



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE CIÊNCIAS CONTÁBEIS E ATUARIAIS

Lista 3

Thales Rousseau Braga Reis Silva

1. Laringe

Um estudo clínico realizado com 90 pacientes do sexo masculino diagnosticados com câncer de laringe. Os dados coletados foram idade dos pacientes no diagnóstico, estágio da doença; (I = Tumor primário - 33 pacientes, II = Envolvimento dos nódulos - 17 pacientes, III = Metástase - 27 pacientes e IV = Combinações dos estágios anteriores - 13 pacientes), além do tempo de acompanhamento e presença de censura.

Tabela 1: Tempos de sobrevivência observados no estudo de câncer de laringe - Grupo 1

Estágio	Idade	Tempo de Acompanhamento
I	43	3.5+
	45	2.4+
	47	5.9
	48	4.5
	51	3.2
	52	4.0+
	53	1.3+
	56	8.1
	57	2.5
	58	3.2+
	58	5.9
	58	9.6
	60	3.5+
	61	6.6
	63	4.0+
	63	3.3
	64	6.2
	66	7.0
	67	6.5+
	68	7.4+
	68	4.5
	68	10.7
	70	5.5
	73	7.4
	73	8.1
	75	6.0+
	76	3.3+
	77	0.6+
	77	6.+
	77	6.1
	79	6.5
	81	5.3+
	86	4.3+

Tabela 2: Tempos de sobrevivência observados no estudo de câncer de laringe - Grupo 2

Estágio	Idade	Tempo de Acompanhamento
II	47	4.3
	50	7.5
	51	3.3
	53	7.6
	61	9.3
	62	7.0+
	63	2.0+
	64	1.8+
	64	4.3
	66	5.0
	67	2.6
	70	3.6+
	71	6.2
	72	3.6
	74	2.2+
	81	4.0+
	86	0.2+

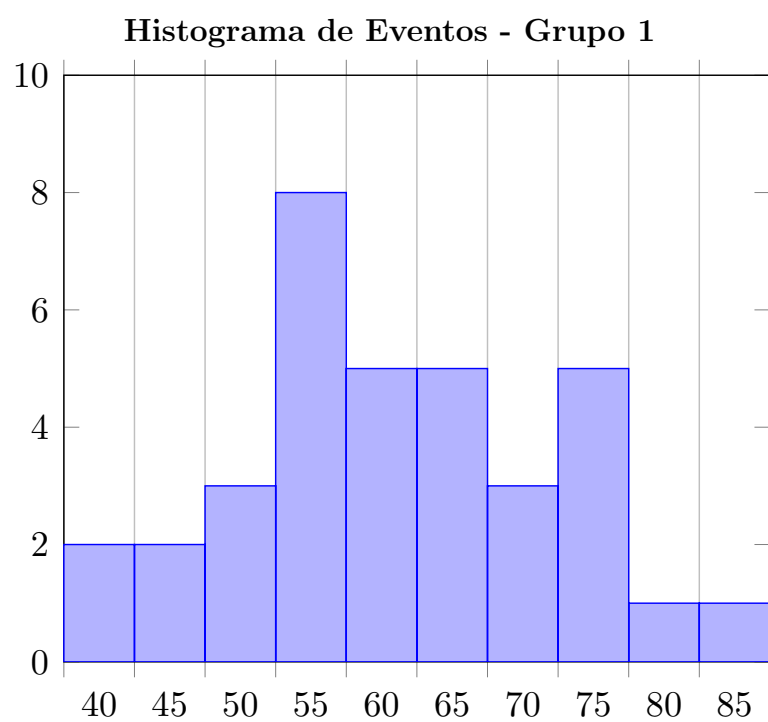
Tabela 3: Tempos de sobrevivência observados no estudo de câncer de laringe - Grupo 3

Estágio	Idade	Tempo de Acompanhamento
III	49	0.3+
	49	1.0+
	49	5.0
	51	10.1
	52	3.7
	53	1.9+
	54	3.2+
	54	4.8
	57	0.5+
	59	5.0+
	60	1.3+
	63	4.8
	64	1.6+
	65	6.4+
	65	6.5
	66	4.5
	68	7.8+
	69	5.1
	69	9.3
	70	6.3+
	71	0.3+
	72	1.9+
	74	1.8+
	78	8.0
	79	0.7+
	81	3.5+
	82	0.8+

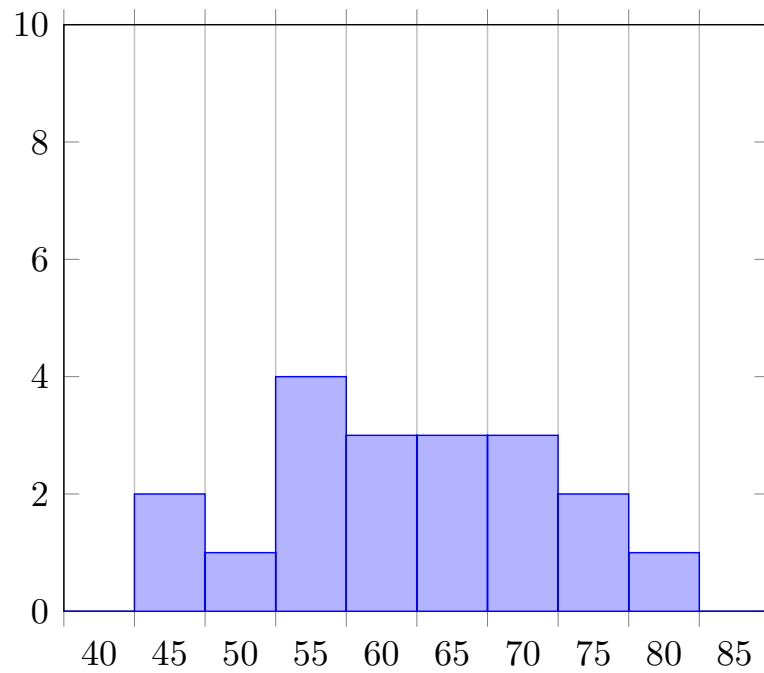
Tabela 3: Tempos de sobrevivência observados no estudo de câncer de laringe - Grupo 4

