HW#7 Survival analysis II
Name: Shih Jia-Han

Summary of section2.3 in

Emura T & Chen YH (2014), Gene selection for survival data under dependent censoring, a copula-based approach, Statistical Methods in Medical Research, DOI: 10.1177/0962280214533378

Section2.3 provides an analytic framework to study the bias when the dependence is modeled via copulas.

The cause-specific hazard

$$h^{\#}(t \mid x_j) = P(t \le T < t + dt, T \le U \mid T \ge t, U \ge u, x_j)/dt$$

describes the "apparent" hazard rate for death in the presence of dependent censoring (Kalbfleisch and Prentice).

We formulate the effect of dependent censoring under the copula models as

$$P(T \ge t, U \ge u \mid x_j) = C_\alpha\{ S_T(t \mid x_j), S_U(u \mid x_j) \},$$

where $S_T(t \mid x_j) = P(T \ge t \mid x_j)$ and $S_T(t \mid x_j) = P(T \ge t \mid x_j)$ are the marginal survival functions, and α is the dependence parameter. As indicated in Rivest and Wells, the casue-specific hazard becomes $h^{\#}_\alpha(t \mid x_j) = r_\alpha(t \mid x_j) h(t \mid x_j)$, where

$$r_\alpha(t \mid x_j) = \frac{C_\alpha^{[1,0]}\{ S_T(t \mid x_j), S_U(t \mid x_j) \} S_T(t \mid x_j)}{C_\alpha\{ S_T(t \mid x_j), S_U(t \mid x_j) \}}$$

and $C_\alpha^{[1,0]}(u, v) = \partial C_\alpha(u, v)/\partial u$.

We can obtain this result by the following computation.

$$
\begin{aligned}
h^{\#}(t \mid x_j) &= P(t \le T < t+dt, T \le U \mid T \ge t, U \ge u, x_j)/dt \\
&= \frac{P(t \le T < t+dt, T \le U, T \ge t, U \ge u \mid x_j)/dt}{P(T \ge t, U \ge u \mid x_j)} \\
&= \frac{P(t \le T < t+dt \mid x_j)P(T \le U, U \ge u \mid x_j)/dt}{P(T \ge t, U \ge u \mid x_j)} \\
&= \frac{P(T \le U, U \ge u \mid x_j)}{P(T \ge t, U \ge u \mid x_j)} P(t \le T < t+dt \mid x_j)/dt \\
&= \frac{P(T \le U, U \ge u \mid x_j)}{P(T \ge t, U \ge u \mid x_j)} P(t \le T < t+dt, T \ge t \mid x_j)/dt \\
&= \frac{P(T \le U, U \ge u \mid x_j)P(T \ge t \mid x_j)}{P(T \ge t, U \ge u \mid x_j)} P(t \le T < t+dt \mid T \ge t, x_j)/dt \\
&= \frac{C_\alpha^{[1,0]}\{S_T(t \mid x_j), S_U(t \mid x_j)\}S_T(t \mid x_j)}{C_\alpha\{S_T(t \mid x_j), S_U(t \mid x_j)\}} h(t \mid x_j) \\
&= r_\alpha(t \mid x_j)h(t \mid x_j).
\end{aligned}
$$

Then we can define the "apparent effect" of gene x_j as

$$
\beta_\alpha^{\#}(t) \equiv \log \frac{h_\alpha^{\#}(t \mid x_j = 1)}{h_\alpha^{\#}(t \mid x_j = 0)} = \log \frac{h(t \mid x_j = 1)}{h(t \mid x_j = 0)} + \log \frac{r_\alpha(t \mid x_j = 1)}{r_\alpha(t \mid x_j = 0)}.
$$

This equation shows that the apparent effect can be partitioned into the true(net) effect and bias due to dependent censoring.

Then we can conduct numerical analysis to gain insight into how dependent censoring affects the apparent effect under several copulas.

Derivation of apparent effect

$$\beta_{\alpha}^{\#}(t) \equiv \log \frac{h(t \mid x_j = 1)}{h(t \mid x_j = 0)} + \log \frac{r_{\alpha}(t \mid x_j = 1)}{r_{\alpha}(t \mid x_j = 0)} .$$

We set marginal distribution as

$S_T(t \mid x_j) = S_T(t \mid 0)^{\exp(\beta_j)}$, $S_U(t \mid x_j) = S_U(t \mid 0)^{\exp(\beta_j)}$ and

$S_U(t \mid x_j) = S_T(t \mid 0)^{p_c/1-p_c}$, where $p_c \times 100\ (\%)$ is the censoring percentage.

