Name: CHAN King Yeung
SID:1155119394
STAT4008 Assignment 3

Question 1

```
> d = sum(c(1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1))
> H = 0.2 * sum(c(12, 15, 17, 17, 18, 19, 20, 20, 20, 21, 24, 27))
> (ts = (d - H) ^ 2 / H)
[1] 31.3913
> (cv = round(qchisq(0.95, 1), 4))
[1] 3.8415
```

$H_0: h(t) = 0.2 \; vs \; H_1: h(t) \neq 0.2 \; \forall \; t \in (0, 27)$

With $\sum d_i = 8$ and $\sum H_0(T_i) = 0.2 \sum T_i = 46$, we have a test statistic $Z(\tau) = \frac{(8-46)^2}{46} \approx 31.3913$ which is greater than the critical value $\chi^2_{0.95}(1) = 3.8415$, thus, we reject H_0 at 5% level of significance.

Question 2

```
> time1 = c(2, 2, 3, 3, 4, 4, 5, 5, 6)
> time2 = c(2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 7)
> cen1 = c(1, 1, 0, 1, 1, 0, 1, 1, 0)
> cen2 = c(1, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1)
> data = data.frame(time = c(time1, time2), censor = c(cen1, cen2), group =
c(rep(1, length(time1)), rep(2, length(time2))))
>
> survGroup = survdiff(Surv(data$time, data$censor) ~ data$group)
> (ts = round(survGroup$chisq, 4))
[1] 0.3357
> (cv = round(qchisq(0.95, 1), 4))
[1] 3.8415
```

$H_0: h_1(t) = h_2(t)\ vs\ H_1: h_1(t) \neq h_2(t)\ \forall\ t$

From the result given by above code, we notice the test statistic (0.3357) which is less than critical value (3.8415), thus, we do not reject H_0 at 5% level of significance.

Question 3

```
> time = c(18, 9, 28, 31, 39, 19, 45, 6, 8, 15, 23, 28, 7, 12, 9, 8, 2, 26, 10,
4, 3, 4, 18, 8, 3, 14, 3, 13, 13, 35)
> censor = c(1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 0)
> group = c(1, 3, 1, 3, 3, 3, 3, 3, 3, 3, 1, 1, 3, 2, 2, 2, 4, 2, 4, 2, 2, 2,
4, 4, 4, 4, 2, 4, 4, 2)
> data = data.frame(time, censor, group)
>
> survGroup = survdiff(Surv(data$time, data$censor) ~ data$group)
> (ts = round(survGroup$chisq, 4))
[1] 5.6891
> (cv = round(qchisq(0.95, 3), 4))
[1] 7.8147
```

There are 4 combinations of prognostic factor in total, hence, we refine individuals into their corresponding group as

Group 1: $x_1 = 0; x_2 = 0$
Group 2: $x_1 = 1; x_2 = 0$
Group 3: $x_1 = 0; x_2 = 1$
Group 4: $x_1 = 1; x_2 = 1$

$H_0: h_i(t) = h_j(t)\ vs\ H_1: h_i(t) \neq h_j(t)\ \forall\ t\ and\ i \neq j$

From the result given by above code, we notice the test statistic (5.6891) which is less than critical value (7.8147), thus, we do not reject H_0 at 5% level of significance.

Question 4

```
> library("readxl")
> data = as.data.frame(read_excel('ass3data.xls'))
```

a)

```
> survGroup = survdiff(Surv(data$Time) ~ data$Group)
> (ts = round(survGroup$chisq, 4))
[1] 6.4561
> (cv = round(qchisq(0.95, length(unique(data$Group)) - 1), 4))
[1] 5.9915
```

$H_0: h_i(t) = h_j(t)\ vs\ H_1: h_i(t) \neq h_j(t)\ \forall\ t\ and\ i \neq j$

From the result given by above code, we notice the test statistic (6.4561) which is greater than critical value (5.9915), thus, we reject H_0 at 5% level of significance.

b)

```
> survTreat = survdiff(Surv(data$Time) ~ data$Treatment)
> (ts = round(survTreat$chisq, 4))
[1] 0.7369
> (cv = round(qchisq(0.95, length(unique(data$Treatment)) - 1), 4))
[1] 5.9915
```

$H_0: h_i(t) = h_j(t)\ vs\ H_1: h_i(t) \neq h_j(t)\ \forall\ t\ and\ i \neq j$

From the result given by above code, we notice the test statistic (0.7369) which is less than critical value (5.9915), thus, we do not reject H_0 at 5% level of significance.

Question 5

```
> data = read.csv('ass3q5.csv')
> set.seed(123457)
> data = data[sample(nrow(data), 100), ]
>
> survStatus = survdiff(Surv(data$Time, data$Status) ~ data$Smoking.Status)
> (ts = round(survStatus$chisq, 4))
[1] 4.2712
> (cv = round(qchisq(0.95, length(unique(data$Smoking.Status)) - 1), 4))
[1] 5.9915
```

$H_0: h_i(t) = h_j(t)\ vs\ H_1: h_i(t) \neq h_j(t)\ \forall\ t\ and\ i \neq j$

From the result given by above code, we notice the test statistic (4.2712) which is less than critical value (5.9915), thus, we do not reject H_0 at 5% level of significance.