Estágio	Idade	Tempo de Acompanhamento
IV	41	1.0+
	48	4.3
	62	2.3+
	65	0.1+
	65	0.8+
	68	1.5+
	69	2.0+
	71	0.3+
	71	3.6+
	74	2.9
	76	0.4+
	78	0.8+
	84	3.8+

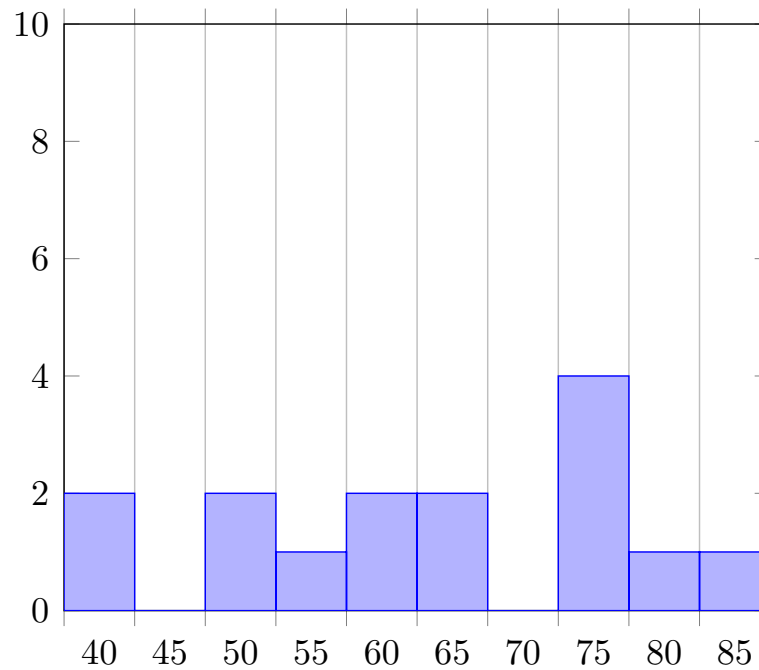
2. Histograma dos dados



Histograma de Sobrevivência - Grupo 1

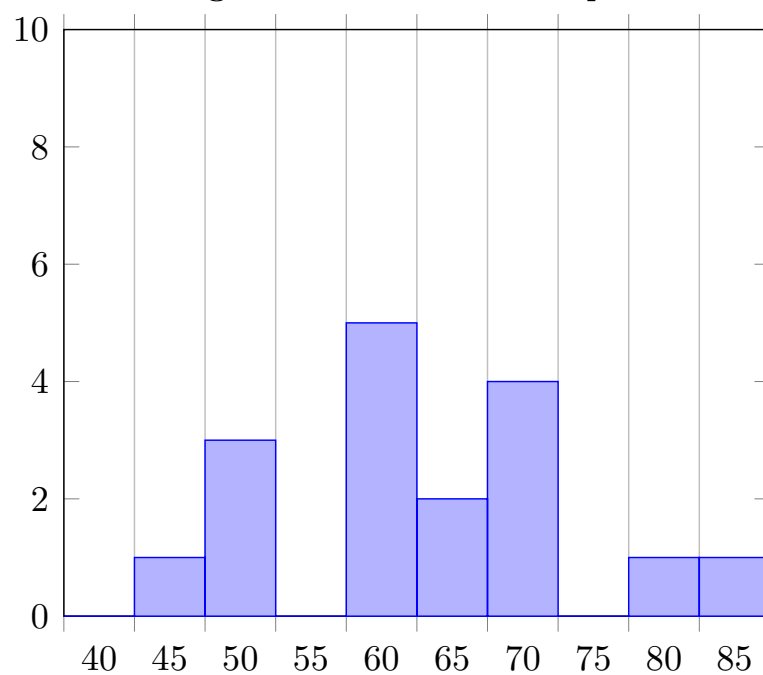


Histograma de Censuras - Grupo 1

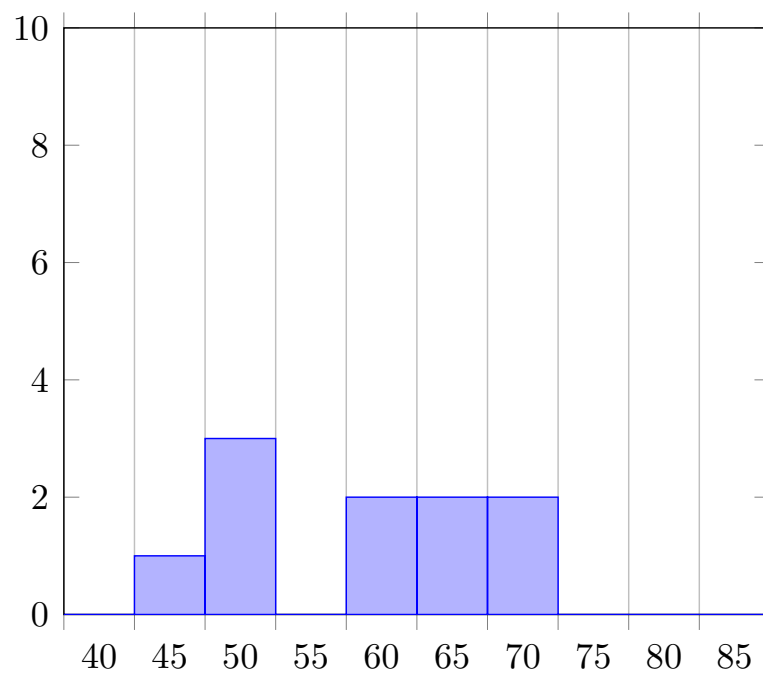


Os dados do grupo com Estágio I de câncer de laringe quando analisado em formato histograma, as observações ficam concentradas na porção central, entre as idades de 55 e 75 anos. Esse padrão ocorre pela de presença do evento morte dos pacientes entre essas faixas etárias, já que os dados de censura se distribuem uniformemente entre as todas as idades.

