

# Análise de Sobrevida

## 0.5 - Aula Prática

Prof. Dr. Eder Angelo Milani

25/04/2025

### E na prática como funciona?

Em geral, na prática temos disponíveis várias variáveis que podem explicar o evento de interesse, com isso, temos que utilizar rotinas para a seleção de variáveis, como por exemplo stepwise. Na sequência as linhas de códigos visam analisar o conjunto de dados trabalhado anteriormente, mas agora utilizando todas as variáveis disponíveis.

### Leitura dos dados

As linhas de código a seguir executam as seguintes tarefas:

- leitura dos dados filtrados para CID C34;
- breve visualização do conjunto de dados;
- modificar a escala de tempo para ano, em vez de dia;
- excluir os tempos iguais a zero.

```
# limpando o que tem na memoria
rm(list=ls())

# local onde esta o arquivo com os dados
setwd("G:\\Meu Drive\\UFG\\Especializacao\\Aulas Análise Sobrevida\\Códigos")

### leitura
dados <- read.csv("cancer_c34.csv")
head(dados)
```

```
##   TOPOGRUP TEMPO  CENSURA ANODIAG IDADE SEXO  CIRURGIA RADIO QUIMIO ECGRUP
## 1      C34   292         1   2014   63    1         0     1     1     III
## 2      C34   132         1   2016   58    2         0     0     0     I
## 3      C34     3         0   2016   61    2         0     0     0     IV
## 4      C34    17         1   2016   67    1         0     0     0     IV
## 5      C34   182         1   2015   57    1         0     0     1     III
## 6      C34   287         1   2015   69    1         0     0     1     IV
```

```
# mudança na variável tempo - de dias para ano
summary(dados$TEMPO)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.0   80.0   253.0   543.7   690.0  2915.0
```

```
dados$TEMPO <- dados$TEMPO/365
```

```

# excluir os tempos iguais a zero
ind_tempo_zero <- which(dados$TEMPO == 0)
ind_tempo_zero

## [1] 256 297 322 374 865 996 1010 1049 1083 1165 1514 1665 1754 2090 5830
## [16] 6012 6196 7079 8030 8049 8383 8495

dados$TEMPO[ind_tempo_zero]

## [1] 0

dados <- dados[-ind_tempo_zero,]

```

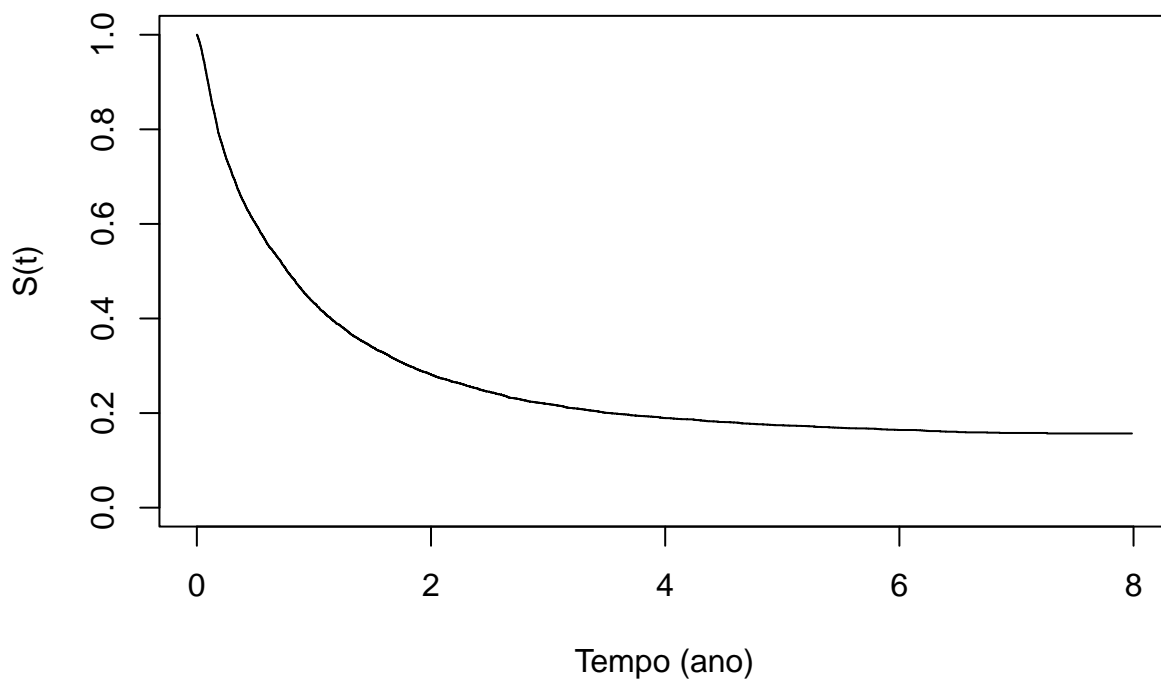
## Iniciando a análise

```

# analisando o tempo de sobrevida
# estimando a funcao de sobrevivencia

library(survival)
ekm <- survfit(Surv(TEMPO, CENSURA) ~ 1, data=dados)
#summary(ekm)
plot(ekm, mark.time = F, conf.int = F, xlab="Tempo (ano)", ylab="S(t)")

