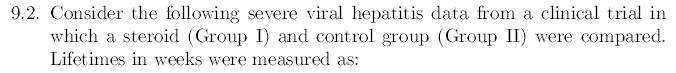
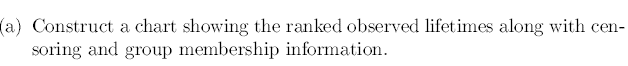
Chapter 9 Cox Proportional Hazards

library(survival)

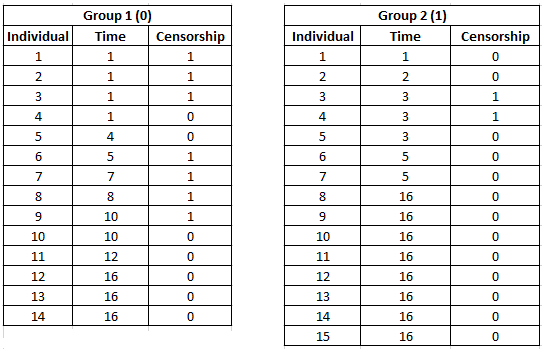
knitr::include\_graphics("92.PNG")



knitr::include\_graphics("92a.PNG")



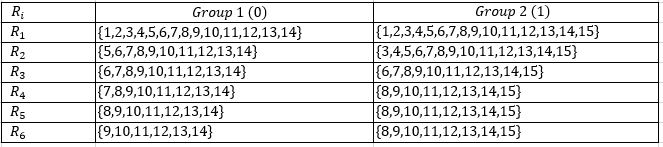
knitr::include\_graphics("92ar.PNG")



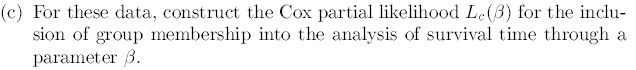
knitr::include\_graphics("92b.PNG")



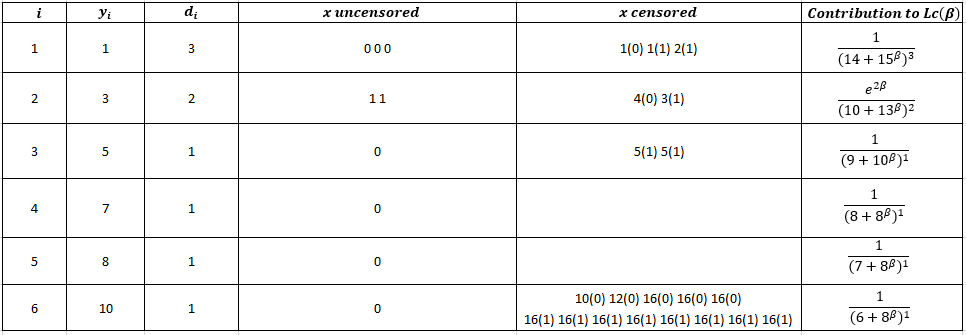
knitr::include\_graphics("92br.PNG")



knitr::include\_graphics("92c.PNG")



knitr::include\_graphics("92cr.PNG")



Simplificando

knitr::include\_graphics("92d.PNG")



hepatitis <- read.csv("hepatitis.csv", header = T, sep =",")  
head(hepatitis)

## time cens group  
## 1 1 1 0  
## 2 1 1 0  
## 3 1 1 0  
## 4 1 0 0  
## 5 4 0 0  
## 6 5 1 0

Compararemos que tan similar es la experiencia de supervivencia para ambos grupos mediante el Logrank test.

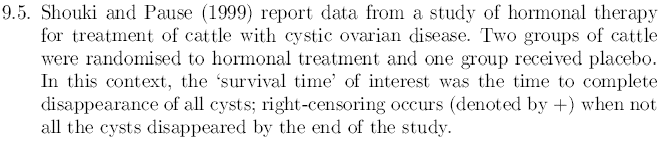
Pra ellos lo primero que debemos hacer es ajustar nuestros datos el modelo cox.ph, con H0: La función de supervivencia es igual para ambos grupos

hepatitis$group <- as.factor(hepatitis$group)  
hepatitis.fit <- coxph(Surv(hepatitis$time, hepatitis$cens) ~ hepatitis$group)  
  
summary(hepatitis.fit)

