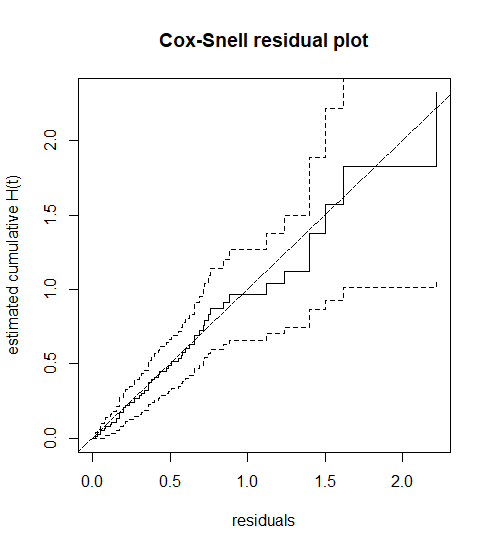
|  |
| --- |
| 15338673 |
| PW Janse van Rensburg |
| Survival Analysis – Assignment 7 |

# Question 1

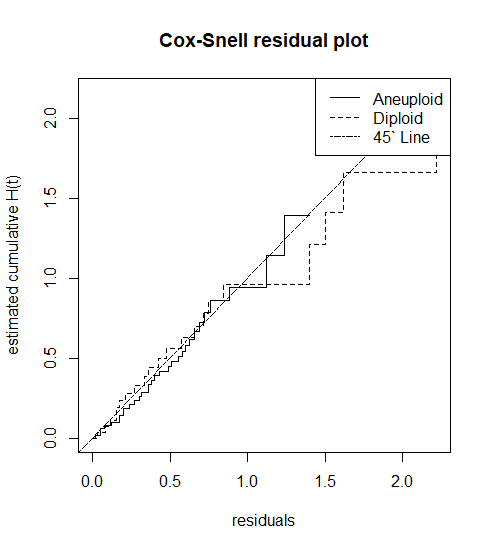
We have the following covariates:

We fit a Cox regression model with the abovementioned covariates.



**Figure 1 – Cox-Snell Residual plot of Times to Death for Patients with Cancer of   
the Tongue**

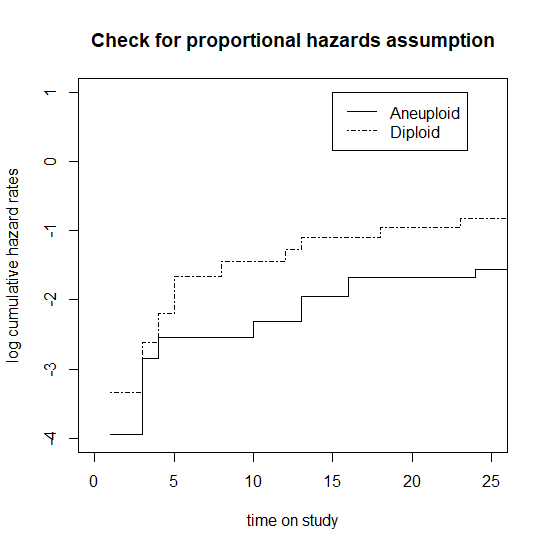
We used the Nelson-Aalen estimate of the cumulative hazard rate to plot the above figure. The residuals were computed and plotted against the estimated cumulative hazard of the residuals. The fitted model appears to follow the 45◦ line, thus the plot suggests that this model does not fit the data too badly.



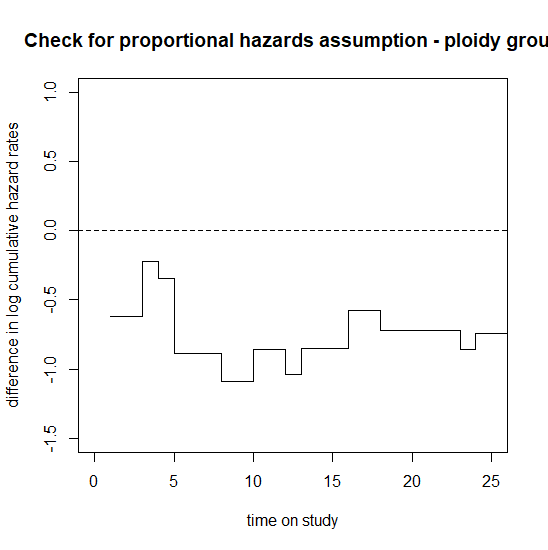
**Figure 2 – Cox-Snell Residual plot of Times to Death for Patients with Cancer of   
the Tongue Stratified on Tumor Type**

We see that for both the Aneuploid and Diploid tumors, their estimated cumulative hazard rates seem to be close to the 45o line, except in the tail where the estimates are quite variable. This analysis suggests that the stratified model fits better than the unstratified model.

