CHAPTER 3

THE LOG RANK TEST

3.1 Introduction

The log rank test is a nonparametric test widely used in clinical trials to compare the survival distributions of individuals in two groups. These groups may be labeled Group I and Group II.

Suppose

(a) there are r distinct death times

$$t_1 < t_2 < t_3 < \ldots < t_r$$

across the two groups,

(b) n_{ij} individuals are at risk of deaths in Group i just before time $t_{(j)}$,

(c) d_{ij} represents the number of deaths in Group i at time $t_{(j)}$.

Let $d_j = d_{1j} + d_{2j}$ and $n_j = n_{1j} + n_{2j}$. Now consider the null hypothesis

H_0 : there is no difference in the survival experiences of the individuals in the two groups.

Under H_0, d_{1j} has a hypergeometric distribution given by

$$\mathrm{P}\begin{pmatrix}\text{number of individuals in} \\ \text{Group I who die at time} \\ t_j \text{ is } d_{1j}\end{pmatrix} = \frac{\binom{d_j}{d_{1j}}\binom{n_j - d_j}{n_{1j} - d_{1j}}}{\binom{n_j}{n_{1j}}}, \qquad (3.1.1)$$

where

$$\binom{d_j}{d_{1j}} = \frac{d_j!}{d_{1j}!\,(d_j - d_{1j})!}. \qquad (3.1.2)$$

The mean of the hypergeometric random variable d_{1j} in Equation (3.1.1) is given by

$$e_{1j} = E\left(d_{1j}\right) = \frac{n_{1j} d_j}{n_j}. \tag{3.1.3}$$

The symbol e_{1j} can be interpreted as the expected number of Group I individuals who die at time $t_{(j)}$. Hence, the sum of differences between the observed and expected Group I individuals can be written as

$$U_L = \sum_{j=1}^{r} \left(d_{1j} - e_{1j}\right), \tag{3.1.4}$$

and it would tend to be near zero under H_0.

Under H_0, the mean of U_L is

$$E\left(U_L\right) = E\left(\sum_{j=1}^{r} \left(d_{1j} - e_{1j}\right)\right) = 0. \tag{3.1.5}$$

and the variance of d_{1j} is

$$v_{1j} = \frac{n_{1j} n_{2j} d_j \left(n_j - d_j\right)}{n_j^2 \left(n_j - 1\right)}. \tag{3.1.6}$$

Then since $d_{11}, d_{12}, ..., d_{1r}$ are independent, the variance of statistic U_L under H_0 is given by

$$V_L = \text{var}\left(U_L\right) = \sum_{j=1}^{r} v_{1j}. \tag{3.1.7}$$

When the number of death times is not too small, it can be shown that under H_0, U_L has approximately a normal distribution and $U_L / \sqrt{V_L}$ has approximately a standard normal distribution, denoted by $N(0, 1)$. We may write

$$\frac{U_L}{\sqrt{V_L}} \overset{approximately}{\sim} N\left(0, 1\right). \tag{3.1.8}$$

In the log-rank test, we reject H_0 at the α level if

$$\left|\frac{U_L}{\sqrt{V_L}}\right| > z_{\alpha/2} \,. \tag{3.1.9}$$

3.2 The Approximate Distribution of U_L

The third and fourth orders central moments of U_L can be expressed in terms of the noncentral moments of d_{1j} as shown below

$$\begin{aligned} E\left[U_L - E(U_L)\right]^3 &= E(U_L^{\,3}) \\ &= E\left[\sum_{j=1}^{r}(d_{1j} - e_{1j})\right]^3 \\ &= E\left[\sum_{j=1}^{r}(d_{1j} - e_{1j})^3\right] \\ &= \sum_{j=1}^{r}\left\{E(d_{1j}^{\,3}) - 3E(d_{1j}^{\,2})\,e_{1j} + 2e_{1j}^{\,3}\right\}, \end{aligned} \tag{3.2.1}$$

$$\begin{aligned} E\left[U_L - E(U_L)\right]^4 &= E(U_L^{\,4}) \\ &= E\left[\sum_{j=1}^{r}(d_{1j} - e_{1j})\right]^4 \\ &= \sum_{j=1}^{r}\left[E(d_{1j}^{\,4}) - 4E(d_{1j}^{\,3})\,e_{1j} + 6E(d_{1j}^{\,2})\,e_{1j}^{\,2} - 3e_{1j}^{\,4}\right] \\ &\quad + 6E\left[\sum_{j=1}^{r-1}\sum_{k=j+1}^{r}(d_{1j} - e_{1j})^2 (d_{1k} - e_{1k})^2\right]. \end{aligned} \tag{3.2.2}$$

The derivation of these moments of d_{1j} is shown in Appendix A.

