BSTT536: Survival Data Analysis

Instructor: Hua Yun Chen, PhD

Division of Epidemiology and Biostatistics
School of Public Health
University of Illinois at Chicago

Table of Content

Example: Myeloma survival functions for male and female

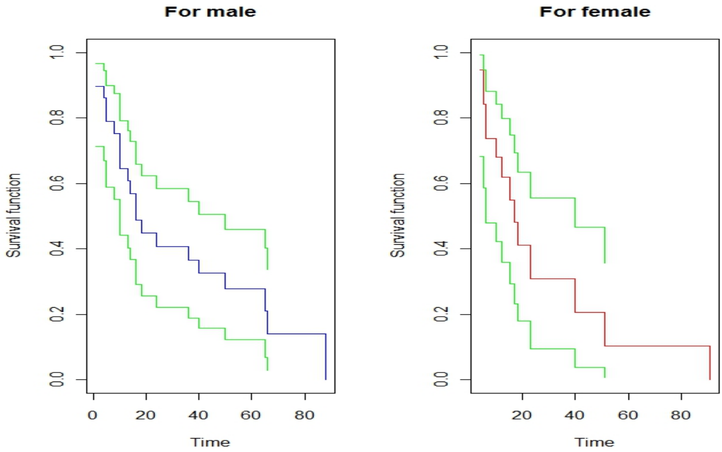

Figure: Survival function estimates for the data on multiple myeloma.

Comparison of two survival functions

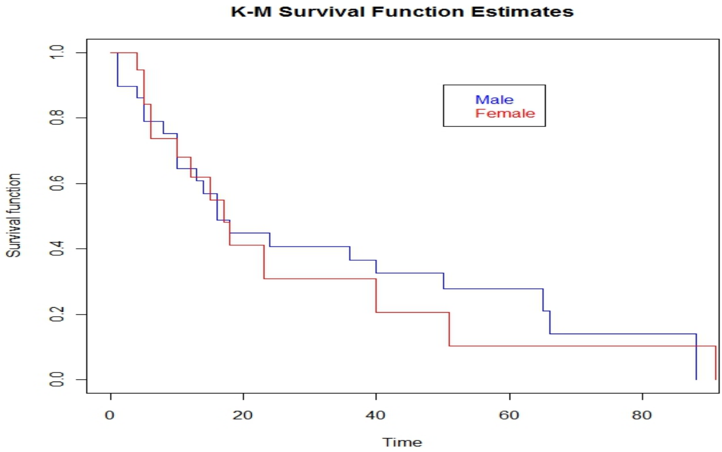

Figure: Comparison of Male and Female KM Survival Curve estimates

Two-sample tests

1. Test of difference between two survival functions need to compare the two functions at all failure time points. All the information should then be assembled to form a test.
2. The null hypothesis is that the two survival functions are the same. $H_0 : S_1(t) = S_2(t)$ for all t.
3. Compare two survival functions at failure time t_k,

Group	At risk	Observed Failures	Expected failures under H_0
1	n_{k1}	d_{k1}	$n_{k1}d_k/n_k$
2	n_{k2}	d_{k2}	$n_{k2}d_k/n_k$
Total	n_k	d_k	d_k

Two-sample tests (continuing 1)

1. The differences between the observed failures and the expected failures under H_0 are respectively

$$d_{k1} - n_{k1}d_k/n_k \text{ and } d_{k2} - n_{k2}d_k/n_k$$

provide evidence for testing the hypothesis.

2. The differences are related as

$$d_{k1} - n_{k1}d_k/n_k = -\left\{d_{k2} - n_{k2}d_k/n_k\right\}.$$

We need only to consider one of the two.

Two-sample tests (continuing 2)

1. The variance for the difference of observed and expected under H_0 can be estimated by

$$\hat{V}_k\left(d_{k1} - n_{k1} d_k / n_k\right) = \frac{n_{k1} n_{k2} d_k (n_k - d_k)}{n_k^2 (n_k - 1)}$$

2. The test statistic is a sum of weighted differences:

$$U_w = \sum_{k=1}^{K} w_k \left(d_{k1} - n_{k1} \frac{d_k}{n_k} \right),$$

where w_k are well-behaved weights.