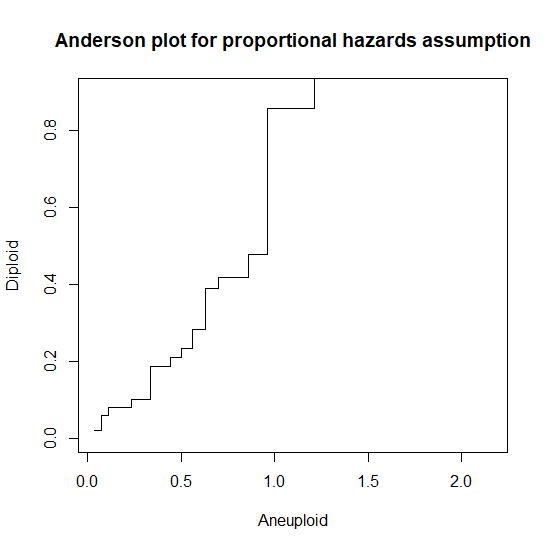
# Question 2



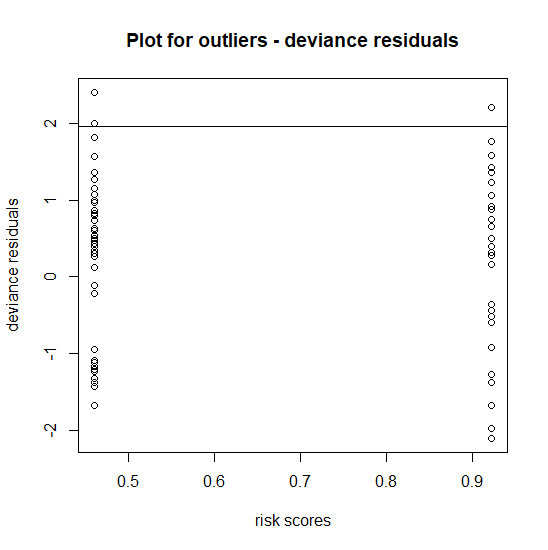
# Question 3



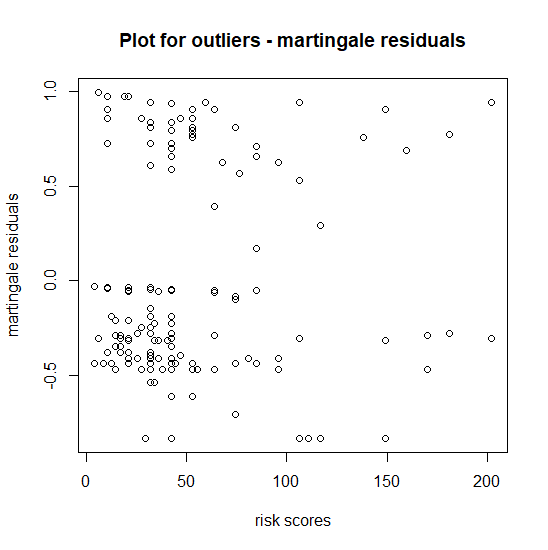
# Question 4



# Question 5



# Question 7



# Question 8

# Appendix A

library(data.table)

library(survival)

section1\_11 <- fread('Section1\_11.dat')

section1\_11\_cox <- coxph(Surv(Death,Censored)~Tumor,data=section1\_11,method="breslow")

res<-residuals(section1\_11\_cox)

cox\_snell<-(section1\_11$Censored-res)

