

Revisão - Modelo de Cox

Estimação, testes e estimação da Função de Sobrevivência

MAE 514 - Introdução à Análise de Sobrevivência e Aplicações

IME-USP



Modelo de Cox - Estimação

Dados: $\{(Z_j, \delta_j, X_j), j = 1, \dots, n\}$, $Z_j = \min(T_j, C_j)$

Verossimilhança Parcial

$$L(\beta) = \prod_{j=1}^n \left(\frac{e^{X_j \beta}}{\sum_{\ell \in \mathcal{R}_j} e^{X_\ell \beta}} \right)^{\delta_j}$$

- Ênfase na posição relativa das falhas com tempos ordenados
- Estimador de máxima verossimilhança *parcial* obtido através da função escore (equação de estimação)
- Solução iterativa (Newton-Raphson)



Modelo de Regressão Semiparamétrico

Cox(1972,1975)

Em sua forma mais simples é um modelo de **riscos proporcionais**

Função de risco - Estrutura semiparamétrica

$$\alpha(t | \mathbf{X}_i) = \alpha_0(t) \exp(\mathbf{X}'_i \beta),$$

$\alpha_0(t)$ arbitrária e não-negativa.

Interpretação: Risco relativo

$$RR(X_1 + 1 | X_1) = e^{\beta_1} \quad (\text{Não depende de } t)$$



Modelo de Cox - Teste de hipóteses

Função de taxa de falha:

$$\alpha(t | \mathbf{X}_i) = \alpha_0(t) \exp(\mathbf{X}'_i \beta),$$

Hipóteses de interesse - Efeito de covariáveis

$$H_0 : \beta = 0$$

$$H_1 : \beta \neq 0$$

Testes assintóticos:

- W : Wald
- R : Rao ou escore
- Λ : Razão de verossimilhanças

Distribuição Assintótica

Sob H_0 ,

$$W, R, \Lambda \xrightarrow{D} \chi_p^2$$



Modelo de Cox - Exemplo - Interpretações

Recorde que

$$RR = e^{\beta}$$

- Para grupo,

$$\widehat{RR}_g = e^{\widehat{\beta}_g} = e^{-1.386} = 0.25 \quad \text{com IC(95\%)} = [0.11, 0.58]$$

isto é, o *risco* de volta dos sintomas para quem toma a droga 6-MP é 25% do *risco* para aqueles recebendo placebo.

- Para logaritmo da contagem de glóbulos brancos,

$$\widehat{RR}_{lg} = e^{\widehat{\beta}_{lg}} = e^{1.691} = 5.42 \quad \text{com IC(95\%)} = [2.81, 10.48]$$

isto é, um aumento de uma unidade no logaritmo da contagem de glóbulos brancos acarreta um aumento no risco de volta dos sintomas de aproximadamente 4.5 vezes (1 + 4.42).



Exemplo - Dados ICESP

Informações obtidas no momento de internação na UTI:

- Idade (anos), Gênero (M/F)
- IMC categorizado (Baixo Peso, Saudável, Sobrepeso)
- Tipo de admissão (Médica, Eletiva, Emergencial)
- Status do câncer (Controlado, Ativo recente, Ativo recaída)
- Extensão do câncer (Limitado, Avançado local, Metástase)
- Cirurgia prévia (Não, Curativa, Paliativa)
- Quimioterapia prévia (Não, Curativa, Paliativa)
- Radioterapia prévia (Não, Curativa, Paliativa)
- Capacidade funcional (0-normal a 4-Preso ao leito)
- Qualidade de vida anterior (Ruim, Razoável, Boa)
- SAPS3cat (Preditor de mortalidade - quanto maior pior)



