

```

### Script do artigo de Mioma (análise de sobrevida)
### Banco Mioma participantes da linha de base 1999 e 2001
### tempo inicial (no caso desta tese - 20 anos para todas),
### tempo final (a idade em 1999 ou a idade da cirurgia de histerectomia por
###          outra causa para os não casos e a idade do desfecho paraos casos),
### tempo (tempo final - tempo inicial) e status (tem ou não o desfecho).
### critério de exclusão; mulheres sem informação sobre MU, histerectomia

library(epiDisplay)
library(sjPlot)

#Abrir o banco
setwd("C:/Prosaude/Replicações alunos/Karine/Artigo mioma")
alasca <- read.csv2("alaskamio.csv") #participantes das fases 1, 2 e 4
names(alasca) # verificar as variáveis do banco
dim(alasca)
attach(alasca)

#####Construção das variáveis#####

#####Desfecho#####
#Diagnóstico médico de mioma uterino auto-relatado.
#Nome da variável nas análises= mio
#Categorias= não=0, sim=1
tab1(status, graph=F)
mio<-ifelse(status==0,0,ifelse(status==1,1,NA)) ##excluindo NAs
tab1(mio, graph=F)

#####Exposição principal#####
## Cor/raça
#Nome da variável nas análises= raca
#Categorias= branca=1, parda=2 e preta=3.
tab1(corecat, graph=F)
raca<-ifelse(corecat==4,NA,corecat) ##excluindo NAs e a categoria amarela
alasca$raca <- factor(raca, levels=c(1,2,3), labels=c("White","Brown","Black"))
tab1(raca, graph=F)

##Idade categorizada
##Nome da variável nas análises= idadecat
##Categorias= menos que 30 anos= 0, 30 a 40= 1, 40 a 50= 2, mais de 50 anos= 3
summary(idadeind)
idadecat<-ifelse(idadeind<30,0,ifelse(idadeind>=30&idadeind<40,1,ifelse(idadeind>=40&
idadeind<50,2,ifelse(idadeind>=50,3,NA))))

```

```

tab1(idadecat, graph=F)
alasca$idadecat <- factor(idadecat, levels=c(0,1,2,3), labels=c("< 30","30-39","40-49","50 or +"))

#####Variáveis marcadoras de PSE ao longo da vida#####
#####PSE na infância#####
##Escolaridade do pai
#Nome da variável nas análises= escpai
#Categorias= 2º grau completo ou mais=0, 1º grau completo=1, até 1º grau incompleto=2
tab1(e4, graph=F)
e4<-ifelse(e4>7,NA,e4)##excluindo NAs
tab1(e4, graph=F)
escpai<-ifelse(e4<3,2,ifelse(e4>2&e4<=4,1,ifelse(e4>4,0,NA)))
tab1(escpai, graph=F)
##Esc pai dicotomizados
escpaidic<-ifelse(e4<3,1,ifelse(e4>2,0,NA))
tab1(escpaidic, graph=F)

##Escolaridade da mãe
#Nome da variável nas análises= escmae
#Categorias= 2º grau completo ou mais=0, 1º grau completo=1, até 1º grau incompleto=2

tab1(e5, graph=F)
e5<-ifelse(e5>7,NA,e5)##excluindo os NAs
tab1(e5, graph=F)
escmae<-ifelse(e5<3,2,ifelse(e5>2&e5<=4,1,ifelse(e5>4,0,NA)))
tab1(escmae, graph=F)
##Esc mae dicotomizados
escmaedic<-ifelse(e5<3,1,ifelse(e5>2,0,NA))
tab1(escmaedic, graph=F)

#Escolaridade dos pais
#Nome da variável nas análises= escpais
#Categorias= pai alta/mãe alta= 0, pai alta/mãe baixa OU pai baixa/mãe alta= 1, pai baixa/mãe baixa= 2

escpais<-ifelse(escpaidic==0&escmaedic==0,0,ifelse(escpaidic==1&escmaedic==0|escpaidic==0&escmaedic==
1,1,ifelse(escpaidic==1&escmaedic==1,2,NA)))
tab1(escpais, graph=F)
alasca$escpais <- factor(escpais, levels=c(0,1,2), labels=c("High","Medium","Low"))

#####PSE no início da vida adulta#####
##Escolaridade
#Nome da variável nas análises= esc
#Categorias= universitário completo= 0, 2º grau completo= 1,até 1º grau completo= 2