```



```

# analisando a quantidade de censura

table(dados$CENSURA)

##

```

```
##      0      1
## 2017 6892

round(prop.table(table(dados$CENSURA)), 2)

##
##      0      1
## 0.23 0.77

# TOPOGRUP

table(dados$TOPOGRUP)

##
##      C34
## 8909

round(prop.table(table(dados$TOPOGRUP)), 2)

##
##      C34
##      1

# nao sera analisada pois e 100% preenchida com C34

# ANODIAG sera analisada como uma variavel categorica
dados$ANODIAG <- factor(dados$ANODIAG)
table(dados$ANODIAG)

##
## 2014 2015 2016
## 2985 2926 2998

round(prop.table(table(dados$ANODIAG)), 2)

##
## 2014 2015 2016
## 0.34 0.33 0.34

is.factor(dados$ANODIAG)

## [1] TRUE

# IDADE e quantitativa

is.numeric(dados$IDADE)

## [1] TRUE

summary(dados$IDADE)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.00   58.00   64.00   64.19   72.00   99.00

# SEXO e qualitativa

dados$SEXO <- factor(dados$SEXO)
table(dados$SEXO)

##
##      1      2
```

```
## 5042 3867
round(prop.table(table(dados$SEXO)), 2)

##
##      1      2
## 0.57 0.43
# CIRURGIA e qualitativa

dados$CIRURGIA <- factor(dados$CIRURGIA)
table(dados$CIRURGIA)

##
##      0      1
## 7302 1607
round(prop.table(table(dados$CIRURGIA)), 2)

##
##      0      1
## 0.82 0.18
# QUIMIO e qualitativa

dados$QUIMIO <- factor(dados$QUIMIO)
table(dados$QUIMIO)

##
##      0      1
## 4029 4880
round(prop.table(table(dados$QUIMIO)), 2)

##
##      0      1
## 0.45 0.55
# RADIO e qualitativa

dados$RADIO <- factor(dados$RADIO)
table(dados$RADIO)

##
##      0      1
## 6217 2692
round(prop.table(table(dados$RADIO)), 2)

##
##      0      1
## 0.7 0.3
# ECGRUP e qualitativa

dados$ECGRUP <- factor(dados$ECGRUP)
table(dados$ECGRUP)

##
##      0      I      II     III     IV      X      Y
##    29   961   484  1645  4883   385   522
```

```
round(prop.table(table(dados$ECGRUP)), 2)
```

```
##
##      0      I      II     III     IV      X      Y
## 0.00 0.11 0.05 0.18 0.55 0.04 0.06
```

```
# veja a volumetria dos niveis 0, X, Y e II
# O estadiamento clinico e uma forma de avaliar a
# extensao do cancer antes do inicio do tratamento principal
```

```
table(dados$CENSURA, dados$ECGRUP)
```

```
##
##      0      I      II     III     IV      X      Y
## 0  23  678  213  376  461   70  196
## 1   6  283  271 1269 4422  315  326
```

```
round(prop.table(table(dados$CENSURA, dados$ECGRUP)), 2)
```

```
##
##      0      I      II     III     IV      X      Y
## 0 0.00 0.08 0.02 0.04 0.05 0.01 0.02
## 1 0.00 0.03 0.03 0.14 0.50 0.04 0.04
```

```
# cuidado com baixa volumetria!
# duvida: sera que nao poderiamos agrupar?
# mas como e so para efeito didatico, mas seguir como e apresentado
```

```
# todas as variaveis ok, estamos prontos para seguir com o ajuste
#install.packages("MASS")
```

```
library(MASS)
```

```
## Warning: pacote 'MASS' foi compilado no R versão 4.4.3
```

```
# modelo inicial com todas as variaveis usando a distribuicao exponencial
modelo_inicial_exp <- survreg(Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE +
                             CIRURGIA + RADIO + QUIMIO + ECGRUP,
                             data = dados, dist = "exponential")
```

```
summary(modelo_inicial_exp)
```

```
##
## Call:
## survreg(formula = Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE + CIRURGIA +
##      RADIO + QUIMIO + ECGRUP, data = dados, dist = "exponential")
##              Value Std. Error      z      p
## (Intercept)  1.76826    0.41550  4.26 2.1e-05
## ANODIAG2015 -0.07605    0.02949 -2.58 0.00992
## ANODIAG2016 -0.03957    0.02950 -1.34 0.17971
## IDADE       -0.00246    0.00110 -2.23 0.02604
## CIRURGIA1    1.35088    0.04518 29.90 < 2e-16
## RADIO1       0.28949    0.02664 10.87 < 2e-16
## QUIMIO1      0.98372    0.02654 37.07 < 2e-16
## ECGRUPI      -0.11603    0.41283 -0.28 0.77866
## ECGRUPII     -1.21322    0.41335 -2.94 0.00333
```

```

## ECGRUPIII -1.92650 0.41041 -4.69 2.7e-06
## ECGRUPIV -2.45918 0.40980 -6.00 2.0e-09
## ECGRUPX -2.02459 0.41312 -4.90 9.6e-07
## ECGRUPY -1.40598 0.41227 -3.41 0.00065
##
## Scale fixed at 1
##
## Exponential distribution
## Loglik(model)= -8369.5 Loglik(intercept only)= -11424.5
## Chisq= 6110.01 on 12 degrees of freedom, p= 0
## Number of Newton-Raphson Iterations: 5
## n= 8909

# vamos agora aplicar a selecao stepwise baseada no AIC
modelo_final_exp <- stepAIC(modelo_inicial_exp, direction = "both")

## Start: AIC=16765.07
## Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE + CIRURGIA + RADIO + QUIMIO +
## ECGRUP
##
## Df AIC
## <none> 16765
## - ANODIAG 2 16768
## - IDADE 1 16768
## - RADIO 1 16885
## - CIRURGIA 1 17912
## - QUIMIO 1 18029
## - ECGRUP 6 19075

# Resumo do modelo final
summary(modelo_final_exp)

##
## Call:
## survreg(formula = Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE + CIRURGIA +
## RADIO + QUIMIO + ECGRUP, data = dados, dist = "exponential")
## Value Std. Error z p
## (Intercept) 1.76826 0.41550 4.26 2.1e-05
## ANODIAG2015 -0.07605 0.02949 -2.58 0.00992
## ANODIAG2016 -0.03957 0.02950 -1.34 0.17971
## IDADE -0.00246 0.00110 -2.23 0.02604
## CIRURGIA1 1.35088 0.04518 29.90 < 2e-16
## RADIO1 0.28949 0.02664 10.87 < 2e-16
## QUIMIO1 0.98372 0.02654 37.07 < 2e-16
## ECGRUPI -0.11603 0.41283 -0.28 0.77866
## ECGRUPII -1.21322 0.41335 -2.94 0.00333
## ECGRUPIII -1.92650 0.41041 -4.69 2.7e-06
## ECGRUPIV -2.45918 0.40980 -6.00 2.0e-09
## ECGRUPX -2.02459 0.41312 -4.90 9.6e-07
## ECGRUPY -1.40598 0.41227 -3.41 0.00065
##
## Scale fixed at 1
##
## Exponential distribution
## Loglik(model)= -8369.5 Loglik(intercept only)= -11424.5
## Chisq= 6110.01 on 12 degrees of freedom, p= 0

