PhUSE 2014

Paper SP07

Examining the tests behind survival curves

Hans-Dieter Spies
inVentiv Health Germany GmbH, Eltville am Rhein, Germany

ABSTRACT

There are several applications of survival curves to show events taking place over time. These curves usually show differences between groups and this paper explores certain underlying tests.
A short introduction to survival analysis will be given using the SAS procedure PROC LIFETEST. To describe variations between two groups, we illustrate the advantages and disadvantages of the nonparametric test “Log-rank” and “Wilcoxon” by appropriate examples.
To analyse more than two groups, the separate SAS procedure PROC PHREG (Proportional Hazards Regression) will be used to perform a Cox regression to differentiate between different model approaches. There are four ways to select models based on the PHREG procedure: forward selection, backward elimination, stepwise selection, and best subset selection.

INTRODUCTION

Survival Analysis typically focuses on time to event data. Typically, survival data are not fully observed, but rather are censored. This can be done in SAS with PROC LIFETEST. We will choose simple data and answer the following questions.

1. Question:
How great is the probability that an event occurs at a particular time.
Answer:
The Kaplan–Meier estimator is an estimator for estimating the survival function from lifetime data. In medical research, it is often used to measure the fraction of patients living for a certain amount of time after treatment. In economics, it can be used to measure the length of time people remain unemployed after a job loss. The estimator was developed in 1958 from Edward L. Kaplan and Paul Meier.

2. Question:
When are 2 survival curves statistically equivalent?
Answer:
We will use the “Log-rank test”. This is a nonparametric test, which does not rely on data belonging to any particular distribution.

Kaplan–Meier method

Let’s start with an example. Have a look at the following test data:

duration of group A	duration of group B
3	15
6*	22
10	28
15	36
32*	40

Tab. 1: Listing of test data

We have two groups with 5 subjects in each group. The numbers represent the time an event occurs for each subject.

If a * (asterisk) comes after a number then the number shows not the time of an event, but the last time of observation.

An important advantage of the Kaplan–Meier method is that censored data can be taken into account, particularly right-censoring, which occurs if a patient withdraws from a study, i.e. is lost from the sample population before the final outcome can be recorded.

Now to the first question. The Kaplan-Meier Estimator (KM-Estimator) also known as the product limit estimator, is used to estimate the survival function from lifetime data.

The Kaplan–Meier estimator $\hat{S}(t)$ is defined as the product.

$$\hat{S}(t) = \prod_{t_{(i)} \le t} \frac{n_i - d_i}{n_i}$$

where $\hat{S}(0) = 1$
d_i = number of events, for which the event occurs at $t_{(i)}$
n_i = number of events under risk at $t_{(i)}$

Let's have a look at our example:

There are 2 subjects in group A with no event during the observation period. In group B all subjects have an event.

Group	duration	Subjects, when time period ends without an event	under risk n_i	event occurs d_i	KM-Estimator S(duration)	
A	3	0	5	1	(5-1)/5 =	0.8000
A	6	1	4	0	$\hat{S}$ (3) * (4-0)/4 =	0.8000
A	10	0	3	1	$\hat{S}$ (6) * (3-1)/3 =	0.5333
A	15	0	2	1	$\hat{S}$ (10) * (2-1)/2 =	0.2667
A	32	1	1	0	$\hat{S}$ (15) * (1-0)/1 =	0.2667
B	15	0	5	1	(5-1)/5 =	0.8000
B	22	0	4	1	$\hat{S}$ (15) * (4-1)/4 =	0.6000
B	28	0	3	1	$\hat{S}$ (22) * (3-1)/3 =	0.4000
B	36	0	2	1	$\hat{S}$ (28) * (2-1)/2 =	0.2000
B	40	0	1	1	$\hat{S}$ (36) * (1-1)/1 =	0.0000