Exemplo - Dados ICESP

Estudo conduzido no Instituto do Câncer do Estado de São Paulo

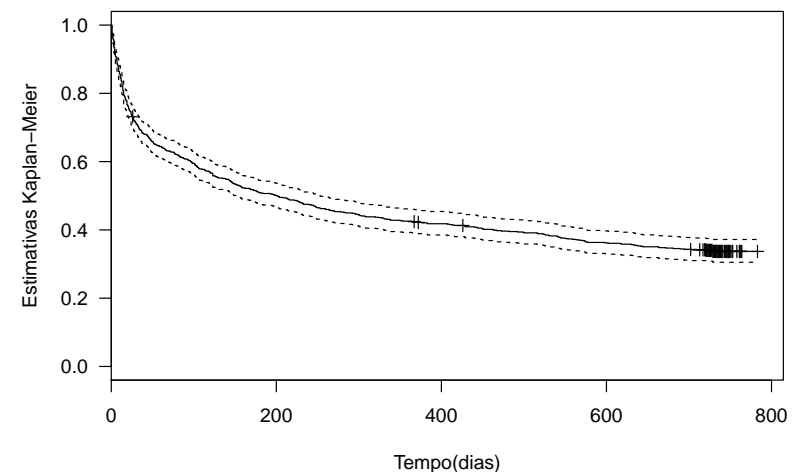
- Avaliar a qualidade de vida de pacientes com câncer internados em UTI
- Combinar índices de qualidade de vida e sobrevida (QALY)

- 793 pacientes oncológicos em estágio avançado
- Início do acompanhamento: Entrada na UTI
- Evento de interesse: Óbito



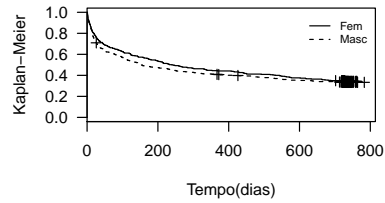
Análise Descritiva Inicial

Sobrevida de Pacientes - ICESP

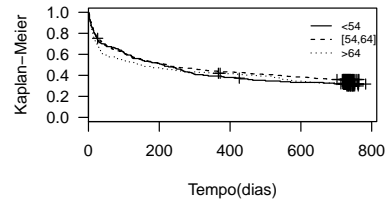


Análise Descritiva Inicial

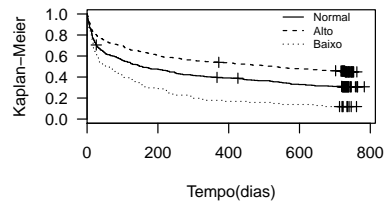
Estimativas por Gênero



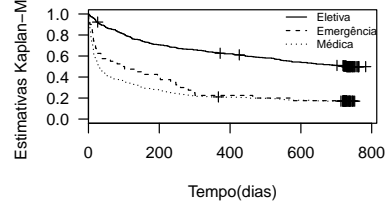
Estimativas por Idade



Estimativas por IMC



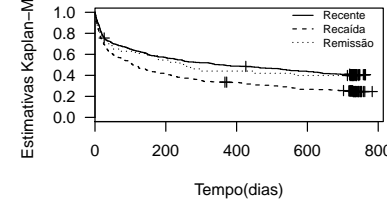
Estimativas por Tipo de Admissão



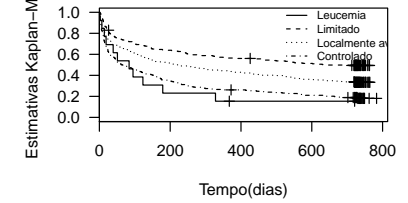
Navigation icons: back, forward, search, etc.