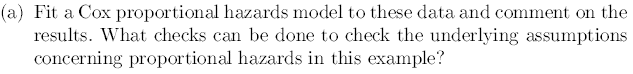
## Call:  
## coxph(formula = Surv(hepatitis$time, hepatitis$cens) ~ hepatitis$group)  
##   
## n= 29, number of events= 9   
##   
## coef exp(coef) se(coef) z Pr(>|z|)   
## hepatitis$group1 -1.4112 0.2439 0.8023 -1.759 0.0786 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## exp(coef) exp(-coef) lower .95 upper .95  
## hepatitis$group1 0.2439 4.101 0.05061 1.175  
##   
## Concordance= 0.664 (se = 0.081 )  
## Likelihood ratio test= 3.78 on 1 df, p=0.05  
## Wald test = 3.09 on 1 df, p=0.08  
## Score (logrank) test = 3.64 on 1 df, p=0.06

Dado que el logrank test p = 0.06 Se rechaza H0, no es decir que las funciones de supervivencia son diferentes para ambos grupos.

knitr::include\_graphics("95.PNG")



knitr::include\_graphics("95a.PNG")



cystic <- read.csv("cystic.csv", header = T, sep =",")  
head(cystic)

## time cens treatment  
## 1 4 1 1  
## 2 6 1 1  
## 3 8 1 1  
## 4 8 1 1  
## 5 9 1 1  
## 6 10 1 1

#### Se cargan los datos

cystic$treatment <- as.factor(cystic$treatment)  
cystic.fit <- coxph(Surv(cystic$time, cystic$cens) ~ cystic$treatment)  
  
summary(cystic.fit)

## Call:  
## coxph(formula = Surv(cystic$time, cystic$cens) ~ cystic$treatment)  
##   
## n= 20, number of events= 14   
##   
## coef exp(coef) se(coef) z Pr(>|z|)   
## cystic$treatment2 -2.113212 0.120849 0.863008 -2.449 0.014339 \*   
## cystic$treatment3 -4.669602 0.009376 1.326075 -3.521 0.000429 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## exp(coef) exp(-coef) lower .95 upper .95  
## cystic$treatment2 0.120849 8.275 0.022266 0.6559  
## cystic$treatment3 0.009376 106.655 0.000697 0.1261  
##   
## Concordance= 0.839 (se = 0.037 )  
## Likelihood ratio test= 21.8 on 2 df, p=2e-05  
## Wald test = 12.8 on 2 df, p=0.002  
## Score (logrank) test = 21.89 on 2 df, p=2e-05

Se puede ver que una persona con el tratamiento 2 es un 12% más propensa a morir que una persona con el tratamiento 3, el tratamiento 1 se comporta como la linea base So(t).

knitr::include\_graphics("95b.PNG")

 Para este ajuste se tiene que la linea base de la funcion de supervivencia para ambos tratamientos es:

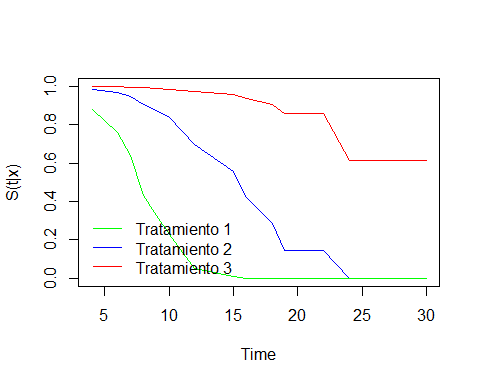
Gracias al ajuste se procede a calcular el riesgo base y la función de supervivencia base

H<-basehaz(cystic.fit,centered=F)  
tpos<-H$time  
H0<-H$hazard  
S0<-exp(-H0)  
round(cbind(tpos,S0,H0),digits=5)

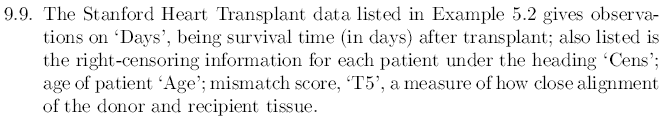
## tpos S0 H0  
## [1,] 4 0.87954 0.12836  
## [2,] 6 0.75910 0.27562  
## [3,] 7 0.63871 0.44831  
## [4,] 8 0.43222 0.83882  
## [5,] 9 0.32913 1.11130  
## [6,] 10 0.22631 1.48585  
## [7,] 12 0.05049 2.98604  
## [8,] 15 0.00817 4.80744  
## [9,] 16 0.00079 7.14291  
## [10,] 18 0.00003 10.39674  
## [11,] 19 0.00000 16.04310  
## [12,] 20 0.00000 16.04310  
## [13,] 22 0.00000 16.04310  
## [14,] 24 0.00000 51.59488  
## [15,] 27 0.00000 51.59488  
## [16,] 30 0.00000 51.59488