Tab. 2: Calculation of Kaplan-Meier Estimator

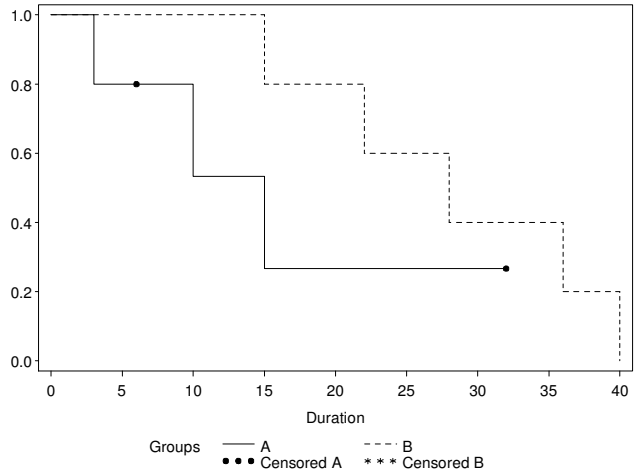


Fig. 1: Kaplan-Meier curve for test data

The curve for group A ends at the level of 0.2667, because the subject with the longest time period is censored.

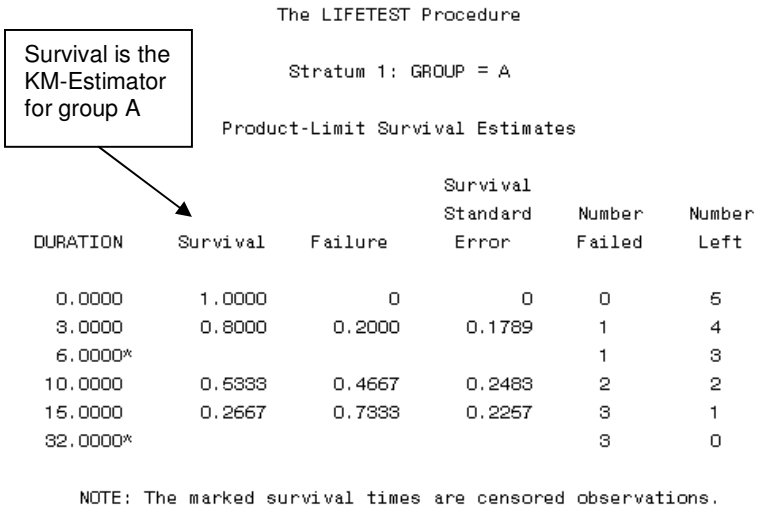
The probability for each group is on the x-axis. When an event occurs the curve jumps to a lower level.

To calculate the KM-Estimator the SAS procedure PROC LIFETEST can be used:

```
PROC LIFETEST DATA= testdata;
    STRATA group /test=LOGRANK;
    TEST status;
    TIME duration*status(0);
quit;
```

Fig. 2: SAS program for PROC LIFETEST

Let's examine the output window of SAS:



The LIFETEST Procedure

Stratum 1: GROUP = A

Product-Limit Survival Estimates

DURATION	Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.0000	1.0000	0	0	0	5
3.0000	0.8000	0.2000	0.1789	1	4
6.0000*				1	3
10.0000	0.5333	0.4667	0.2483	2	2
15.0000	0.2667	0.7333	0.2257	3	1
32.0000*				3	0

NOTE: The marked survival times are censored observations.

Fig. 3: Output of SAS program for PROC LIFETEST

The same information is available for group B.

The Log-Rank Test

The log-rank test is the most commonly used statistical test for comparing the survival distributions of two or more groups (e.g. different treatment groups in a clinical trial).