Análise Descritiva Inicial

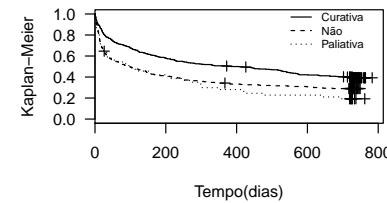
Estimativas por Status do Câncer



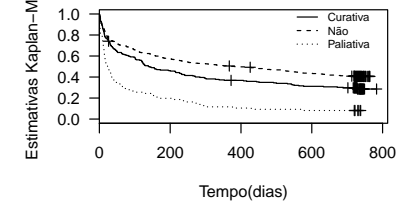
Estimativas por Extensão do Câncer



Estimativas por Cirurgia Prévia



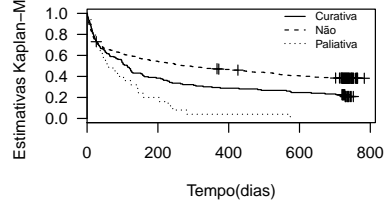
Estimativas por Químio prévia



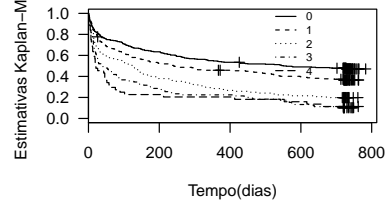
Navigation icons: back, forward, search, etc.

Análise Descritiva Inicial

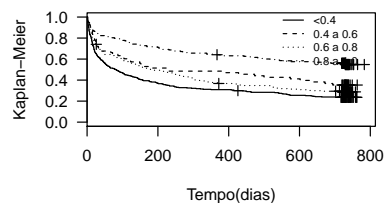
Estimativas por Rádio prévia



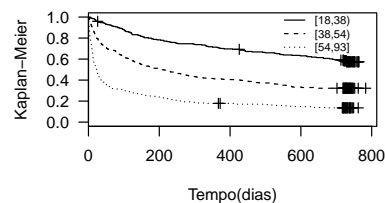
Estimativas por Capacidade Funcional



Estimativas por QV Prévia



Estimativas por SAPS3



Navigation icons: back, forward, search, etc.

Dados ICESP - Avaliação Individual de Fatores

Níveis descritivos (valores p)

Fator	Log-rank	Gehan-Wilcoxon
Gênero	0.369	0.223
Idade	0.137	0.224
IMC	<0.001	<0.001
Tipo de admissão	<0.001	<0.001
Status do câncer	<0.001	<0.001
Extensão do câncer	<0.001	<0.001
Cirurgia prévia	<0.001	<0.001
Químio prévia	<0.001	<0.001
Rádio prévia	<0.001	0.001
Capac.Funcional	<0.001	<0.001
Qualidade de vida	<0.001	<0.001
SAPS3	<0.001	<0.001

Navigation icons: back, forward, search, etc.

Dados ICESP - Modelo de Cox

Modelo Completo

```
modelo1 <- coxph(Surv(tempo,delta)~
  idade + sexo + imc + tipoadm +
  status_cancer + extensao_cancer +
  cirurgiap + quimio + radio +
  capac_func_ecog + QV.0 + SAPS3_Pontos,
  data=karina)
modelo1
```

Navigation icons

Dados ICESP - Modelo de Cox

Modelo Reduzido

```
modelo2 <- coxph(Surv(tempo,delta)~imc+tipoadm+quimio
  +capac_func_ecog+SAPS3_Pontos, data=karina)
modelo2
```

	coef	exp(coef)	se(coef)	z	p
imc	-0.0531	0.948	0.00966	-5.50	3.9e-08
tipoadmEmergencia	0.5261	1.692	0.19585	2.69	7.2e-03
tipoadmMedica	0.4034	1.497	0.12725	3.17	1.5e-03
quimionao	0.2092	1.233	0.10747	1.95	5.2e-02
quimiopaliativa	0.5215	1.684	0.14084	3.70	2.1e-04
capac_func_ecog	0.1911	1.211	0.03471	5.51	3.7e-08
SAPS3_Pontos	0.0303	1.031	0.00399	7.59	3.2e-14

Likelihood ratio test=317 on 7 df,p=0 n= 791,number of events=521
(2 observations deleted due to missingness)