Let ε be a random variable which has a quadratic-normal distribution with parameters 0 and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)^T$ (see Equation (2.3.10)). We find $\boldsymbol{\lambda}$ such that

$$E(\varepsilon^k) = E\left[U_L - E(U_L)\right]^k, \qquad k = 2, 3, 4. \tag{3.2.3}$$

Then U_L has approximately a quadratic-normal distribution with parameters E(U_L) and $\boldsymbol{\lambda}$, and we may reject H_0 at the α level if

$$U_L < A_L \text{ or } U_L > A_U , \tag{3.2.4}$$

where

$$A_U = E(U_L) + \lambda_1 z_{\alpha/2} + \lambda_2 \left(\left(z_{\alpha/2} \right)^2 - \frac{1+\lambda_3}{2} \right), \tag{3.2.5}$$

and

$$A_L = E(U_L) + \lambda_1 \left(-z_{\alpha/2} \right) + \lambda_2 \left(\lambda_3 \left(-z_{\alpha/2} \right)^2 - \frac{1+\lambda_3}{2} \right). \tag{3.2.6}$$

3.3 Numerical Results

We perform a simulation in which we set the number of deaths to be r = 10 and d_j =1 for j = 1, 2, …, r. When an individual dies at time t_j the probability (under H_0) that the individual is from Group i (i.e. d_{ij} = 1) is n_{ij}/n_j. This probability enables us to generate the value of $\mathbf{d}_1 = \left(d_{11}, d_{12}, ..., d_{1r} \right)$. Suppose a total of N values of $\mathbf{d}_1$ are generated. For each generated values of $\mathbf{d}_1$, we determine whether H_0 is rejected or not when each of the following two tests is performed:

Test 1: Test based on normal distribution (see Equation (3.1.9)).

Test 2: Test based on quadratic-normal distribution (see Equation (3.2.4)).

Table 3.3.1 shows the rejection probabilities of the test based on quadratic-normal distribution and the test based on normal distribution. We observed from the table that the rejection probability of the test based on quadratic-normal distribution tends to be closer to the target value 0.05 when $n_{11} > n_{21}$. When $n_{11} < n_{21}$, the two tests have comparable rejection probabilities.

Table 3.3.1 Rejection probabilities of the tests based on quadratic-normal and normal distributions (α = 0.05, r = 10, d_j = 1, $1 \le j \le r$, N = 10000).

n_{11}	n_{21}	Rejection Probability	
		Normal Distribution	Quadratic-normal Distribution
50	11	0.128	0.049
49	12	0.064	0.053
48	13	0.091	0.045
47	14	0.106	0.033
46	15	0.114	0.034
45	16	0.072	0.067
44	17	0.066	0.060
43	18	0.079	0.063
42	19	0.086	0.049
41	20	0.108	0.043
40	21	0.049	0.061
39	22	0.042	0.052
38	23	0.080	0.060
37	24	0.096	0.050
36	25	0.097	0.037
35	26	0.039	0.037
34	27	0.025	0.033
33	28	0.040	0.061
32	29	0.059	0.064
31	30	0.083	0.055
30	31	0.065	0.026
29	32	0.030	0.031
28	33	0.029	0.037
27	34	0.035	0.047
26	35	0.053	0.062
25	36	0.071	0.057
24	37	0.094	0.065
23	38	0.024	0.036
22	39	0.023	0.025
21	40	0.034	0.044
20	41	0.050	0.046
19	42	0.047	0.047
18	43	0.047	0.047
17	44	0.049	0.049
16	45	0.073	0.065
15	46	0.036	0.097
14	47	0.022	0.069
13	48	0.017	0.030
12	49	0.013	0.010
11	50	0.036	0.025