## C.4 Modern Survival Problem sheet 4: Nonparametric testing and semiparametric models

(1) The object `tongue` in the package `KMsurv` lists survival or right-censoring times in weeks after diagnosis for 80 patients with tongue tumours. The `type` random variable is 1 or 2, depending as the tumour was aneuploid or diploid respectively.

(a) Use the log-rank test to test whether the difference in survival distributions is significant at the 0.05 level.

We give below `R` code for computing this in two different ways: Using the function `survdiff`, which does the computation automatically, and by extracting the relevant quantities from the survival object and doing the computation directly.

We get $Z = -1.67$, which corresponds to a $p$-value of 0.09.

Using `survdiff` we get the same result, but it is reported as a chi-squared statistic of 2.8 (which is $1.67^2$) on 1 degree of freedom.

SURVDIFF CODE

```
> require('survival')
> require('KMsurv')
> data(tongue)
> attach(tongue)

>
  > tongue.surv=Surv(time,delta)
> tongue.fit=survfit(tongue.surv~type)
  > tdiff=survdiff(tongue.surv~type)
  > tdiff
Call:
  survdiff(formula = tongue.surv ~ type)

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= 0.0949
```

DIRECT COMPUTATION

```
# Problem sheet 4, question 1
require('survival')
```

```
require('KMsurv')
data(tongue)
attach(tongue)

tongue.surv=Surv(time,delta)
tongue.fit=survfit(tongue.surv~type)


n1=tongue.fit$strata[1]
n2=tongue.fit$strata[2]

# Input two vectors of times t1,t2, and
# numbers at risk n1,n2 whose length is 1 longer than the t's
#   Output four vectors I1, I2, (of same length as t1,t2) and Y1,Y2
#   I1[k] gives an index of I2 corresponding to
#   the last time in t2 that precedes t1[k]
#       Thus, we have t2[I1[k]]<=t1[k] < t2[I1[k]+1],
#   and r2[I1[k]+1] is the number of type 2 individuals at risk
#   at the time t1[k] (when there are r1[k] type 1 individuals)
#   Y1=r1[I1]

crossrisk=function(t1,t2,r1,r2){
  I1=rep(0,length(t1))
  I2=rep(0,length(t2))
  for(i in seq(length(t1))){
    I1[i]=1+sum(t1[i]>t2)
  }
  for(i in seq(length(t2))){
    I2[i]=1+sum(t2[i]>t1)
  }
  list(I1,I2,r1[I2],r2[I1])
}

r1=tongue.fit$n.risk[seq(n1)]
r2=tongue.fit$n.risk[seq(n1+1,n1+n2)]

r1=c(r1,r1[n1]-tongue.fit$n.event[n1]-tongue.fit$n.censor[n1])
r2=c(r2,r2[n2]-tongue.fit$n.event[n1+n2]-tongue.fit$n.censor[n1+n2])
t1=tongue.fit$time[seq(n1)]
t2=tongue.fit$time[seq(n1+1,n1+n2)]
```

```
cr=crossrisk(t1,t2,r1,r2)

Y1=c(r1[-n1],cr[[3]])
Y2=c(cr[[4]],r2[-n2])
  # Note: r1 and r2 had an extra count added on to make crossrisk work
d1=c(tongue.fit$n.event[seq(n1)],rep(0,n2))
d2=c(rep(0,n1),tongue.fit$n.event[seq(n1+1,n1+n2)])

t=c(t1,t2)

# We have to deal with the problem of ties between times for the two groups


dup1=which(duplicated(t,fromLast=TRUE))
dup2=which(duplicated(t))
ndup=length(dup1)

# Type 2 Event counts are removed from the second appearance
# and placed in the first appearance
d2[dup1]=d2[dup2]
d2=d2[-dup2]
d1=d1[-dup2]

# Type 2 at-risk counts are removed from the second appearance
# and placed in the first appearance
Y2[dup1]=Y2[dup2]
Y2=Y2[-dup2]
Y1=Y1[-dup2]
t=t[-dup2]

tord=order(t)
t=t[tord]  #put times in order
## Now put everything else in the same order
Y=Y[tord]
Y1=Y1[tord]
Y2=Y2[tord]
d=d[tord]
d1=d1[tord]
d2=d2[tord]
```

```
Y=Y1+Y2
d=d1+d2

# Product of number at risk
atriskprod=Y1*Y2
includes=(atriskprod>0)&(d>0)
# We only get contributions if someone's at risk and events occurred at that time

Y=Y[includes]
Y1=Y1[includes]
Y2=Y2[includes]
d=d[includes]
d2=d2[includes]
d1=d1[includes]

t=t[includes]

wLR=Y1*Y2/Y
p=1
q=0

S=c(1,cumprod((Y-d)/Y))[-length(Y)]  #K-M estimator for survival
wFH=(1-S)^q*S^p*wLR

# Now compute the test statistic

w=wLR

M=w*(d1/Y1-d2/Y2)
sigma=w*w*d*(Y-d)/Y2/Y1/(Y-1)
sK=d*Y1*Y2*(Y-d)/Y^2/(Y-1)

Z=sum(M)/sqrt(sum(sigma))

> Z
[1] -1.670246
```