```

```
## Number of Newton-Raphson Iterations: 5
## n= 8909
AIC(modelo_final_exp)

## [1] 16765.07
# modelo inicial com todas as variaveis usando a distribuicao weibull
modelo_inicial_w <- survreg(Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE +
                           CIRURGIA + RADIO + QUIMIO + ECGRUP,
                           data = dados, dist = "weibull")

summary(modelo_inicial_w)

##
## Call:
## survreg(formula = Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE + CIRURGIA +
##       RADIO + QUIMIO + ECGRUP, data = dados, dist = "weibull")
##
##              Value Std. Error      z      p
## (Intercept)  1.89628    0.51418  3.69 0.00023
## ANODIAG2015 -0.05358    0.03648 -1.47 0.14188
## ANODIAG2016 -0.00623    0.03649 -0.17 0.86435
## IDADE       -0.00223    0.00138 -1.62 0.10483
## CIRURGIA1    1.50651    0.05611 26.85 < 2e-16
## RADIO1       0.31307    0.03293  9.51 < 2e-16
## QUIMIO1      1.07447    0.03286 32.69 < 2e-16
## ECGRUPI      -0.16246    0.51061 -0.32 0.75036
## ECGRUPII     -1.44507    0.51137 -2.83 0.00472
## ECGRUPIII    -2.21368    0.50779 -4.36 1.3e-05
## ECGRUPIV     -2.78408    0.50706 -5.49 4.0e-08
## ECGRUPX      -2.37715    0.51124 -4.65 3.3e-06
## ECGRUPY      -1.68049    0.51012 -3.29 0.00099
## Log(scale)   0.21265    0.00915 23.24 < 2e-16
##
## Scale= 1.24
##
## Weibull distribution
## Loglik(model)= -8064.6   Loglik(intercept only)= -10251.9
## Chisq= 4374.76 on 12 degrees of freedom, p= 0
## Number of Newton-Raphson Iterations: 5
## n= 8909
# vamos agora aplicar a selecao stepwise baseada no AIC
modelo_final_w <- stepAIC(modelo_inicial_w, direction = "both")

## Start:  AIC=16157.12
## Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE + CIRURGIA + RADIO + QUIMIO +
##       ECGRUP
##
##              Df    AIC
## - ANODIAG    2 16156
## <none>         16157
## - IDADE      1 16158
## - RADIO      1 16248
## - CIRURGIA   1 17057
## - QUIMIO     1 17135
```

```
## - ECGRUP      6 17894
##
## Step:  AIC=16155.67
## Surv(TEMPO, CENSURA) ~ IDADE + CIRURGIA + RADIO + QUIMIO + ECGRUP
##
##           Df    AIC
## <none>      16156
## - IDADE      1 16156
## + ANODIAG    2 16157
## - RADIO      1 16246
## - CIRURGIA   1 17059
## - QUIMIO     1 17134
## - ECGRUP     6 17890
```

```
# Resumo do modelo final
summary(modelo_final_w)
```

```
##
## Call:
## survreg(formula = Surv(TEMPO, CENSURA) ~ IDADE + CIRURGIA + RADIO +
##      QUIMIO + ECGRUP, data = dados, dist = "weibull")
##           Value Std. Error      z      p
## (Intercept)  1.87806     0.51390  3.65 0.00026
## IDADE        -0.00226     0.00138 -1.65 0.09992
## CIRURGIA1     1.50881     0.05611 26.89 < 2e-16
## RADIO1        0.31315     0.03293  9.51 < 2e-16
## QUIMIO1       1.07409     0.03287 32.68 < 2e-16
## ECGRUPI      -0.16492     0.51079 -0.32 0.74679
## ECGRUPII     -1.44504     0.51156 -2.82 0.00473
## ECGRUPIII    -2.21437     0.50798 -4.36 1.3e-05
## ECGRUPIV     -2.78252     0.50724 -5.49 4.1e-08
## ECGRUPX      -2.37895     0.51141 -4.65 3.3e-06
## ECGRUPY      -1.67970     0.51030 -3.29 0.00100
## Log(scale)   0.21304     0.00914 23.31 < 2e-16
##
## Scale= 1.24
##
## Weibull distribution
## Loglik(model)=-8065.8  Loglik(intercept only)= -10251.9
## Chisq= 4372.21 on 10 degrees of freedom, p= 0
## Number of Newton-Raphson Iterations: 5
## n= 8909
```