Navigation icons

Dados ICESP - Modelo de Cox

	coef	exp(coef)	se(coef)	z	p
idade	-0.000365	1.000	0.00338	-0.1080	9.1e-01
sexoM	0.131286	1.140	0.09147	1.4352	1.5e-01
imc	-0.048520	0.953	0.00978	-4.9624	7.0e-07
tipoadmEmergencia	0.494766	1.640	0.20484	2.4154	1.6e-02
tipoadmMedica	0.387125	1.473	0.13965	2.7721	5.6e-03
status_cancerrecaida	0.010060	1.010	0.11114	0.0905	9.3e-01
status_cancerremissao	-0.309645	0.734	0.20254	-1.5288	1.3e-01
extensao_cancerLimitado	0.145922	1.157	0.33060	0.4414	6.6e-01
extensao_cancerLocal av	0.401507	1.494	0.32296	1.2432	2.1e-01
extensao_cancerMetastase	0.278095	1.321	0.31673	0.8780	3.8e-01
cirurgiano	0.090537	1.095	0.10096	0.8968	3.7e-01
cirurgiapaliativa	-0.060160	0.942	0.17364	-0.3465	7.3e-01
quimionao	0.272285	1.313	0.12092	2.2517	2.4e-02
quimiopaliativa	0.515762	1.675	0.15820	3.2602	1.1e-03
radionao	-0.187709	0.829	0.11745	-1.5982	1.1e-01
radiopaliativa	-0.049353	0.952	0.24054	-0.2052	8.4e-01
capac_func_ecog	0.140149	1.150	0.05109	2.7432	6.1e-03
QV.0	-0.260541	0.771	0.14638	-1.7799	7.5e-02
SAPS3_Pontos	0.031026	1.032	0.00487	6.3723	1.9e-10

Likelihood ratio test=333 on 19 df,p=0 n= 786,number of events= 517
(7 observations deleted due to missingness)

Navigation icons

Dados ICESP - Modelo de Cox - Interpretações

	coef	exp(coef)	se(coef)	z	p
imc	-0.0531	0.948	0.00966	-5.50	3.9e-08

- Um acréscimo de uma unidade no IMC diminui a taxa de óbito em aprox. 5% (outras características constantes).

tipoadmEmergencia	0.5261	1.692	0.19585	2.69	7.2e-03
tipoadmMedica	0.4034	1.497	0.12725	3.17	1.5e-03

- Admissão emergencial leva a uma taxa de óbito aprox. 70% maior comparativamente à admissão Eletiva.
- A taxa de óbito de pacientes internados por indicação médica é, aproximadamente, 50% maior do que aquela associada com pacientes internados eletivamente.

Navigation icons

Dados ICESP - Modelo de Cox - Interpretações

	coef	exp(coef)	se(coef)	z	p
quimionao	0.2092	1.233	0.10747	1.95	5.2e-02
quimiopaliativa	0.5215	1.684	0.14084	3.70	2.1e-04

- Pacientes não submetidos à quimioterapia têm taxa de óbito 23% superior a pacientes que fizeram quimio curativa
- A taxa de óbito daqueles que fizeram quimio paliativa é 68% superior daqueles com quimio curativa.

	coef	exp(coef)	se(coef)	z	p
capac_func_ecog	0.1911	1.211	0.03471	5.51	3.7e-08

- O aumento de 1 grau na escala ECOG implica um aumento de, aprox., 21% na taxa de óbito.

	coef	exp(coef)	se(coef)	z	p
SAPS3_Pontos	0.0303	1.031	0.00399	7.59	3.2e-14

- Aumentando-se a classificação SAPS3 em 1 unidade ocorre um aumento de 3% na taxa de óbito.

Dados ICESP - Modelo de Cox - Interpretações

Transformações para tornar o modelo interpretável

- IMC.12 = IMC - 12 (menor IMC nos dados)
- SAPS3.16 = SAPS3 - 16 (menor valor possível)

$\alpha_0(t)$: taxa de óbito para pacientes com:

- IMC igual a 12
- Admissão eletiva
- Submetidos a quimio prévia com finalidade curativa
- Capacidade funcional normal
- Escore SAPS3 igual a 16

Dados ICESP - Modelo de Cox - Interpretações

Grupo referência:

$$\alpha(t | \mathbf{X}^T \boldsymbol{\beta}) = \alpha_0(t) \exp\{\mathbf{X}^T \boldsymbol{\beta}\}$$

- $\alpha_0(t)$ representa a taxa de óbito de um paciente em que $\mathbf{X} = \mathbf{0}$.

