SURVIVAL ANALYSIS

Chapter 2. Kaplan-Meier Survival Curves and the Log-Rank Test

Instructor: Seokho Lee

2013Fall at HUFS

An Example of Kaplan-Meier Curves

EXAMPLE

The data: remission times (weeks) for two groups of leukemia patients

Group 1 ($n = 21$) treatment	Group 2 ($n = 21$) placebo
6, 6, 6, 7, 10,	1, 1, 2, 2, 3,
13, 16, 22, 23,	4, 4, 5, 5,
6+, 9+, 10+, 11+,	8, 8, 8, 8,
17+, 19+, 20+,	11, 11, 12, 12,
25+, 32+, 32+,	15, 17, 22, 23
34+, 35+,	

Note: + denotes censored

	# failed	# censored	Total
Group 1	9	12	21
Group 2	21	0	21

Descriptive statistics:

$\overline{T}_1$ (ignoring + 's) = 17.1, $\overline{T}_2 = 8.6$

$\overline{h}_1 = .025$, $\overline{h}_2 = .115$, $\dfrac{\overline{h}_2}{\overline{h}_1} = 4.6$

- The data for this example derive from a study of the remission times in weeks for two groups of leukemia patients, with 21 patients in each group. Group 1 is the treatment group and group 2 is the placebo group. The basic question of interest concerns comparing the survival experience of the two groups.
- Of the 21 persons in group 1, 9 failed during the study period and 12 were censored. In contrast, none of the data in group 2 are censored; that is, all 21 persons in the placebo group went out of remission during the study period.
- In Chapter 1, we observed for this data set that group 1 appears to have better survival prognosis than group 2, suggesting that the treatment is effective. This conclusion was supported by descriptive statistics for the average survival time and average hazard rate shown. Note, however, that descriptive statistics provide overall comparisons but do not compare the two groups at different times of follow-up.

An Example of Kaplan-Meier Curves

EXAMPLE (continued)

Ordered failure times:

Group 1 (treatment)

$t_{(j)}$	n_j	m_j	q_j
0	21	0	0
6	21	3	1
7	17	1	1
10	15	1	2
13	12	1	0
16	11	1	3
22	7	1	0
23	6	1	5
>23	—	—	—

Group 2 (placebo)

$t_{(j)}$	n_j	m_j	q_j
0	21	0	0
1	21	2	0
2	19	2	0
3	17	1	0
4	16	2	0
5	14	2	0
8	12	4	0
11	8	2	0
12	6	2	0
15	4	1	0
17	3	1	0
22	2	1	0
23	1	1	0

- A table of ordered failure times is shown here for each group. These tables provide the basic information for the computation of KM curves.

- Each table begins with a survival time of zero, even though no subject actually failed at the start of follow-up. The reason for the zero is to allow for the possibility that some subjects might have been censored before the earliest failure time.

- Also, each table contains a column denoted as n_j that gives the number of subjects in the risk set at the start of the interval. Given that the risk set is defined as the collection of individuals who have survived at least to time $t_{(j)}$, it is assumed that n_j includes those persons failing at time $t_{(j)}$. In other words, n_j counts those subjects at risk for failing instantaneously prior to time $t_{(j)}$.

An Example of Kaplan-Meier Curves

Group 2 (placebo)

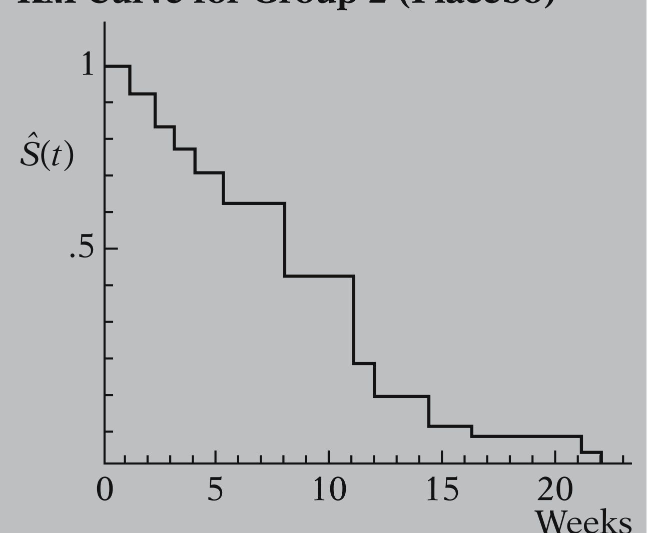
$t_{(j)}$	n_j	m_j	q_j
0	21	0	0
1	21	2	0
2	19	2	0
3	17	1	0
4	16	2	0
5	14	2	0
8	12	4	0
11	8	2	0
12	6	2	0
15	4	1	0
17	3	1	0
22	2	1	0
23	1	1	0

Group 2: no censored subjects

Group 2 (placebo)

$t_{(j)}$	n_j	m_j	q_j	$\hat{S}(t_{(j)})$
0	21	0	0	1
1	21	2	0	19/21 = .90
2	19	2	0	17/21 = .81
3	17	1	0	16/21 = .76
4	16	2	0	14/21 = .67
5	14	2	0	12/21 = .57
8	12	4	0	8/21 = .38
11	8	2	0	6/21 = .29
12	6	2	0	4/21 = .19
15	4	1	0	3/21 = .14
17	3	1	0	2/21 = .10
22	2	1	0	1/21 = .05
23	1	1	0	0/21 = .00

KM Curve for Group 2 (Placebo)



$S(t) = \Pr(T > t)$

- We now describe how to compute the KM curve for the table for group 2. The computations for group 2 are quite straightforward because there are no censored subjects for this group.
- The table of ordered failure times for group 2 is presented here again with the addition of another column that contains survival probability estimates. These estimates are the KM survival probabilities for this group. We will discuss the computations of these probabilities shortly.
- A plot of the KM survival probabilities corresponding to each ordered failure time is shown here for group 2. Empirical plots such as this one are typically plotted as a step function that starts with a horizontal line at a survival probability of 1 and then steps down to the other survival probabilities as we move from one ordered failure time to another.

An Example of Kaplan-Meier Curves

Group 2 (placebo)

$t_{(j)}$	n_j	m_j	q_j	$\hat{S}(t_{(j)})$
0	21	0	0	1
1	21	2	0	19/21 = .90
2	19	2	0	17/21 = .81
3	17	1	0	16/21 = .76
4	16	2	0	14/21 = .67
5	14	2	0	12/21 = .57
8	12	4	0	8/21 = .38
11	8	2	0	6/21 = .29
12	6	2	0	4/21 = .19
15	4	1	0	3/21 = .14
17	3	1	0	2/21 = .10
22	2	1	0	1/21 = .05
23	1	1	0	0/21 = .00

$$\hat{S}(t_{(j)}) = \frac{\#\text{ surviving past } t_{(j)}}{21}$$

No censorship in group 2

Alternative formula: KM approach

- Thus, considering the group 2 data, the probability of surviving past zero is unity, as it will always be for any data set.
- Next, the probability of surviving past the first ordered failure time of one week is given by 19/21 (or .90) because 2 people failed at one week, so that 19 people from the original 21 remain as survivors past one week.
- Similarly, the next probability concerns subjects surviving past two weeks, which is 17/21 (or .81) because 2 subjects failed at one week and 2 subjects failed at two weeks leaving 17 out of the original 21 subjects surviving past two weeks.
- The remaining survival probabilities in the table are computed in the same manner, that is, we count the number of subjects surviving past the specified time being considered and divide this number by 21, the number of subjects at the start of follow-up.
- Recall that no subject in group 2 was censored, so the q column for group 2 consists entirely of zeros. If some of the q's had been nonzero, an alternative formula for computing survival probabilities would be needed. This alternative formula is called the Kaplan–Meier (KM) approach and can be illustrated using the group 2 data even though all values of q are zero.