Test construction:
Denote the distinct times of events as $t_1 < t_2 < \cdots < t_k$, and define

- d_{ij} = number of subjects in group i with event (uncensored) at t_j (i = 1; 2; j = 1; 2; : : : ; k)
- d_j = $d_{1j} + d_{2j}$ = total number of events at t_j
- $n_i(t_j)$ = number of subjects in group i who are at risk at t_j (i = 1; 2; j = 1; 2; : : : ; k)
- $n(t_j)$ = $n_1(t_j) + n_2(t_j)$ = total number of persons in both groups who are at risk at t_j

The information at time t_j can be summarized in the following 2x2 table:

	observed event at t_j		at risk at t_j
group A	d_{1j}	$n_1(t_j) - d_{1j}$	$n1(t_j)$
group B	d_{2j}	$n_2(t_j) - d_{2j}$	$n2(t_j)$

This leads to the Log-rank statistic for two groups which is X^2 distributed:

$$\frac{\left(\Sigma\left(d_{ij} - (n_i(t_j) * d_{ij})\right)\right)^2}{\mathrm{Var}\left(\Sigma\left(d_{ij} - (n_i(t_j) * d_{ij})\right)\right)} \sim X^2$$

Now we will calculate the value of Chi-square for the example:

The Log-Rank Test for test data

We have to consider all time points t_j at which events take place. These are listed in the first column.

Then we have to list all events at each of the time points t_j for each group. These are listed in the next column.

When we count the subjects under risk we have to consider the censored cases. Consider $t_j = 10$. In group A 1 subject has an event at $t_j = 3$ and for another subject the observation period ends at $t_j = 6$. So 2 subjects are no longer at risk at $t_j = 10$. For $t_j = 10$ is n1=5-2=3. In group B all subjects are under risk at $t_j = 10$: n2=5.

Next the expected numbers for each t_j have to be calculated. To calculate expected numbers we have to set the observed number in relation to all numbers that are available at this time.
The formula for group A is $d_j * n1_j/n_j$. Note that the numbers of the expected numbers of both groups sum up to the number of events of both groups. For example, for the time point $t_j = 15$ the number of events is 2 and the expected numbers are 0.5714 + 1.4286 = 2.

The variance for each time point t_j is defined as (dj*(nj-dj)*n1j*n2j) / (nj*nj*(nj-1)). The variance over all time points is the sum = 1.1720.

Time t_j	Events at t_j			subjects under risk short before t_j			Expected numbers		Observed - expected		Variance
/group	A	B	both	A	B	both	A	B	A	B	both
	$d1_j$	$d2_j$	$d_j = d1_j+d2_j$	$n1_j$	$n2_j$	$n_j = n1_j+n2_j$	$e1_j = d_j*n1_j/n_j*$	$e2_j = d_j*n2_j/n_j$	$d1_j-e1_j$	$d2_j-e2_j$	$(d_j*(n_j-dj)*n1_j*n2_j) / (n_j*n_j*(n_j-1))$
3	1	0	1	5	5	10	0.5000	0.5000	0.5000	-0.5000	0.2500
10	1	0	1	3	5	8	0.3750	0.6250	0.6250	-0.6250	0.2344
15	1	1	2	2	5	7	0.5714	1.4286	0.4286	-0.4286	0.3401
22	0	1	1	1	4	5	0.2000	0.8000	-0.2000	0.2000	0.1600
28	0	1	1	1	3	4	0.2500	0.7500	-0.2500	0.2500	0.1875
36	0	1	1	0	2	2	0	1	0	0	0
40	0	1	1	0	1	1	0	1	0	0	0
					Sum:		1.8964	6.1036	1.1036	-1.1036	1.1720

Tab. 3: Calculation

Now we have to build the log-rank characteristic:

	Sum of events d_g	Sum of expected e_g	chi-square $(d_g-e_g)^2 / var(d_g-e_g)$
Group A	3	1.8964	1.0391
Group B	5	6.1036	1.0391

Tab. 4: Calculation

The probability can be calculated with SAS function PROBCHI().

```
data chi;
    x1=1-probchi(1.0391,1);
run;
```

Fig. 4: SAS program to calculate the probability

The result is x1 = 0.3080308805

Now let us compare the results with the output of the SAS procedure PROC LIFETEST:

The LIFETEST Procedure

Testing Homogeneity of Survival Curves for DURATION over Strata

Rank Statistics

GROUP	Log-Rank
A	1.1036
B	-1.1036

Covariance Matrix for the Log-Rank Statistics

GROUP	A	B
A	1.17201	-1.17201
B	-1.17201	1.17201

Test of Equality over Strata

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	1.0391	1	0.3080

Fig. 5: Output of SAS program for PROC LIFETEST

- This colour shows the sum of number of observed – expected values.
- This colour shows the variance.
- This colour shows characteristic value of the chi-square distribution.
- This colour shows (1 minus “the probability the characteristic value of the chi-square distribution”).