(b) Repeat the above with a test that emphasises differences shortly after diagnosis.
This can also be done with `survdiff`.

```
> tdiff2=survdiff(tongue.surv~type,rho=1)
# rho=1 corresponds to p=1 for Fleming-Harrington weights
> tdiff2
Call:
  survdiff(formula = tongue.surv ~ type, rho = 1)

N Observed Expected (O-E)^2/E (O-E)^2/V
type=1 52     20.2     24.4     0.731       3.3
type=2 28     15.1     10.9     1.643       3.3

Chisq= 3.3  on 1 degrees of freedom, p= 0.0694
> w=wFH
> M=w*(d1/Y1-d2/Y2)
> sigma=w*w*d*(Y-d)/Y2/Y1/(Y-1)
> sK=d*Y1*Y2*(Y-d)/Y^2/(Y-1)
>
  > Z=sum(M)/sqrt(sum(sigma))
> Z
[1] -1.805118
> Z^2
[1] 3.25845
```

(c) Calculate and plot the estimated excess mortality for aneuploid compared with diploid. In this case there are no nuisance covariates, so the excess mortality is simply the difference between the Nelson–Aalen estimators, with increments at time $t$ by a predictable function $k(t)$:

$$\hat{\Gamma}(t) = \sum_{t_i \le t} k(t_i) \left( \frac{G(t_i}{Y(1;t_i)} - \frac{1-G_i}{Y(0;t_i)} \right).$$

Here $G_i = 1$ when there is an aneuploid-type event at time $t_i$, and 0 when there is a diploid-type event. Because there are ties in the data, we change this slightly, to

$$\hat{\Gamma}(t) = \sum_{t_i \le t} k(t_i) \left( \frac{d_i^{aneup}}{Y^{aneup}(t_i)} - \frac{d^{dip}}{Y^{dip}(t_i)} \right).$$

We could choose anything for $k$, but we just take $k \equiv 1$ here. One approach, then, would be to take the differences between the estimators calculated by `survfit` in the previous part. Extracting the components from the `survfit` object would be fairly opaque, though, so we give code below that does the computation directly.

```
etimes=tongue$time[tongue$delta==1]  #Event times
aneup=subset(tongue,type==1)
```

```
dip=subset(tongue,type==2)
## For each event time, count number of events of each type at that time
d.aneup=sapply(etimes,function(t) sum(aneup$delta[aneup$time==t]))
d.dip=sapply(etimes,function(t) sum(dip$delta[dip$time==t]))
## For each event time, count number at risk of each type at that time
Y.aneup=sapply(etimes,function(t) sum(aneup$delta[aneup$time>t]))
Y.dip=sapply(etimes,function(t) sum(dip$delta[dip$time>t]))
# Stop when we run out of individuals
tmax=max(which(Y.aneup*Y.dip>0))

Gammaincrement=(d.aneup/Y.aneup-d.dip/Y.dip)[1:tmax]
Gamma=cumsum(Gammaincrement)
varinc=(d.aneup/Y.aneup^2+d.dip/Y.aneup^2)[1:tmax]
vargamma=cumsum(varinc)
sdgamma=sqrt(vargamma)

conflevel=.95
z=-qnorm((1-conflevel)/2)
upper=Gamma+sdgamma*z
lower=Gamma-sdgamma*z
```