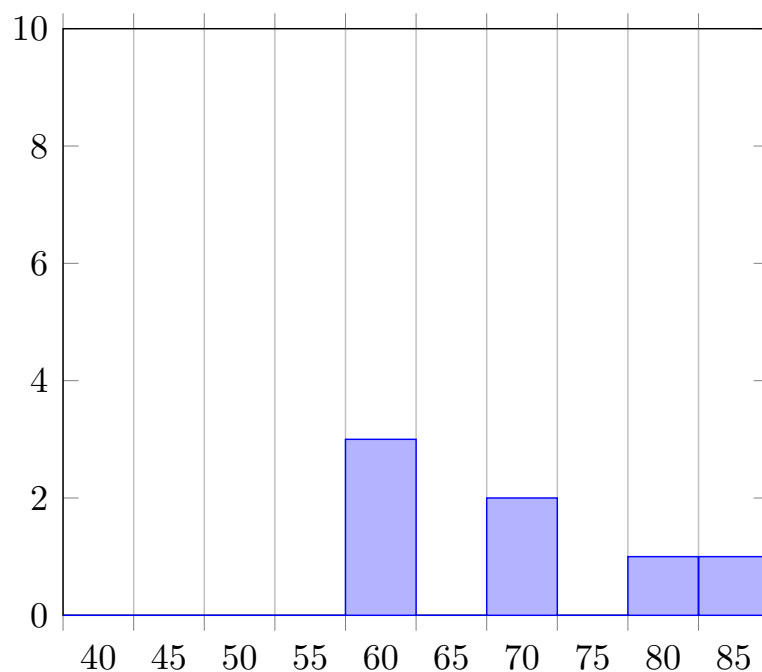
Histograma de Eventos - Grupo 2



Histograma de Sobrevivência - Grupo 2

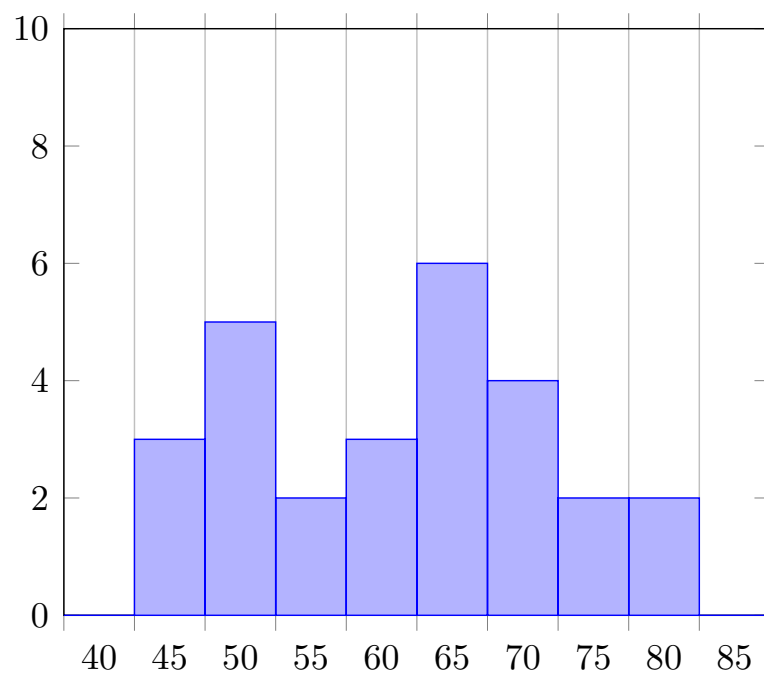


Histograma de Censuras - Grupo 2

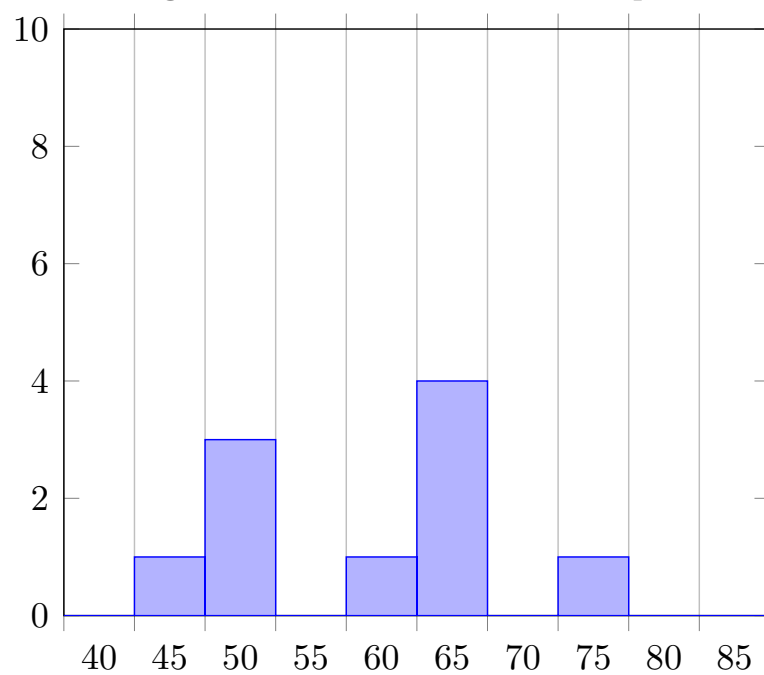


As observações do grupo de pacientes com Estágio II de câncer de laringe, embora se concentrem levemente ao centro. Assim, analisando o histograma de ocorrência do evento morte, as observações parecem acontecer de maneira constante entre as faixas etárias até 75 anos, enquanto os eventos de censura ocorrem a partir dos 60 anos, assim ao visualizar o histograma geral, essas duas curvas se sobrepõem criando a maior ocorrência de eventos no centro do histograma.