An Example of Kaplan-Meier Curves

EXAMPLE

$$\hat{S}(4) = 1 \times \frac{19}{21} \times \frac{17}{19} \times \frac{16}{17} \times \frac{14}{16} = \frac{14}{21} = .67$$

$$\Pr\left(T > t_{(j)} \mid T \geq t_{(j)}\right)$$

$$\hat{S}(4) = 1 \times \frac{19}{21} \times \frac{17}{19} \times \frac{16}{17} \times \frac{14}{16} = \frac{14}{21} = .67$$

$$\frac{19}{21} = \Pr(T > 1 | T \geq 1)$$

$$\frac{16}{17} = \Pr(T > 3 | T \geq 3)$$

17 = # in risk set at week 3

$$\hat{S}(4) = 1 \times \frac{19}{21} \times \frac{17}{19} \times \frac{16}{17} \times \frac{14}{16}$$

$$\hat{S}(8) = 1 \times \frac{19}{21} \times \frac{17}{19} \times \frac{16}{17} \times \frac{14}{16} \times \frac{12}{14} \times \frac{8}{12}$$

KM formula = product limit formula

- For example, an alternative way to calculate the survival probability of exceeding four weeks for the group 2 data can be written using the KM formula shown here. This formula involves the product of conditional probability terms. That is, each term in the product is the probability of exceeding a specific ordered failure time $t_{(j)}$ given that a subject survives up to that failure time.
- Thus, in the KM formula for survival past four weeks, the term 19/21 gives the probability of surviving past the first ordered failure time, one week, given survival up to the first week. Note that all 21 persons in group 2 survived up to one week, but that 2 failed at one week, leaving 19 persons surviving past one week.
- Similarly, the term 16/17 gives the probability of surviving past the third ordered failure time at week 3, given survival up to week 3. There were 17 persons who survived up to week 3 and one of these then failed, leaving 16 survivors past week 3. Note that the 17 persons in the denominator represents the number in the risk set at week 3.
- Notice that the product terms in the KM formula for surviving past four weeks stop at the fourth week with the component 14/16. Similarly, the KM formula for surviving past eight weeks stops at the eighth week.
- More generally, any KM formula for a survival probability is limited to product terms up to the survival week being specified. That is why the KM formula is often referred to as a “product-limit” formula.

An Example of Kaplan-Meier Curves

Group 1 (treatment)

$t_{(j)}$	n_j	m_j	q_j	$\hat{S}(t_{(j)})$
0	21	0	0	1
6	21	3	1	$1 \times \frac{18}{21}$
		.		
		.		
		.		

- Next, we consider the KM formula for the data from group 1, where there are several censored observations.
- The estimated survival probabilities obtained using the KM formula are shown here for group 1.
- The first survival estimate on the list is $\hat{S}(0) = 1$, as it will always be, because this gives the probability of surviving past time zero.

An Example of Kaplan-Meier Curves

EXAMPLE (continued)

Group 1 (treatment)

$t_{(j)}$	n_j	m_j	q_j	$\hat{S}(t_{(j)})$
0	21	0	0	1
6	21	3	1	$1 \times \frac{18}{21} = .8571$
7	17	1	1	$.8571 \times \frac{16}{17} = .8067$
10	15	1	2	$.8067 \times \frac{14}{15} = .7529$
13	12	1	0	$.7529 \times \frac{11}{12} = .6902$
16	11	1	3	$.6902 \times \frac{10}{11} = .6275$
22	7	1	0	$.6275 \times \frac{6}{7} = .5378$
23	6	1	5	$.5378 \times \frac{5}{6} = .4482$

Fraction at $t_{(j)}$: $\Pr(T > t_{(j)} \mid T \geq t_{(j)})$

Not available at $t_{(j)}$: failed prior to $t_{(j)}$
or
censored prior to $t_{(j)}$

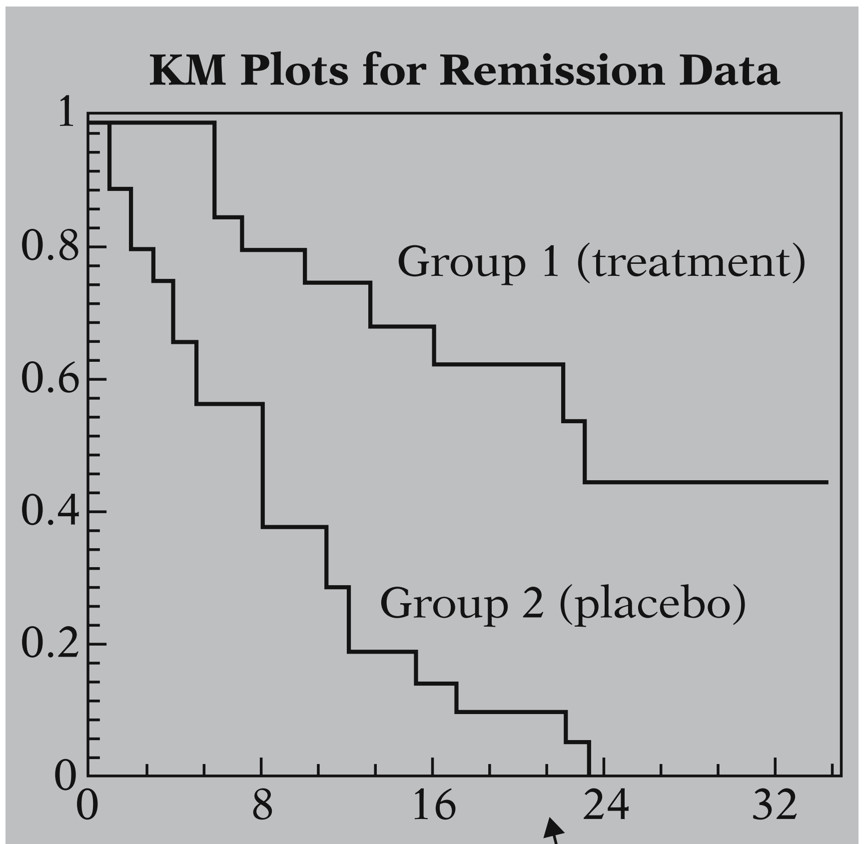
group 1 only

- The other survival estimates are calculated by multiplying the estimate for the immediately preceding failure time by a fraction. For example, the fraction is 18/21 for surviving past week 6, because 21 subjects remain up to week 6 and 3 of these subjects fail to survive past week 6. The fraction is 16/17 for surviving past week 7, because 17 people remain up to week 7 and one of these fails to survive past week 7. The other fractions are calculated similarly.