Una vez con estos datos generamos el gráficos de la probabilidad estimada para cada uno de los tratamientos.

plot(tpos,S0^(exp(-2.1132)),ylim=c(0,1),type="l",xlab="Time",ylab="S(t|x)", col="blue")  
lines(tpos,S0^(exp(-4.6696)), col="red")  
lines(tpos,S0, col="green")  
legend("bottomleft", legend=c("Tratamiento 1", "Tratamiento 2", "Tratamiento 3"), col = c("green","blue","red"), lty = 1, bty = "n")

 Se puede ver claramente que el mejor tratamiento es el tratamiento 3, pues los tiempos de sobrevivencia son mayores.

knitr::include\_graphics("99.PNG")

 #### Se cargan los datos

heart <- read.csv("heart\_data.csv", header= T, sep=";")  
head(heart)

## Days Cens Age T5  
## 1 15 1 54.3 1.11  
## 2 3 1 40.4 1.66  
## 3 624 1 51.0 1.32  
## 4 46 1 42.5 0.61  
## 5 127 1 48.0 0.36  
## 6 64 1 54.6 1.89

knitr::include\_graphics("99a.PNG")



### Ajuste del modelo.

cox.heart <- coxph(Surv(heart$Days,heart$Cens)~ heart$Age + heart$T5)  
summary(cox.heart)

## Call:  
## coxph(formula = Surv(heart$Days, heart$Cens) ~ heart$Age + heart$T5)  
##   
## n= 69, number of events= 45   
##   
## coef exp(coef) se(coef) z Pr(>|z|)   
## heart$Age 5.603e-02 1.058e+00 2.313e-02 2.423 0.0154 \*  
## heart$T5 6.465e-05 1.000e+00 5.286e-05 1.223 0.2213   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## exp(coef) exp(-coef) lower .95 upper .95  
## heart$Age 1.058 0.9455 1.011 1.107  
## heart$T5 1.000 0.9999 1.000 1.000  
##   
## Concordance= 0.616 (se = 0.05 )  
## Likelihood ratio test= 7.93 on 2 df, p=0.02  
## Wald test = 6.98 on 2 df, p=0.03  
## Score (logrank) test = 7.08 on 2 df, p=0.03

knitr::include\_graphics("99b.PNG")

 ### Construccion de los residuales Cox-Snell para CoxPH

Se calcula como la diferencia entre los residuales martingale y la censura. La pendiente debe ser de aproximadamente 45 grados.

### Captura de los resiguales de martingale

martinagale<- residuals(cox.heart, type = "martingale")

### Calculo de los residuos CoxSnell

coxsnell <- -(martinagale - heart$Cens)

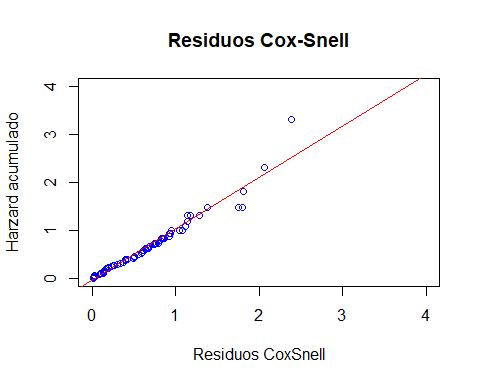
A continuación, el residuo de Cox-Snell para ajustar un modelo nulo de Cox y obtener el estimador de riesgo acumulativo de Nelson-Aalen.

fit\_coxsnell <- coxph(formula = Surv(coxsnell, Cens) ~ 1,  
 data = heart,  
 ties = c("efron","breslow","exact")[1])  
  
base\_haz <- basehaz(fit\_coxsnell, centered = FALSE)  
  
head(base\_haz)

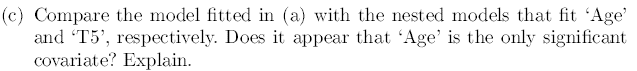
## hazard time  
## 1 0.00000000 0.01357663  
## 2 0.01470588 0.01445762  
## 3 0.01470588 0.02447342  
## 4 0.02985740 0.02488801  
## 5 0.04524201 0.02752381  
## 6 0.06086701 0.02945246

