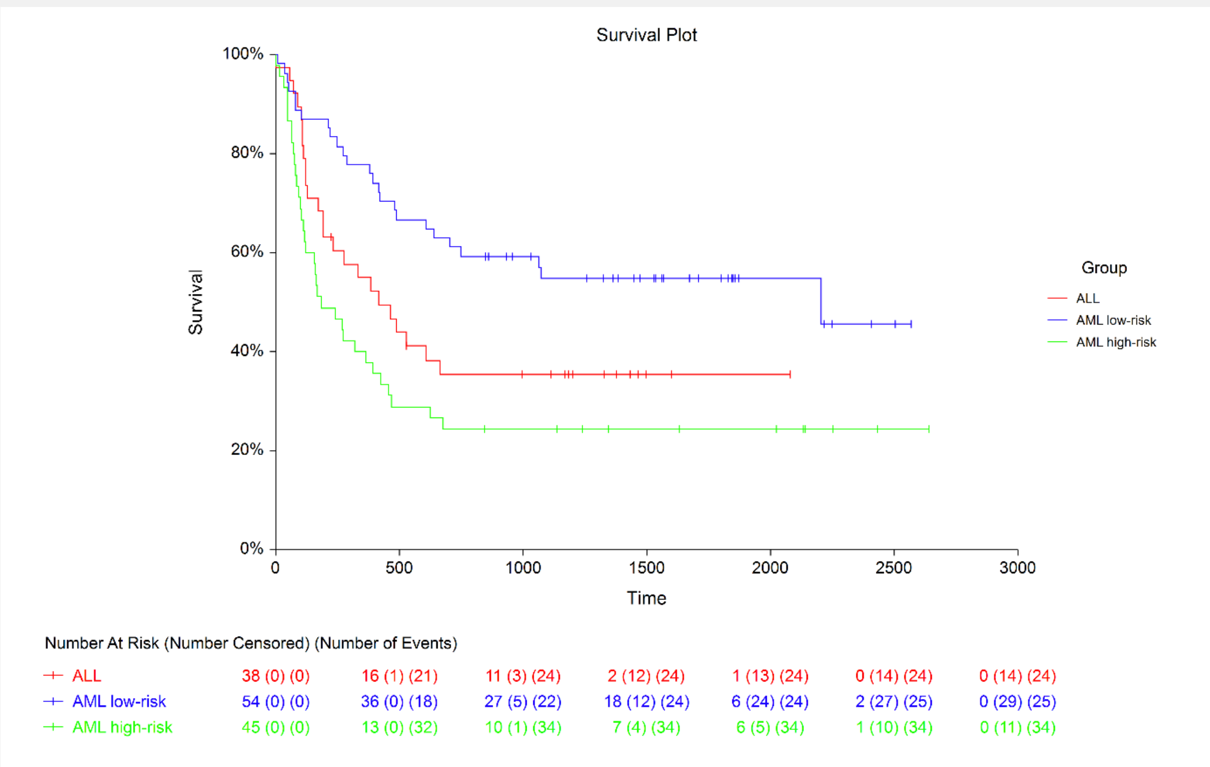
General Sub-Tab
Show At-Risk Table.......................... **Checked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

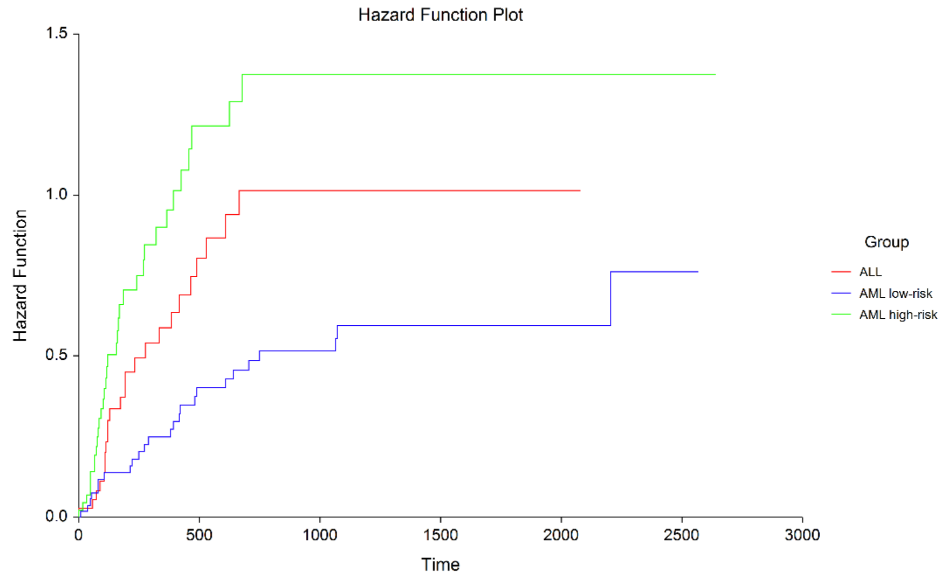
Output

Kaplan-Meier Survival Curve(s)



The modified survival plot is displayed with numbers at risk, cumulative censoring, and cumulative events at each reference time point. You can further modify the plot as required to suit your needs.

Cumulative Hazard Function Plot(s)



Number At Risk

Group = ALL	38	16	11	2	1	0	0
Group = AML low-risk	54	36	27	18	6	2	0
Group = AML high-risk	45	13	10	7	6	1	0

This plot is displayed with default at-risk table settings to give you an idea of what you get if you make no modifications to the table.