Two-sample tests (continuing 3)

1. The variance of U_w can be estimated by

$$\hat{V}_w = \sum_{k=1}^{K} w_k^2 \hat{V}_k$$

   See the derivation part for how to obtain this result.

2. The test is

$$\frac{U_w}{\sqrt{V_w}} \sim N(0,1).$$

   or

$$U_w V_w^{-1} U_w \sim \chi_1^2.$$

Log-rank test and Wilcoxon test for two-sample comparison

1. The log-rank test is obtained if $w_k = 1$ for all k. This means that the log-rank test is powerful if the differences between the two groups are in the same direction. Otherwise, the differences can be cancelled when the summation is taken.
2. If $w_k = n_k$ for all k. The test is called the Wilcoxon test. It gives more weight to the difference in the groups with a large number of subjects at risk.
3. The general test is called weighted log-rank test.

Example: Compare male and female survival curves using log-rank test

Formulas: $e_{k1} = n_{k1} d_k / n_k$ and
$V_k = n_{k1} n_{k2} d_k (n_k - d_k) / \{ n_k^2 (n_k - 1) \}$.

Time	d_{k1}/n_{k1}	d_{k2}/n_{k2}	d_k/n_k	e_{k1}	$d_{k1} - e_{k1}$	V_k
1	3/29	0/19	3/48	1.81	1.19	0.6869
4	1/25	1/19	2/44	1.14	-0.14	0.4793
5	2/24	2/18	4/42	2.29	-0.29	0.9079
6	0/22	2/16	2/38	1.16	-1.16	0.4744
8	1/22	0/13	1/35	0.63	0.37	0.2335
10	3/21	1/13	4/34	2.47	0.53	0.8588
12	0/17	1/11	1/28	0.61	-0.61	0.2385
13	1/17	0/9	1/26	0.65	0.35	0.2263
14	1/16	0/9	1/25	0.64	0.36	0.2304
15	0/15	1/9	1/24	0.63	-0.63	0.2344
16	2/14	0/8	2/22	1.27	0.73	0.4408
17	0/12	1/8	1/20	0.60	-0.60	0.2400

Example: Compare male and female survival curves using log-rank test (continuing)

Formulas: $e_{k1} = n_{k1} d_k / n_k$ and
$V_k = n_{k1} n_{k2} d_k (n_k - d_k) / \{n_k^2 (n_k - 1)\}$.

Time	d_{k1}/n_{k1}	d_{k2}/n_{k2}	d_k/n_k	e_{k1}	$d_{k1} - e_{k1}$	V_k
18	1/12	1/7	2/19	1.26	-0.26	0.4395
23	0/11	1/4	1/15	0.73	-0.73	0.1956
24	1/11	0/3	1/14	0.79	0.21	0.1684
36	1/10	0/3	1/13	0.77	0.23	0.1775
40	1/9	1/3	2/12	1.50	-0.50	0.3409
50	1/7	0/2	1/9	0.78	0.22	0.1728
51	0/6	1/2	1/8	0.75	-0.75	0.1875
65	1/4	0/1	1/5	0.80	0.20	0.1600
66	1/3	0/1	1/4	0.75	0.25	0.1875
88	1/1	0/1	1/2	0.50	0.50	0.2500
91	0/0	1/1	1/1	0.00	0.00	NA
Total					-0.5196	7.53

Example: Compare male and female survival curves using Wilcoxon test

Formulas: $e_{k1} = n_{k1} d_k / n_k$ and
$V_k = n_{k1} n_{k2} d_k (n_k - d_k) / \{n_k^2 (n_k - 1)\}$.