Note that when we calculate the probability, we set “x1=1-probchi(1.0391,1)”. In the end this indicates that the two curves are different. .

A second example

For the second example we choose data where both groups are identical to demonstrate a result for a simple:

duration of group A	duration of group B
3	3
6	6
12	12
25	25
35	35

Tab. 5: Listing of data with equal groups

Again we have 5 subjects in each group. The duration of all events are identical in both groups.

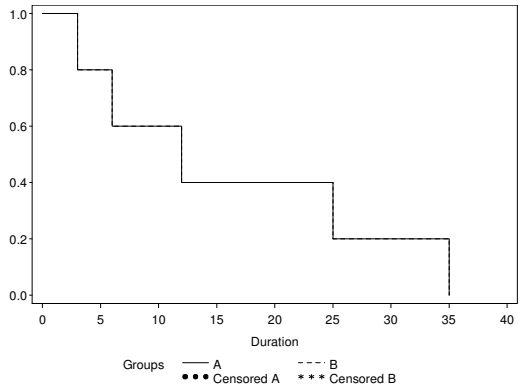


Fig. 6: Kaplan-Meier curve for data with equal groups

Because the events in both groups are identical the curve of group B lags behind the curve of group A.

The LIFETEST Procedure

Testing Homogeneity of Survival Curves for DURATION over Strata

Rank Statistics

GROUP	Log-Rank
A	0
B	0

Covariance Matrix for the Log-Rank Statistics

GROUP	A	B
A	1.60635	-1.60635
B	-1.60635	1.60635

Test of Equality over Strata

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	0.0000	1	1.0000

Chi-square = 0 indicates no difference between the groups

Fig. 7: Output of SAS program for PROC LIFETEST data with equal groups

Some further aspects

- Log-rank test: is almost appropriate. The test statistic follows the chi-square distribution.
- Wilcoxon test: is similar to the log-rank test, but will place higher priority in the beginning of the curves. The test statistic follows also roughly the chi-square distribution.
- Likelihood ratio test: this assumes that the survival time of each group is exponentially distributed. Also this test statistic follows roughly the chi-square distribution.

To analyse more than two groups, the separate SAS procedure PROC PHREG (Proportional Hazards Regression) can be used to perform a Cox regression to differentiate between different model approaches. Whereas the Kaplan-Meier method with log-rank test is useful for comparing survival curves in two or more groups, Cox regression (or proportional hazards regression) allows analyzing the effect of several risk factors on survival.

CONCLUSION

First a simple example of Survival Analysis was shown to explain how to handle censored cases.
To get the statistical characteristics, many calculations are needed. First the Kaplan-Meier-Estimator was calculated. This belongs to the descriptive analyses which is performed most often.
In next step, we demonstrated how the chi-square value arises from the Log-rank-test. The chi-square distribution is used to find the probability that the two curves are significantly different or not.
Finally the procedure PROC LIFETEST was found to be very useful in obtaining all of these values.

CONTACT INFORMATION

Hans-Dieter Spies
inVentiv Health GermanyGmbH
Grosse Hub 10d
65344 Eltville-am-Rhein
Tel: +49 (0) 6123 70437 13
hansdieter.spies@inventivhealth.com



>> TRANSFORMING PROMISING IDEAS
INTO COMMERCIAL REALITY

Join us on Facebook and Twitter

Brand and product names are trademarks of their respective companies.