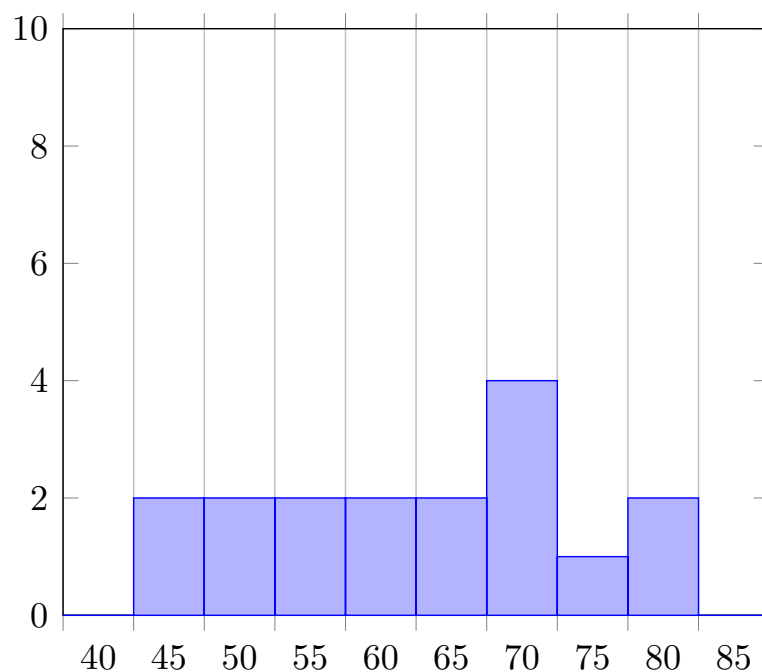
Histograma de Eventos - Grupo 3



Histograma de Sobrevivência - Grupo 3

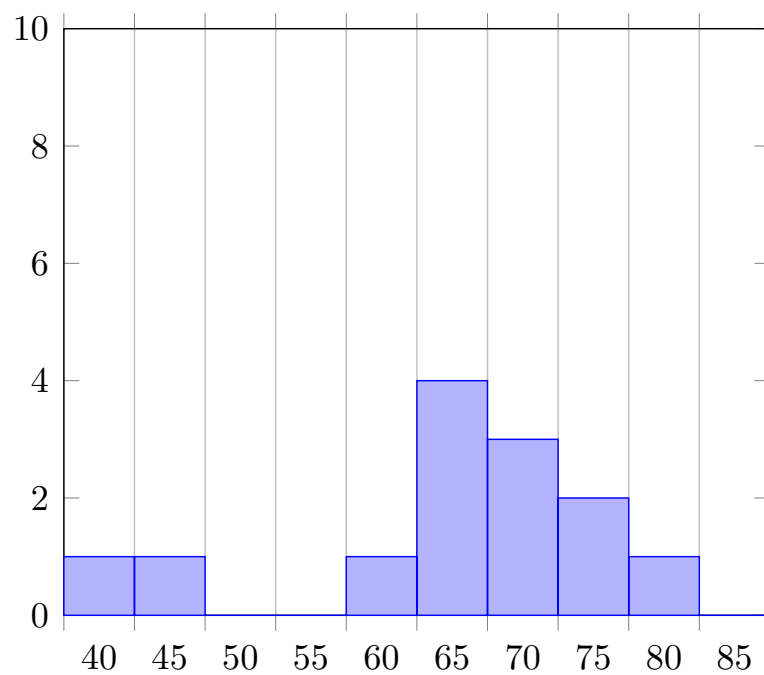


Histograma de Censuras - Grupo 3

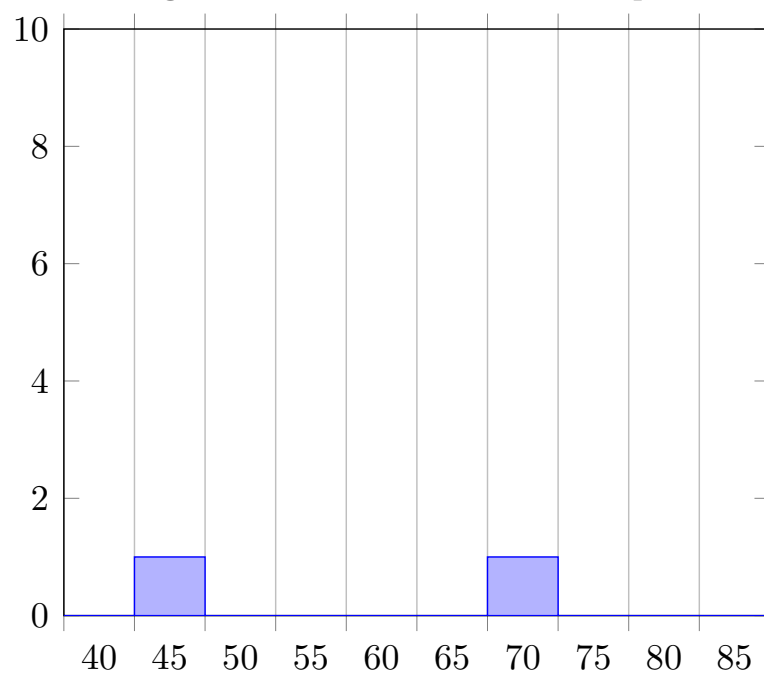


As observações dos ocorrência de eventos do grupo de Estágio III de câncer de laringe se aproxima de uma distribuição normal. Assim, separar as censuras da ocorrência do falecimentos dos pacientes, as observações das censuras se distribuem igualmente entre as faixas etárias, desse modo, o evento morte parece ser responsável pela distribuição dos dados gerias se aproximarem de uma normal.

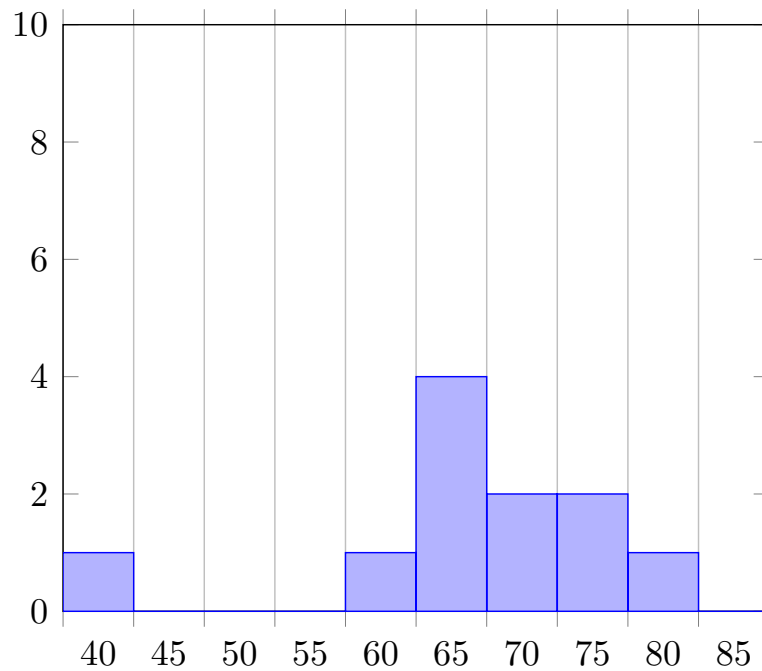
Histograma de Eventos - Grupo 4



Histograma de Sobrevivência - Grupo 4



Histograma de Censuras - Grupo 4



Nas observações dos grupo de Estágio IV de câncer de laringe, os eventos de morte quase não foram observados, possivelmente, pelo número reduzido de pacientes em relação aos demais grupos. Assim, a curva geral é muito similar a curva de censuras.

