Homework 3

Denis Ostroushko

Problem 1

(i)

Figure 1 shows the two curves for two groups. It appears that the two curves are somewhat different, but there is a lot of overlap in the confidence bands for two groups.

```
Call:
survdiff(formula = Surv(time, delta) ~ type, data = tongue, rho = 0)

        N Observed Expected (O-E)^2/E (O-E)^2/V
type=1 52       31     36.6     0.843      2.79
type=2 28       22     16.4     1.873      2.79

 Chisq= 2.8  on 1 degrees of freedom, p= 0.09
```

- Null hypothesis: There is no difference in survival between the two groups being compared.
- Alternative hypothesis: There is a difference in survival between the two groups being compared.
- Test statistic: 2.8
- Use χ^2_1 random variable and its 95th percentile to find critical value. Critical value: 3.841
- Conclusion: Survival rates for Aneuploid tumor type appear to be higher than the Diploid group. P-value suggests that there may be a true difference between the two groups, however, the difference is not statistically significant at the 95% significance level. We do not reject the null hypothesis due to a lack of statistical evidence.

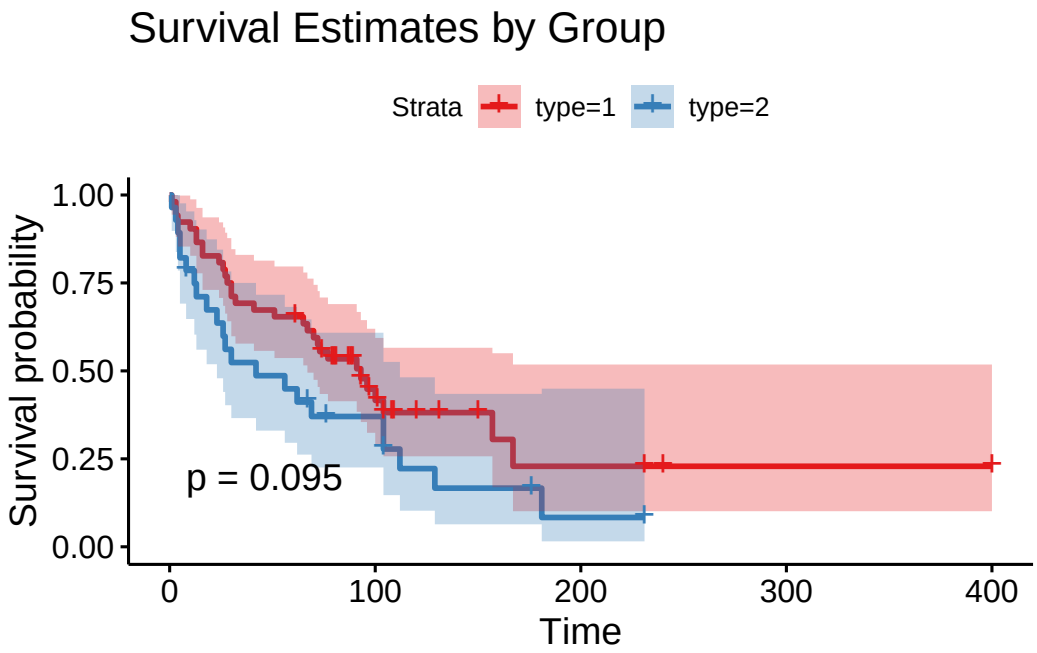

Figure 1: Time to tongue cancer between two types of tumor groups

(ii)

I am using a function below to conduct desired test:

```
FHtestrcp(Surv(time, delta) ~ type, data = tongue, type = "two-sided", alpha = 0.05)
```

```
    Two-sample test for right-censored data

Parameters: rho=0, lambda=0
Distribution: permutation approach (asymptotic approximation using central limit theorem)

Data: Surv(time, delta) by type

        N   O-E
type=1 52 -5.55
type=2 28  5.55

Statistic Z= 1.6, p-value= 0.104
Alternative hypothesis: survival functions not equal
```

- Null hypothesis: There is no difference in survival between the two groups being compared.
- Alternative hypothesis: There is a difference in survival between the two groups being compared.
- Test statistic: 1.6; p-value: 0.104
- Conclusion: urvival rates for Aneuploid tumor type appear to be higher than the Diploid group. P-value suggests that there may be a true difference between the two groups, however, the difference is not statistically significant at the 95% significance level. We do not reject the null hypothesis due to a lack of statistical evidence.

