**Statistics 641, Fall 2013**
**Homework #7**
**Solutions**

1. Suppose that we have conducted a trial with two interim analyses and a final analysis, using the one-sided error spending function $g(t) = 0.025t^2$.

   Data from each analysis have been compiled into the data file `data7.csv` which contains the following variables:

   - `z`: Treatment variable (0,1)
   - `time1`: Follow-up time at analysis 1
   - `dead1`: Death indicator at analysis 1 (0=alive, 1=dead)
   - `time2`: Follow-up time at analysis 2
   - `dead2`: Death indicator at analysis 2 (0=alive, 1=dead)
   - `time3`: Follow-up time at analysis 3
   - `dead3`: Death indicator at analysis 3 (0=alive, 1=dead)

   For subjects who were still to be enrolled at the time of each interim analysis, the follow-up time is zero. The primary analysis is the comparison of treatment groups using the log-rank test.

   (a) Compute the number of events required to achieve 90% power for a hazard ratio of 0.825 (17.5% reduction in hazard) with (two-sided) $\alpha = 0.05$. Calculate expected information at the conclusion of the trial assuming the target number of event is reached.

   $$\frac{4 \times (1.96 + 1.28)^2}{(\log 0.825)^2} = 1135$$

   Expected Fisher information is $1135/4 = 283.75$, however because we ultimately only care about information fractions, we can rescale information as we wish and use 1135 as full information.

   (b) Compute the critical values expected if the interim analyses occur at 1/3 and 2/3 of full information, and using the $\alpha$-spending function $g(t)$ given above.

```
> b <- bounds(1:3/3, iuse=c(3,3), phi=c(2,2), alpha=c(.025,.025))
> summary(b)
...
Boundaries:
      Time    Lower   Upper   Exit pr.  Diff. pr.
1  0.33333  -2.7729  2.7729  0.0055556  0.0055556
2  0.66667  -2.3472  2.3472  0.0222222  0.0166667
3  1.00000  -2.0619  2.0619  0.0500000  0.0277778
```

The critical values are $\pm 2.7729, \pm 2.3472$ and $\pm 2.0619$.

(c) For each interim analysis and the final analysis, compute the log-rank $Z$-statistic. Use the convention that $Z$ is positive if the difference favors group $z = 1$.

```
> data <- read.csv("data7.csv")
> survdiff(Surv(time1, dead1)~z,data=data)
Call:
survdiff(formula = Surv(time1, dead1) ~ z, data = data)

       N Observed Expected (O-E)^2/E (O-E)^2/V
z=0 1122      172      162     0.556      1.09
z=1 1122      159      169     0.536      1.09

 Chisq= 1.1  on 1 degrees of freedom, p= 0.295
> survdiff(Surv(time2, dead2)~z,data=data)
Call:
survdiff(formula = Surv(time2, dead2) ~ z, data = data)

       N Observed Expected (O-E)^2/E (O-E)^2/V
z=0 1122      439      423     0.620      1.23
z=1 1122      417      433     0.605      1.23

 Chisq= 1.2  on 1 degrees of freedom, p= 0.268
> survdiff(Surv(time3, dead3)~z,data=data)
Call:
survdiff(formula = Surv(time3, dead3) ~ z, data = data)

       N Observed Expected (O-E)^2/E (O-E)^2/V
z=0 1122      673      632      2.72      5.34
z=1 1122      617      658      2.61      5.34

 Chisq= 5.3  on 1 degrees of freedom, p= 0.0209
```

For each analysis, the $Z$ score is the square root of the chi-square statistic, with the appropriate sign. Since the problem statement indicates that a positive $Z$ corresponds to benefit for treatment 1, we use the observed and expected values from the above to determine the sign. In each case, the observed number of deaths in group 1 is smaller than the expected, indicating that the observed treatment difference favors group 1. Hence the $Z$ statistics are

| Analysis time | $Z$ |
|---|---|
| 1 | $\sqrt{1.094} = 1.046$ |
| 2 | $\sqrt{1.227} = 1.108$ |
| 3 | $\sqrt{5.339} = 2.311$ |

(The printed output only gives 1 decimal place, but the $Z$ can be calculated more precisely by extracting the '`chisq`' element directly from the survdiff object.)