- For a specified failure time $t_{(j)}$, the fraction may be generally expressed as the conditional probability of surviving past time $t_{(j)}$, given availability (i.e., in the risk set) at time $t_{(j)}$. This is exactly the same formula that we previously used to calculate each product term in the product limit formula used for the group 2 data.

- Note that a subject might not be available at time $t_{(j)}$ for one of two reasons: (1) either the subject has failed prior to $t_{(j)}$, or (2) the subject has been censored prior to $t_{(j)}$. Group 1 has censored observations, whereas group 2 does not. Thus, for group 1, censored observations have to be taken into account when determining the number available at $t_{(j)}$.

An Example of Kaplan-Meier Curves



Obtain KM plots from computer package, e.g., SAS, Stata, SPSS R

- Plots of the KM curves for groups 1 and 2 are shown here on the same graph. Notice that the KM curve for group 1 is consistently higher than the KM curve for group 2. These figures indicate that group 1, which is the treatment group, has better survival prognosis than group 2, the placebo group. Moreover, as the number of weeks increases, the two curves appear to get farther apart, suggesting that the beneficial effects of the treatment over the placebo are greater the longer one stays in remission.

- The KM plots shown above can be easily obtained from most computer packages that perform survival analysis, including SAS, Stata, SPSS, and R. All the user needs to do is provide a KM computer program with the basic data layout and then provide appropriate commands to obtain plots.

General Features of KM Curves

General KM formula:

$$\hat{S}(t_{(j)}) = \hat{S}(t_{(j-1)}) \times \hat{\Pr}(T > t_{(j)} | T \geq t_{(j)})$$

KM formula = product limit formula

$$\hat{S}(t_{(j-1)}) = \prod_{i=1}^{j-1} \hat{\Pr}(T > t_{(i)} | T \geq t_{(i)})$$

EXAMPLE

$$\hat{S}(10) = .8067 \times \frac{14}{15} = .7529$$
$$= \boxed{\frac{18}{21} \times \frac{16}{17}} \times \frac{14}{15}$$

$$\hat{S}(16) = .6902 \times \frac{10}{11}$$
$$= \boxed{\frac{18}{21} \times \frac{16}{17} \times \frac{14}{15} \times \frac{11}{12}} \times \frac{10}{11}$$

- The general formula for a KM survival probability at failure time $t_{(j)}$ is shown here. This formula gives the probability of surviving past the previous failure time $t_{(j-1)}$, multiplied by the conditional probability of surviving past time $t_{(j)}$, given survival to at least time $t_{(j)}$.
- The above KM formula can also be expressed as a product limit if we substitute for the survival probability $\hat{S}(t_{(j-1)})$, the product of all fractions that estimate the conditional probabilities for failure times $t_{(j-1)}$ and earlier.
- For example, the probability of surviving past ten weeks is given in the table for group 1 by .8067 times 14/15, which equals .7529. But the .8067 can be alternatively written as the product of the fractions 18/21 and 16/17. Thus, the product limit formula for surviving past 10 weeks is given by the triple product shown here.
- Similarly, the probability of surviving past sixteen weeks can be written either as .6902 × 10/11, or equivalently as the five-way product of fractions shown here.

General Features of KM Curves

$$\hat{S}(t_{(j)}) = \prod_{i=1}^{j} \hat{\Pr}[T > t_{(i)} | T \geq t_{(i)}]$$
$$= \hat{S}(t_{(j-1)}) \times \hat{\Pr}(T > t_{(i)} | T \geq t_{(i)})$$

Math proof:

$$\Pr(\text{A and B}) = \Pr(\text{A}) \times \Pr(\text{B} \mid \text{A})$$
always

$$\text{A} = \text{“}T \geq t_{(j)}\text{”} \rightarrow \text{A and B} = \text{B}$$
$$\text{B} = \text{“}T > t_{(j)}\text{”}$$
$$\Pr(\text{A and B}) = \Pr(\text{B}) = \mathcal{S}(t_{(j)})$$

No failures during $t_{(j-1)} < \text{T} < t_{(j)}$

$$\Pr(\text{A}) = \Pr(T > t_{(j-1)}) = \mathcal{S}(t_{(j-1)})$$

$$\Pr(\text{B}|\text{A}) = \Pr(T > t_{(j)} | T \geq t_{(j)})$$

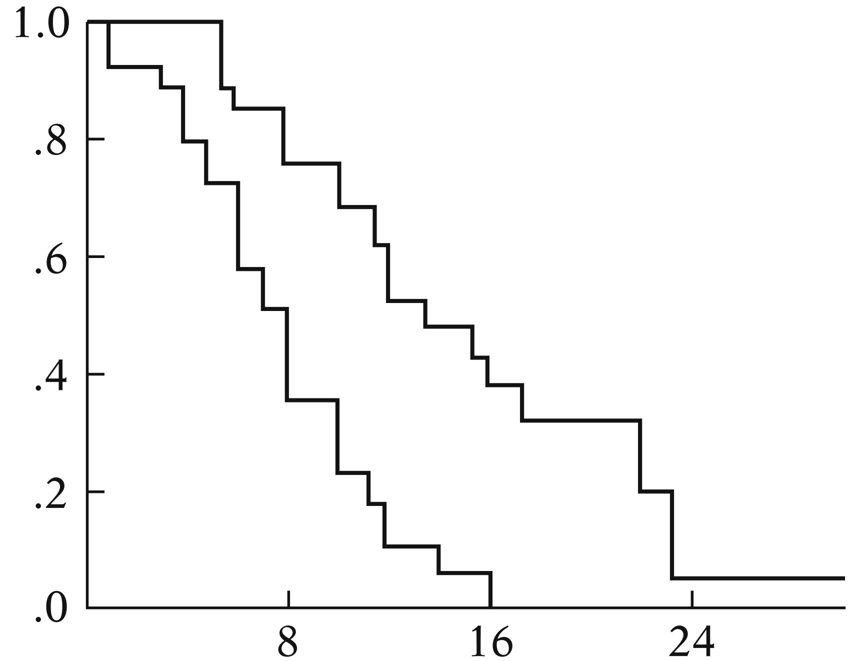
Thus, from Pr(A and B) formula,

$$\Pr(\text{A and B}) = \Pr(\text{A}) \times \Pr(\text{B} \mid \text{A})$$
$$\mathcal{S}(t_{(j)}) = \mathcal{S}(t_{(j-1)}) \times \Pr(T > t_{(j)} | T \geq t_{(j)})$$

- The general expression for the product limit formula for the KM survival estimate is shown here together with the general KM formula given earlier. Both expressions are equivalent.
- A simple mathematical proof of the KM formula can be described in probability terms. One of the basic rules of probability is that the probability of a joint event, say A and B, is equal to the probability of one event, say A, times the conditional probability of the other event, B, given A.
- If we let A be the event that a subject survives to at least time $t_{(j)}$ and we let B be the event that a subject survives past time $t_{(j)}$, then the joint event A and B simplifies to the event B, which is inclusive of A. It follows that the probability of A and B equals the probability of surviving past time $t_{(j)}$.
- Also, because $t_{(j)}$ is the next failure time after $t_{(j\text{-}1)}$, there can be no failures after time $t_{(j\text{-}1)}$ and before time $t_{(j)}$. Therefore, the probability of A is equivalent to the probability of surviving past the (j–1)th ordered failure time.
- Furthermore, the conditional probability of B given A is equivalent to the conditional probability in the KM formula.
- Thus, using the basic rules of probability, the KM formula can be derived.