Time	d_{k1}/n_{k1}	d_{k2}/n_{k2}	d_k/n_k	e_{k1}	$n_k(d_{k1} - e_{k1})$	$n_k^2 V_k$
1	3/29	0/19	3/48	1.81	57	1583
4	1/25	1/19	2/44	1.14	-6	928
5	2/24	2/18	4/42	2.29	-12	1602
6	0/22	2/16	2/38	1.16	-44	685
8	1/22	0/13	1/35	0.63	13	286
10	3/21	1/13	4/34	2.47	18	993
12	0/17	1/11	1/28	0.61	-17	187
13	1/17	0/9	1/26	0.65	9	153
14	1/16	0/9	1/25	0.64	9	144
15	0/15	1/9	1/24	0.63	-15	135
16	2/14	0/8	2/22	1.27	16	213
17	0/12	1/8	1/20	0.60	-12	96

Example: Compare male and female survival curves using Wilcoxon test (continuing)

Formulas: $e_{k1} = n_{k1} d_k / n_k$ and
$V_k = n_{k1} n_{k2} d_k (n_k - d_k)/\{n_k^2 (n_k - 1)\}$.

Time	d_{k1}/n_{k1}	d_{k2}/n_{k2}	d_k/n_k	e_{k1}	$n_k(d_{k1} - e_{k1})$	$n_k^2 V_k$
18	1/12	1/7	2/19	1.26	-5	159
23	0/11	1/4	1/15	0.73	-11	44
24	1/11	0/3	1/14	0.79	3	33
36	1/10	0/3	1/13	0.77	3	30
40	1/9	1/3	2/12	1.50	-6	49
50	1/7	0/2	1/9	0.78	2	14
51	0/6	1/2	1/8	0.75	-6	12
65	1/4	0/1	1/5	0.80	1	4
66	1/3	0/1	1/4	0.75	1	3
88	1/1	0/1	1/2	0.50	1	1
91	0/0	1/1	1/1	0.00	0	NA
Total					-1	7353

Log-rank test and Wilcoxon test for two-sample comparison

1. The log-rank test for comparing survival functions for male and female in the myeloma data

$$\frac{(-0.5196)^2}{7.53} = 0.0359.$$

$p\chi^2(0.0359) = 0.15$. $p - value = 1 - 0.15 = 0.85$.

2. The Wilcoxon test for comparing survival functions for male and female in the myeloma data

$$\frac{(-1)^2}{7353} = 0.00014.$$

$p\chi^2(0.0014) = 0.01$. $p - value = 1 - 0.01 = 0.99$.

Comparing survival functions for more than two groups

1. For more than two groups, compute

$$e_{kg} = n_{kg} d_{k+}/n_{k+}, g = 1, \cdots, G-1,$$

   where G is the total number of groups.

2. Form a vector

$$\Delta_k = (d_{kg} - e_{kg}, g = 1, \cdots, G-1).$$

3. The difference for the last group is excluded because the sum of all the differences equals zero.

Comparing survival functions for more than two groups (continuing)

1. The variance covariance estimate is

$$V_{kgg'} = \frac{n_{kg} d_{k+}(n_{k+} - d_{k+})}{n_{k+}(n_{k+} - 1)} \left(\delta_{gg'} - \frac{n_{kg'}}{n_{k+}} \right).$$

   where $\delta_{gg'} = 1$ if $g = g'$, $\delta_{gg'} = 0$ otherwise. Let $V_k = (V_{kgg'}, g, g' = 1, \cdots, G - 1)$, a $(G - 1) \times (G - 1)$ matrix.

2. The log-rank test statistic

$$(\sum_k \Delta_k)^t \left(\sum_k V_k \right)^{-1} \sum_k \Delta_k \sim \chi^2_{G-1}.$$

Comparison of survival curves within stratum

1. If the survival distribution is affected by the stratum variables and the difference among groups within each stratum are the same across strata, the stratified weighted log-rank test can be applied.
2. First, compute the sum of differences and the variance within each stratum by the applying the weighted log-rank test to the data in each stratum.
3. Then sum over the differences and the variance across strata. Form a χ^2 test by squared the sum of differences across strata, and then divided by the variance.
4. The χ^2 statistic has the same degree of freedom as having a single stratum.