```
AIC(modelo_final_w)
```

```
## [1] 16155.67
```

```
# modelo inicial com todas as variaveis usando a distribuicao log-normal
modelo_inicial_ln <- survreg(Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE +
      CIRURGIA + RADIO + QUIMIO + ECGRUP,
      data = dados, dist = "lognormal")

summary(modelo_inicial_ln)
```

```
##
## Call:
```



```
## survreg(formula = Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE + CIRURGIA +
##      RADIO + QUIMIO + ECGRUP, data = dados, dist = "lognormal")
##              Value Std. Error      z      p
## (Intercept)  0.64904    0.35432  1.83 0.06698
## ANODIAG2015  0.02673    0.03740  0.71 0.47482
## ANODIAG2016  0.06511    0.03725  1.75 0.08049
## IDADE        -0.00142    0.00145 -0.98 0.32786
## CIRURGIA1    1.45362    0.04691 30.99 < 2e-16
## RADIO1       0.40586    0.03393 11.96 < 2e-16
## QUIMIO1      1.25617    0.03339 37.62 < 2e-16
## ECGRUPI      0.03938    0.34539  0.11 0.90923
## ECGRUPII     -1.16680    0.34822 -3.35 0.00081
## ECGRUPIII    -1.79683    0.34382 -5.23 1.7e-07
## ECGRUPIV     -2.31711    0.34244 -6.77 1.3e-11
## ECGRUPX      -2.26762    0.34909 -6.50 8.3e-11
## ECGRUPY      -1.60496    0.34694 -4.63 3.7e-06
## Log(scale)   0.31797    0.00880 36.12 < 2e-16
##
## Scale= 1.37
##
## Log Normal distribution
## Loglik(model)= -7607.1   Loglik(intercept only)= -9674
##  Chisq= 4133.73 on 12 degrees of freedom, p= 0
## Number of Newton-Raphson Iterations: 4
## n= 8909
```

```
# vamos agora aplicar a seleção stepwise baseada no AIC
modelo_final_ln <- stepAIC(modelo_inicial_ln, direction = "both")
```

```
## Start:  AIC=15242.2
## Surv(TEMPO, CENSURA) ~ ANODIAG + IDADE + CIRURGIA + RADIO + QUIMIO +
##      ECGRUP
##
##              Df    AIC
## - IDADE        1 15241
## - ANODIAG       2 15241
## <none>          15242
## - RADIO        1 15382
## - CIRURGIA     1 16184
## - QUIMIO       1 16534
## - ECGRUP       6 16777
##
## Step:  AIC=15241.16
## Surv(TEMPO, CENSURA) ~ ANODIAG + CIRURGIA + RADIO + QUIMIO +
##      ECGRUP
##
##              Df    AIC
## - ANODIAG       2 15240
## <none>          15241
## + IDADE         1 15242
## - RADIO         1 15383
## - CIRURGIA     1 16207
## - QUIMIO       1 16567
## - ECGRUP       6 16779
##
```

```
## Step: AIC=15240.2
## Surv(TEMPO, CENSURA) ~ CIRURGIA + RADIO + QUIMIO + ECGRUP
##
##           Df    AIC
## <none>      15240
## + ANODIAG   2 15241
## + IDADE     1 15241
## - RADIO     1 15382
## - CIRURGIA  1 16205
## - QUIMIO    1 16565
## - ECGRUP    6 16779

# Resumo do modelo final
summary(modelo_final_ln)