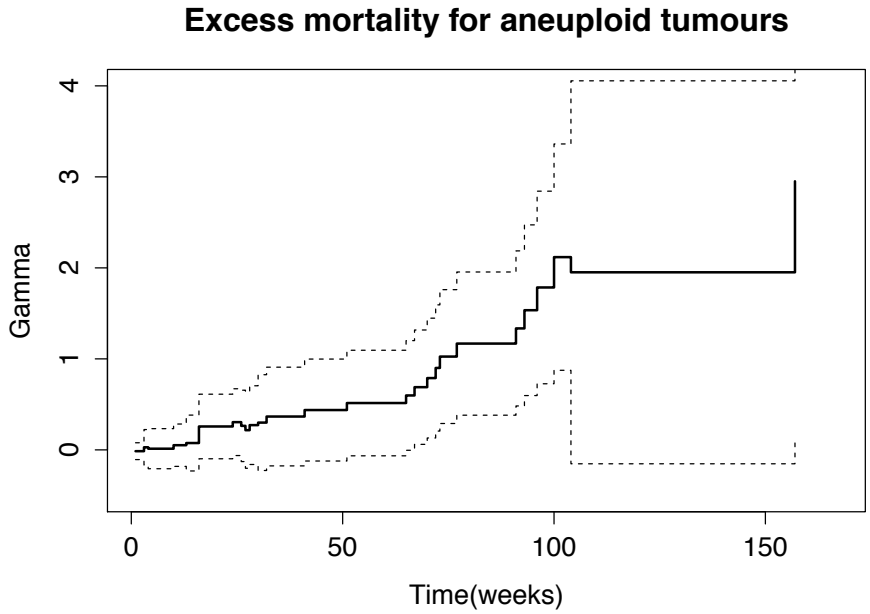


Figure C.2: Excess mortality for aneuploid vs diploid tumours from `tongue` dataset.

(2) Show that the class of non-parametric test statistics that we have defined in section 8.1 includes the Wilcoxon rank-sum statistic, in the special case where there is no censoring or truncation. What weight function do we need to take to recover the rank-sum statistic? Derive the sampling distribution of the rank-sum statistic. *Optional: Try out the two statistics on some simulated data. They may be drawn from any distribution you like.*

When computing the rank-sum statistic for $n$ subjects, with $n_i$ from group $i$ ($i = 0, 1$) we start by assigning ranks to each of the observations $T_1, \ldots, T_n$, and defining $R_i$ to be the sum of the ranks of subjects of type $i$. We then define

$$U_i := n_0 n_1 + \frac{n_i(n_i+1)}{2} - R_i.$$

Either one of these may be used as the rank-sum statistic, corresponding to the two different tails of the distribution.

Observe that when there is no censoring or truncation, the rank of an observation is simply the number of individuals still at risk; that is, $Y_\cdot(t_i)$. Thus

$$R_0 = \sum_{i:G_i=0} Y_0(t_i) + \sum_{i:G_i=0} Y_1(t_i).$$

Note that $Y_0$ is decremented by one at each time $t_i$ such that $G_i = 0$, so

$$\sum_{i:G_i=0} Y_0(t_i) = \frac{n_0(n_0+1)}{2},$$

and

$$\begin{aligned} U_1 - U_0 &= \sum_{i=1}^{n} \frac{(-1)^{G_i} w(t_i)}{Y_{G_i}(t_i)} \\ U_1 + U_0 &= n_0 n_1, \end{aligned}$$

where $w(t_i) = Y_1(t_i)Y_0(t_i)$. Thus $\mathbb{E}[U_j] = n_0 n_1/2$ and the variance may be computed as the expected value of the predictable variation

