Reliability Data and Survival Analysis
Homework 3

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

1

You are given a small data sets on survival times of subjects in two groups: group 1: $2, 3+, 4, 4+$ and group 2 : $3, 5, 5+, 6$, where $+$ means a censored observation. Conduct the standard log-rank test (weight function = 1) by hand to compare the difference in the survival distribution. Which group has better survival?

solution:

Keep writing habits, we use group 0&1 instead of group 1&2. Suppose $H_0 : F_0(\cdot) = F_1(\cdot)$ holds. Then $k = 5$ and $\tau_1, \tau_2, \tau_3, \tau_4, \tau_5 = 2, 3, 4, 5, 6$.The information at time τ_j can be summarized in the following 2x2 tables as the following form:

$Group$	Yes	No	$Total$
0	d_{0j}	$r_{0j} - d_{0j}$	r_{0j}
1	d_{1j}	$r_{1j} - d_{1j}$	r_{1j}
$Total$	d_j	$r_j - d_j$	r_j

$O_j = d_{1j}$	0	1	0	1	1
$E_j = \left(\frac{r_{1j}}{r_j}\right) d_j$	1/2	4/7	3/5	1	1
$V_j = \frac{r_{0j} r_{1j} d_j (r_j - d_j)}{r_j^2 (r_j - 1)}$	1/4	12/49	6/25	0	0

1	3	4
0	4	4
1	7	8

0	3	3
1	3	4
1	6	7

1	1	2
0	3	3
1	4	5

0	0	0
1	2	3
1	2	3

0	0	0
1	0	1
1	0	1

Hence

$$O = 3, \quad E = 3.671429, \quad V = 0.734898 \Rightarrow Z^2 = \frac{(O-E)^2}{V} = 0.6134415$$

According to $Z^2 \overset{apx}{\sim} \chi^2_1$, we get $p - value = 0.433495$.We do not reject H_0 , which means we suppose that two groups have the same survival at significance level 5%.

2

One of the goals of recent research is to explore the efficacy of triple-drug combinations of antiretroviral therapy for treatment of HIVinfected patients. Because of limitations on potency and the continuing emergence of drug resistance seen with the use of currently available antiretroviral agents in monotherapy and two-drug regimens, triplecombination regimens should represent a more promising approach to maximize antiviral activity, maintain long-term efficacy, and reduce the incidence of drug resistance. Towards this end, investigators performed a randomized study comparing AZT+ zalcitabine (ddC) versus AZT + zalcitabine (ddC) + saquinavir. The data, time from administration of treatment (in days) until the CD4 count reached a prespecified level, is given below for the two groups.

AZT + *zalcitabine*(ddC) : $85, 32, 38+, 45, 4+, 84, 49, 180+, 87, 75, 102, 39, 12, 11, 80, 35, 6$

AZT + *zalcitabine*(ddC) + *saquinavir* : $22, 2, 48, 85, 160, 238, 56+, 94+, 51+, 12, 171, 80, 180, 4, 90, 180+, 3$

Use the log rank statistic to test if there is a difference in the distribution of the times at which patient's CD4 reaches the prespecified level for the two treatments.
solution:

```
library(survival)
tt<-c(85, 32, 38, 45, 4, 84, 49, 180, 87, 75, 102, 39, 12, 11, 80, 35, 6,
22, 2, 48, 85, 160, 238, 56, 94, 51, 12, 171, 80, 180, 4, 90, 180, 3)
delta<-c(1,1,0,1,0,1,1,0,1,1,
1,1,1,1,1,1,1,1,1,1,
1,1,1,0,0,0,1,1,1,1,
1,1,0,1)
trt<-rep(c(1,2),each=17)
survdiff(Surv(tt,delta)~trt)
```

```
Call:
survdiff(formula = Surv(tt, delta) ~ trt)
        N Observed Expected (O-E)^2/E (O-E)^2/V
trt=1 17       14     10.5     1.140      2.05
trt=2 17       13     16.5     0.729      2.05

 Chisq= 2  on 1 degrees of freedom, p= 0.2
```

So we conclude that there is a no difference in the distribution of the times at which patient's CD4 reaches the prespecified level for the two treatments at significance level 5%.

3

Consider the kidney infection data from HW3_data. Do the following:

(a) Plot the Kaplan-Meier (KM) estimators of survival curves for the two treatment methods, respectively (in the same figure).