No exemplo anterior o grupo de referência é um paciente com

- IMC=0(!!!!)
- Tipo de admissão = Eletiva
- Submetido à quimioterapia prévia de caráter curativo
- Com capacidade funcional (ECOG) = 0
- Com valor SAPS3 = 0 (O menor valor possível é 16!!!!)

Dados ICESP - Modelo de Cox - Interpretações

Modelo Final

```
karina$imc.12<-karina$imc-12
karina$SAPS3.16<-karina$SAPS3_Pontos-16
final<- coxph(Surv(tempo,delta)~imc.12
             +tipoadm + quimio
             +capac_func_ecog+SAPS3.16, data=karina)
final
```

	coef	exp(coef)	se(coef)	z	p
imc.12	-0.0531	0.948	0.00966	-5.50	3.9e-08
tipoadmEmergencia	0.5261	1.692	0.19585	2.69	7.2e-03
tipoadmMedica	0.4034	1.497	0.12725	3.17	1.5e-03
quimionao	0.2092	1.233	0.10747	1.95	5.2e-02
quimiopaliativa	0.5215	1.684	0.14084	3.70	2.1e-04
capac_func_ecog	0.1911	1.211	0.03471	5.51	3.7e-08
SAPS3.16	0.0303	1.031	0.00399	7.59	3.2e-14

Não ocorrem alterações nas estimativas e significâncias!

Modelo de Cox - Intervalos de Confiança

Seja $\hat{\beta}$ o EMVP para β no modelo de Cox

Distribuição Assintótica de $\hat{\beta}$

$$\frac{\hat{\beta} - \beta}{\sqrt{\widehat{\text{Var}}(\hat{\beta})}} \xrightarrow{D} \mathcal{N}(0, 1)$$

em que $\widehat{\text{Var}}(\hat{\beta})$ é obtido da matriz de informação observada. Logo,

$$IC(\beta, \gamma) = \left[\hat{\beta} - z_{1-\frac{\gamma}{2}} \sqrt{\widehat{\text{Var}}(\hat{\beta})}; \hat{\beta} + z_{1-\frac{\gamma}{2}} \sqrt{\widehat{\text{Var}}(\hat{\beta})} \right]$$

Intervalo de Confiança para o Risco Relativo

$$IC(\beta, \gamma) = \left[e^{\hat{\beta} - z_{1-\frac{\gamma}{2}} \sqrt{\widehat{\text{Var}}(\hat{\beta})}}; e^{\hat{\beta} + z_{1-\frac{\gamma}{2}} \sqrt{\widehat{\text{Var}}(\hat{\beta})}} \right]$$

Intervalos de Confiança - ICESP

	exp(coef)	exp(-coef)	lower .95	upper .95
imc.12	0.9483	1.0545	0.9305	0.9664
tipoadmEmergencia	1.6923	0.5909	1.1528	2.4842
tipoadmMedica	1.4970	0.6680	1.1665	1.9210
quimionao	1.2327	0.8112	0.9986	1.5218
quimiopaliativa	1.6845	0.5937	1.2781	2.2200
capac_func_ecog	1.2106	0.8260	1.1310	1.2958
SAPS3.16	1.0307	0.9702	1.0227	1.0388

Concordance= 0.742 (se = 0.013)
 Rsquare= 0.33 (max possible= 1)
 Likelihood ratio test= 317.3 on 7 df, p=0
 Wald test = 325.9 on 7 df, p=0
 Score (logrank) test = 351.9 on 7 df, p=0

- Verifique se os IC's incluem o valor 1
- O comando `summary` também fornece os testes
 - Razão de Verossimilhanças
 - Wald
 - Rao/Score

