Reliability Data and Survival Analysis
Homework 4

Chenghua Liu
liuch18@mails.tsinghua.edu.cn
Department of Computer Science
Tsinghua University

1

An investigator asked you to help design a clinical trial for comparing a new treatment to the standard treatment for patients with some kind of cancer. Suppose the mean survival time of the standard treatment is 3 years and the new treatment is expected to extend the mean survival time to 5 years. For design purpose, let us assume the survival times for each treatment have exponential distribution. We would like to use the log-rank test for testing the survival difference at level $\alpha = 0.01$ and the investigator wants to have 90% power to detect the above difference. Assume equal number of patients will be allocated to each treatment. Do the following:

(a) What is the expected total number of deaths we have to observe in order for the log-rank test to have the desired power to detect the difference we expect?

(b) Suppose accrual rate is constant, and accrual period is 2 years, and follow for an additional 3 years. The patients allocation between new treatment and standard treatment is 2 : 1. How many patients should be enrolled?

(c) Derive the sample size formular if the survival times for each treatment have Weibull distribution.

solution:

(a)

According to the assumption of exponential distribution, the hazard rate is

$$\lambda_0 = \frac{1}{m_0} = \frac{1}{3} \quad \lambda_1 = \frac{1}{m_1} = \frac{1}{5}$$

So the hazard ratio is $\frac{\lambda_1}{\lambda_0} = \frac{3}{5}$ and $\beta_A = \log(\frac{3}{5})$ .The expected total number of death from both treatments must be equal to $(\theta = 1/2, \alpha = 0.01, \gamma = 1 - 0.9 = 0.1)$

$$d = \frac{\left(z_{\alpha/2} + z_\gamma\right)^2}{\left(\beta_A\right)^2 * \theta(1-\theta)} = \frac{4(2.575829 + 1.281552)^2}{\log(3/5)^2} = 228$$

(b)

We have known that accural period $A = 2$ years,an additional $F = 3$ years ,so $L = A + F =$ 5years. Note that the patients allocation between new treatment and standard treatment is $2:1$, the expected number of death would be equal to $D_0 + D_1$,where

$$D_j = \frac{(1+j)n}{3}\left\{1 - \frac{\mathrm{e}^{-\lambda_j L}}{\lambda_j A}\left(\mathrm{e}^{\lambda_j A} - 1\right)\right\}, \quad j = 0, 1$$

For our problem, the specific expected number of deaths is

$$\begin{aligned} D_1 + D_0 &= \frac{2n}{3}\left\{1 - \frac{\mathrm{e}^{-0.2*5}}{0.2*2}\left(\mathrm{e}^{0.2*2} - 1\right)\right\} + \frac{n}{3}\left\{1 - \frac{\mathrm{e}^{-0.33*5}}{0.33*2}\left(\mathrm{e}^{0.33*2} - 1\right)\right\} \\ &= \frac{2n}{3} * 0.548 + \frac{n}{3} * 0.731 = 0.609 * n \end{aligned}$$

Thus if we want the expected number of deaths to equal to $d \quad (\theta = \frac{1}{3})$

$$d = \frac{\left(z_{\alpha/2} + z_\gamma\right)^2}{\left(\beta_A\right)^2 * \theta(1-\theta)} = \frac{9(2.575829 + 1.281552)^2}{2\log(3/5)^2} = 256.5974$$

then

$$0.609 * n = 256.5974 \Longleftrightarrow n = 421.3422 \approx 422$$

So 422 patients should be enrolled.

(c)

Suppose the entry rate follows a uniform distribution in $[0, A]$. That is

$$Q_E(u) = P[E \le u] = \begin{cases} 0 & \text{if } u \le 0 \\ \frac{u}{A} & \text{if } 0 < u \le A \\ 1 & \text{if } u > A \end{cases}$$

Consequently,

$$H_C(u) = Q_E(L-u) = \begin{cases} 1 & \text{if } u \le L - A \\ \frac{L-u}{A} & \text{if } L - A < u \le L \\ 0 & \text{if } u > L \end{cases}$$

Recall for $W(p,\lambda)$, we have $\lambda(t) = p\lambda(\lambda t)^{p-1}$ and $S(t) = e^{-(\lambda t)^p}$ .Hence,

$$
\begin{aligned}
P[\Delta = 1] &= \int_0^L \lambda_T(u) S_T(u) H_C(u) du \\
&= \int_0^{L-A} p\lambda(\lambda u)^{p-1} e^{-(\lambda u)^p} du + \int_{L-A}^{L} p\lambda(\lambda u)^{p-1} e^{-(\lambda u)^p} \frac{L-u}{A} du \\
&= \int_0^{L-A} p\lambda(\lambda u)^{p-1} e^{-(\lambda u)^p} du + \frac{L}{A}\int_{L-A}^{L} p\lambda(\lambda u)^{p-1} e^{-(\lambda u)^p} du - \frac{1}{A}\int_{L-A}^{L} p(\lambda u)^p e^{-(\lambda u)^p} du \\
&= -e^{-(u\lambda)^p}\Big|_0^{L-A} - \frac{L}{A} e^{-(u\lambda)^p}\Big|_{L-A}^{L} + \frac{1}{A\lambda}\text{Gamma}\left(1+\frac{1}{p}, (u\lambda)^p\right)\Big|_{L-A}^{L}
\end{aligned}
$$