```

```

tab1(e35, graph=F)
e35<-ifelse(e35>7,NA,e35) ##excluindo os NAs
tab1(e35, graph=F)
esc<-ifelse(e35<4,2,ifelse(e35>=4&e35<6,1,ifelse(e35>=6,0,NA)))
tab1(esc, graph=F)
alasca$esc <- factor(esc, levels=c(0,1,2), labels=c("College or more","Secondary education","Primary education"))

#####PSE ao longo da vida#####
##PSE cumulativa contínua (pai - participante)
#Nome da variável nas análises= sespai
sespai<-ifelse(esc+escpai==0,0,ifelse(esc+escpai==1,1,ifelse(esc+escpai==2,2,ifelse(esc+escpai==3,3,
ifelse(esc+escpai==4,4,NA))))))
tab1(sespai, graph=F)

##PSE cumulativa categórica (pai - participante)
#Nome da variável nas análises= SESpai
#Categorias= 0 ou 1= 0,2= 1, 3 - 4=2
SESpai<-ifelse(sespai<2,0,ifelse(sespai==2,1,ifelse(sespai>2,2,NA)))
tab1(SESpai, graph=F)

##PSE cumulativa contínua (mãe - participante)
#Nome da variável nas análises= sesmae
sesmae<-ifelse(esc+escmae==0,0,ifelse(esc+escmae==1,1,ifelse(esc+escmae==2,2,ifelse(esc+escmae==3,3,
ifelse(esc+escmae==4,4,NA))))))
tab1(sesmae, graph=F)

##PSE cumulativa categórica (mãe - participante)
#Nome da variável nas análises= SESmae
#Categorias= 0 ou 1= 0,2= 1, 3 - 4=2
SESmoe<-ifelse(sesmae<2,0,ifelse(sesmae==2,1,ifelse(sesmae>2,2,NA)))
tab1(SESmoe, graph=F)

##PSE cumulativa contínua (mãe - participante)
#Nome da variável nas análises= sespais
sespais<-ifelse(esc+escpais==0,0,ifelse(esc+escpais==1,1,ifelse(esc+escpais==2,2,ifelse(esc+escpais==3,3,
ifelse(esc+escpais==4,4,NA))))))
tab1(sespais, graph=F)

##PSE cumulativa categórica (pais - participante)
#Nome da variável nas análises= SESpais
#Categorias= 0 ou 1= 0,2= 1, 3 - 4=2
SESpais<-ifelse(sespais<2,0,ifelse(sespais==2,1,ifelse(sespais>2,2,NA)))
tab1(SESpais, graph=F)

```

```

alasca$SESpais <- factor(SESpais, levels=c(0,1,2), labels=c("High","Medium","Low"))

##Mobilidade social (pai - participante)
#Nome da variável nas análises= mob
#Categorias= pai alta/participante alta= 0, pai baixa/ participante alta= 1, pai alta/
#participante baixa= 2, pai baixa/ participante baixa=3
mob.part<-ifelse(e35 < 4,1, ifelse(e35>=4,2,NA)) ##dicotomizando a escolaridade do participante
tab1(mob.part, graph=F)

mob.pai <- ifelse(e4 < 4,1, ifelse(e4 >= 4,2,NA)) ##dicotomizando a escolaridade do pai
tab1(mob.pai, graph=F)

mob.mae <- ifelse(e5 < 4,1, ifelse(e5>=4,2,NA)) ##dicotomizando a escolaridade da mãe
tab1(mob.mae, graph=F)

mob<-ifelse(mob.pai==2&mob.part==2,0,ifelse(mob.pai==1&mob.part==2,1,ifelse(mob.pai==2&mob.part==1,2,
ifelse(mob.pai==1&mob.part==1,3,NA))))
tab1(mob, graph=F)

##Mobilidade social (mae - participante)
#Nome da variável nas análises= mob1
#Categorias= mãe alta/participante alta=0, mãe baixa/ participante alta= 1,mãe alta/ participante
#baixa= 2,mãe baixa/ participante baixa= 3
mob1<-ifelse(mob.mae==2&mob.part==2,0,ifelse(mob.mae==1&mob.part==2,1,ifelse(mob.mae==2&mob.part==1,2,
ifelse(mob.mae==1&mob.part==1,3,NA))))
tab1(mob1, graph=F)

#####Variáveis marcadoras de acesso e utilização de serviços de saúde#####
##Plano de saúde
#Nome da variável nas análises= plano
#Categorias= sim=0, não= 1
table(b1, graph=F)
b1<-ifelse(b1>3,NA,b1) ##excluindo NAs
plano<-ifelse(b1<3,0,ifelse(b1==3,1,NA))
tab1(plano, graph=F)
alasca$plano <- factor(plano, levels=c(0,1), labels=c("Yes","No"))

##Realização de teste Papanicolau
#Nome da variável nas análises= papa
#Categorias= realizou há menos de 3 anos= 0, nunca realizou/realizou há mais de 3 anos= 1
tab1(m7, graph=F)
m7<-ifelse(m7>4,NA,m7) ##excluindo NAs
papa<-ifelse(m7<3,1,ifelse(m7>=3,0,NA))