The Log-Rank Test for Two Groups

Are KM curves statistically equivalent?



- Chi-square test
- Overall comparison of KM curves
- Observed versus expected counts
- Categories defined by ordered failure times

- We now describe how to evaluate whether or not KM curves for two or more groups are statistically equivalent. In this section we consider two groups only. The most popular testing method is called the log–rank test.
- When we state that two KM curves are "statistically equivalent," we mean that, based on a testing procedure that compares the two curves in some "overall sense," we do not have evidence to indicate that the true (population) survival curves are different.
- The log–rank test is a large-sample chi-square test that uses as its test criterion a statistic that provides an overall comparison of the KM curves being compared. This (log–rank) statistic, like many other statistics used in other kinds of chi-square tests, makes use of observed versus expected cell counts over categories of outcomes. The categories for the log–rank statistic are defined by each of the ordered failure times for the entire set of data being analyzed.

The Log-Rank Test for Two Groups

EXAMPLE

Remission data: $n = 42$

	# failures		# in risk set	
$t_{(j)}$	m_{1j}	m_{2j}	n_{1j}	n_{2j}
1	0	2	21	21
2	0	2	21	19
3	0	1	21	17
4	0	2	21	16
5	0	2	21	14
6	3	0	21	12
7	1	0	17	12
8	0	4	16	12
10	1	0	15	8
11	0	2	13	8
12	0	2	12	6
13	1	0	12	4
15	0	1	11	4
16	1	0	11	3
17	0	1	10	3
22	1	1	7	2
23	1	1	6	1

- As an example of the information required for the log–rank test, we again consider the comparison of the treatment (group 1) and placebo (group 2) subjects in the remission data on 42 leukemia patients.
- Here, for each ordered failure time, $t_{(j)}$, in the entire set of data, we show the numbers of subjects (m_{ij}) failing at that time, separately by group (i), followed by the numbers of subjects (n_{ij}) in the risk set at that time, also separately by group.
- Thus, for example, at week 4, no subjects failed in group 1, whereas two subjects failed in group 2. Also, at week 4, the risk set for group 1 contains 21 persons, whereas the risk set for group 2 contains 16 persons.
- Similarly, at week 10, one subject failed in group 1, and no subjects failed at group 2; the risk sets for each group contain 15 and 8 subjects, respectively.

The Log-Rank Test for Two Groups

Expected cell counts:

$$e_{1j} = \left(\frac{n_{1j}}{n_{1j} + n_{2j}}\right) \times (m_{1j} + m_{2j})$$

$$\underset{\substack{\text{Proportion}\\\text{in risk set}}}{\uparrow} \qquad \underset{\substack{\text{\# of failures over}\\\text{both groups}}}{\uparrow}$$

$$e_{2j} = \left(\frac{n_{2j}}{n_{1j} + n_{2j}}\right) \times (m_{1j} + m_{2j})$$

$$\begin{gathered}\text{\# of failure times}\\ \downarrow\end{gathered}$$

$$O_i - E_j = \sum_{j=1}^{17} (m_{ij} - e_{ij}),$$

$$i = 1, 2$$

EXAMPLE

$O_1 - E_1 = -10.26$

$O_2 - E_2 = \ 10.26$

- We now expand the previous table to include expected cell counts and observed minus expected values for each group at each ordered failure time. The formula for the expected cell counts is shown here for each group. For group 1, this formula computes the expected number at time j (i.e., e_{1j}) as the proportion of the total subjects in both groups who are at risk at time j, that is, $n_{1j}/(n_{1j}+n_{2j})$, multiplied by the total number of failures at that time for both groups (i.e., $m_{1j} + m_{2j}$). For group 2, e_{2j} is computed similarly.

- When two groups are being compared, the log–rank test statistic is formed using the sum of the observed minus expected counts over all failure times for one of the two groups. In this example, this sum is -10.26 for group 1 and 10.26 for group 2. We will use the group 2 value to carry out the test, but as we can see, except for the minus sign, the difference is the same for the two groups.

The Log-Rank Test for Two Groups

EXAMPLE

Expanded Table (Remission Data)

		# failures		# in risk set		# expected		Observed–expected	
j	$t_{(j)}$	m_{1j}	m_{2j}	n_{1j}	n_{2j}	e_{1j}	e_{2j}	$m_{1j} - e_{1j}$	$m_{2j} - e_{2j}$
1	1	0	2	21	21	$(21/42) \times 2$	$(21/42) \times 2$	–1.00	1.00
2	2	0	2	21	19	$(21/40) \times 2$	$(19/40) \times 2$	–1.05	1.05
3	3	0	1	21	17	$(21/38) \times 1$	$(17/38) \times 1$	–0.55	0.55
4	4	0	2	21	16	$(21/37) \times 2$	$(16/37) \times 2$	–1.14	1.14
5	5	0	2	21	14	$(21/35) \times 2$	$(14/35) \times 2$	–1.20	1.20
6	6	3	0	21	12	$(21/33) \times 3$	$(12/33) \times 3$	1.09	–1.09
7	7	1	0	17	12	$(17/29) \times 1$	$(12/29) \times 1$	0.41	–0.41
8	8	0	4	16	12	$(16/28) \times 4$	$(12/28) \times 4$	–2.29	2.29
9	10	1	0	15	8	$(15/23) \times 1$	$(8/23) \times 1$	0.35	–0.35
10	11	0	2	13	8	$(13/21) \times 2$	$(8/21) \times 2$	–1.24	1.24
11	12	0	2	12	6	$(12/18) \times 2$	$(6/18) \times 2$	–1.33	1.33
12	13	1	0	12	4	$(12/16) \times 1$	$(4/16) \times 1$	0.25	–0.25
13	15	0	1	11	4	$(11/15) \times 1$	$(4/15) \times 1$	–0.73	0.73
14	16	1	0	11	3	$(11/14) \times 1$	$(3/14) \times 1$	0.21	–0.21
15	17	0	1	10	3	$(10/13) \times 1$	$(3/13) \times 1$	–0.77	0.77
16	22	1	1	7	2	$(7/9) \times 2$	$(2/9) \times 2$	–0.56	0.56
17	23	1	1	6	1	$(6/7) \times 2$	$(1/7) \times 2$	–0.71	0.71
Totals		9	21			19.26	10.74	–10.26	–10.26

The Log-Rank Test for Two Groups

Two groups:
$O_2 - E_2 =$ summed observed minus expected score for group 2

$$\boxed{\text{Log–rank statistic} = \frac{(O_2 - E_2)^2}{\text{Var}(O_2 - E_2)}}$$

- For the two-group case, the log–rank statistic, shown here at the left, is computed by dividing the square of the summed observed minus expected score for one of the groups—say, group 2—by the variance of the summed observed minus expected score.

$$\text{Var}(O_i - E_i) = \sum_j \frac{n_{1j} n_{2j} (m_{1j} + m_{2j})(n_{1j} + n_{2j} - m_{1j} - m_{2j})}{(n_{1j} + n_{2j})^2 (n_{1j} + n_{2j} - 1)}$$

$i = 1, 2$

- The expression for the estimated variance is shown here. For two groups, the variance formula is the same for each group. This variance formula involves the number in the risk set in each group (n_{ij}) and the number of failures in each group (m_{ij}) at time j. The summation is over all distinct failure times.

