

Lista de Exercícios

Análise de Sobrevivência

Tailine J. S. Nonato

20 de julho de 2025

Índice

Orientações	2
Entrada dos dados	4
Exercício 1: Modelo de Cox	6
(i) Ajuste do modelo	6
(ii) Estimativa da função de sobrevivência	6
(iii) Estimativa da média e mediana do tempo de sobrevivência	7
(iv) Estimativa da vida média residual em $T=50$	8
(v) Verificação gráfica da qualidade do ajuste	8
Exercício 2: Ajuste de modelo paramétrico	10
Análise Preliminar	10
Modelos paramétricos (Weibull e Log-Normal)	11
(i.1) Ajuste do modelo	11
(ii) Estimativa da função de sobrevivência	13

Orientações

Para cada um dos modelos especificados abaixo:

- (i) Ajuste o modelo solicitado, apresentando as estimativas dos parâmetros do modelo escolhido e coeficientes de regressão com seus respectivos intervalos de confiança. Para os parâmetros positivos, apresente o intervalo do tipo log e para os parâmetros limitados em $[0-1]$ o intervalo do tipo log-log.
- (ii) Segundo o modelo ajustado no item (i), estime a função de sobrevivência de um indivíduo do grupo Ag+ que apresente WBC=5.000. Calcule um intervalo de confiança com uma transformação que garanta que os limites do intervalo não ultrapassem o intervalo $[0-1]$ para função de sobrevivência obtida.
- (iii) Segundo o modelo ajustado no item (i), calcule a estimativa da média e da mediana do tempo de sobrevivência de um indivíduo do grupo Ag+ que apresente WBC=5.000. Calcule o intervalo de confiança do tipo log para essas estimativas (média e mediana).
- (iv) Segundo o modelo ajustado no item (i), qual é a estimativa da vida média residual em $T=50$ de um indivíduo do grupo Ag+ que apresente WBC=5.000?
- (v) Verifique graficamente a qualidade do ajuste do modelo ajustado confrontando a função de sobrevivência estimada empírica com a ajustada pelo modelo (dica: categorize a covariável numérica).

Nota: Caso adote um conjunto de dados diferente da Giolo, considere $T = t$ um valor arbitrário para cada uma das duas covariáveis nos itens (ii, iii e iv).

Exercício 1. O modelo semi-paramétrico de riscos proporcionais de Cox

Exercício 2. Um modelo de regressão paramétrico

Nota: Ajuste ao menos dois modelos de regressão (que não seja exponencial) paramétricos para esse conjunto de dados e decida pela utilização de um deles.

Exercício 3. “Discretize” os tempos de sobrevivência e ajuste um modelo de regressão discreto

Nota: Ajuste um modelo de regressão discreto correspondente ao modelo contínuo adotado como melhor modelo no Exercício 2. Como discretizar: transforme os tempos para meses completos (“ $\text{floor}(t/30)$ ”). Por exemplo. Se $t=4$, considere que $t=0$ mês completo; se $t=5$,

considere que $t^*=1$ mês completo e assim sucessivamente. Caso adote um conjunto de dados diferente da Giolo, proponha um procedimento para discretizar os tempos.

Exercício 4. Um modelo chances de sobrevivência proporcionais

Nota: Ajuste um modelo de regressão Chance de Sobrevivências Proporcionais correspondente ao modelo adotado como melhor modelo no Exercício 2.

Entrada dos dados

Tabela 1: Primeiras linhas do conjunto de dados

t	WBC	logWBC	grupo	censura
65	2300	3.361728	Ag+	1
156	750	2.875061	Ag+	1
100	4300	3.633469	Ag+	1
134	2600	3.414973	Ag+	1
16	6000	3.778151	Ag+	1
108	10000	4.000000	Ag+	1
121	10000	4.000000	Ag+	1
4	17000	4.230449	Ag+	1
39	5400	3.732394	Ag+	1
143	7000	3.845098	Ag+	1
56	9400	3.973128	Ag+	1
26	32000	4.505150	Ag+	1
22	35000	4.544068	Ag+	1
1	100000	5.000000	Ag+	1
1	100000	5.000000	Ag+	1
5	52000	4.716003	Ag+	1
65	100000	5.000000	Ag+	1
56	4400	3.643453	Ag-	1
65	3000	3.477121	Ag-	1
17	4000	3.602060	Ag-	1
7	1500	3.176091	Ag-	1
16	9000	3.954242	Ag-	1
22	5300	3.724276	Ag-	1
3	10000	4.000000	Ag-	1
4	19000	4.278754	Ag-	1
2	27000	4.431364	Ag-	1
3	28000	4.447158	Ag-	1
8	31000	4.491362	Ag-	1
4	26000	4.414973	Ag-	1
3	21000	4.322219	Ag-	1
30	79000	4.897627	Ag-	1