```

```

tab1(papa, graph=F)
alasca$papa <- factor(papa, levels=c(0,1), labels=c("Within the past 3 years", "Never/more than 3 years ago"))

##Realização de exame de mama
#Nome da variável nas análises= mama
#Categorias= realizou há menos de 3 anos= 0, nunca realizou/realizou há mais de 3 anos= 1
tab1(m8, graph=F)
m8<-ifelse(m8>4,NA,m8)##excluindo NAs
mama<-ifelse(m8<3,1,ifelse(m8>=3,0,NA))
tab1(mama, graph=F)
alasca$mama <- factor(mama, levels=c(0,1), labels=c("Within the past 3 years", "Never/more than 3 years ago"))

tempo1<-tempo/365.25
table(tempo1)

#####
#####Características das participantes segundo cor/raça#####
#Tabela 1
tableStack(vars=c(idadecat,escpais,esc,SESpais,plano,papa,mama), total.column=T, by=raca, dataFrame = alasca)

#####Número de pessoas-ano e hazard ratio de Cor/raça e mioma
#Tabela 2
tabela2 <- table(alasca$raca[mio==1])
tabela2 <- c(sum(tabela2), tabela2)
names(tabela2)[1] <- "Overall"
tabela2b <- round(c(sum(tempo1), sum(tempo1[alasca$raca=="White"]), sum(tempo1[alasca$raca=="Brown"]),
sum(tempo1[alasca$raca=="Black"]))), 3)
tabela2c <- round(tabela2 / tabela2b * 1000, 2)
tabela2d <- cox.display(coxph(y~as.factor(raca),data=alasca))$table
tabela2d <- c("-", 1, paste(round(tabela2d[1,1], 2), " (", round(tabela2d[1,2], 2), " - ",
round(tabela2d[1,3], 2), ") ", sep=""),
paste(round(tabela2d[2,1], 2), " (", round(tabela2d[2,2], 2), " - ",
round(tabela2d[2,3], 2), ") ", sep=""))
tabela2 <- cbind.data.frame(tabela2, tabela2b, tabela2c, tabela2d)
colnames(tabela2) <- c("N° Cases", "N° Person-Years", "Crude IR", "Crude HR (95% CI)")
tab_df(tabela2, encoding = "Windows-1252", show.rownames = T)

#####Curvas de Risco acumulado (Kaplan-Meier) e p-valor dos testes Log-Rank#####
##### e Peto para o diagnóstico de mioma uterino ajustadas pela cor/raça das participantes####

y<-Surv(tempo1,status)
KMraca<-survfit(y~raca,data=alasca)
survdiff(y~raca,data=alasca)##Log rank

```

```

survdiff(y~raca,data=alasca,rho=1)##Peto
plot(KMraca,conf.int=F,xlab="dias",ylab="Incidence",
lty=c(1:3),fun="cumhaz")
legend(x="topleft",legend=c("White","Bronw-Mixed Race","Black"),
lty=c(1:3),title="Color/Race",bty="n")
legend(x="bottomright",legend=c("Log-Rank - p=0,004","Peto- p=0,003"),bty="n")
title("Cumulative Risk - UL diagnosis" )

```

```

#####
#####Modelos multivariados de Cox#####

```

```
#Tabela 3
```

```

#Modelo de Cox somente ajustado por cor/raça
modbruto<-coxph(y~as.factor(raca),data=alasca,x=T)
summary(modbruto)
tab_model(modbruto)

```

```

#Modelo de Cox ajustado somente por idade
mod0<-coxph(y~as.factor(raca)+idadeind,data=alasca,x=T)
summary(mod0)
tab_model(mod0)

```

```

#Modelo de Cox ajustado por idade e variáveis de acesso a serviços de saúde
mod<-coxph(y~as.factor(raca)+idadeind+as.factor(plano)+as.factor(papa)+as.factor(mama),data=alasca,x=T)
summary(mod)
tab_model(mod)