Then consider the bias part,

$$r_{\alpha}(t \mid x_j = 0) = \frac{C_{\alpha}^{[1,0]}\{ S_T(t \mid x_j = 0), S_U(t \mid x_j = 0) \} S_T(t \mid x_j = 0)}{C_{\alpha}\{ S_T(t \mid x_j = 0), S_U(t \mid x_j = 0) \}} .$$

Under the Clayton copula model and only consider $\alpha > 0$, we have

$$\begin{aligned} C_{\alpha}\{ S_T(t \mid x_j = 0), S_U(t \mid x_j = 0) \} &= C_{\alpha}\{ S_T(t \mid x_j = 0), S_T(t \mid x_j = 0)^{p_c/1-p_c} \} \\ &= \{ S_T(t \mid 0)^{-\alpha} + S_T(t \mid 0)^{-\alpha p_c/1-p_c} - 1 \}^{\frac{-1}{\alpha}} \end{aligned}$$

and

$$\begin{aligned} & C_{\alpha}^{[1,0]}\{ S_T(t \mid x_j = 0), S_U(t \mid x_j = 0) \} \\ &= \frac{d}{dS_T(t \mid 0)} \{ S_T(t \mid 0)^{-\alpha} + S_U(t \mid 0)^{-\alpha} - 1 \}^{\frac{-1}{\alpha}} \\ &= \frac{-1}{\alpha} \{ S_T(t \mid 0)^{-\alpha} + S_U(t \mid 0)^{-\alpha} - 1 \}^{\frac{-1}{\alpha} - 1} \cdot \{ -\alpha S_T(t \mid 0)^{-\alpha - 1} \} \\ &= \{ S_T(t \mid 0)^{-\alpha} + S_T(t \mid 0)^{-\alpha p_c/1-p_c} - 1 \}^{\frac{-1}{\alpha} - 1} S_T(t \mid 0)^{-\alpha - 1}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} r_{\alpha}(t \mid x_j = 0) &= \frac{\{ S_T(t \mid 0)^{-\alpha} + S_T(t \mid 0)^{-\alpha p_c/1-p_c} - 1 \}^{\frac{-1}{\alpha} - 1} S_T(t \mid 0)^{-\alpha - 1} S_T(t \mid 0)}{\{ S_T(t \mid 0)^{-\alpha} + S_T(t \mid 0)^{-\alpha p_c/1-p_c} - 1 \}^{\frac{-1}{\alpha}}} \\ &= \frac{S_T(t \mid 0)^{-\alpha}}{S_T(t \mid 0)^{-\alpha} + S_T(t \mid 0)^{-\alpha p_c/1-p_c} - 1}. \end{aligned}$$

Similarly, we can obtain

$$C_\alpha\{ S_T(t\,|\,x_j=1), S_U(t\,|\,x_j=1)\} = \{ S_T(t\,|1)^{-\alpha} + S_T(t\,|1)^{-\alpha p_c/1-p_c} - 1\}^{\frac{-1}{\alpha}}$$

and

$$\begin{aligned} & C_\alpha^{[1,0]}\{ S_T(t\,|\,x_j=1), S_U(t\,|\,x_j=1)\} \\ &= \frac{-1}{\alpha}\{ S_T(t\,|1)^{-\alpha} + S_U(t\,|1)^{-\alpha} - 1\}^{\frac{-1}{\alpha}-1} \cdot \{-\alpha S_T(t\,|1)^{-\alpha-1}\} \\ &= \{ S_T(t\,|1)^{-\alpha} + S_T(t\,|1)^{-\alpha p_c/1-p_c} - 1\}^{\frac{-1}{\alpha}-1} S_T(t\,|1)^{-\alpha-1}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} r_\alpha(t\,|\,x_j=1) &= \frac{\{ S_T(t\,|1)^{-\alpha} + S_T(t\,|1)^{-\alpha p_c/1-p_c} - 1\}^{\frac{-1}{\alpha}-1} S_T(t\,|1)^{-\alpha-1} S_T(t\,|1)}{\{ S_T(t\,|1)^{-\alpha} + S_T(t\,|1)^{-\alpha p_c/1-p_c} - 1\}^{\frac{-1}{\alpha}}} \\ &= \frac{S_T(t\,|1)^{-\alpha}}{S_T(t\,|1)^{-\alpha} + S_T(t\,|1)^{-\alpha p_c/1-p_c} - 1}. \end{aligned}$$

Hence

$$\frac{r_\alpha(t\,|\,x_j=1)}{r_\alpha(t\,|\,x_j=0)} = \frac{S_T(t\,|1)^{-\alpha}\{ S_T(t\,|\,0)^{-\alpha} + S_T(t\,|\,0)^{-\alpha p_c/1-p_c} - 1\}}{S_T(t\,|\,0)^{-\alpha}\{ S_T(t\,|1)^{-\alpha} + S_T(t\,|1)^{-\alpha p_c/1-p_c} - 1\}}.$$

Then we fix t by setting $S_T(t\,|\,0) = 0.5$ and $S_T(t\,|1) = 0.5^{\exp(1)}$.

Finally we have

$$\frac{r_\alpha(t\,|\,x_j=1)}{r_\alpha(t\,|\,x_j=0)} = \frac{0.5^{-\alpha\exp(1)}\{ 0.5^{-\alpha} + (0.5)^{-\alpha p_c/1-p_c} - 1\}}{0.5^{-\alpha}\{ 0.5^{-\alpha\exp(1)} + (0.5)^{-\alpha\exp(1)\,p_c/1-p_c} - 1\}}.$$

Therefore we can analysis the relation between apparent effect and α, under several p_c.