# Does not want to accept surv without a formula arg, so we jippo it a bit

section1\_11\_temp <- section1\_11

section1\_11\_temp$Tumor <- 1

surv\_obj <- Surv(cox\_snell,section1\_11$Censored)~section1\_11\_temp$Tumor

aa<-survfit(surv\_obj,conf.type = 'none')

plot(aa,fun="cumhaz",main="Cox-Snell residual plot",xlab="residuals",ylab="estimated cumulative H(t)")

abline(0,1,lty=6)

# Stratifying on tumor type

surv\_obj\_strat <- Surv(cox\_snell,section1\_11$Censored)~section1\_11$Tumor

aa\_strat <-survfit(surv\_obj\_strat,conf.type = 'none')

plot(aa\_strat,fun="cumhaz",main="Cox-Snell residual plot",xlab="residuals",ylab="estimated cumulative H(t)",lty=c(1,2))

abline(0,1,lty=6)

legend(legend=c('Aneuploid','Diploid','45` Line'),lty = c(1,2,6),'topright')

#Data prep for two tumor types

aneuploid<-summary(survfit(Surv(section1\_11$Death,section1\_11$Censored)~section1\_11$Tumor,subset=(Tumor==1),data=section1\_11,conf.type="none"))

diploid<-summary(survfit(Surv(section1\_11$Death,section1\_11$Censored)~section1\_11$Tumor,subset=(Tumor==2),data=section1\_11,conf.type="none"))

naan<-cumsum(aneuploid$n.event/aneuploid$n.risk)

nadi<-cumsum(diploid$n.event/diploid$n.risk)

logan<-log(naan)

logdi<-log(nadi)

plot(sort(aneuploid$time),sort(logan),type="s",main="Check for proportional hazards assumption",

xlab="time on study",ylab="log cumulative hazard rates",ylim=c(-4,1),xlim=c(0,25))

lines(sort(diploid$time),sort(logdi),type="s",lty=4)

legend(x=15,y=1,legend=c("Aneuploid","Diploid"),lty=c(1,4))

all.times<-sort(unique(c(aneuploid$time,diploid$time)))

log.H<-matrix(0,nrow=length(all.times),ncol=2)

res<-cbind(all.times,log.H)

dimnames(res)<-list(c(),c("all times","logH.aneuploid","logH.diploid"))

res[match(aneuploid$time,all.times),"logH.aneuploid"]<-logan

res[match(diploid$time,all.times),"logH.diploid"]<-logdi

for(i in 1:length(all.times)){

if(res[i,"logH.aneuploid"]==0)

res[i,"logH.aneuploid"]<-res[(i-1),"logH.aneuploid"]

if(res[i,"logH.diploid"]==0)

res[i,"logH.diploid"]<-res[(i-1),"logH.diploid"]

}

plot(x=all.times,y=(res[,"logH.aneuploid"]-res[,"logH.diploid"]),type="s",xlim=c(0,25),ylim=c(-1.5,1),

main="Check for proportional hazards assumption - ploidy group",xlab="time on study",

ylab="difference in log cumulative hazard rates")

abline(h=0,lty=2)

plot(x=exp(res[,"logH.aneuploid"]),y=exp(res[,"logH.diploid"]),type="s",ylim=c(0,0.9),

main="Anderson plot for proportional hazards assumption",xlab="Aneuploid",ylab="Diploid")

abline(0,1,lty=7)

risk<-0.4609544\*section1\_11$Tumor

dev.res<-residuals(coxph(Surv(section1\_11$Death,section1\_11$Censored)~section1\_11$Tumor,method="breslow"),type="deviance")

plot(x=risk,y=dev.res,main="Plot for outliers - deviance residuals",xlab="risk scores",ylab="deviance residuals")

abline(h=1.96,lty=7)

section1\_6 <- fread('Section1\_6.dat')

section1\_6\_cox <- coxph(Surv(V17,V18)~V5,data=section1\_6,method="breslow")

risk\_burn <-2.127598\*section1\_6$V5

mart\_res\_burn <- residuals(coxph(Surv(section1\_6$V17,section1\_6$V18)~section1\_6$V4,method="breslow"),type="martingale")

plot(x=risk\_burn,y=mart\_res\_burn,main="Plot for outliers - martingale residuals",xlab="risk scores",ylab="martingale residuals")

abline(h=1.96,lty=7)