3. Ajuste de Modelos

3.1 Modelo 1

	Coef	Exp(Coef)	se(Coef)	z	Pr(> z)
Fator (Estágio 2)	0.06576	1.06797	0.45844	0.143	0.8859
Fator (Estágio 3)	0.61206	1.84423	0.35520	1.723	0.0849
Fator (Estágio4)	1.72284	5.60040	0.41966	4.105	4.04e-05 ***

	Exp(Coef)	Exp(-Coef)	IC _{95%} Inferior	IC _{95%} Superior
Fator (Estágio 2)	1.068	0.9364	0.4348	2.623
Fator (Estágio 3)	1.844	0.5422	0.9193	3.700
Fator (Estágio 4)	5.600	0.1786	2.4604	12.748

Concordância = 0.668 (se = 0.037)

Teste de Razão de Verossimilhança = 16.26 com 3 GL, p=0.001

Teste Wald = 18.95 com 3 GL, p=3e-04

Teste Logrank = 22.46 com 3 GL, p=5e-05

O modelo I considera apenas a variável Estágio da doença. Assim, submetendo o modelo aos testes, obtemos o valor de 0,668 como coeficiente de correlação de concordância. Já o Teste de Razão de Verossimilhança foi de 16,26 para um valor-p de 0,001. O teste Wald com coeficiente de 18,95 para um valor-p de 0,0003 e o teste de logrank com 22,46 de coeficiente um p valor de 0,00005, valores obtidos com 3 graus de liberdade e com significância estatística.

3.2 Modelo 2

	Coef	Exp(Coef)	se(Coef)	z	Pr(> z)
Fator (Estágio 2)	0.13856	1.14862	0.46231	0.300	0.764
Fator (Estágio 3)	0.63835	1.89335	0.35608	1.793	0.073
Fator (Estágio 4)	1.69306	5.43607	0.42221	4.010	6.07e-05 ***
Idade	0.01890	1.01908	0.01425	1.326	0.185

	Exp(Coef)	Exp(-Coef)	IC_{95%} Inferior	IC_{95%} Superior
Fator (Estágio 2)	1.149	0.8706	0.4642	2.842
Fator (Estágio 3)	1.893	0.5282	0.9422	3.805
Fator (Estágio 4)	5.436	0.1840	2.3763	12.436
Idade	1.019	0.9813	0.9910	1.048

Concordância = 0.682 (se = 0.039)

Teste de Razão de Verossimilhança = 18.07 com 4 GL, p=0.001

Teste Wald = 20.82 com 4 GL, p=3e-04

Teste Logrank = 24.33 com 4 GL, p=7e-05

O modelo 2 acrescenta a variável Idade ao modelo. Ao rodarmos os teste no modelo, obtemos os valores de concordância de 0,682, no Teste de Razão de Verossimilhança, o valor-p continua igual a 0,001, mas com 4 graus de liberdade. Os teste de Wald e Logrank, ficam iguais 0,0003 e 0,00007 com 4 graus de liberdade, respectivamente. Demonstrando a significância da inclusão da variável.

3.3 Modelo 3

	Coef	Exp(Coef)	se(Coef)	z	Pr(> z)
Fator (Estágio 2)	-7.946142	0.000354	3.678209	-2.160	0.0307*
Fator (Estágio 3)	-0.122500	0.884706	2.468331	-0.050	0.9604
Fator (Estágio 2)	0.846986	2.332605	2.425717	0.349	0.7270
Idade	-0.002559	0.997444	0.026051	-0.098	0.9218
Fator (Estágio 2):Idade	0.120254	1.127783	0.052307	2.299	0.0215*
Fator (Estágio 3):Idade	0.011351	1.011416	0.037449	0.303	0.7618
Fator (Estágio 4):Idade	0.013673	1.013767	0.035967	0.380	0.7038

	Exp(Coef)	Exp(-Coef)	IC _{95%} Inferior	IC _{95%} Superior
Fator (Estágio 2)	0.000354	2824.6562	2.619e-07	0.4786
Fator (Estágio 3)	0.884705	1.1303	7.011e-03	111.6467
Fator (Estágio 4)	2.332605	0.4287	2.009e-02	270.7790
Idade	0.997444	1.0026	9.478e-01	1.0497
Fator (Estágio 2):Idade	1.127783	0.8867	1.018e+00	1.2495
Fator (Estágio 3):Idade	1.011416	0.9887	9.398e-01	1.0884
Fator (Estágio 4):Idade	1.013767	0.9864	9.448e-01	1.0878

Concordância = 0.694 (se = 0.039)