```

```

#Modelo de Cox ajustado por idade, variáveis de acesso e escolaridade dos pais
mod1<-coxph(y~as.factor(raca)+idadeind+as.factor(plano)+as.factor(papa)+as.factor(mama)+as.factor(escpais),data=alasca,
,x=T)
summary(mod1)
tab_model(mod1)

```

```

#Modelo de Cox ajustado por idade, variáveis de acesso e escolaridade da participante
mod4<-coxph(y~as.factor(raca)+idadeind+as.factor(plano)+as.factor(papa)+as.factor(mama)+as.factor(esc),data=alasca,x=T)
summary(mod4)
tab_model(mod4)

```

```

#Modelo de Cox ajustado por idade, variáveis de acesso e PSE cumulativa (pai - participante)
mod5<-coxph(y~as.factor(raca)+idadeind+as.factor(plano)+as.factor(papa)+as.factor(mama)+as.factor(sespais),data=alasca,
,x=T)
summary(mod5)
tab_model(mod5)

```

```
#Modelo de Cox ajustado por idade, variáveis de acesso e PSE cumulativa (mãe - participante)
mod6<-coxph(y~as.factor(raca)+idadeind+as.factor(plano)+as.factor(papa)+as.factor(mama)+as.factor(SESpais),data=alasca
,x=T)
summary(mod6)
```

```
#Fim da tabela 3
```

```
#####
#####Análise de Resíduosde Schoenfeld#####
```

```
source("Rfunmar.r")
#Modelo de Cox ajustado por idade, variáveis de acesso e escolaridade dos pais
zph<-cox.zph(mod1)
zph
##Gráficos de Sch pelas covariáveis
mod1$call
par(mfrow=c(3,3))
plot(zph[1],main="Cor/raça Parda")
abline(h=mod1$coef[1],lty=3)
plot(zph[2],main="Cor/raça Preta")
abline(h=mod1$coef[2],lty=3)
plot(zph[3],main="Idade")
abline(h=mod1$coef[3],lty=3)
plot(zph[4],main="Plano de saúde")
abline(h=mod1$coef[4],lty=3)
plot(zph[5],main="Exame de mama")
abline(h=mod1$coef[5],lty=3)
plot(zph[6],main="Teste Papanicolaou")
abline(h=mod1$coef[6],lty=3)
plot(zph[7],main="Escolaridade do pai1")
abline(h=mod1$coef[7],lty=3)
plot(zph[8],main="escolaridade do pai2")
abline(h=mod1$coef[8],lty=3)
```

```
#Modelo de Cox ajustado por idade, variáveis de acesso e escolaridade da participante
zph3<-cox.zph(mod4)
zph3
##Gráficos de Sch pelas covariáveis
mod4$call
par(mfrow=c(3,3))
plot(zph3[1],main="Cor/raça Parda")
abline(h=mod4$coef[1],lty=3)
```

```
plot(zph3[2],main="Cor/raça Preta")
abline(h=mod4$coef[2],lty=3)
plot(zph3[3],main="Idade")
abline(h=mod4$coef[3],lty=3)
plot(zph3[4],main="Plano de saúde")
abline(h=mod4$coef[4],lty=3)
plot(zph3[5],main="Exame de mama")
abline(h=mod4$coef[5],lty=3)
plot(zph3[6],main="Teste Papanicolaou")
abline(h=mod4$coef[6],lty=3)
plot(zph3[7],main="Escolaridade1")
abline(h=mod4$coef[7],lty=3)
plot(zph3[8],main="Escolaridade2")
abline(h=mod4$coef[8],lty=3)
```

```
#Modelo de Cox ajustado por idade, variáveis de acesso e PSE cumulativa
zph5<-cox.zph(mod5)
zph5
```

```
##Gráficos de Sch pelas covariáveis
mod6$call
par(mfrow=c(3,3))
plot(zph5[1],main="Cor/raça Parda")
abline(h=mod6$coef[1],lty=3)
plot(zph5[2],main="Cor/raça Preta")
abline(h=mod6$coef[2],lty=3)
plot(zph5[3],main="Idade")
abline(h=mod6$coef[3],lty=3)
plot(zph5[4],main="Plano de saúde")
abline(h=mod6$coef[4],lty=3)
plot(zph5[5],main="Exame de mama")
abline(h=mod6$coef[5],lty=3)
plot(zph5[6],main="Teste Papanicolaou")
abline(h=mod6$coef[6],lty=3)
plot(zph5[7],main="PSE mãe/participante")
abline(h=mod6$coef[7],lty=3)
```