$$\begin{aligned} \frac{1}{4}\sum_{i=1}^{n} \mathbb{E}\left[Y_{1-G_i}(t_i)^2 \,\middle|\, \mathcal{F}_{t_i-}\right] &= \frac{1}{4}\sum_{i=1}^{n} \frac{Y_0(t_i)^2 Y_1(t_i)}{Y_0(t_i)+Y_1(t_i)} + \frac{Y_1(t_i)^2 Y_0(t_i)}{Y_0(t_i)+Y_1(t_i)} \\ &= \frac{1}{4}\sum_{i=1}^{n} Y_0(t_i)Y_1(t_i). \end{aligned}$$

The variance is the expected value of this sum. Since $Y_1(t_i) = n - i + 1 - Y_0(t_i)$, this is

$$\mathrm{Var}(U_1) = \frac{1}{4}\sum_{i=1}^{n}(n-i+1)\mathbb{E}[Y_0(t_i)] - \frac{1}{4}\sum_{i=1}^{n}\mathbb{E}[Y_0(t_i)]^2 - \frac{1}{4}\sum_{i=1}^{n}\mathrm{Var}(Y_0(t_i)).$$

We observe now that under the null hypothesis, the $n-i+1$ survivors up to time $t_i$ are a uniform random pick from the $n$ subjects. Thus $Y_0(t_i)$ has hypergeometric distribution with parameters $(n, n_0, n-i+1)$, so

$$\mathbb{E}\big[Y_0(t_i)\big] = \frac{(n-i+1)n_0}{n},$$
$$\mathrm{Var}\big(Y_0(t_i)\big) = \frac{(n-i+1)n_0n_1(i-1)}{n^2(n-1)}.$$

(Properties of the hypergeometric distribution may be found at http://en.wikipedia.org/wiki/Hypergeometric_distribution.) Thus

$$\begin{aligned}\mathrm{Var}(U_j) &= \frac{n_0}{4n}\sum_{i=1}^{n}\left(i^2 - \frac{i^2n_0}{n} - \frac{n_1i(n-i)}{n(n-1)}\right)\\ &= \frac{n_0n_1}{4n(n-1)}\sum_{i=1}^{n}(i^2-i)\\ &= \frac{n_0n_1(n+1)}{12}\end{aligned}$$

after some algebra.

Thus, when $n_0$ and $n_1$ are both large, we may use

$$\frac{U_j - n_0n_1/2}{\sqrt{n_0n_1(n+1)/12}}$$

as a test statistic, assuming it should have standard normal distribution if the null hypothesis holds.

(3) Suppose we have an additive-hazards model where an individual has covariates $(X_1, \ldots, X_p)$ and the individual hazards are then

$$\alpha(t) = \beta_0(t) + \beta_1(t)X_1 + \cdots + \beta_p(t)X_p,$$

where the $X_k$ are random variables. An observation consists of a single right-censored event.

(a) Suppose the variable $X_p$ is not observed, so is not included in the model. If the random variables $X_k$ are all independent, show that the remaining model is still an additive-hazards model with a different baseline hazard $\beta_0(t)$.

Consider a single individual under observation, producing a time $T$ that is either an event time ($\delta = 1$) or a censoring time ($\delta = 0$).

Let $\mathcal{F}_t$ be the $\sigma$-algebra for the observations up to time $t$ including the fixed covariates $X_1, \ldots, X_{p-1}$, and $\mathcal{G}_t$ be the extension of $\mathcal{F}_t$ to include the covariate $X_p$. We are given that the fully observed process has intensity for individual $i$

$$\alpha(t) = \alpha^{\mathcal{G}}(t) = Y_i(t)\left(\beta_0(t) + \beta_1(t)X_1 + \cdots + \beta_p(t)X_p\right).$$