Teste de Razão de Verossimilhança = 24.27 com 7 GL, p=0.001

Teste Wald = 24.11 com 7 GL, p=0.001

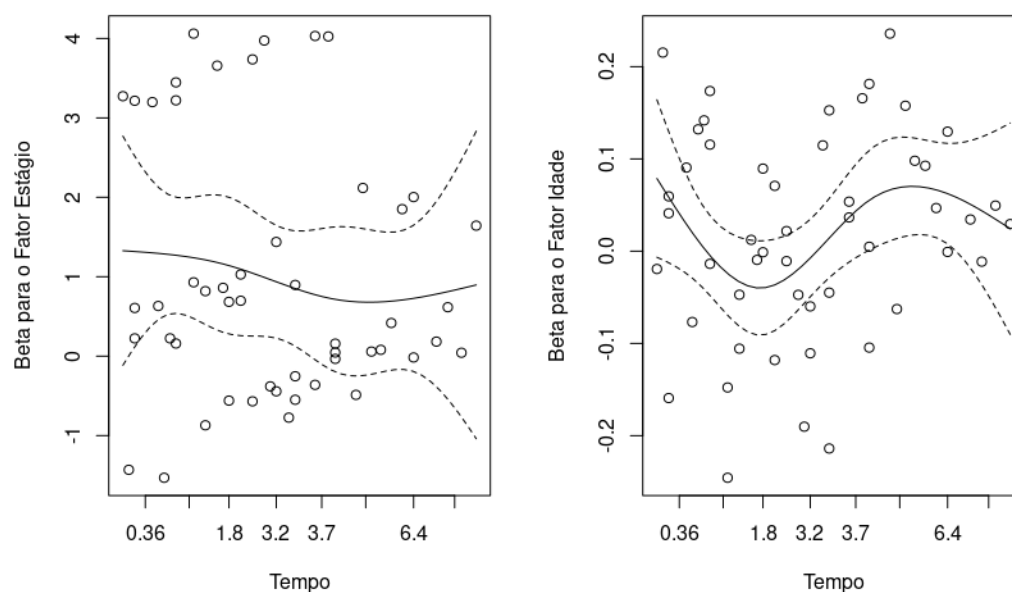
Teste Logrank = 28.59 com 7 df, p=2e-04

O modelo 3 adiciona uma terceira variável teste de Concordância foi de 0,694, o maior entre todos os modelos. Por sua vez, o Teste de Razão de Verossimilhança foi de 24,27 com 7 graus de liberdade para um valor-p igual aos demais modelos, ou seja, 0,001. Ainda, no Teste Wald o coeficiente obtido foi 24,11 para um p-valor de 0,001 e no Teste Logrank o coeficiente foi 28,59 para um p-valor igual a 0,002 com 7 graus de liberdade. Todos os teste apresentando significância estatística. Portanto, a adição da variável tem significância ao modelo.

4. Resíduos de Schoenfeld

4.1 Modelo 2

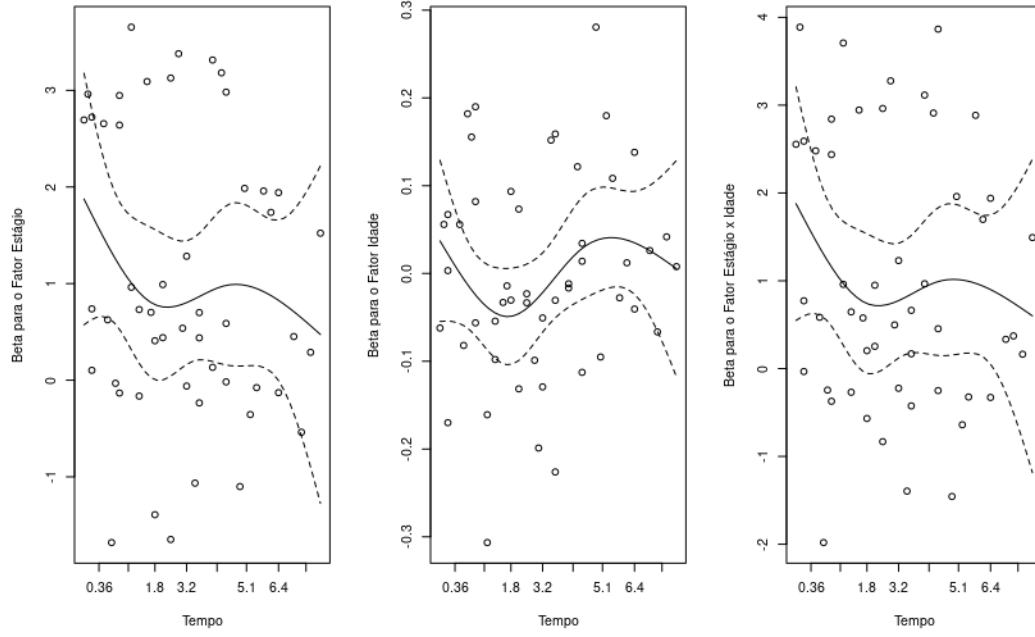
Valores de Beta ao Longo do Tempo



O primeiro gráfico mostra interação do estágio da doença fator tempo com o estágio da doença, já o segundo gráfico introduz a variável Idade. O primeiro gráfico é apresenta uma homocedasticidade maior, o que causa uma variação menor ao longo do tempo que no segundo, indicando um ajuste melhor do valor de beta com base na suposição de proporcionalidade ao longo do tempo. Assim, o fator estágio da doença é mais propocional ao longo do tempo que o fator idade.

4.2 Modelo 3

Valores de Beta ao Longo do Tempo



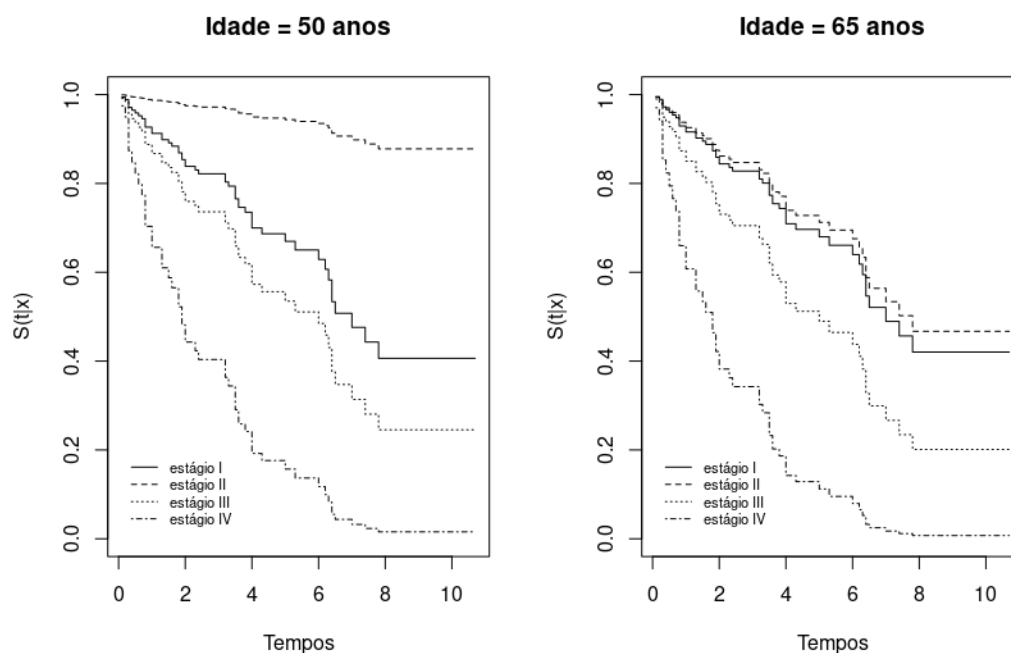
Os valores dos betas ao longo do tempo para as variáveis Estágio e Estágio x Idade tem um formato similar, e sofrem menos variações se comparadas aos betas da variável Idade.