(d) Calculate the information fraction at each of the two interim analyses based on the expected number of events from (a) as full information. Compute the critical values used at these analysis times based on observed information and expected information at trial completion. (Note that the actual observed information will not be known until the completion of the trial). Would the stopping boundaries be crossed at either interim analysis?

---

Expected full information, expressed as the number of events, is 1134.

| Analysis time | Deaths | Information Fraction |
|---|---|---|
| 1 | $172 + 159 = 331$ | $331/1134 = 0.2919$ |
| 2 | $439 + 417 = 856$ | $856/1134 = 0.7549$ |

```
> b2 <- bounds(c(331,856,1134)/1134, iuse=c(3,3), phi=c(2,2),
+   alpha=c(.025,.025))
> summary(b2)
...
Boundaries:
      Time    Lower   Upper   Exit pr.  Diff. pr.
1  0.29189  -2.8582  2.8582  0.0042599  0.0042599
2  0.75485  -2.2293  2.2293  0.0284899  0.0242300
3  1.00000  -2.0822  2.0822  0.0500000  0.0215101
```

The new critical values are $\pm 2.8582, \pm 2.2293$ and $\pm 2.0822$. The first is larger than planned because the information fraction is smaller than planned and so less alpha is spent. The second is smaller than planned because the information fraction is larger than planned and the increment in alpha spent is larger. Finally, the final critical value is larger primarily because the alpha remaining is less than planned (0.0215 versus 0.0278).

Neither observed $Z$ exceeds its critical value, so the early stopping boundaries would not be crossed.

---

(e) Compute the information fractions at the two interim analysis based on *observed* information at trial completion. Compute the critical value at the final analysis given the actual information observed at trial completion.

---

At trial completion, we have 1290 events rather than the target of 1134. The new information fractions are

$$\begin{aligned} \frac{331}{1290} &= .2566 \\ \frac{856}{1290} &= .6636 \end{aligned}$$

Now comes the tricky part. The critical values determined using expected information are the ones that we used, so we need to calculate the final critical value based on observed full information, but using the critical values based on expected information. The easiest way to do this is to used the second information fraction ('`t2`') in the `bounds` function:

```
> bf <- bounds(c(331,856,1134)/1134, t2=c(331,856,1290)/1290,
+  iuse=c(3,3), phi=c(2,2), alpha=c(.025,.025))
> summary(bf)
...
Boundaries:
      Time   Time 2    Lower   Upper   Exit pr.  Diff. pr.
1  0.29189  0.25659  -2.8582  2.8582  0.0042599  0.0042599
2  0.75485  0.66357  -2.2293  2.2293  0.0284899  0.0242300
3  1.00000  1.00000  -2.1175  2.1175  0.0500000  0.0215101
```

Note that the first two critical values are the same as before, but the last is $\pm 2.1175$, which is larger in absolute value than before. This value increases because the increment in information is the same (0.0215), but because the second interim occurs earlier when using observed information, there is less probability stripped from the tail of the final statistic, so the tail is thicker and the critical value raised to maintain the same tail probability.

Alternatively, one could use the `drift` function to find the correct critical value, either by trial and error, or a root-finding method.

```
## try 2.11
> drift(t=c(331,856,1290)/1290, zb=c(b2$upper[1:2], 2.11), drft=0)$power
[1] 0.05051074   ## too large
## try 2.12
> drift(t=c(331,856,1290)/1290, zb=c(b2$upper[1:2], 2.12), drft=0)$power
[1] 0.0498287   ## too small
## try 2.1175
> drift(t=c(331,856,1290)/1290, zb=c(b2$upper[1:2], 2.1175), drft=0)$power
[1] 0.04999755  ## just right (or close enough)
```

(f) Find the information fractions based on observed information at trial completion based on the Wald test and compare to those in part (e).