(We take the regression coefficients $\beta_k(t)$ to be nonrandom.) We need to show that the hazard rate conditioned on reduced information $\alpha^{\mathcal{F}}$ also fits into the additive-hazards model. By the Innovation Theorem (Theorem 3.1),

$$\begin{aligned}\alpha^{\mathcal{F}}(t) &= \mathbb{E}\big[\alpha_i^{\mathcal{G}}(t) \,\big|\, \mathcal{F}_{t-}\big] \\ &= \sum_{k=0}^{p} \mathbb{E}\big[\beta_k(t) X_k \,\big|\, \mathcal{F}_{t-}\big] \\ &= \sum_{k=0}^{p-1} \beta_k(t) X_k + \mathbb{E}\big[X_p \,\big|\, \mathcal{F}_{t-}\big]\beta_p(t).\end{aligned}$$

We only need to consider this conditioning on the event $T \geq t$, since the intensity is 0 on the event $\{T < t\}$.
Since we have assumed that the covariates $X_k$ are independent, on the event $\{Y_i(t) = 1\}$

$$\begin{aligned}\mathbb{E}\big[X_p \,\big|\, \mathcal{F}_{t-}\big] &= \mathbb{E}\big[X_p \,\big|\, \big(T \wedge t, \delta \mathbf{1}_{\{T<t\}}\big]\big] \\ &= \mathbb{E}\left[X_p \frac{S(t|X_p)}{S(t)}\right]\end{aligned}$$

by Bayes' Law, where $S(t|x) = \int_t^\infty f(s|x)ds$ is the conditional survival function.
We have

$$S(t|x) = \exp\left\{-\int_0^t \alpha(s)ds\right\} = \exp\left\{-B_0(t) - \sum_{k=1}^{p-1} X_k B_k(t) - x B_p(t)\right\};$$

thus, by independence of $X_k$,

$$S(t) = \mathbb{E}\left[e^{-X_p B_p(t)}\right] \exp\left\{-B_0(t) - \sum_{k=1}^{p-1} X_k B_k(t)\right\},$$

and

$$\frac{S(t|x)}{S(t)} = \frac{e^{-xB_p(t)}}{M_{X_p}(-B_p(t))},$$

where $M_{X_p}$ is the moment generating function of $X_p$.
Thus, the hazard rate for the reduced model is

$$\alpha^{\mathcal{F}}(t) = \left(\beta_0(t) + \frac{\mathbb{E}[X_p e^{-X_p B_p(t)}]}{\mathbb{E}[e^{-X_p B_p(t)}]}\beta_p(t)\right) + \sum_{k=1}^{p-1} \beta_k(t) X_k.$$

The first term in brackets is the new baseline hazard.

(b) Suppose the random variables $X_k$ are multivariate normal (but not independent). How does the model change when $X_p$ is dropped?
We may assume without loss of generality that the $X_k$ have mean 0. We can represent $X_p = c_p Z + \sum_{k=1}^{p-1} c_k X_k$, where $Z$ is standard normal independent of $X_1, \ldots, X_{p-1}$, and $c_1, \ldots, c_p$. We have then

$$\mathbb{E}\big[X_p \,\big|\, \mathcal{F}_{t-}\big] = \sum_{k=1}^{p-1} c_k X_k + c_p \mathbb{E}\left[Z \frac{S(t|Z)}{S(t)}\right]$$

By the same argument as before, we have

$$\frac{S(t|z)}{S(t)} = \frac{e^{-z c_p B_p(t)}}{M_Z(-c_p B_p(t))} = e^{-z c_p B_p(t) - c_p^2 B_p(t)^2/2}.$$

Thus the hazard rate for the reduced model is

$$\alpha^{\mathcal{F}}(t) = \left(\beta_0(t) + \mathbb{E}[Z e^{-Z c_p B_p(t)}] c_p \beta_p(t) e^{-c_p^2 B_p(t)^2/2}\right) + \sum_{k=1}^{p-1} \left(\beta_k(t) + c_k \beta_p(t)\right) X_k.$$

Thus, we still have an additive-hazards model, but now all the coefficients have changed.