Para visualizar los residuales CoxSnell tenemos que:

plot(base\_haz$time,base\_haz$hazard,ylim=c(0,4), xlim=c(0,4),ylab="Harzard acumulado", xlab="Residuos CoxSnell", col="blue", main="Residuos Cox-Snell")  
abline(lm(base\_haz$hazard~base\_haz$time), col ="red")

 Este grafico muestra un intercepto cernad

knitr::include\_graphics("99c.PNG")

 Para este caso crearemos otros 2 modelos que compararemos con el modelo que tiene en cuenta tanto la edad como T5

### Modelo coxph con edad

cox.heart.age <- coxph(Surv(heart$Days,heart$Cens)~ heart$Age)  
summary(cox.heart.age)

## Call:  
## coxph(formula = Surv(heart$Days, heart$Cens) ~ heart$Age)  
##   
## n= 69, number of events= 45   
##   
## coef exp(coef) se(coef) z Pr(>|z|)   
## heart$Age 0.05455 1.05607 0.02254 2.42 0.0155 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## exp(coef) exp(-coef) lower .95 upper .95  
## heart$Age 1.056 0.9469 1.01 1.104  
##   
## Concordance= 0.609 (se = 0.054 )  
## Likelihood ratio test= 6.67 on 1 df, p=0.01  
## Wald test = 5.86 on 1 df, p=0.02  
## Score (logrank) test = 5.79 on 1 df, p=0.02

### Modelo coxph con T5

cox.heart.T5 <- coxph(Surv(heart$Days,heart$Cens)~ heart$T5)  
summary(cox.heart.T5)

## Call:  
## coxph(formula = Surv(heart$Days, heart$Cens) ~ heart$T5)  
##   
## n= 69, number of events= 45   
##   
## coef exp(coef) se(coef) z Pr(>|z|)  
## heart$T5 6.329e-05 1.000e+00 5.271e-05 1.201 0.23  
##   
## exp(coef) exp(-coef) lower .95 upper .95  
## heart$T5 1 0.9999 1 1  
##   
## Concordance= 0.574 (se = 0.049 )  
## Likelihood ratio test= 1.22 on 1 df, p=0.3  
## Wald test = 1.44 on 1 df, p=0.2  
## Score (logrank) test = 1.49 on 1 df, p=0.2

### Comparacion mediante anova.

anova(cox.heart,cox.heart.age,cox.heart.T5 , test = "LRT")

## Analysis of Deviance Table  
## Cox model: response is Surv(heart$Days, heart$Cens)  
## Model 1: ~ heart$Age + heart$T5  
## Model 2: ~ heart$Age  
## Model 3: ~ heart$T5  
## loglik Chisq Df P(>|Chi|)   
## 1 -157.28   
## 2 -157.91 1.2640 1 0.2609   
## 3 -160.63 5.4473 0 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Esta comparación nos muestra que no existe un diferencia significativa entre el modelo con las variables edad y t5, y el modelo que solo tiene en cuenta a T5, dado que P es muy pequeño. Esto nos da a entender que el modelo más simple que solo se relaciona con T5 es más optimo para seleccionarce.

### Verificación del modelo a elegir

Finalmente verificamos el resultado anterior mediante el AIC

require("MASS")

## Loading required package: MASS

Reg\_paso<-stepAIC(cox.heart)

## Start: AIC=318.55  
## Surv(heart$Days, heart$Cens) ~ heart$Age + heart$T5  
##   
## Df AIC  
## - heart$T5 1 317.81  
## <none> 318.55  
## - heart$Age 1 323.26  
##   
## Step: AIC=317.81  
## Surv(heart$Days, heart$Cens) ~ heart$Age  
##   
## Df AIC  
## <none> 317.81  
## - heart$Age 1 322.48

Reg\_paso

## Call:  
## coxph(formula = Surv(heart$Days, heart$Cens) ~ heart$Age)  
##   
## coef exp(coef) se(coef) z p  
## heart$Age 0.05455 1.05607 0.02254 2.42 0.0155  
##   
## Likelihood ratio test=6.67 on 1 df, p=0.009817  
## n= 69, number of events= 45

Como vemos el menor AIC es el pertenece al modelo que solo tiene como covariable a T5, por ende este es el mejor modelo y es el indicado para seleccionarse, por otro lado verificamos que la Edad no es una variable regresora significativa.