##
## Call:
## survreg(formula = Surv(TEMPO, CENSURA) ~ CIRURGIA + RADIO + QUIMIO +
##          ECGRUP, data = dados, dist = "lognormal")
##           Value Std. Error      z      p
## (Intercept)  0.5900      0.3416  1.73 0.08412
## CIRURGIA1    1.4572      0.0464 31.37 < 2e-16
## RADIO1       0.4073      0.0338 12.03 < 2e-16
## QUIMIO1      1.2603      0.0331 38.12 < 2e-16
## ECGRUPI      0.0305      0.3451  0.09 0.92950
## ECGRUPII    -1.1763      0.3479 -3.38 0.00072
## ECGRUPIII   -1.8026      0.3436 -5.25 1.5e-07
## ECGRUPIV    -2.3211      0.3423 -6.78 1.2e-11
## ECGRUPX     -2.2752      0.3489 -6.52 7.0e-11
## ECGRUPY     -1.6067      0.3468 -4.63 3.6e-06
## Log(scale)   0.3180      0.0088 36.12 < 2e-16
##
## Scale= 1.37
##
## Log Normal distribution
## Loglik(model)=-7609.1  Loglik(intercept only)= -9674
## Chisq= 4129.73 on 9 degrees of freedom, p= 0
## Number of Newton-Raphson Iterations: 4
## n= 8909

AIC(modelo_final_ln)

## [1] 15240.2

# AIC do modelo exponencial = 16765
# AIC do modelo weibull = 16155
# AIC do modelo log-normal = 15240
# o modelo escolhido e o com distribuicao log-normal
# as variaveis IDADE e ANODIAG nao sao estatisticamente significativas
# para explicar a variavel tempo ate a morte

modelo_final <- modelo_final_ln

# intervalo de confianca
confint(modelo_final) # de 95% de Wald
```

```
##           2.5 %      97.5 %
## (Intercept) -0.07948421  1.2595509
## CIRURGIA1    1.36617296  1.5482589
## RADIO1       0.34090243  0.4735906
## QUIMIO1      1.19550692  1.3250909
## ECGRUPI      -0.64588137  0.7069474
## ECGRUPII     -1.85815177 -0.4944163
## ECGRUPIII    -2.47598584 -1.1292032
## ECGRUPIV     -2.99191680 -1.6502916
## ECGRUPX      -2.95895318 -1.5913926
## ECGRUPY      -2.28640769 -0.9270419
```

```
#confint(modelo_final, level = .9) # de 90% de Wald
```

```
# Adequacao do modelo ajustado
```

```
matriz_modelo <- model.matrix(~ CIRURGIA + RADIO + QUIMIO + ECGRUP, data = dados)
head(matriz_modelo)
```

```
##      (Intercept) CIRURGIA1 RADIO1 QUIMIO1 ECGRUPI ECGRUPII ECGRUPIII ECGRUPIV
## 1             1         0      1         1         0         0         1         0
## 2             1         0      0         0         1         0         0         0
## 3             1         0      0         0         0         0         0         1
## 4             1         0      0         0         0         0         0         1
## 5             1         0      0         1         0         0         1         0
## 6             1         0      0         1         0         0         0         1
##      ECGRUPX ECGRUPY
## 1          0        0
## 2          0        0
## 3          0        0
## 4          0        0
## 5          0        0
## 6          0        0
```

```
x_beta <- matriz_modelo %*% modelo_final$coefficients
```

```
cox_snell_modelo_final <- - log(1 - pnorm((log(dados$TEMPO)-x_beta) / modelo_final$scale) )
```

```
summary(cox_snell_modelo_final)
```

```
##           V1
## Min.      :0.000001
## 1st Qu.:0.267995
## Median :0.574661
## Mean    :0.771428
## 3rd Qu.:1.058726
## Max.    :5.829345
```