t	WBC	logWBC	grupo	censura
4	100000	5.000000	Ag-	1
43	100000	5.000000	Ag-	1

Exercício 1: Modelo de Cox

(i) Ajuste do modelo

Tabela 2: Estimativas dos parâmetros do modelo de Cox

	coef	exp(coef)	se(coef)	z	Pr(> z)
logWBC	0.8480101	2.334996	0.3131691	2.707834	0.0067724
grupoAg-	1.0673215	2.907581	0.4293041	2.486166	0.0129128

Tabela 3: Razões de risco e ICs para o modelo de Cox

	RR	IC_inf	IC_sup
logWBC	2.335	1.264	4.314
grupoAg-	2.908	1.253	6.745

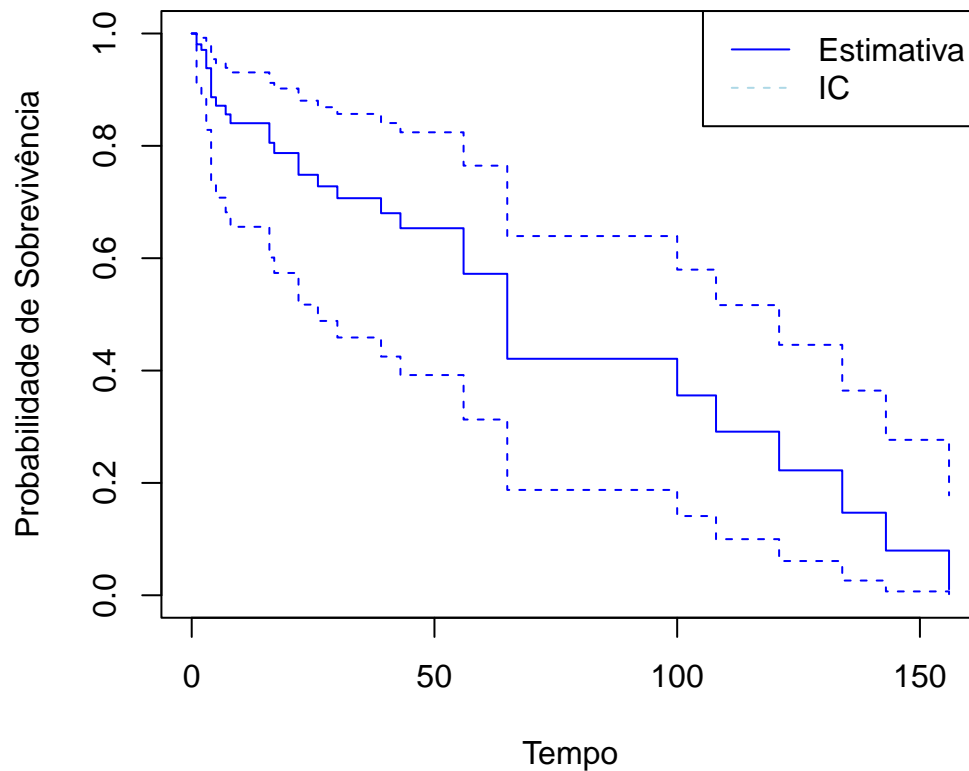
(ii) Estimativa da função de sobrevivência