(4) In section 9.5.1 we describe fitting the Aalen additive hazards model for the special case of a single (possibly time-varying) covariate. Suppose we constrain the assumptions further, to assume that $x_i$ is constant in time, and takes on only the values 0 and 1. Explain how this is related to the excess mortality model. Compare the results we would obtain from the methods described in this section, to those obtained from the methods of section 7.2.

Assume there are no ties. Since $x_i(t) = x_i$ is 0 or 1, we may write $Y_1(t) = \#R(t) \cdot \mu_1(t)$ the number of individuals in group 1, and $Y_0(t) = \#R(t) \cdot (1 - \mu_1(t))$. Also $\mu_2(t) = \mu_1(t)$. We have then by (9.10)

$$\begin{pmatrix} \hat{B}_0(t) \\ \hat{B}_1(t) \end{pmatrix} = \sum_{t_j \le t} \frac{1}{Y_0(t_j) Y_1(t_j)} \begin{pmatrix} Y_1(t_j) x_j \\ -(1 - x_j) Y_1(t_j) + x_j Y_0(t_j) \end{pmatrix}.$$

If we think of this as an excess mortality model, $B_1$ is the same as what was called $\Gamma$. We have

$$\hat{B}_1(t) = \sum_{t_j \le t} -\frac{\mathbf{1}_{\{G_j = 0\}}}{Y_0(t_j)} + \frac{\mathbf{1}_{\{G_j = 1\}}}{Y_1(t_j)}.$$

This is the same as the estimator we worked out for the two-sample case for excess mortality, where the weight function is 1.

(5) Refer to the AML study, which is described at length in Example 8.1.4 and analysed with the Cox model in section 11.3. Using the data described in those places, estimate the difference in cumulative hazard to 20 weeks between the two groups by

(a) The nonparametric method described in section 7.2;

In the terminology of section 8.1.4 there is no nuisance categorisation, so by (7.5) the difference may be estimated by the difference between the Nelson–Aalen estimators:

$$\hat{\Gamma}(t) = \int_0^t \frac{dN_1(s)}{Y_1(s)} - \int_0^t \frac{dN_0(s)}{Y_0(s)}$$

Calling the Maintenance group number 1, and Nonmaintenance number 0, we read off of Table 8.2 $\hat{A}_1(20) = \hat{A}_1(18) = 0.32$, and $\hat{A}_0(20) = 0.49$, yielding

$$\hat{\Gamma}(20) = -0.17.$$

The variance will be the sum of the variances of the two estimators (since they are independent). As long as there are no ties between events from different groups, this may be estimated by

$$\sum_{t_i \le t} \sum_{k=0}^{d_i} \big(Y(G_i; t_i) - k\big)^{-2}.$$

From the Table we can see that this is

$$\sigma_1^2(20) + \sigma_0^2(20) = \frac{1}{12^2} + \frac{1}{11^2} + \frac{1}{10^2} + \frac{1}{9^2} + \frac{1}{8^2} + \frac{1}{10^2} + \frac{1}{8^2} = 0.0788.$$

Thus, an approximate 95% confidence interval for $\hat{\Gamma}(20)$ would be

$$-0.17 \pm 0.28 \cdot 1.96 = -0.17 \pm 0.550.$$

(b) The semiparametric method based on the relative-risk regression.

The Cox model fit by `coxph` produced the outcome

| coxph(formula = Surv(time, status) ~ x, data = aml) | | | | | |
|---|---|---|---|---|---|
| | coef | exp(coef) | se(coef) | z | p |
| ×Nonmaintained | 0.916 | 2.5 | 0.512 | 1.79 | 0.074 |
| Likelihood ratio test=3.38 on 1 df | | p=0.0658 | n= 23 | | |

In Table 11.2 we tabulated the estimators for the baseline hazard, obtaining $\hat{A}_0(18) = 0.254$. A central estimate for the difference in cumulative hazard between the two groups would be

$$(1 - e^{\hat{\beta}})\hat{A}_0(18) = -1.5 \cdot 0.254 = -0.38.$$

We see that this is a substantially larger estimate than we made in the nonparametric model. This is consistent with the plot in Figure 11.3, where the purple circles and blue crosses (representing the survival estimates from the proportional hazards model for the two groups) are further apart at $t_i = 18$ than the black and red lines (representing the Kaplan–Meier estimators). This reflects that fact that the separate Kaplan–Meier estimators are cruder, making larger jumps at less frequent intervals.