(b) Perform the standard log-rank test, Gehan's Wilcoxon test and Peto- Prentice's Wilcoxon test to determine if there is any difference in survival between the two treatment methods. Which test is more powerful? Could you given some explanations on the power of these tests based on the KM plots obtained in (a)?

solution:
(a)

```
data<-read.table("/Users/liuchenghua/Downloads/HW3_data.txt",header=T)
tt<-as.numeric(unlist(data["month"]))
delta<-as.numeric(unlist(data["status"]))
delta = as.numeric(delta==0)
trt<-rep(c(1,2),each=22)
km_fit<-survfit(Surv(tt,delta)~trt)
library("survminer")
ggsurvplot(
km_fit,    data,
size = 1,# change line size
palette =c("#E7B800", "#2E9FDF"),# custom color palettes
pval = TRUE,# Add p-value
legend.labs =c("new drug", "standard treatment"),# Change legend labels
ggtheme = theme_bw()# Change ggplot2 theme
)
```

We get the figure 1.

(b)

```
> library(survMisc)
> comp(ten(km_fit))$tests$lrTests
                        Q         Var       Z     pNorm
1               5.38276     6.24552  2.1539  0.031250 *
n             204.00000  6396.00000  2.5508  0.010748 *
sqrtN          33.84760   189.84564  2.4566  0.014027 *
S1              4.38630     3.21014  2.4481  0.014360 *
S2              4.28834     3.01786  2.4685  0.013567 *
FH_p=1_q=1      0.86459     0.22080  1.8400  0.065771 .
```

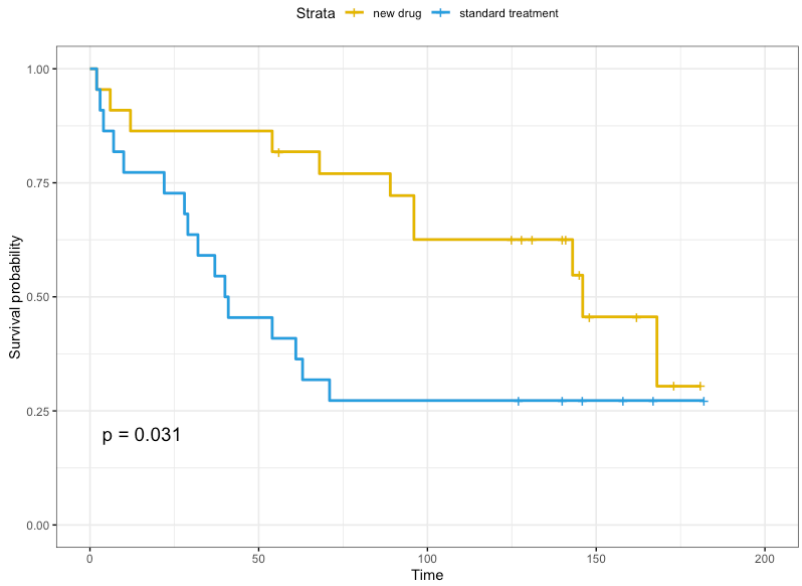

图 1: Kaplan-Meier estimators of survival curves

Through the help document"help(comp)", we find the corresponding value

$$P_{log-rank} = 0.031250$$
$$P_{Gehan'sWilcoxon} = 0.010748$$
$$P_{Peto-Prentice'sWilcoxon} = 0.014360$$

Gehan's Wilcoxon test and Peto- Prentice's Wilcoxon test are more powerful in this case.Note that both tests place more emphasis on earlier failure times compared to the log-rank test.We can observe the significant difference in earlier failure times from the figure1,which explains why they are more powerful in this case.

4

A study was performed to determine the efficacy of boron neutron capture therapy (BNCT) in treating the therapeutically refractory F98 glioma, using boronophenylalanine (BPA) as the capture agent. F98 glioma cells were implanted into the brains of rats. Three groups of rats were studied. One group went untreated, another was treated only with radiation, and the third group received radiation plus an appropriate concentration of BPA. The data for the three groups lists the death times (in days) and is given below (+Censored observation):

Untreated	Radiated	Radiated + BPA
20	26	31
21	28	32
23	29	34
24	29	35
24	30	36
26	30	38
26	31	38
27	31	39
28	32	42^+
30	35^+	42^+

(a) Compare the survival curves for the three groups.

(b) Perform pairwise tests to determine if there is any difference in survival between pairs of groups.

(c) There is a priori evidence that, if there is a difference in survival, there should be a natural ordering, namely, untreated animals will have the worst survival, radiated rats will have slightly improved survival, and the radiated rats + BPA should have the best survival. Perform the test for trend which would test this ordered hypothesis.

solution:

(a)

```
library("ggpubr")
library("survminer")
tt<-c(20,21,23,24,24,26,26,27,28,30,
26,28,29,29,30,30,31,31,32,35,
31,32,34,35,36,38,38,39,42,42)
delta<-c(1,1,1,1,1,1,1,1,1,1,
1,1,1,1,1,1,1,1,1,0,
1,1,1,1,1,1,1,1,0,0)
trt<-rep(c(1,2,3),each=10)
km_fit<-survfit(Surv(tt,delta)~trt)
ggsurvplot(
    km_fit,    data,
    size = 1,# change line size
    palette =c("#E7B800", "#2E9FDF", "#FFB6C1"),# custom color palettes
    pval = TRUE,# Add p-value
    legend.labs =c("Untreated", "Radiated","Radiated + BPA"),# Change legend labels
    ggtheme = theme_bw()# Change ggplot2 theme
)
survdiff(Surv(tt,delta)~trt)
```

```
Call:
survdiff(formula = Surv(tt, delta) ~ trt)
        N Observed Expected (O-E)^2/E (O-E)^2/V
trt=1 10       10     2.66    20.300    27.365
trt=2 10        9     7.60     0.258     0.431
trt=3 10        8    16.74     4.566    16.587

 Chisq= 33.4  on 2 degrees of freedom, p= 6e-08
```

We get the figure 2 and conclude that there is a a big difference in the three survival curves at significance level 5%.

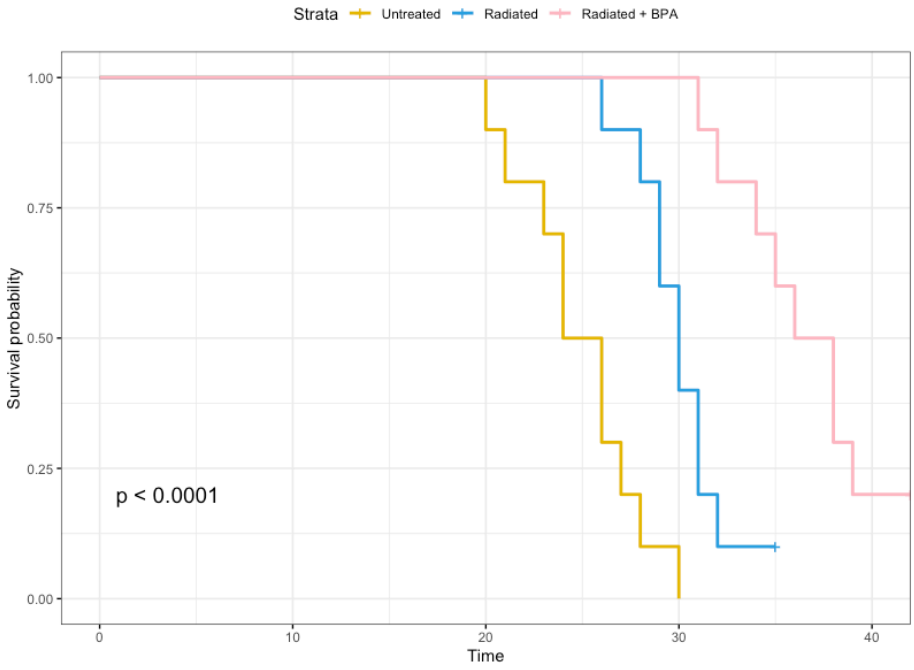

图 2: Kaplan-Meier estimators of survival curves

(b)

```
d=cbind(tt,delta,trt)
d<-as.data.frame(d)
library(survminer)
pairwise_survdiff(Surv(tt,delta)~trt,data=d)
```

```
> pairwise_survdiff(Surv(tt,delta)~trt,data=d)

        Pairwise comparisons using Log-Rank test

data:  d and trt

  1       2
```

```
8  2 0.0011  -
9  3 9.7e-06 0.0014
10
11 P value adjustment method: BH
```

The Bonferroni multiple comparison procedure would test at the $0.05/3 = 0.0167$ level.We conclude that all pairwise comparisons are significant at significance level 5%.

(c)

```
1 > surv_pvalue(fit,data = d, method = "survdiff",test.for.trend = TRUE)
2  variable pval        method    pval.txt
3  1        trt     0 Log-rank, tft p < 0.0001
```

So we conclude that there exist trend.

5

Consider the ovarian data from Survival library. Test the hypothesis that there is no difference in the survival times between the treatment groups using a log- rank statistic stratified on ECOG performance status.

https://stat.ethz.ch/R-manual/R-devel/library/survival/html/ovarian.html

solution:

```
1  > library(survival)
2  > survdiff(Surv(futime, fustat)~rx+strata(ecog.ps), ovarian)
3  Call:
4  survdiff(formula = Surv(futime, fustat) ~ rx + strata(ecog.ps),
5      data = ovarian)
6
7         N Observed Expected (O-E)^2/E (O-E)^2/V
8  rx=1 13        7      5.5     0.409     0.768
9  rx=2 13        5      6.5     0.346     0.768
10
11  Chisq= 0.8  on 1 degrees of freedom, p= 0.4
```

Not significant, hence we can't reject null hypothesis.