Fig.1

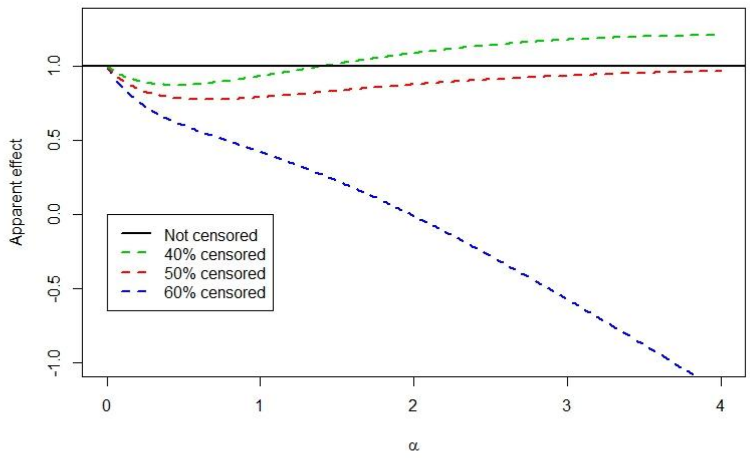

Fig.2

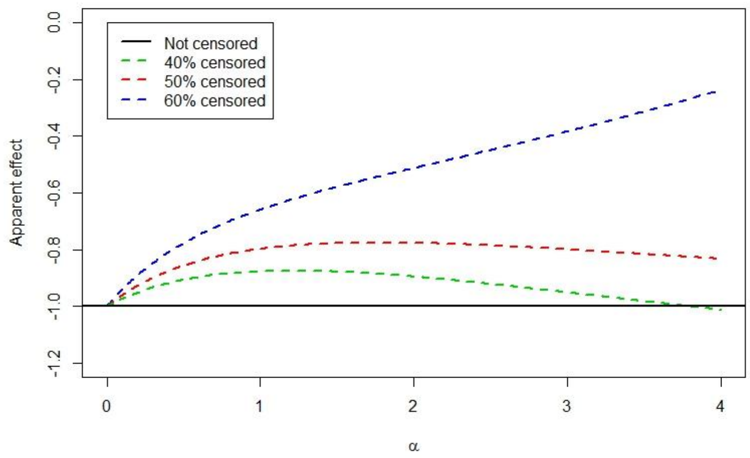

Fig.1 is set as net effect is 1 and Fig.2 is set as net effect is -1. From the result we can noticed that if the censoring probability is high (60%), the apparent effect is significantly different form the true(net) effect.

R code

```
########## Apparent effect function ##########

beta_ae=function(alpha,net_e,pc) {


    ST1=0.5^exp(net_e)

    ST0=0.5

    P=pc/(1-pc)

    r1=ST1^-alpha/(ST1^-alpha+ST1^(-alpha*P)-1)

    r0=ST0^-alpha/(ST0^-alpha+ST0^(-alpha*P)-1)


    net_e+log(r1/r0)


}
```

```
########## Plot net effect=1 ##########
q=seq(0,4,by=0.01)
beta_alpha1=c()
beta_alpha2=c()
beta_alpha3=c()

for (i in 1:length(q)){

   beta_alpha1[i]=beta_ae(q[i],1,0.4)
   beta_alpha2[i]=beta_ae(q[i],1,0.5)
   beta_alpha3[i]=beta_ae(q[i],1,0.6)

}

plot(q,beta_alpha3,type="l",ylim=c(-1,1.3),col=4,lty=2,main="Fig.1",xlab=expressio
n(alpha),ylab="Apparent effect",lwd=2)
lines(q,beta_alpha2,col=2,lty=2,lwd=2)
lines(q,beta_alpha1,col=3,lty=2,lwd=2)
abline(h=1,lwd=2)
legend(0,0,c("Not censored","40% censored","50% censored","60%
censored"),lty=c(1,2,2,2),lwd=2,col=c(1,3,2,4))
```

```
########## Plot net effect=-1 ##########
q=seq(0,4,by=0.01)
beta_alpha1=c()
beta_alpha2=c()
beta_alpha3=c()

for (i in 1:length(q)){

   beta_alpha1[i]=beta_ae(q[i],-1,0.4)
   beta_alpha2[i]=beta_ae(q[i],-1,0.5)
   beta_alpha3[i]=beta_ae(q[i],-1,0.6)

}

plot(q,beta_alpha3,type="l",ylim=c(-1.2,0),col=4,lty=2,main="Fig.2",xlab=expressio
n(alpha),ylab="Apparent effect",lwd=2)
lines(q,beta_alpha2,col=2,lty=2,lwd=2)
lines(q,beta_alpha1,col=3,lty=2,lwd=2)
abline(h=-1,lwd=2)
legend(0,0,c("Not censored","40% censored","50% censored","60%
censored"),lty=c(1,2,2,2),lwd=2,col=c(1,3,2,4))
```