H_0: no difference between survival curves

Log–rank statistic $\sim \chi^2$ with 1 df under H_0

- The null hypothesis being tested is that there is no overall difference between the two survival curves. Under this null hypothesis, the log–rank statistic is approximately chi-square with one degree of freedom. Thus, a P-value for the log–rank test is determined from tables of the chi-square distribution.

The Log-Rank Test for Two Groups

Computer programs:

R: surfit() and **survdiff()**

- descriptive statistics for KM curves
- log–rank statistic
- Alternative statistics to log–rank statistic

- Several computer programs are available for calculating the log–rank statistic. For example the survival package of the **R** program has functions, one of which is called "**survfit()**" that computes descriptive information about Kaplan–Meier curves and the other is "**survdiff()**" that computes the log–rank statistic and alternative statistics to the log–rank statistic, to be described later. Other packages, like **SAS**, **SPSS**, and **Stata** have procedures that provide results similar to those of **R**.

EXAMPLE

Using R: Remission Data

Group	Events observed	Events expected
1	9	19.25
2	21	10.75
Total	30	30.00

Log-rank = chi2(2) = 16.79
P-value = Pr > chi2 = 0.000

- For the remission data, the edited printout from using the **R** "**survfit()**" and "**surfdiff()**" function is shown here. The log–rank statistic is 16.79 and the corresponding P-value is zero to three decimal places. This P-value indicates that the null hypothesis should be rejected. We can therefore conclude that the treatment and placebo groups have significantly different KM survival curves.

The Log-Rank Test for Two Groups

EXAMPLE

$$O_2 - E_2 = 10.26$$

$$\text{Var}(O_2 - E_2) = 6.2685$$

$$\text{Log–rank statistic} = \frac{(O_2 - E_2)^2}{\widehat{\text{Var}}(O_2 - E_2)}$$

$$= \frac{(10.26)^2}{6.2685} = 16.793$$

- Although the use of a computer is the easiest way to calculate the log–rank statistic, we provide here some of the details of the calculation. We have already seen from earlier computations that the value of $O_2 - E_2$ is 10.26. The estimated variance of $O_2 - E_2$ is computed from the variance formula above to be 6.2685. The log–rank statistic then is obtained by squaring 10.26 and dividing by 6.285, which yields 16.793, as shown on the computer printout.

Approximate formula:

$$X^2 \approx \sum_{i}^{\#\text{ of groups}} \frac{(O_i - E_i)^2}{E_i}$$

- An approximation to the log–rank statistic, shown here, can be calculated using observed and expected values for each group without having to compute the variance formula. The approximate formula is of the classic chi-square form that sums over each group being compared the square of the observed minus expected value divided by the expected value.

EXAMPLE

$$X^2 = \frac{(-10.26)^2}{19.26} + \frac{(10.26)^2}{10.74}$$

$$= 15.276$$

$$\text{Log–rank statistic} = 16.793$$

- The calculation of the approximate formula is shown here for the remission data. The expected values are 19.26 and 10.74 for groups 1 and 2, respectively. The chi-square value obtained is 15.276, which is slightly smaller than the log–rank statistic of 16.793.

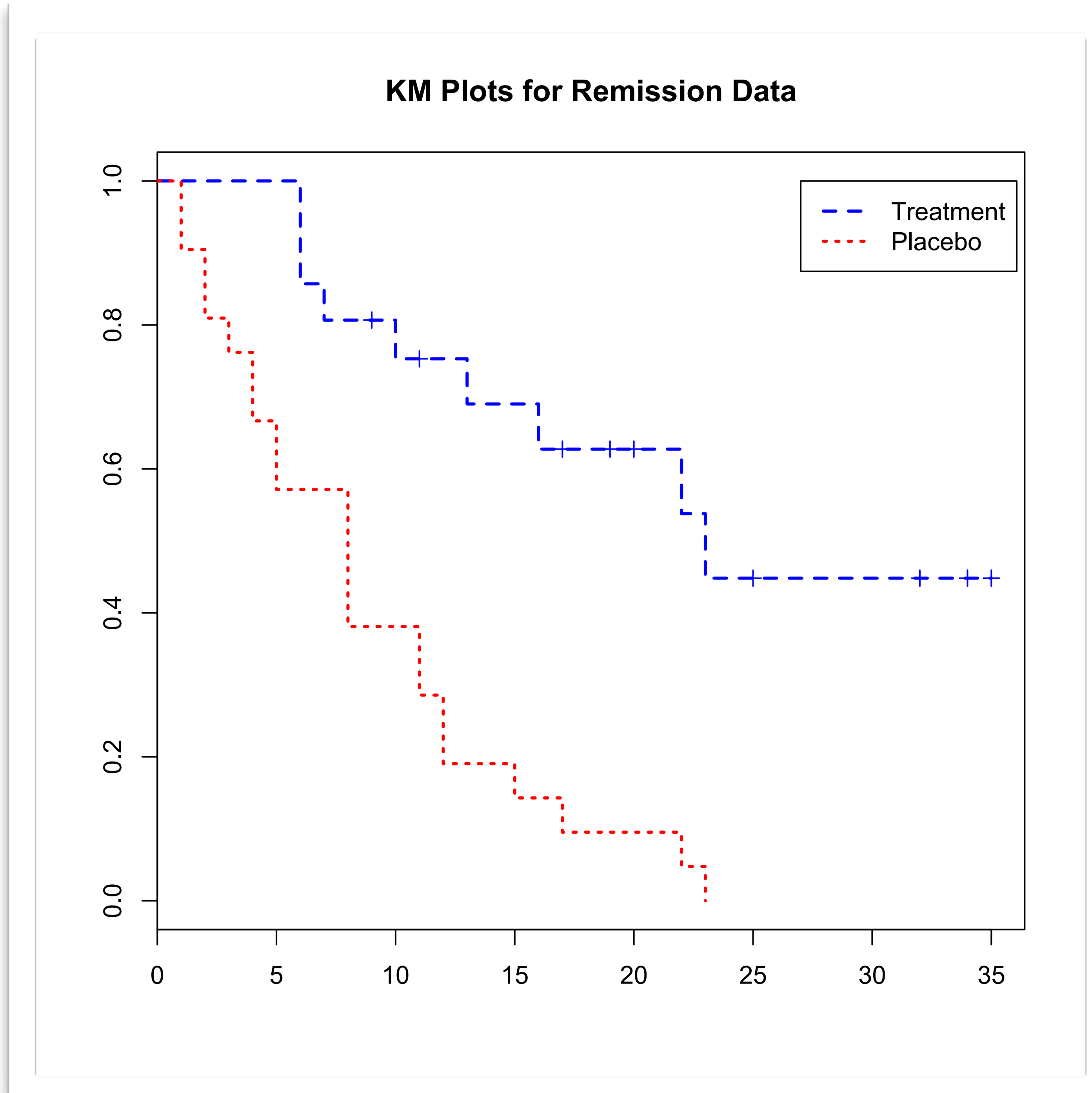
R Implementation for Remission Data

```
> res1<-survfit(remission~x1)
>
> res1
Call: survfit(formula = remission ~ x1)

     records n.max n.start events median 0.95LCL 0.95UCL
x1=0      21    21      21      9     23      16      NA
x1=1      21    21      21     21      8       4      12
> summary(res1)
Call: survfit(formula = remission ~ x1)

                x1=0
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
    6     21       3    0.857  0.0764        0.720        1.000
    7     17       1    0.807  0.0869        0.653        0.996
   10     15       1    0.753  0.0963        0.586        0.968
   13     12       1    0.690  0.1068        0.510        0.935
   16     11       1    0.627  0.1141        0.439        0.896
   22      7       1    0.538  0.1282        0.337        0.858
   23      6       1    0.448  0.1346        0.249        0.807

                x1=1
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
    1     21       2   0.9048  0.0641      0.78754        1.000
    2     19       2   0.8095  0.0857      0.65785        0.996
    3     17       1   0.7619  0.0929      0.59988        0.968
```

```
    4     16       2   0.6667  0.1029      0.49268        0.902
    5     14       2   0.5714  0.1080      0.39455        0.828
    8     12       4   0.3810  0.1060      0.22085        0.657
   11      8       2   0.2857  0.0986      0.14529        0.562
   12      6       2   0.1905  0.0857      0.07887        0.460
   15      4       1   0.1429  0.0764      0.05011        0.407
   17      3       1   0.0952  0.0641      0.02549        0.356
   22      2       1   0.0476  0.0465      0.00703        0.322
   23      1       1   0.0000     NaN           NA           NA
```