(iii)

- 56-Week restricted mean survival time for Aneuploid Tumor group: 43.69
- 56-Week restricted mean survival time for Diploid Tumor group: 35.31
- Null Hypothesis: $\hat{\mu}_A = \hat{\mu}_D$
- Alternative Hypothesis: $\hat{\mu}_A \neq \hat{\mu}_D$
- Test statistic calculated using code below:

```
fit_1 <- survfit(Surv(time, delta) ~1 ,data = tongue %>% filter(type == 1))

s.km.fit = summary(fit_1, rmean=56)

fit_2 <- survfit(Surv(time, delta) ~1 ,data = tongue %>% filter(type == 2))

s.km.fit2 = summary(fit_2, rmean=56)


test_stat <- abs(s.km.fit2$table["rmean"] - s.km.fit$table["rmean"] - 0)/
  sqrt(s.km.fit$table["se(rmean)"] ^ 2 * length(which(tongue$type == 1)) / nrow(tongue)
       + s.km.fit2$table["se(rmean)"] ^ 2* length(which(tongue$type == 2)) / nrow(tongue))

# qt(0.975, df = nrow(tongue) - 2)
```

- Test statistic: 2.586; P-value: 0.00578
- Conclusion: small p-value gives us enough statistical evidence to reject null hypothesis. Therefore, we can conclude that the 56-week restricted mean survival time for people with Aneuploid tumor is consistently higher than that of people with Diploid tumor.

Problem 2

(i)

- Null hypothesis: all four survial rates are the same, i.e. there is no difference between the four groups.
- Althernative hypothesis: some groups haev different survival rates
- Test output:

```
Call:
survdiff(formula = Surv(time, delta) ~ stage, data = larynx,
    rho = 0)

         N Observed Expected (O-E)^2/E (O-E)^2/V
stage=1 33       15    22.57     2.537     4.741
stage=2 17        7    10.01     0.906     1.152
stage=3 27       17    14.08     0.603     0.856
stage=4 13       11     3.34    17.590    19.827

 Chisq= 22.8  on 3 degrees of freedom, p= 5e-05
```

- Test - statistic: 22.8; p-value < 0.0005.
- Conclusion: we have enough evidence to conclude that some droups have differing survival rates. Therefore, we may proceed to conduct pairwise tests and visual exploration of the data to report which groups have higher survival rates for the purpose of statistical inference.

Comparison with the book results

The book reports results using Z-value, and a standard normal underlying distribution for the hypothesis testing. Therefore, test statistic is different from the one we saw in `R` output. `R` output using Chi-square statistic with 3 degrees of freedom.

However, both test produce extremely small p-values, therefore, we arrive at the same conclusion of rejecting null hypothesis.

Problem 3

(i)

- Null hypothesis: There is no difference in survival rates between the 4 groups after stratifying for diagnosis year
- Alternative hypothesis: There is a difference in survival rates between the 4 groups after stratifying for diagnosis year
- Test output:

```
Call:
survdiff(formula = Surv(time, delta) ~ stage + strata(strat),
    data = larynx, rho = 0)

         N Observed Expected (O-E)^2/E (O-E)^2/V
stage=1 33       15    22.19     2.327     4.376
stage=2 17        7    10.25     1.030     1.338
stage=3 27       17    14.12     0.587     0.835
stage=4 13       11     3.44    16.580    19.421

 Chisq= 22.3  on 3 degrees of freedom, p= 6e-05
```

- Test statistic: chi-square = 22.3 on 3 degrees of freedom; p-value < 0.0001
- Conclusion: we reject null hypothesis and conclude that the survival rates for varying stage groups are different, after adjusting for the diagnosis year.

(ii)

Test 1

- Null hypothesis: survival rates for 4 stages are not different for people diagnosed prior to 1975
- Alternative hypothsis: some survival rates are different for patients diagnosed prior to 1975
- Test output:

```
Call:
survdiff(formula = Surv(time, delta) ~ stage, data = larynx %>%
    filter(strat == "Prior to 1975"), rho = 0)

         N Observed Expected (O-E)^2/E (O-E)^2/V
stage=1 21       13   16.474     0.733     1.487
stage=2  7        4    5.802     0.560     0.693
stage=3 16       12    9.769     0.509     0.737
stage=4  4        4    0.954     9.728    10.402

 Chisq= 12  on 3 degrees of freedom, p= 0.007
```

- Chi-square test statistic with 3 degrees of freedom: 12; p-value: 0.007
- Conclusion: we reject the null hypothesis, some survival rates are different for varying stages for patients diagnosed prior to 1975.

Test 2

- Null hypothesis: survival rates for 4 stages are not different for people diagnosed in 1975 and later
- Alternative hypothsis: some survival rates are different for patients diagnosed in 1975 and later

```
Call:
survdiff(formula = Surv(time, delta) ~ stage, data = larynx %>%
    filter(strat != "Prior to 1975"), rho = 0)

         N Observed Expected (O-E)^2/E (O-E)^2/V
stage=1 12        2     5.71    2.4114     3.743
stage=2 10        3     4.45    0.4706     0.653
stage=3 11        5     4.35    0.0963     0.132
stage=4  9        7     2.49    8.1691     9.933

 Chisq= 11.6  on 3 degrees of freedom, p= 0.009
```

- Chi-square test statistic with 3 degrees of freedom: 11.6; p-value: 0.009
- Conclusion: we reject the null hypothesis, some survival rates are different for varying stages for patients diagnosed in 1975 and later.