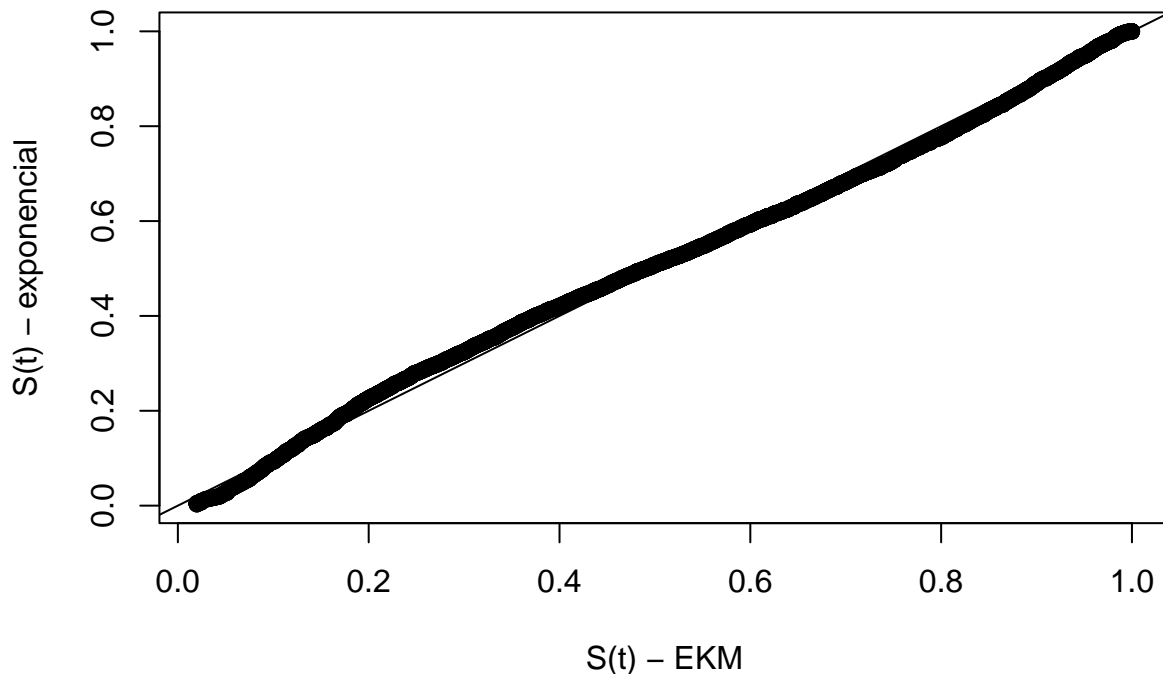
```
# grafico
```

```
ekm_cos_snell_final <- survfit(Surv(cox_snell_modelo_final, dados$CENSURA) ~ 1)
```

```
exp_cox_snell_modelo_final <- exp(- ekm_cos_snell_final$time)
```

```
plot(ekm_cos_snell_final$surv, exp_cox_snell_modelo_final, main= "Resíduos de Cox-Snell", ylab="S(t) - e",
abline(a=0, b=1))
```