R Implementation for Remission Data

```
> plot(res1,main='KM Plots for Remission Data', lty=c(2,3), col=c('blue','red'))
> legend(27, 1, c('Treatment','Placebo'), lty=c(2,3), col=c('blue','red'))
```



R Implementation for Remission Data

```
> res2<-survdiff(remission~x1)
> res2
Call:
survdiff(formula = remission ~ x1)

      N Observed Expected (O-E)^2/E (O-E)^2/V
x1=0 21        9     19.3      5.46      16.8
x1=1 21       21     10.7      9.77      16.8

 Chisq= 16.8  on 1 degrees of freedom, p= 4.17e-05
```

The Log-Rank Test for Several Groups

H_0: All survival curves are the same.

Log–rank statistics for > 2 groups involves variances and covariances of $O_i - E_i$.

G (≥ 2) groups:
log–rank statistic $\sim\chi^2$ with $G - 1$ df

Approximation formula:

$$X^2 = \sum_{i}^{\#\text{ of groups}} \frac{(O_i - E_i)^2}{E_i}$$

Not required because computer program calculates the exact log–rank statistic

- The log–rank test can also be used to compare three or more survival curves. The null hypothesis for this more general situation is that all survival curves are the same.
- Although the same tabular layout can be used to carry out the calculations when there are more than two groups, the test statistic is more complicated mathematically, involving both variances and covariances of summed observed minus expected scores for each group. A convenient mathematical formula can be given in matrix terms. We present the matrix formula for the interested reader at the end of this chapter.
- We will not describe further details about the calculation of the log–rank statistic, because a computer program can easily carry out the computations from the basic data file. Instead, we illustrate the use of this test with data involving more than two groups.
- If the number of groups being compared is $G(\geq 2)$, then the log–rank statistic has approximately a large sample chi-square distribution with $G-1$ degrees of freedom. Therefore, the decision about significance is made using chi-square tables with the appropriate degrees of freedom.
- The approximate formula previously described involving only observed and expected values without variance or covariance calculations can also be used when there are more than two groups being compared. However, practically speaking, the use of this approximate formula is not required as long as a computer program is available to calculate the exact log–rank statistic.

The Log-Rank Test for Several Groups

EXAMPLE

vets.dat: survival time in days,

$n = 137$

Veteran's Administration Lung Cancer Trial

Column 1:	Treatment (standard = 1, test = 2)
Column 2:	Cell type 1 (large = 1, other = 0)
Column 3:	Cell type 2 (adeno = 1, other = 0)
Column 4:	Cell type 3 (small = 1, other = 0)
Column 5:	Cell type 4 (squamous = 1, other = 0)
Column 6:	Survival time (days)
Column 7:	Performance Status (0 = worst . . . 100 = best)
Column 8:	Disease duration (months)
Column 9:	Age
Column 10:	Prior therapy (none = 0, some = 1)
Column 11:	Status (0 = censored, 1 = died)

- We now provide an example to illustrate the use of the log–rank statistic to compare more than two groups.
- The data set "vets.dat" considers survival times in days for 137 patients from the Veteran's Administration Lung Cancer Trial cited by Kalbfleisch and Prentice in their text (*The Statistical Analysis of Survival Time Data*, John Wiley, pp. 223–224, 1980). A complete list of the variables is shown here. Failure status is defined by the status variable (column 11).
- Among the variables listed, we now focus on the performance status variable (column 7). This variable is a continuous variable, so before we can obtain KM curves and the log–rank test, we need to categorize this variable.

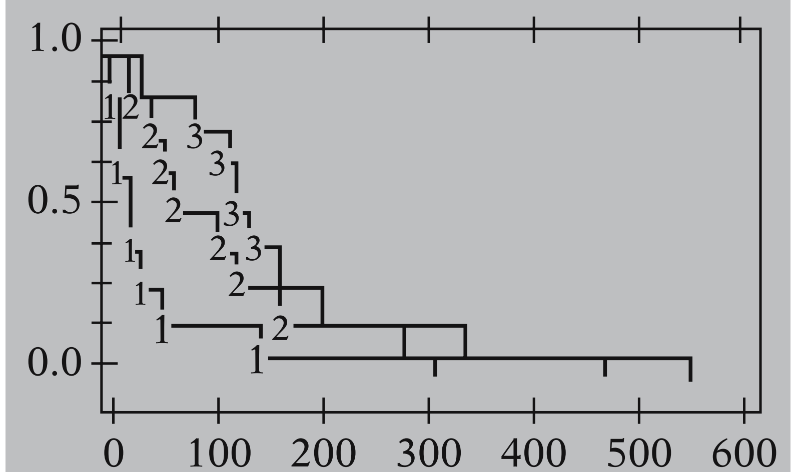
T… Test for Several Groups

EXAMPLE (continued)