Intervalos de Confiança - ICESP

`summary`(final)

	coef	exp(coef)	se(coef)	z	Pr(> z)			
imc.12	-0.053106	0.948280	0.009663	-5.496	3.88e-08	***		
tipoadmEmergencia	0.526089	1.692301	0.195848	2.686	0.007227	**		
tipoadmMedica	0.403430	1.496951	0.127245	3.170	0.001522	**		
quimionao	0.209242	1.232744	0.107474	1.947	0.051545	.		
quimiopaliativa	0.521458	1.684482	0.140844	3.702	0.000214	***		
capac_func_ecog	0.191122	1.210607	0.034711	5.506	3.67e-08	***		
SAPS3.16	0.030282	1.030745	0.003990	7.590	3.21e-14	***		

Signif. codes:	0	'***'	0.001	'**'	0.01	'*' 0.05	'.' 0.1	' ' 1

Estimação da Função de Sobrevida - Modelo de Cox

Modelo de riscos proporcionais

$$S(t | X) = e^{-\int_0^t \alpha_0(s) e^{X\beta} ds} = \left(e^{-\int_0^t \alpha_0(s) ds} \right) e^{X\beta} = \left(S_0(t) \right) e^{X\beta}$$

com $S_0(t)$ a função de sobrevivência basal ou *baseline*.

Logo, um estimador para a $S(t)$ é

$$\hat{S}(t | X) = \left(\hat{S}_0(t) \right) e^{X\hat{\beta}}$$

Para o modelo de Cox,

- $\hat{\beta}$ é o EMVP
- $S_0(t)$ é, em geral, estimada em um único passo.

Estimação de $S_0(t)$ - Modelo de Cox

Estimação em duas etapas:

- 1 β é estimado usando a verossimilhança parcial
- 2 Dado $\hat{\beta}$ de (1), $S_0(t)$ é estimada por um dos seguintes métodos (biblioteca `Survival`)

Estimador de Breslow

$$\hat{S}_0(t) = \prod_{i:t_i < t} \left(1 - \frac{d_i}{\sum_{j \in \mathcal{R}_j} e^{X_j \hat{\beta}}}\right)$$

- \mathcal{R}_j : conjunto de índices associados com indivíduos *em risco* no instante t_j
- d_j : total de indivíduos *falhando* no instante t_j

Estimação de $S(t)$ - Dados do ICESP

$\alpha_o(t)$: Função de risco de um paciente com IMC igual a 12, admitido na UTI em caráter eletivo, submetido previamente à quimio curativa, com capacidade funcional normal e escore SAPS igual a 16.

Estimação de $S_0(t)$

`summary` (final)

```

                coef exp(coef)  se(coef)      z Pr(>|z|)
imc.12          -0.053106  0.948280  0.009663 -5.496 3.88e-08 ***
tipoadmEmergencia 0.526089  1.692301  0.195848  2.686 0.007227 **
tipoadmMedica     0.403430  1.496951  0.127245  3.170 0.001522 **
quimionao        0.209242  1.232744  0.107474  1.947 0.051545 .
quimiopaliativa  0.521458  1.684482  0.140844  3.702 0.000214 ***
capac_func_ecog  0.191122  1.210607  0.034711  5.506 3.67e-08 ***
SAPS3.16         0.030282  1.030745  0.003990  7.590 3.21e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

Estimação de $S_0(t)$ - Modelo de Cox

Estimador de Kalbfleisch e Prentice

$$\hat{S}_0(t) = \prod_{i:t_i < t} \hat{h}_i$$

em que

$$\hat{h}_i = \left(1 - \frac{e^{X_i \hat{\beta}}}{\sum_{j \in \mathcal{R}_i} e^{X_j \hat{\beta}}}\right) e^{-X_i \hat{\beta}}$$

- Expressão anterior supõe ausência de empates
- Ocorrência de empates $\rightarrow \hat{h}_i$ precisam ser avaliados numericamente
- Esse estimador é baseado em uma verossimilhança não paramétrica completa

Intervalos de Confiança - ICESP