## Resíduos de Cox-Snell



*# Encontramos um modelo que se ajustou bem aos dados*

A análise feita até o momento mostrou que o modelo log-normal com as covariáveis cirurgia, radio, quimio e ecrup se ajustaram bem aos dados. Com isso, podemos obter algumas interpretações.

Seguindo a análise anterior, vamos utilizar a função de sobrevivência e a razão do tempo mediano para as interpretações, com o objetivo de comparar dois pacientes, sendo que o paciente 1 fez cirurgia, fez radio, fez quimio e é ECGRUP = II, enquanto que o paciente 2 fez cirurgia, fez radio, fez quimio e é do ECGRUP = IV, portanto a diferença é apenas o estadiamento.

*# Interpretacoes*

*# Considere dois pacientes*

*# paciente 1 - Fez cirurgia, fez radio, fez quimio e ECGRUP = II - x=(1, 1, 1, 1, 0, 1, 0, 0, 0, 0)*

*# paciente 2 - Fez cirurgia, fez radio, fez quimio e ECGRUP = IV - x=(1, 1, 1, 1, 0, 0, 0, 1, 0, 0)*

*# vamos calcular a funcao de sobrevivencia no tempo de 1 ano - log(1)*

*# Para o paciente 1*

```
x_beta_p1 <- c(1, 1, 1, 1, 0, 1, 0, 0, 0, 0) %*% modelo_final$coefficients
```

```
sobre_p1 <- pnorm((-log(1)+x_beta_p1) / modelo_final$scale)
```

```
sobre_p1
```

```
##           [,1]
```

```
## [1,] 0.9676297
```