Performance Status Categories

Group #	Categories	Size
1	0–59	52
2	60–74	50
3	75–100	35

KM curves for performance status groups



Group	Events observed	Events expected
1	50	26.30
2	47	55.17
3	31	46.53
Total	128	128.00

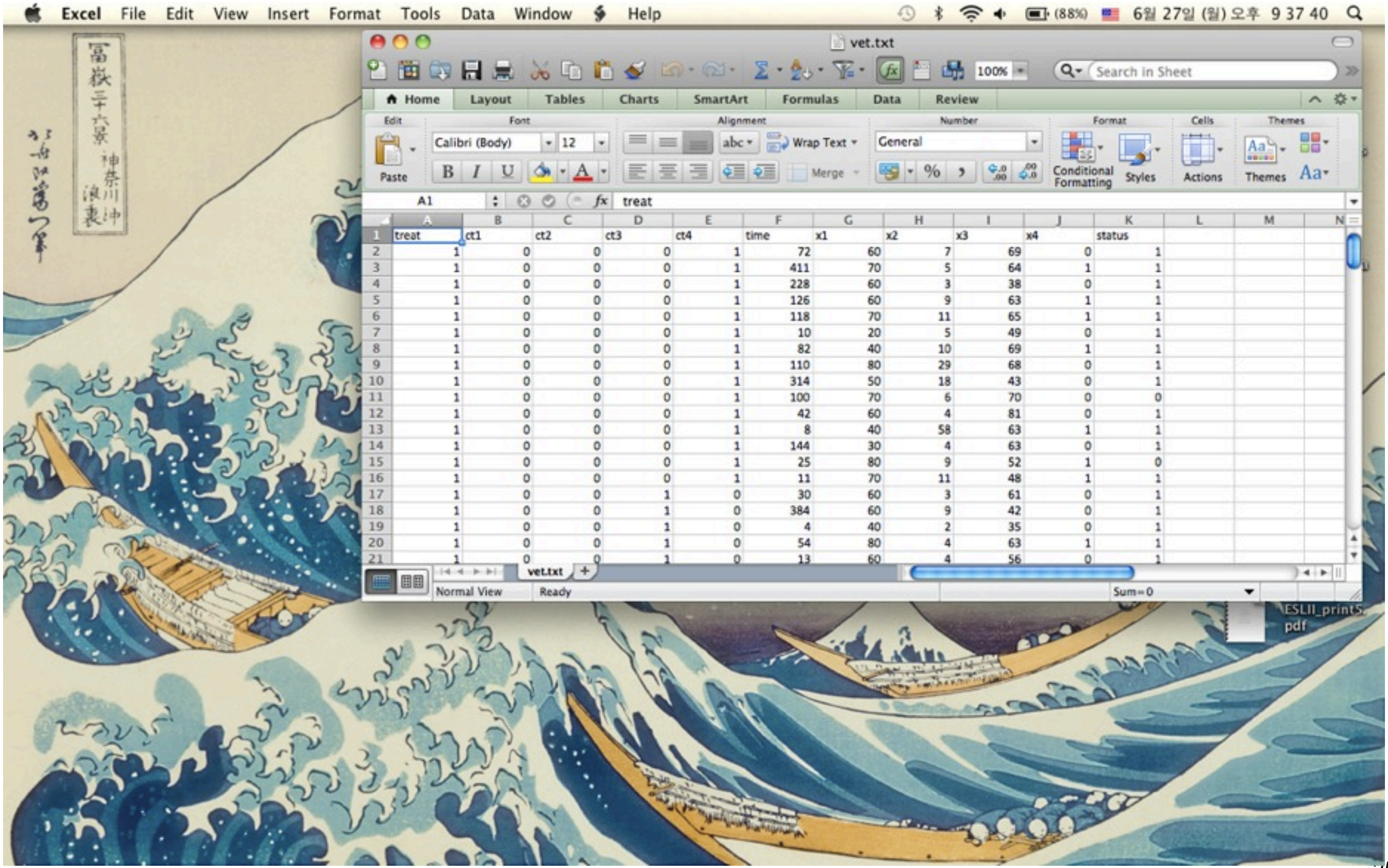
Log–rank = chi2(2) = 29.18
P-value = Pr > chi2 = 0.0000
G = 3 groups; $df = G - 1 = 2$

Log–rank test is highly significant.

Conclude significant difference among three survival curves.

- If, for the performance status variable, we choose the categories 0–59, 60–74, and 75–100, we obtain three groups of sizes 52, 50, and 35, respectively.
- The KM curves for each of three groups are shown here. Notice that these curves appear to be quite different. A test of significance of this difference is provided by the log–rank statistic.
- An edited printout of descriptive information about the three KM curves together with the log–rank test results are shown here. These results were obtained using the R program.
- Because three groups are being compared here, $G = 3$ and the degrees of freedom for the log–rank test is thus $G - 1$, or 2. The log–rank statistic is computed to be 29.181, which has a P-value of zero to three decimal places. Thus, the conclusion from the log–rank test is that there is a highly significant difference among the three survival curves for the performance status groups.