5. Estimativas de $\hat{S}_0(t)$ e $\hat{\Lambda}_0(t)$

Tempos	S_0	H_0
0.1	0.99377	0.00625
0.2	0.98739	0.01269
0.3	0.96737	0.03318
0.4	0.96039	0.04041
0.5	0.95319	0.04794
0.6	0.94596	0.05555
0.7	0.93875	0.06321
0.8	0.91713	0.08650
1.0	0.90154	0.10365
1.3	0.88552	0.12158
1.5	0.87745	0.13073
1.6	0.86907	0.14033
1.8	0.85231	0.15980
1.9	0.83549	0.17974
2.0	0.81848	0.20030
2.2	0.81848	0.20030
2.3	0.80945	0.21140
2.4	0.80004	0.22309
2.5	0.80004	0.22309
2.6	0.80004	0.22309
2.9	0.80004	0.22309
3.2	0.7765	0.24891
3.3	0.76923	0.26236
3.5	0.73805	0.30374
3.6	0.71695	0.33274
3.7	0.71695	0.33274
3.8	0.70498	0.34959

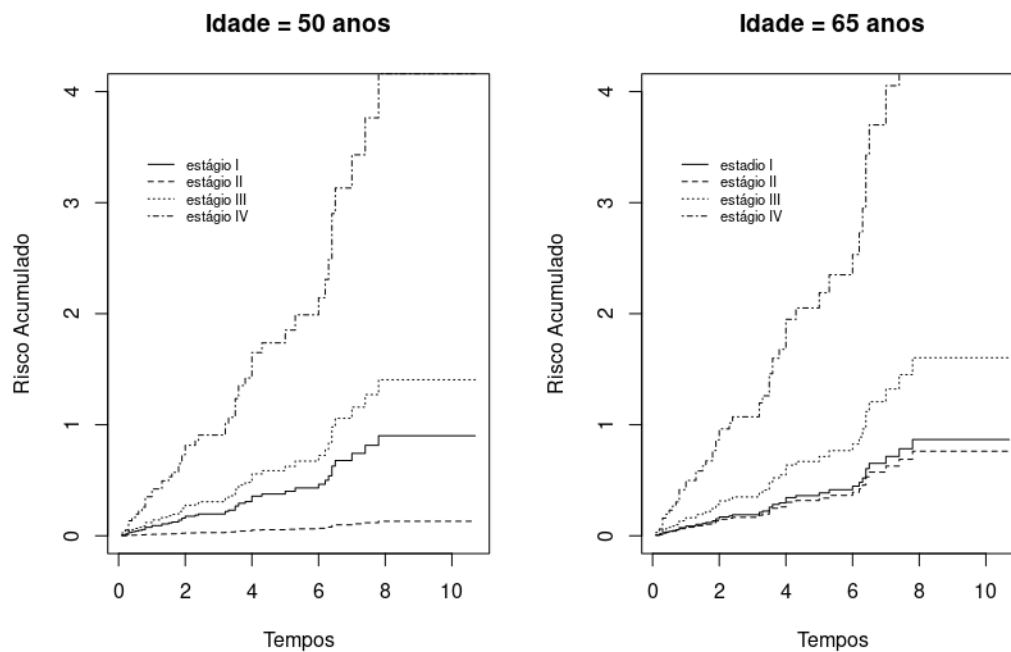
Tempos	S_0	H_0
4.0	0.66650	0.40572
4.3	0.65242	0.42707
4.5	0.65242	0.42707
4.8	0.65242	0.42707
5.0	0.63406	0.45561
5.1	0.63406	0.45561
5.3	0.61308	0.48925
5.5	0.61308	0.48925
5.9	0.61308	0.48925
6.0	0.59024	0.52723
6.1	0.59024	0.52723
6.2	0.56680	0.56775
6.3	0.54126	0.61386
6.4	0.48988	0.71360
6.5	0.46291	0.77022
6.7	0.46291	0.77022
7.0	0.43005	0.84384
7.4	0.39625	0.92571
7.5	0.39625	0.92571
7.6	0.39625	0.92571
7.8	0.35937	1.02341
8.0	0.35937	1.02341
8.1	0.35937	1.02341
9.3	0.35937	1.02341
9.6	0.35937	1.02341
10.1	0.35937	1.02341
10.7	0.35937	1.02341

6. Curvas de Sobrevida



A Curva de Sobrevivência do Estágio I é ligeiramente diferente entre os grupos de 50 e 65 anos, tendo uma maior índice de eventos no primeiro grupo. Já para o Estágio II, a Curva de Sobrevivência é bastante discrepante em favor do grupo de pacientes com 50 anos, sendo a sobrevivência quase o dobro do grupo de 65 anos. Já para os demais grupos, Estágio III e Estágio IV, em ambos os casos a distribuição é parecida tanto para pacientes com 50 anos como para pacientes com 65 anos. Assim, para esses Estágios da doença, a idade não parece ser um componente com grande influência na sobrevivência dos indivíduos. Em ambos os gráficos, os pacientes no Estágio II sobrevivem mais que os pacientes do Estágio I, que por sua vez, sobrevivem mais que os pacientes do Estágio III, que ainda apresentam maior sobrevivência que as pessoas com Estágio IV da doença.

7. Curvas de Taxa de Falhas Acumulada



Os gráficos das Taxas de Falha Acumuladas fornecem conclusões similares as obtidas pelos gráficos das Curvas de Sobrevivência, visto que, que são funções inversas. Assim, os Estágios I, III e IV, tem curvas similares entre si, enquanto o Estágio II, difere bastante entre as idades 50 e 65 anos. Outro ponto importante, é a similaridade entre as Taxas de Falha Acumulada do Estágio I e Estágio II aos 65 anos. Portanto, a mudança no gráfico do Estágio II, sugere que a idade tem uma influência diretamente proporcional quando se trata desse estágio.

