Survival Analysis HW 3

Victor Feagins

10/1/2021

### Question 5.8

Using the data from Table 3.1 compare the age distribution of the two groups

Table :First 10 Observations of Melanoma Patients

| PATIENT | AGE | SEX | TYPE2A | TYPE3A | TYPE3B | TYPE4A | TREATMENT | REMTIME | RELAPSE | LIFETIME | DEATH |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 59 | 0 | 0 | 0 | 1 | 0 | 1 | 33.7 | 0 | 33.7 | 0 |
| 2 | 50 | 0 | 0 | 0 | 1 | 0 | 1 | 3.8 | 1 | 3.9 | 1 |
| 3 | 76 | 1 | 0 | 0 | 1 | 0 | 1 | 6.3 | 1 | 10.5 | 1 |
| 4 | 66 | 0 | 0 | 0 | 1 | 0 | 1 | 2.3 | 1 | 5.4 | 1 |
| 5 | 33 | 1 | 0 | 0 | 1 | 0 | 1 | 6.4 | 1 | 19.5 | 1 |
| 6 | 23 | 0 | 0 | 0 | 1 | 0 | 1 | 23.8 | 0 | 23.8 | 0 |
| 7 | 40 | 0 | 0 | 0 | 1 | 0 | 1 | 1.8 | 1 | 7.9 | 1 |
| 8 | 34 | 1 | 0 | 0 | 1 | 0 | 1 | 5.5 | 1 | 16.9 | 0 |
| 9 | 34 | 1 | 0 | 0 | 1 | 0 | 1 | 16.6 | 0 | 16.6 | 0 |
| 10 | 38 | 0 | 1 | 0 | 0 | 0 | 1 | 33.7 | 0 | 33.7 | 0 |

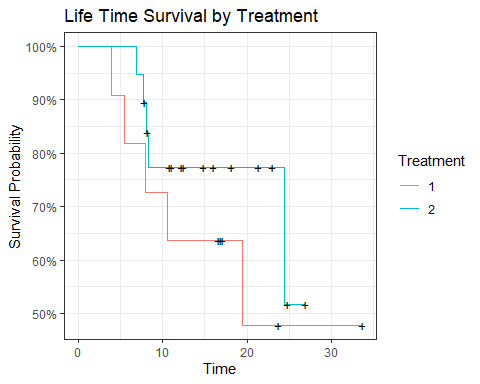
#### Lifetime

In this problem we will be looking at lifetime of a patient comparing treatments.

surv.data.5.8 <- Surv(df.5.8$LIFETIME, df.5.8$DEATH)  
km.5.8 <- survfit(surv.data.5.8 ~ TREATMENT, data = df.5.8)  
summary(km.5.8)

## Call: survfit(formula = surv.data.5.8 ~ TREATMENT, data = df.5.8)  
##   
## TREATMENT=1   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 3.9 11 1 0.909 0.0867 0.754 1.000  
## 5.4 10 1 0.818 0.1163 0.619 1.000  
## 7.9 9 1 0.727 0.1343 0.506 1.000  
## 10.5 8 1 0.636 0.1450 0.407 0.995  
## 19.5 4 1 0.477 0.1755 0.232 0.981  
##   
## TREATMENT=2   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 6.9 19 1 0.947 0.0512 0.852 1.000  
## 7.7 18 1 0.895 0.0704 0.767 1.000  
## 8.0 16 1 0.839 0.0854 0.687 1.000  
## 8.3 13 1 0.774 0.1003 0.601 0.998  
## 24.4 3 1 0.516 0.2211 0.223 1.000

autoplot(km.5.8, conf.int = F) +  
 labs(title = "Life Time Survival by Treatment",  
 x = "Time",  
 y = "Survival Probability",  
 colour = "Treatment")



Testing to see if the groups are different?

survdiff(surv.data.5.8 ~ TREATMENT, data = df.5.8)

## Call:  
## survdiff(formula = surv.data.5.8 ~ TREATMENT, data = df.5.8)  
##   
## N Observed Expected (O-E)^2/E (O-E)^2/V  
## TREATMENT=1 11 5 3.68 0.469 0.747  
## TREATMENT=2 19 5 6.32 0.274 0.747  
##   
## Chisq= 0.7 on 1 degrees of freedom, p= 0.4

When looking at the log rank test we have a p value of .4 which means we fail to reject the null that the two treatments result in different lifetime survival times.