```

                exp(coef) exp(-coef) lower .95 upper .95
imc.12          0.9483      1.0545      0.9305      0.9664
tipoadmEmergencia 1.6923      0.5909      1.1528      2.4842
tipoadmMedica     1.4970      0.6680      1.1665      1.9210
quimionao        1.2327      0.8112      0.9986      1.5218
quimiopaliativa  1.6845      0.5937      1.2781      2.2200
capac_func_ecog  1.2106      0.8260      1.1310      1.2958
SAPS3.16         1.0307      0.9702      1.0227      1.0388
    
```

```

Concordance= 0.742 (se = 0.013 )
Rsquare= 0.33 (max possible= 1 )
Likelihood ratio test= 317.3 on 7 df, p=0
Wald test = 325.9 on 7 df, p=0
Score (logrank) test = 351.9 on 7 df, p=0
    
```

- Verifique se os IC's incluem o valor 1
- O comando `summary` também fornece os testes
 - Razão de Verossimilhanças
 - Wald
 - Rao/Escore

Estimação da Função de Sobrevida - Modelo de Cox

Modelo de riscos proporcionais

$$S(t | X) = e^{-\int_0^t \alpha_0(s) e^{X\beta} ds} = \left(e^{-\int_0^t \alpha_0(s) ds} \right)^{e^{X\beta}} = \left(S_0(t) \right)^{e^{X\beta}}$$

com $S_0(t)$ a função de sobrevivência basal ou *baseline*.

Logo, um estimador para a $S(t)$ é

$$\hat{S}(t | X) = \left(\hat{S}_0(t) \right)^{e^{X\hat{\beta}}}$$

Para o modelo de Cox,

- $\hat{\beta}$ é o EMVP
- $S_0(t)$ é, em geral, estimada em um por passo.

Estimação de $S_0(t)$ - Modelo de Cox

Estimador de Kalbfleisch e Prentice

$$\hat{S}_0(t) = \prod_{i:t_i < t} \hat{h}_i$$

em que

$$\hat{h}_i = \left(1 - \frac{e^{X_i \hat{\beta}}}{\sum_{j \in \mathcal{R}_i} e^{X_j \hat{\beta}}} \right) e^{-X_i \hat{\beta}}$$

- Expressão anterior supõe ausência de empates
- Ocorrência de empates $\rightarrow \hat{h}_i$ precisam ser avaliados numericamente
- Esse estimador é baseado em uma verossimilhança não paramétrica completa
- Estimadores de Breslow e Kalbfleisch-Prentice fornecem estimativas similares (exceto para muitos empates).