To perform the Wald test, we fit Cox proportional hazards models:

```
> coxph(Surv(time1, dead1)~z,data=data)
Call:
coxph(formula = Surv(time1, dead1) ~ z, data = data)


    coef exp(coef) se(coef)     z   p
z -0.115     0.891     0.11 -1.04 0.3

Likelihood ratio test=1.09  on 1 df, p=0.296  n= 2244, number of events= 331
## extract just variance
> coxph(Surv(time1, dead1)~z,data=data)$var
           [,1]
[1,] 0.01210758
```

Note that $0.11^2 = 0.0121$, so it's easier to extract the `var` component directly. Information is the reciprcal of the variance:

```
> 1/coxph(Surv(time1, dead1)~z,data=data)$var
         [,1]
[1,] 82.59292    ## Fisher information at the estimated log-HR
## multiply by 4 to get number of events:
> 1/coxph(Surv(time1, dead1)~z,data=data)$var*4
         [,1]
[1,] 330.3717    ## very close to the observed number of events (331)
> 1/coxph(Surv(time2, dead2)~z,data=data)$var
         [,1]
[1,] 213.8128
> 1/coxph(Surv(time3, dead3)~z,data=data)$var
         [,1]
[1,] 321.8442
```

| Analysis time | $1/\text{Var}\left(\widehat{\beta}\right)$ | Information Fraction |
|---|---|---|
| 1 | 82.59 | $82.59/321.84 = 0.2566$ |
| 2 | 213.81 | $213.81/321.84 = 0.6643$ |
| 3 | 321.84 | 1 |

These information fractions are nearly identical to those based on numbers of deaths.

(g) What is the conclusion from this trial?

The $Z$ statistics at each of the interim analyses are within the monitoring boundaries, so the trial would not have stopped early on that basis. At the final analysis, the $Z$ crosses the upper boundary, so we would have demonstrated that treatment 1 is superior to treatment 0.

(h) Calculate an adjusted $p$-value at the conclusion of the trial using both the stage-wise and likelihood ratio orderings.

For SW, we need the probability of stopping early plus the probablity of going to the end and observing a final $Z$ at least as large as that observed.

```
> drift(t=c(331,856,1290)/1290, zb=c(b2$upper[1:2], 2.311), drft=0)$power
[1] 0.03979745
```

To get LR, note that the final $Z$ is smaller than the stage 1 critical value, but larger than the stage 2 critical value. We can get the probability of a $Z$ at least as large in either stage 1 or stage 2 at one time:

```
> drift(t=c(331,856)/1290, zb=c(b2$upper[1], 2.311), drft=0)$power
[1] 0.02368678
```

The probability of a larger $Z$ at the final is

```
> drift(t=c(331,856,1290)/1290, zb=c(b2$upper[1:2], 2.311), drft=0)$exit[3]
[1] 0.01130732
```

The sum is 0.02369 + 0.01131 = 0.03499. Note that this is slightly smaller than the SW, but larger than the "nominal" from part (c) of 0.0209.

(i) Calculate the conditional power at each interim analysis given the alternative hypothesis that the log-hazard ratio is 0.825 favoring group $z = 1$. If the trial should stop for futility if conditional power falls below 20%, would the trial continue to its planned conclusion?

Conditional power is given by

$$1 - \Phi\left(\frac{b_K - B(t) - (1-t)\theta}{\sqrt{1-t}}\right)$$

where $K = 3$, $b_K = 2.0822$ (from part (d)), $B(t) = \sqrt{t}Z(t)$ with $t = 0.29189$ or $0.75485$ at analyses 1 and 2 respectively, and $Z(t)$ is the observed $Z$ at analysis time $t$. We compute as $\theta = \beta\sqrt{I_K} = -\log 0.825\sqrt{1135/4} = 3.24$ (we use the minus sign because of the convention that a positive value of $Z$ favors group 1.) Note that because of the way the study was powered, this is exactly 1.96+1.28 from the numerator in part (a)

Plugging into the formula we have at analysis 1:

$$1 - \Phi\left(\frac{2.0822 - \sqrt{.2919} \times 1.046 - (1 - 0.2919) \times 3.24}{\sqrt{1 - 0.2919}}\right) = 1 - \Phi(-0.9227) = 0.822$$

At analysis 2:

$$1 - \Phi\left(\frac{2.0822 - \sqrt{.7549} \times 1.108 - (1 - 0.7549) \times 3.24}{\sqrt{1 - 0.7549}}\right) = 1 - \Phi(.6577) = 0.255$$

Both of these are above 20%, so based on conditional power, the trial would continue to its planned conclusion.