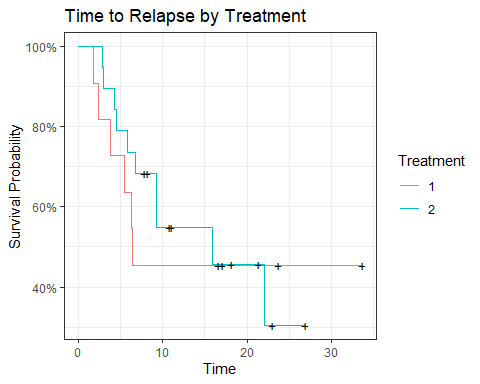
#### Remission Time

We did not see a significant change in the lifetimes between treatments but perhaps we can see a difference in the remission times.

surv.data.5.8 <- Surv(df.5.8$REMTIME, df.5.8$RELAPSE)  
km.5.8 <- survfit(surv.data.5.8 ~ TREATMENT, data = df.5.8)  
summary(km.5.8)

## Call: survfit(formula = surv.data.5.8 ~ TREATMENT, data = df.5.8)  
##   
## TREATMENT=1   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 1.8 11 1 0.909 0.0867 0.754 1.000  
## 2.3 10 1 0.818 0.1163 0.619 1.000  
## 3.8 9 1 0.727 0.1343 0.506 1.000  
## 5.5 8 1 0.636 0.1450 0.407 0.995  
## 6.3 7 1 0.545 0.1501 0.318 0.936  
## 6.4 6 1 0.455 0.1501 0.238 0.868  
##   
## TREATMENT=2   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 2.8 19 1 0.947 0.0512 0.852 1.000  
## 3.0 18 1 0.895 0.0704 0.767 1.000  
## 4.3 17 1 0.842 0.0837 0.693 1.000  
## 4.5 16 1 0.789 0.0935 0.626 0.996  
## 5.8 15 1 0.737 0.1010 0.563 0.964  
## 6.8 14 1 0.684 0.1066 0.504 0.929  
## 9.2 10 2 0.547 0.1215 0.354 0.846  
## 15.9 6 1 0.456 0.1311 0.260 0.801  
## 22.1 3 1 0.304 0.1518 0.114 0.809

autoplot(km.5.8, conf.int = F) +  
 labs(title = "Time to Relapse by Treatment",  
 x = "Time",  
 y = "Survival Probability",  
 colour = "Treatment")



survdiff(surv.data.5.8 ~ TREATMENT, data = df.5.8)

## Call:  
## survdiff(formula = surv.data.5.8 ~ TREATMENT, data = df.5.8)  
##   
## N Observed Expected (O-E)^2/E (O-E)^2/V  
## TREATMENT=1 11 6 5.55 0.0366 0.0573  
## TREATMENT=2 19 10 10.45 0.0194 0.0573  
##   
## Chisq= 0.1 on 1 degrees of freedom, p= 0.8

In time to relapse we also don’t see a significant difference in time to relapse with a p value of .8.

### Question 5.14

Sixty-six patients are separated into three treatment groups. The initial remission times are given in weeks. Are the three treatment groups equally effective in prolonging remission? Discuss your findings!

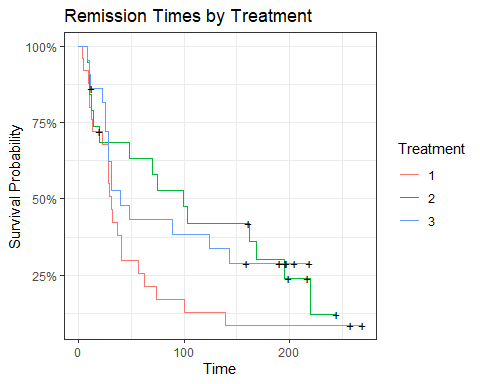
Table :9 Observations of Remission Times

| treat | n | time | cens |
| --- | --- | --- | --- |
| 1 | 25 | 28 | 1 |
| 1 | 25 | 13 | 1 |
| 1 | 25 | 41 | 1 |
| 2 | 19 | 169 | 1 |
| 2 | 19 | 70 | 1 |
| 2 | 19 | 199 | 0 |
| 3 | 22 | 48 | 1 |
| 3 | 22 | 196 | 0 |
| 3 | 22 | 40 | 1 |

surv.data.5.14 <- Surv(df.5.14$time, abs(df.5.14$cens))  
km.5.14 <- survfit(surv.data.5.14 ~ treat, data = df.5.14)  
summary(km.5.14)