Test of trend in ordered groups (more powerful)

1. If the G groups can be ordered, a one-degree of freedom trend test instead of the $G-1$ degree of freedom test can be formed by assigning scores to each group.
2. Let u_g be the score for group g. Let
$$U_T = \sum_{g=1}^{G} u_g(d_{+g} - e_{+g}).$$
3. The variance of U_T can be estimated by
$$V_T = \sum_{g=1}^{G} (u_g - u_+)^2 e_{+g},$$
where $u_+ = \sum_g u_g e_{+g} / \sum_g e_{+g}$.
4. Test of trend can be perform by examing
$$\frac{U_T^2}{V_T} \sim \chi_1^2.$$

*Derivation of the variance formula

1. At failure time t_k, the observed data are

Group	Failures	non-Failures	at Risk
1	d_{k1}	$n_{k1} - d_{k1}$	n_{k1}
2	d_{k2}	$n_{k2} - d_{k2}$	n_{k2}
Total	d_k	$n_k - d_k$	n_k

2. The distribution of the number of death at time t_k for groups 1 and 2 are respectively

$$P(D_{kj} = d_{kj} | n_{kj}) = \binom{n_{kj}}{d_{kj}} p_{kj}^{d_{kj}} (1 - p_{kj})^{n_{kj} - d_{kj}}, j = 1, 2.$$

*Derivation of the variance formula (continuing 1)

1. The groups are independent. So the joint probability are the product of the two.
2. For fixed marginals: n_k, d_k, and n_{k1} and n_{k2}, (that is, conditional on the marginals are fixed), the distribution $P(D_{k1} = d_{k1} \mid d_k, n_k, n_{k1}, n_{k2})$ follows

$$\frac{\begin{pmatrix} n_{k1} \\ d_{k1} \end{pmatrix} p_{k1}^{d_{k1}} (1 - p_{k1})^{n_{k1} - d_{k1}} \begin{pmatrix} n_{k2} \\ d_{k2} \end{pmatrix} p_{k2}^{n_k - d_{k1}} (1 - p_{k2})^{n_{k2} - d_{k2}}}{\sum_d \begin{pmatrix} n_{k1} \\ d \end{pmatrix} p_{k1}^{d} (1 - p_{k1})^{n_{k1} - d} \begin{pmatrix} n_{k2} \\ d_k - d \end{pmatrix} p_{k2}^{n_k - d} (1 - p_{k2})^{n_{k2} - d_k + d}}.$$

where $\max(0, d_k - n_{k2}) \leq d \leq \min(d_{k1}, n_{k1})$.

*Derivation of the variance formula (continuing 2)

1. Under $H_0 : p_{k1} = p_{k2}$,

$$
\begin{aligned}
P(D_{k1} = d_{k1} | d_k, n_{k1}, n_{k2}) &= \frac{\begin{pmatrix} n_{k1} \\ d_{k1} \end{pmatrix} \begin{pmatrix} n_{k2} \\ d_{k2} \end{pmatrix}}{\sum_d \begin{pmatrix} n_{k1} \\ d \end{pmatrix} \begin{pmatrix} n_{k2} \\ d_k - d \end{pmatrix}} \\
&= \frac{\begin{pmatrix} n_{k1} \\ d_{k1} \end{pmatrix} \begin{pmatrix} n_{k2} \\ d_{k2} \end{pmatrix}}{\begin{pmatrix} n_k \\ d_k \end{pmatrix}}.
\end{aligned}
$$

2. The mean for D_{k1} under H_0 is $n_{k1} d_k / n_k$.
3. The variance for D_{k1} under H_0 is

$$
\frac{n_{k1} n_{k2} d_k (n_k - d_k)}{n_k^2 (n_k - 1)}.
$$

*Derivation of the variance formula (continuing 3)

1. The conditional variance for $D_{k1} - E(D_{k1}|d_k, n_{k1}, n_{k2})$ is the same as D_{k1}.
2. Let $U_k = D_{k1} - n_{k1}d_k/n_k$ and

$$U_w = \sum_k w_k U_k,$$

   where w_k is a weight assigned to U_k.
3. The variance for U_w can be obtained by conditional argument as,

$$V(U_w) = \sum_{k=1}^{K} w_k^2 V(U_k).$$

4. There is a restriction on what types of weights can be used. *The weight w need to be predictable with respect to the natural filtration in order for this formula to be correct.