To estimate the standard error, we begin by assuming (with little justification) that the estimators $\hat{\beta}$ and $\hat{A}_0(t_i)$ are approximately independent. Then we can use the delta method to estimate the variance. Let $\sigma_\beta^2$ be the variance of $\hat{\beta}$, and $\sigma_A^2$ the variance of $\hat{A}(18)$. So we can represent

$$\hat{\beta} \approx \beta_0 + \sigma_\beta Z, \qquad \hat{A}_0(18) = A_0(18) + \sigma_A Z',$$

where $Z$ and $Z'$ are standard normal (also approximately independent). We already have the estimate $\hat{\sigma}_\beta \approx 0.512$. We haven't given a formula for an estimator of $\sigma_A(18)$, but we can easily compute it with `R`.

```
require(survival)

cp=coxph(Surv(time,status)~x,data=aml)

aml.fit=survfit(cp)

aml.fit$std.err[aml.fit$time==18]
[1] 0.150247
```

Then our estimator for the difference in cumulative hazard is

$$\begin{aligned}(1 - e^{\hat{\beta}})\hat{A}_0(18) &\approx \left(1 - e^{\beta_0 + \sigma_\beta Z}\right)\left(A_0(18) + \sigma_A Z'\right) \\ &\approx \left(1 - e^{\beta_0}\left(1 + \sigma_\beta Z\right)\right)\left(A_0(18) + \sigma_A Z'\right) \\ &\approx \left(1 - e^{\beta_0}\right) A_0(18) - e^{\beta_0}\sigma_\beta A_0(18) Z + \left(1 - e^{\beta_0}\right)\sigma_A Z' - e^{\beta_0}\sigma_\beta\sigma_A Z Z'.\end{aligned}$$

(Note that the approximation in the first line is based on assuming $\sigma_\beta$ is much smaller than $\beta_0$, which isn't really very true here.) As long as we are assuming independence of $Z$ and $Z'$, the variance will be approximately

$$\left(e^{\beta_0}\sigma_\beta A_0(18)\right)^2 + \left(\left(1 - e^{\beta_0}\right)\sigma_A\right)^2 = 0.325^2 + .225^2 = 0.156,$$

so the standard error is about 0.395.

A better estimate, also taking into account the dependence between $\hat{\beta}$ and $\hat{A}_0$, could be obtained by not using the delta method, but instead treating the normal distribution of $\hat{\beta}$

as a Bayesian posterior distribution on $\beta_0$. For a range of possible $\beta_0$ we can compute an approximate mean and variance for $\hat{A}_0$, and then compute a Monte Carlo estimator of the variance of $\hat{\Gamma}$.

(c) Using the proportional hazards method, suppose an individual were to switch from maintenance to non-maintenance after 10 weeks, and suppose the hazard rates change instantaneously. Estimate the difference in cumulative hazard to 20 weeks between that individual and one who had always been in the non maintenance group.
We let $x_0(t)$ be the covariate trajectory for this individual, so recalling that the maintained group is the baseline this means that

$$x_0(t) = \begin{cases} 0 & \text{if } t \leq 10, \\ 1 & \text{if } t > 10. \end{cases}$$

Using the formula (10.9) we estimate for this individual

$$\begin{aligned} \hat{A}\big(20 \,\big|\, x_0\big) &= \int_0^t e^{\beta x_0(u)} d\hat{A}_0(u) \\ &= \int_0^{10} d\hat{A}_0(u) + \int_{10}^{20} e^{\beta} d\hat{A}_0(u) \\ &\approx \hat{A}_0(10) + 2.5\big(\hat{A}_0(20) - \hat{A}_0(10)\big) \\ &= 0.14 + 2.5(0.114) \\ &= 0.425. \end{aligned}$$