R Implementation for Veteran Data



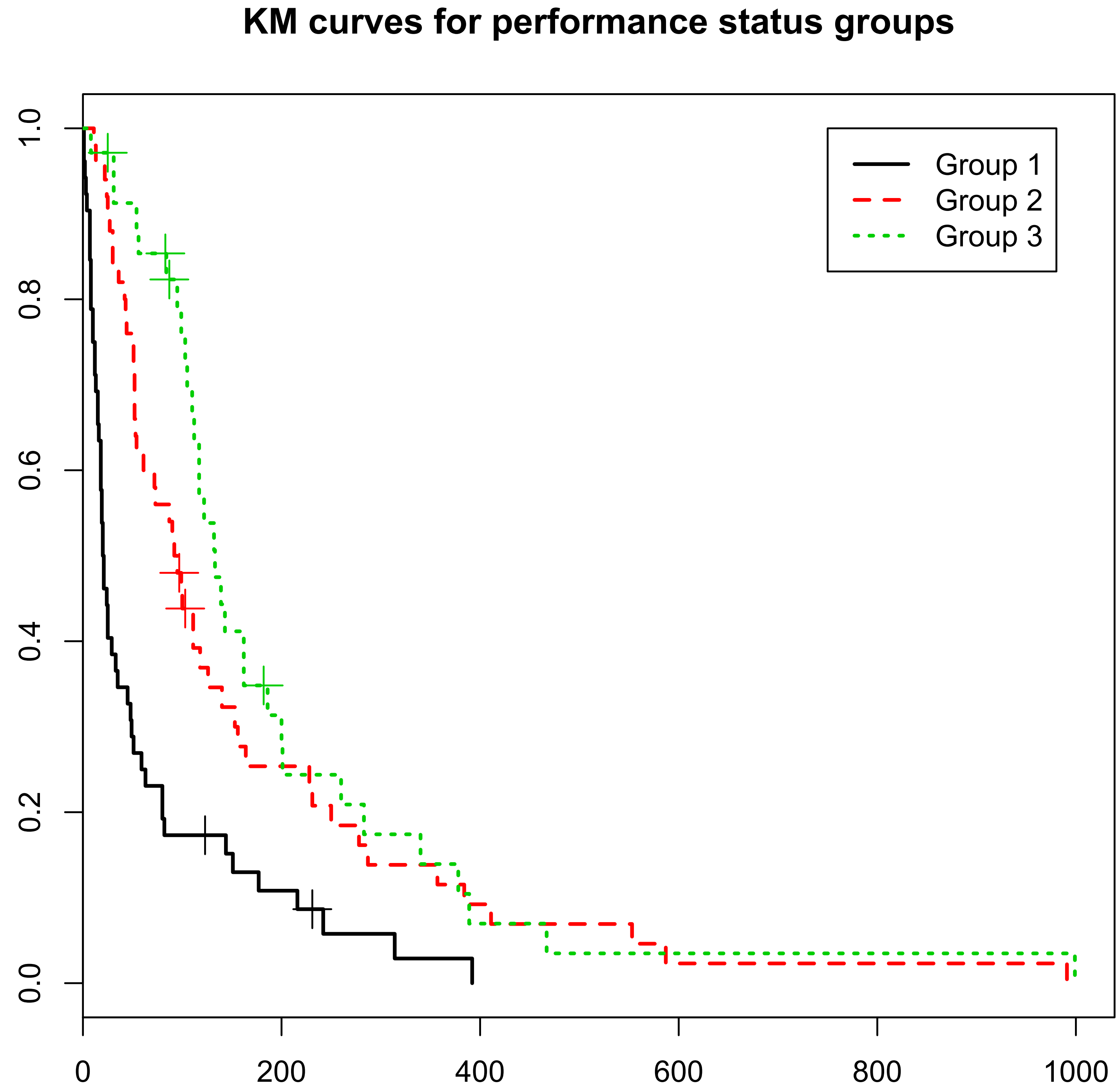
treat	ct1	ct2	ct3	ct4	time	x1	x2	x3	x4	status
1	0	0	0	1	72	60	7	69	0	1
1	0	0	0	1	411	70	5	64	1	1
1	0	0	0	1	228	60	3	38	0	1
1	0	0	0	1	126	60	9	63	1	1
1	0	0	0	1	118	70	11	65	1	1
1	0	0	0	1	10	20	5	49	0	1
1	0	0	0	1	82	40	10	69	1	1
1	0	0	0	1	110	80	29	68	0	1
1	0	0	0	1	314	50	18	43	0	1
1	0	0	0	1	100	70	6	70	0	0
1	0	0	0	1	42	60	4	81	0	1
1	0	0	0	1	8	40	58	63	1	1
1	0	0	0	1	144	30	4	63	0	1
1	0	0	0	1	25	80	9	52	1	0
1	0	0	0	1	11	70	11	48	1	1
1	0	0	1	0	30	60	3	61	0	1
1	0	0	1	0	384	60	9	42	0	1
1	0	0	1	0	4	40	2	35	0	1
1	0	0	1	0	54	80	4	63	1	1
1	0	0	1	0	13	60	4	56	0	1

R Implementation for Veteran Data

```
> ### Rcode2-2.r
>
> library(survival) # loading survival package
>
> vet.dat<-read.table('vet.dat',sep='\t',header=TRUE)
> #vet.dat<-read.table(file.choose(),sep='\t',header=TRUE)
> head(vet.dat)
  treat ct1 ct2 ct3 ct4 time x1 x2 x3 x4 status
1     1   0   0   0   1   72 60  7 69  0      1
2     1   0   0   0   1  411 70  5 64  1      1
3     1   0   0   0   1  228 60  3 38  0      1
4     1   0   0   0   1  126 60  9 63  1      1
5     1   0   0   0   1  118 70 11 65  1      1
6     1   0   0   0   1   10 20  5 49  0      1
>
> vet<-Surv(vet.dat[,6],vet.dat[,11])
>
> grp1<-which(vet.dat[,7]<=59)
> grp2<-which(vet.dat[,7]>=60 & vet.dat[,7]<=74)
> grp3<-which(vet.dat[,7]>=75)
> length(grp1)
[1] 52
> length(grp2)
[1] 50
> length(grp3)
[1] 35
```

```
> grp<-rep(1,137)
> grp[grp2]<-2
> grp[grp3]<-3
> res<-survfit(vet~grp)
> plot(res,main='KM curves for performance status groups',
lty=1:3, col=1:3, lwd=2, cex=2)
> legend(750,1,c('Group 1', 'Group 2', 'Group 3'), lty=1:3,
col=1:3, lwd=2)
> res2<-survdiff(vet~grp)
> res2
Call:
survdiff(formula = vet ~ grp)

       N Observed Expected (O-E)^2/E (O-E)^2/V
grp=1 52       50     26.3     21.36     28.22
grp=2 50       47     55.2      1.21      2.19
grp=3 35       31     46.5      5.18      8.44

 Chisq= 29.2  on 2 degrees of freedom, p= 4.61e-07
```

R Implementation for Veteran Data



The Log-Rank Test for Several Groups

Chapters

1. Introduction

✓ 2. Kaplan–Meier Survival Curves and the Log–Rank Test

Next:

3. The Cox Proportional Hazards Model and Its Characteristics

- This chapter is now complete. Chapter 3 introduces the Cox proportional hazards (PH) model, which is the most popular mathematical modeling approach for estimating survival curves when considering several explanatory variables simultaneously.

Appendix: Matrix Formula for the Log-Rank Statistic for Several Groups

For $i = 1, 2, \ldots, G$ and $j = 1, 2, \ldots, k$, where $G = \#$ of groups and $k = \#$ of distinct failure times,

$n_{ij} = \#$ at risk in ith group at jth ordered failure time

$m_{ij} =$ observed # of failures in ith group at jth ordered failure time

$e_{ij} =$ expected # of failures in ith group at jth ordered failure time

$$= \left(\frac{n_{ij}}{n_{1j} + \cdots + n_{Gj}}\right)(m_{1j} + \cdots + m_{Gj})$$

$$n_j = \sum_{i=1}^{G} n_{ij}$$

$$m_j = \sum_{i=1}^{G} m_{ij}$$

$$O_i - E_i = \sum_{j=1}^{k} (m_{ij} - e_{ij})$$

$$\mathrm{Var}(O_i - E_i) = \sum_{j=1}^{k} \frac{n_{ij}(n_j - n_{ij})m_j(n_j - m_j)}{n_j^2(n_j - 1)}$$

$\mathbf{d} = (O_1 - E_1, O_2 - E_2, \ldots, O_{G-1} - E_{G-1})'$

$\mathbf{V} = ((v_{il}))$

where $v_{ii} = \mathrm{Var}\,(O_i - E_i)$ and $v_{il} = \mathrm{Cov}\,(O_i - E_i, O_l - E_l)$ for $i = 1, 2, \ldots, G-1$; $l = 1, 2, \ldots, G-1$.

Then, the log–rank statistic is given by the matrix product formula:

$$\boxed{\text{Log–rank statistic} = \mathbf{d}'\mathbf{V}^{-1}\mathbf{d}}$$

which has approximately a chi-square distribution with $G - 1$ degrees of freedom under the null hypothesis that all G groups have a common survival curve.