Problem 4

(i)

- Null hypothesis: survival rates for 4 combinations of race and gender are not different
- Alternative hypothesis: some survival rates for 4 combinations of race and gender are not different
- Test output:

```
Call:
survdiff(formula = Surv(time, delta) ~ race + gender, data = kidtran)

                   N Observed Expected (O-E)^2/E (O-E)^2/V
race=1, gender=1 432       73    69.25    0.2025    0.4013
race=1, gender=2 280       39    47.39    1.4860    2.2531
race=2, gender=1  92       14    14.52    0.0184    0.0205
race=2, gender=2  59       14     8.84    3.0173    3.2245

 Chisq= 4.7  on 3 degrees of freedom, p= 0.2
```

- P-value: 0.2
- Conclusion: p-value is greater than the accepted significance level with $\alpha = 0.05$, therefore we can not reject null hypothesis. We do not have enough evidence to conclude that the survival rates for the four combinations are different.

(ii)

Males subset

- Null hypothesis: survival rates are not different between races for males.
- Alternative hypothesis: survival rates are different between races for males.
- Test output:

```
Call:
survdiff(formula = Surv(time, delta) ~ race, data = kidtran %>%
    filter(gender == 1))

         N Observed Expected (O-E)^2/E (O-E)^2/V
race=1 432       73     71.9    0.0168     0.097
```

```
race=2  92       14     15.1    0.0801     0.097

 Chisq= 0.1  on 1 degrees of freedom, p= 0.8
```

- Conclusion: chi-square statistic is small and p-value is large. Therefore, we can not reject null hypothesis. So, there is no strong evidence that te survival rates for two races are different for the sub-population of males.

Females subset

- Null hypothesis: survival rates are not different between races for females
- Alternative hypothesis: survival rates are different between races for females
- Test output:

```
Call:
survdiff(formula = Surv(time, delta) ~ race, data = kidtran %>%
    filter(gender == 2))

         N Observed Expected (O-E)^2/E (O-E)^2/V
race=1 280       39    44.79     0.748      4.85
race=2  59       14     8.21     4.076      4.85

 Chisq= 4.8  on 1 degrees of freedom, p= 0.03
```

- Conclusion: we reject the null hypothesis and conclude that the survival rates differ between races when considering the population of females.

(iii)

- Null hypothesis: There is no difference in survival rates between races after stratifying for gender stratification
- Alternative hypothesis: There is a difference in survival rates between races after stratifying for gender stratification
- Test output:

```
Call:
survdiff(formula = Surv(time, delta) ~ race + strata(gender),
    data = kidtran)

         N Observed Expected (O-E)^2/E (O-E)^2/V
```

```
race=1 712      112    116.7     0.188      1.13
race=2 151       28     23.3     0.942      1.13

 Chisq= 1.1  on 1 degrees of freedom, p= 0.3
```

- Conclusion: there are no differences in survival rates between two races when adjusting with stratification by gender. This result is contradictory to the previous part. This result highlights the importance of data exploration and familiarity with sample and effect sizes to detect true effects of covariates.

Problem 5

(i)

- Null hypothesis: There is no difference in survival rates between risk groups
- Alternative hypothesis: There is some difference in survival rates between risk groups
- Test output:

```
Call:
survdiff(formula = Surv(t1, d1) ~ group, data = bmt)

         N Observed Expected (O-E)^2/E (O-E)^2/V
group=1 38       24     21.3     0.337     0.462
group=2 54       23     38.8     6.413    12.562
group=3 45       34     20.9     8.190    11.162

 Chisq= 15.2  on 2 degrees of freedom, p= 5e-04
```

- Conclusion: High test statistic value and low p-value suggest that the survival rates are different for some risk groups.

(ii)

- Null hypothesis: There is no difference in survival rates between risk groups when stratifying by hospital site
- Alternative hypothesis: There is some difference in survival rates between risk groups when stratifying by hospital site
- Test output:

```
Call:
survdiff(formula = Surv(t1, d1) ~ group + strata(z9), data = bmt)

         N Observed Expected (O-E)^2/E (O-E)^2/V
group=1 38       24     26.4     0.211     0.349
group=2 54       23     33.4     3.239     6.127
group=3 45       34     21.2     7.664    10.584

 Chisq= 11.4  on 2 degrees of freedom, p= 0.003
```

- Conclusion: High test statistic value and low p-value suggest that the survival rates are different for some risk groups when accounting for stratification by hospital site.

(iii)

P-values for pairwise comparisons of survival rates are given below:

```r
library(emmeans)

pairwise_survdiff(Surv(t1, d1) ~ group, data = bmt, p.adjust.method = "bonferroni")
```

```
    Pairwise comparisons using Log-Rank test

data:  bmt and group

  1       2
2 0.04559 -
3 0.41420 0.00044

P value adjustment method: bonferroni
```

- Labels: `1` = ALL; `2` = AML Low Risk; `3` = AML High Risk
- It apprears that All-group survival rate is statistically significantly different from AML Low Risk-group survival rate, after adjuting for multiple comparisons.
- It apprears that AML Low Risk-group survival rate is statistically significantly different from AAML High Risk- group survival rate, after adjuting for multiple comparisons.