Estimação de $S_0(t)$ - Modelo de Cox

Estimação em duas etapas:

- 1 β é estimado usando a verossimilhança parcial
- 2 Dado $\hat{\beta}$ de (1), $S_0(t)$ é estimada por um dos seguintes métodos (biblioteca *Survival*)

Estimador de Breslow

$$\hat{S}_0(t) = \prod_{i:t_i < t} \left(1 - \frac{d_i}{\sum_{j \in \mathcal{R}_j} e^{X_j \hat{\beta}}} \right)$$

- \mathcal{R}_j : conjunto de índices associados com indivíduos *em risco* no instante t_j
- d_j : total de indivíduos *falhando* no instante t_j

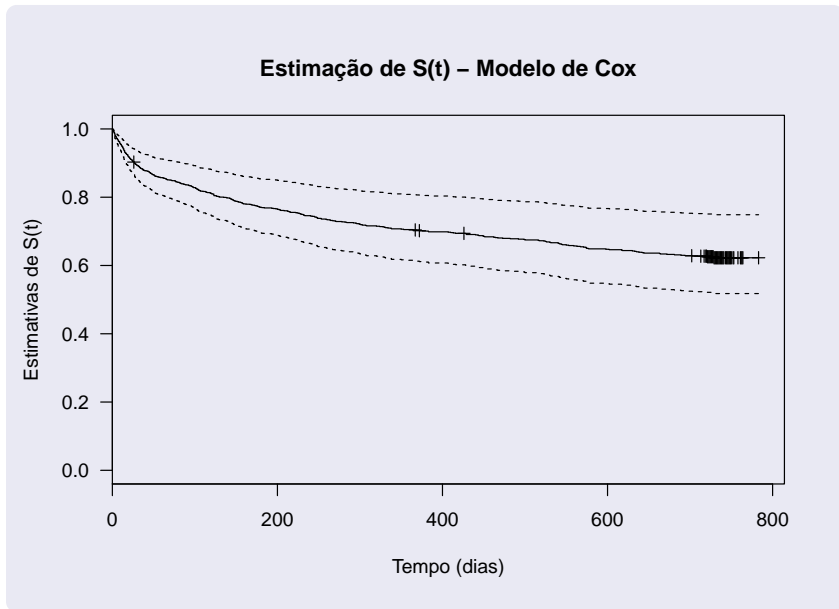
Estimação de $S(t)$ - Dados do ICESP

$\alpha_0(t)$: Risco de um paciente com IMC, admitido na UTI em caráter eletivo, submetido previamente a quimio curativa, capacidade funcional normal e escore SAPS 16.

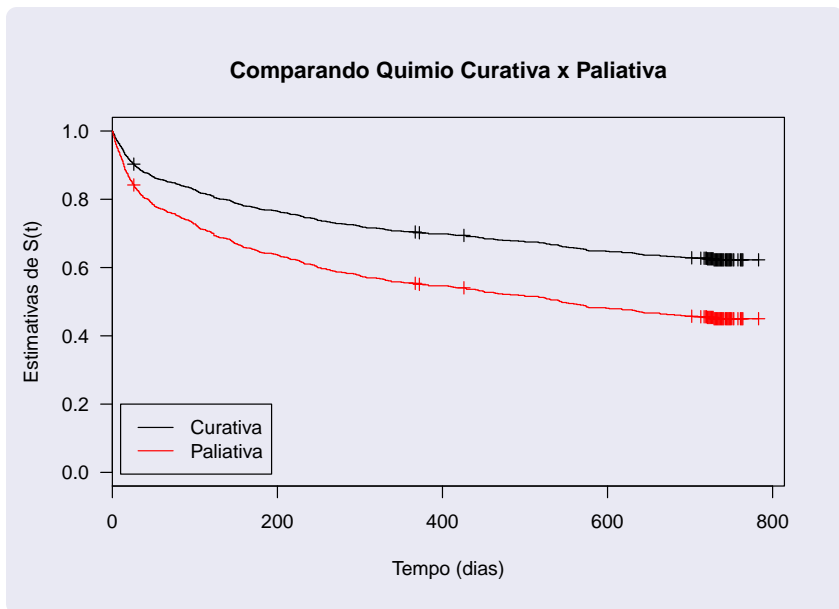
Estimação de $S_0(t)$

```
plot(surfit(final, type="breslow",
           newdata=data.frame(imc.12=0,
                              tipoadm="Elective",
                              quimio="curative",
                              capac_func_ecog=0,
                              SAPS3.16=0),
           add=TRUE),
     las=1,
     main="Estimação de S(t) - Modelo de Cox",
     ylab="Estimativas de S(t)",
     xlab="Tempo (dias)")
```

Estimação de $S(t)$ - Dados do ICESP



Estimação de $S(t)$ - Dados do ICESP



Estimação de $S(t)$ - Dados do ICESP

```
plot(survfit(final, type="breslow",
            newdata=data.frame(imc.12=0,
                                tipoadm="Elective",
                                quimio=c("curativa", "paliativa"),
                                capac_func_ecog=0,
                                SAPS3.16=0),
            add=TRUE),
     col=c(1, 2), las=1,
     main="Comparando Químio Curativa x
          Paliativa",
     ylab="Estimativas de S(t)",
     xlab="Tempo (dias)",
     legend(10, 0.2, legend=c("Curativa", "Paliativa"),
           lty=1, col=c(1, 2))
```