*# Para o paciente 2*

```
x_beta_p2 <- c(1, 1, 1, 1, 0, 0, 0, 1, 0, 0) %*% modelo_final$coefficients
```

```
sobre_p2 <- pnorm((-log(1)+x_beta_p2) / modelo_final$scale)
```

```

sobre_p2

##           [,1]
## [1,] 0.8447235
# vamos calcular a funcao de sobrevivencia no tempo de 10 ano - log(10)

# Para o paciente 1
sobre_p1 <- pnorm((-log(10)+x_beta_p1) / modelo_final$scale)
sobre_p1

##           [,1]
## [1,] 0.5681483
# Para o paciente 2
sobre_p2 <- pnorm((-log(10)+x_beta_p2) / modelo_final$scale)
sobre_p2

##           [,1]
## [1,] 0.2542034
# MTTF - tempo medio
# modelo log-normal - exp(mu + sigma2/2)
# mu = x%%beta
# sigma = scale

# Para o paciente 1
mttf_1 <- exp(x_beta_p1 + modelo_final$scale^2 / 2)
mttf_1

##           [,1]
## [1,] 32.55556
# Para o paciente 2
mttf_2 <- exp(x_beta_p2 + modelo_final$scale^2 / 2)
mttf_2

##           [,1]
## [1,] 10.36182
# razao dos tempos medianos

exp(c(1, 1, 1, 1, 0, 1, 0, 0, 0, 0) %% modelo_final$coefficients -
    c(1, 1, 1, 1, 0, 0, 0, 1, 0, 0) %% modelo_final$coefficients)

##           [,1]
## [1,] 3.141876
# calculo do tempo mediano

paciente1 <- data.frame(
  CIRURGIA = factor(1, levels = levels(dados$CIRURGIA)),
  RADIO = factor(1, levels = levels(dados$RADIO)),
  QUIMIO = factor(1, levels = levels(dados$QUIMIO)),
  ECGRUP = factor("II", levels = levels(dados$ECGRUP))
)

predict(modelo_final, newdata = paciente1, type = "quantile", p = 0.5)

```

```
##          1
## 12.6608

paciente2 <- data.frame(
  CIRURGIA = factor(1, levels = levels(dados$CIRURGIA)),
  RADIO = factor(1, levels = levels(dados$RADIO)),
  QUIMIO = factor(1, levels = levels(dados$QUIMIO)),
  ECGRUP = factor("IV", levels = levels(dados$ECGRUP))
)

predict(modelo_final, newdata = paciente2, type = "quantile", p = 0.5)
```

```
##          1
## 4.029694
# verificando a razao do tempo mediano

12.66/4.029
```

```
## [1] 3.142219
# podemos utilizar a funcao predict para quaisquer percentil

# Percentil 0.25

predict(modelo_final, newdata = paciente1, type = "quantile", p = 0.25)
```

```
##          1
## 5.010404

predict(modelo_final, newdata = paciente2, type = "quantile", p = 0.25)
```

```
##          1
## 1.594717
```

A função de sobrevivência do paciente 1 no tempo de 1 ano é igual a 0,97, enquanto que do paciente 2 é igual a 0,84. Ou seja, espera-se que 97% dos pacientes que apresentam as mesmas características do paciente 1 estejam vivos após 1 ano do diagnóstico, já para pacientes com características iguais ao do paciente 2, é esperado 84%.

Quando analisamos as funções de sobrevivência no tempo de 10 anos, temos 0,57 e 0,25, para os pacientes 1 e 2, respectivamente.

Obtemos que a razão do tempo mediano do paciente 1 pelo paciente 2 é aproximadamente igual a 3,14, ou seja, o tempo mediano do paciente 1 (12,66 anos) é 3,14 vezes o tempo mediano do paciente 2 (4,03 anos).

Além disso, calculamos o tempo percentil 0,25, ou seja, para pacientes com as mesmas covariáveis que o paciente 1, é esperado que 75% deles sobrevivam além do tempo de 5,01 anos. Já para pacientes com as mesmas covariáveis que o paciente 2, é esperado que 75% deles sobrevivam além do tempo de 1,59 anos.