8. Códigos no R

```
# Importando bibliotecas
library('survival')

# Definindo diretório de trabalho
setwd(getwd())

# Leitura dos dados
laringe<-read.table("laringe.txt", h=T)
attach(laringe)
```

```

# Ajuste de três modelos
# Modelo 1
fit2<-coxph(Surv(tempo,cens) ~ factor(estagio), data=laringe, x = T, method="breslow")
summary(fit2)
fit2$loglik

# Modelo 2
fit3<- coxph(Surv(tempo,cens) ~ factor(estagio)+ idade, data=laringe, x = T, method="breslow")
summary(fit3)
fit3$loglik

# Modelo 3
fit4<-coxph(Surv(tempo,cens) ~ factor(estagio) + idade + factor(estagio)*idade, data=laringe,
x = T, method="breslow")
summary(fit4)
fit4$loglik

# Resíduos de Schoenfeld, verificação da suposição de riscos proporcionais (Modelo fit 3)
resid(fit3,type="scaledsch")
cox.zph(fit3, transform="identity")
par(mfrow=c(1,2))
plot(cox.zph(fit3), xlab = "Tempo", ylab = c("Beta para o Fator Estágio", "Beta para o
Fator Idade"))
mtext("Valores de Beta", side = 3, line = - 2, outer = TRUE)

# Resíduos de Schoenfeld, verificação da suposição de riscos proporcionais (Modelo fit 4)
resid(fit4,type="scaledsch")
cox.zph(fit4, transform="identity")
par(mfrow=c(1,3))
plot(cox.zph(fit4), xlab = "Tempo", ylab = c("Beta para o Fator Estágio", "Beta para o
Fator Idade", "Beta para o Fator Estágio x Idade"))
mtext("Valores de Beta", side = 3, line = - 2, outer = TRUE)

# Estimando a função sobrevida e a função taxa de falha acumulada Ht<-basehaz(fit4,centered=F)
tempo<-Ht$time
H0<-Ht$hazard
S0<- exp(-H0)
round(cbind(tempo, S0,H0),digits=5)

# Gráfico curvas de sobrevivência por grau da doença e idade - 50 e 65 anos
fit4<-coxph(Surv(tempo,cens) ~ factor(estagio) + idade + factor(estagio)*idade, data=laringe,
x = T, method="breslow")
Ht<-basehaz(fit4,centered=F)
tempo<-Ht$time
H0<-Ht$hazard

```

```

S0<- exp(-H0)
round(cbind(tempos,S0,H0),digits=5)
tt<-sort(tempos)
aux1<-as.matrix(tt)
n<-nrow(aux1)
aux2<-as.matrix(cbind(tempos,S0))
S00<-rep(max(aux2[,2]),n)
for(i in 1:n)
if(tt[i]> min(aux2[,1]))
i1<- aux2[,1]<= tt[i]
S00[i]<-min(aux2[i1,2])
ts0<-cbind(tt,S00)
ts0
b<-fit4$coefficients
id<-50
st1<- S00*(exp(b[4]*id)) # S(t|x) estágio I e idade = 50 anos
st2<- S00*(exp(b[1]+((b[4]+b[5])*id))) # S(t|x) estágio II e idade = 50 anos
st3<- S00*(exp(b[2]+((b[4]+b[6])*id))) # S(t|x) estágio III e idade = 50 anos
st4<- S00*(exp(b[3]+((b[4]+b[7])*id))) # S(t|x) estágio IV e idade = 50 anos
id<- 65
st11<- S00*(exp(b[4]*id)) # S(t|x) estágio I e idade = 65 anos
st21<- S00*(exp(b[1]+((b[4]+b[5])*id))) # S(t|x) estágio II e idade = 65 anos
st31<- S00*(exp(b[2]+((b[4]+b[6])*id))) # S(t|x) estágio III e idade = 65 anos
st41<- S00*(exp(b[3]+((b[4]+b[7])*id))) # S(t|x) estágio IV e idade = 65 anos
par(mfrow=c(1,2))
plot(tt,st1,type="s",ylim=range(c(0,1)),xlab="Tempos",ylab="S(t|x)",lty=1)
lines(tt,st2,type="s",lty=2)
lines(tt,st3,type="s",lty=3)
lines(tt,st4,type="s",lty=4)
legend(0,0.2,lty=c(1,2,3,4),c("estágio I","estágio II","estágio III","estágio IV"), lwd=1,bty="n",cex=
title("Idade = 50 anos")
plot(tt,st11,type="s",ylim=range(c(0,1)),xlab="Tempos",ylab="S(t|x)",lty=1)
lines(tt,st21,type="s",lty=2)
lines(tt,st31,type="s",lty=3)
lines(tt,st41,type="s",lty=4)
legend(0,0.2,lty=c(1,2,3,4),c("estágio I","estágio II","estágio III","estágio IV"), lwd=1,bty="n",cex=
title("Idade = 65 anos")

# Gráfico taxa de falha por grau da doença e idade (50 e 65 anos) Ht1<- -log(st1)
Ht2<- -log(st2)
Ht3<- -log(st3)
Ht4<- -log(st4)
Ht11<- -log(st11)
Ht21<- -log(st21)
Ht31<- -log(st31)
Ht41<- -log(st41)
par(mfrow=c(1,2))

```

```

plot(tt,Ht1,type="s",ylim=range(c(0,4)),xlab="Tempos",ylab="Risco Acumulado",lty=1)
lines(tt,Ht2,type="s",lty=2)
lines(tt,Ht3,type="s",lty=3)
lines(tt,Ht4,type="s",lty=4)
legend(0.5,3.5, lty=c(1,2,3,4),c("estágio I","estágio II","estágio III","estágio IV"), lwd=1,bty="n",ce
title("Idade = 50 anos")
plot(tt,Ht11,type="s",ylim=range(c(0,4)),xlab="Tempos",ylab="Risco Acumulado",lty=1)
lines(tt,Ht21,type="s",lty=2)
lines(tt,Ht31,type="s",lty=3)
lines(tt,Ht41,type="s",lty=4)
legend(0.5,3.5, lty=c(1,2,3,4),c("estadio I","estágio II","estágio III","estágio IV"), lwd=1,bty="n",ce
title("Idade = 65 anos")

```