Therefore, if we accrue n patients uniformly over A years, who fail according to an Weibull distribution with hazard λ, and follow them for an additional F years, then the expected number of deaths in the sample is

$$
n * \left\{ -e^{-(u\lambda)^p}\Big|_0^{F} - \frac{L}{A} e^{-(u\lambda)^p}\Big|_{F}^{L} + \frac{1}{A\lambda}\text{Gamma}\left(1+\frac{1}{p}, (u\lambda)^p\right)\Big|_{F}^{L} \right\}
$$

2

Suppose we are planning a randomized trial involving a new experimental treatment which the investigators believe will decrease the hazard of liver failure by 20% compared with the standard treatment.

(a) how many subjects will the investigators need to enroll in each arm of the trial in order to achieve 80% power? Assuming the survival time follows exponential distribution. You can make appropriate assumptions about any values you need.

(b) Carry out a simulation to estimate the empirical power of the log-rank test in this case:

  (i) generate n observations for each group from the exponential distribution (i.e., one data set)

  (ii) carry out a log-rank test, where n is the sample size from (a).

  (iii) Repeat steps (i) and (ii) 1000 times (i.e., You will need to generate 1,000 independent data sets). What is the empirical power?

solution:

(a)

According to the assumption of exponential distribution, the hazard ratio is $\frac{\lambda_1}{\lambda_0} = \frac{4}{5}$ and $\beta_A = \log(\frac{4}{5})$ .The expected total number of death from both treatments must be equal to (assume $\theta = 1/2, \alpha = 0.05$, and we know $\gamma = 1 - 0.8 = 0.2$)

$$
d = \frac{\left(z_{\alpha/2} + z_\gamma\right)^2}{\left(\beta_A\right)^2 * \theta(1-\theta)} = \frac{4(1.959964 + 0.8416212)^2}{\log(4/5)^2} = 630.5201
$$

We assume that accural period $A = 2$ years,an additional $F = 3$ years ,so $L = A + F =$ 5years. Note that the patients allocation between new treatment and standard treatment is $1 : 1$, the expected number of death would be equal to $D_0 + D_1$,where

$$D_j = \frac{n}{2}\left\{1 - \frac{\mathrm{e}^{-\lambda_j L}}{\lambda_j A}\left(\mathrm{e}^{\lambda_j A} - 1\right)\right\}, \quad j = 0, 1$$

We assume $\lambda_0 = 0.25$, then $\lambda_1 = 0.2$. The specific expected number of deaths is

$$\begin{aligned} D_1 + D_0 &= \frac{n}{2}\left\{1 - \frac{\mathrm{e}^{-0.25*5}}{0.25*2}\left(\mathrm{e}^{0.25*2} - 1\right)\right\} + \frac{n}{2}\left\{1 - \frac{\mathrm{e}^{-0.2*5}}{0.2*2}\left(\mathrm{e}^{0.2*2} - 1\right)\right\} \\ &= \frac{n}{2} * 0.6282765 + \frac{n}{2} * 0.5476695 = 0.587973 * n \end{aligned}$$

Thus if we want the expected number of deaths to equal d, then

$$0.587973 * n = 630.5201 \Longleftrightarrow n = 1072.362$$

So 1073 patients should be enrolled.

(b)

By reading documentation for Package 'npsurvSS',

(https://cran.r-project.org/web/packages/npsurvSS/npsurvSS.pdf)

We can simulate to estimate the empirical power of the log-rank test easily.

```
install.packages("npsurvSS")
library(npsurvSS)
arm0 <- create_arm(size=537, accr_time=2, surv_shape=1,surv_scale=0.25,loss_shape =1,
loss_scale=0, follow_time=3)
arm1 <- create_arm(size=537, accr_time=2, surv_shape=1,surv_scale=0.2,loss_shape =1,
loss_scale=0, follow_time=3)
power_two_arm(arm0, arm1, list(test="weighted logrank",
                               weight="1",
                               mean.approx="generalized schoenfeld",
                               var.approx="1"))
```

We get the output 0.799542 which is closed to 80%.