## Call: survfit(formula = surv.data.5.14 ~ treat, data = df.5.14)  
##   
## treat=1   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 4 25 1 0.9600 0.0392 0.8862 1.000  
## 5 24 1 0.9200 0.0543 0.8196 1.000  
## 9 23 1 0.8800 0.0650 0.7614 1.000  
## 10 22 2 0.8000 0.0800 0.6576 0.973  
## 12 20 1 0.7600 0.0854 0.6097 0.947  
## 13 19 1 0.7200 0.0898 0.5639 0.919  
## 23 17 1 0.6776 0.0940 0.5164 0.889  
## 28 16 3 0.5506 0.1010 0.3843 0.789  
## 29 13 1 0.5082 0.1017 0.3433 0.752  
## 31 12 1 0.4659 0.1017 0.3037 0.715  
## 32 11 1 0.4235 0.1009 0.2655 0.676  
## 37 10 1 0.3812 0.0993 0.2288 0.635  
## 41 9 2 0.2965 0.0936 0.1597 0.550  
## 57 7 1 0.2541 0.0893 0.1277 0.506  
## 62 6 1 0.2118 0.0838 0.0975 0.460  
## 74 5 1 0.1694 0.0770 0.0695 0.413  
## 100 4 1 0.1271 0.0684 0.0442 0.365  
## 139 3 1 0.0847 0.0572 0.0225 0.319  
##   
## treat=2   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 8 19 1 0.947 0.0512 0.8521 1.000  
## 10 18 2 0.842 0.0837 0.6931 1.000  
## 12 16 1 0.789 0.0935 0.6259 0.996  
## 14 15 1 0.737 0.1010 0.5632 0.964  
## 20 14 1 0.684 0.1066 0.5041 0.929  
## 48 13 1 0.632 0.1107 0.4480 0.890  
## 70 12 1 0.579 0.1133 0.3946 0.850  
## 75 11 1 0.526 0.1145 0.3435 0.806  
## 99 10 1 0.474 0.1145 0.2949 0.761  
## 103 9 1 0.421 0.1133 0.2485 0.713  
## 162 7 1 0.361 0.1119 0.1965 0.663  
## 169 6 1 0.301 0.1082 0.1486 0.609  
## 195 5 1 0.241 0.1019 0.1049 0.552  
## 220 2 1 0.120 0.0992 0.0239 0.605  
##   
## treat=3   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 8 22 1 0.955 0.0444 0.871 1.000  
## 10 21 1 0.909 0.0613 0.797 1.000  
## 11 20 1 0.864 0.0732 0.732 1.000  
## 23 18 1 0.816 0.0834 0.668 0.997  
## 25 17 2 0.720 0.0973 0.552 0.938  
## 28 15 2 0.624 0.1054 0.448 0.869  
## 31 13 2 0.528 0.1088 0.352 0.791  
## 40 11 1 0.480 0.1090 0.307 0.749  
## 48 10 1 0.432 0.1082 0.264 0.705  
## 89 9 1 0.384 0.1062 0.223 0.660  
## 124 8 1 0.336 0.1032 0.184 0.613  
## 143 7 1 0.288 0.0990 0.147 0.565

autoplot(km.5.14, conf.int = F) +  
 labs(title = "Remission Times by Treatment",  
 x = "Time",  
 y = "Survival Probability",  
 colour = "Treatment")



survdiff(surv.data.5.14 ~ treat, data = df.5.14)

## Call:  
## survdiff(formula = surv.data.5.14 ~ treat, data = df.5.14)  
##   
## N Observed Expected (O-E)^2/E (O-E)^2/V  
## treat=1 25 22 15.4 2.865 4.335  
## treat=2 19 15 18.7 0.730 1.195  
## treat=3 22 15 17.9 0.482 0.764  
##   
## Chisq= 4.3 on 2 degrees of freedom, p= 0.1

Over all treatments we see a significance difference in at least one treatment.

survminer::pairwise\_survdiff(Surv(time, cens) ~ treat, data = df.5.14, p.adjust.method = "bonferroni")

##   
## Pairwise comparisons using Log-Rank test   
##   
## data: df.5.14 and treat   
##   
## 1 2   
## 2 0.16 -   
## 3 0.27 1.00  
##   
## P value adjustment method: bonferroni

When doing pairwise comparisons we see that treatment 1 is the most different which we could have inferred from the graph but none of the treatments are significantly different from each other. We will need more information in treatment 1 study to possibly arrive at statistical significance.

survminer::pairwise\_survdiff(Surv(time, cens) ~ treat, data = df.5.14, p.adjust.method = "none")

##   
## Pairwise comparisons using Log-Rank test   
##   
## data: df.5.14 and treat   
##   
## 1 2   
## 2 0.055 -   
## 3 0.091 0.868  
##   
## P value adjustment method: none

It is important to recognize that without adjusting for the p.value we do see a significant difference showing the importance for adjusting and possible reasoning to continue examining treatment 1.