Tabela 4: Estimativa da função de sobrevivência para WBC=5000

Tempo	Sobrevivencia	SE	IC_inf	IC_sup
1	0.981	0.015	0.911	0.996
2	0.971	0.020	0.892	0.992
3	0.938	0.033	0.828	0.979
4	0.887	0.051	0.734	0.954
5	0.871	0.056	0.708	0.947
7	0.856	0.061	0.682	0.939
8	0.840	0.066	0.656	0.931
16	0.805	0.076	0.601	0.912
17	0.787	0.081	0.574	0.902
22	0.748	0.091	0.517	0.880
26	0.728	0.096	0.488	0.869

Tempo	Sobrevivencia	SE	IC_inf	IC_sup
30	0.707	0.101	0.459	0.857
39	0.680	0.107	0.425	0.841
43	0.653	0.112	0.392	0.824
56	0.572	0.119	0.313	0.765
65	0.421	0.123	0.188	0.639
100	0.356	0.120	0.141	0.580
108	0.291	0.114	0.100	0.517
121	0.222	0.106	0.061	0.446
134	0.147	0.092	0.026	0.364
143	0.080	0.070	0.007	0.277
156	0.011	0.024	0.000	0.178

Função de Sobrevivência Estimada WBC=5000



(iii) Estimativa da média e mediana do tempo de sobrevivência

Tabela 5: Estimativas da média e mediana do tempo de sobrevivência com ICs

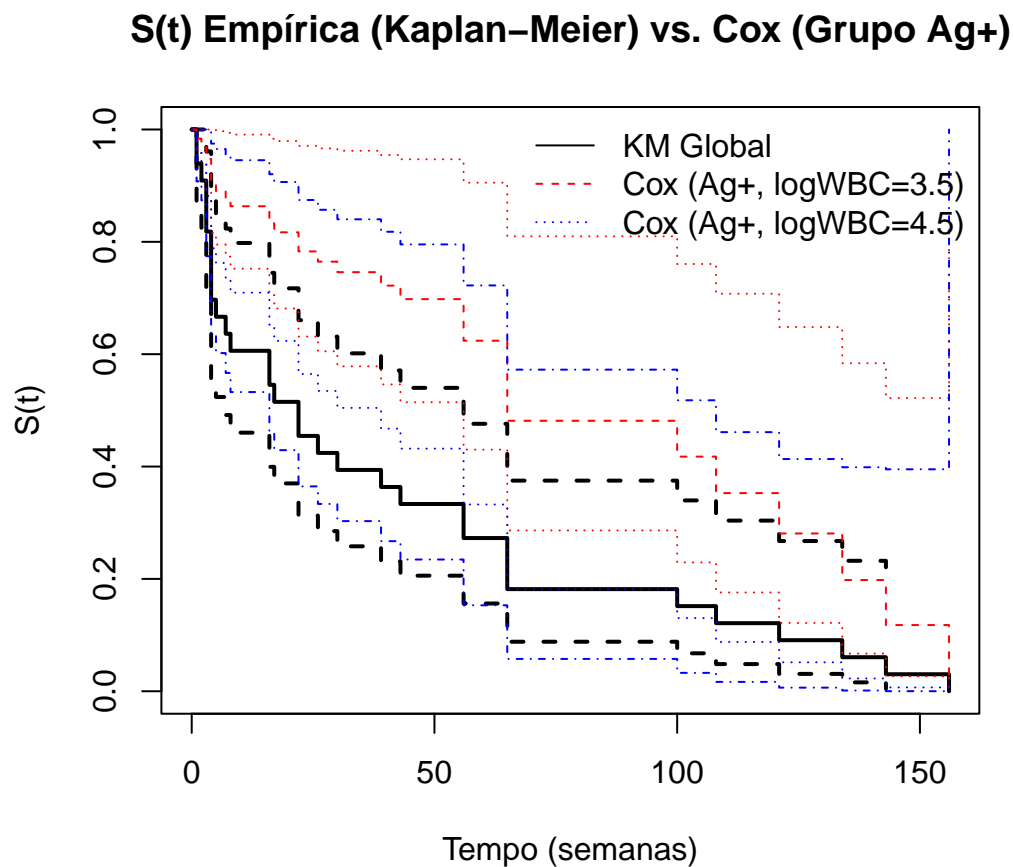
Item	Estimativa	IC_inf	IC_sup
Média	50.273	39.742	63.932
Mediana	28.000	22.154	35.388

(iv) Estimativa da vida média residual em $T=50$

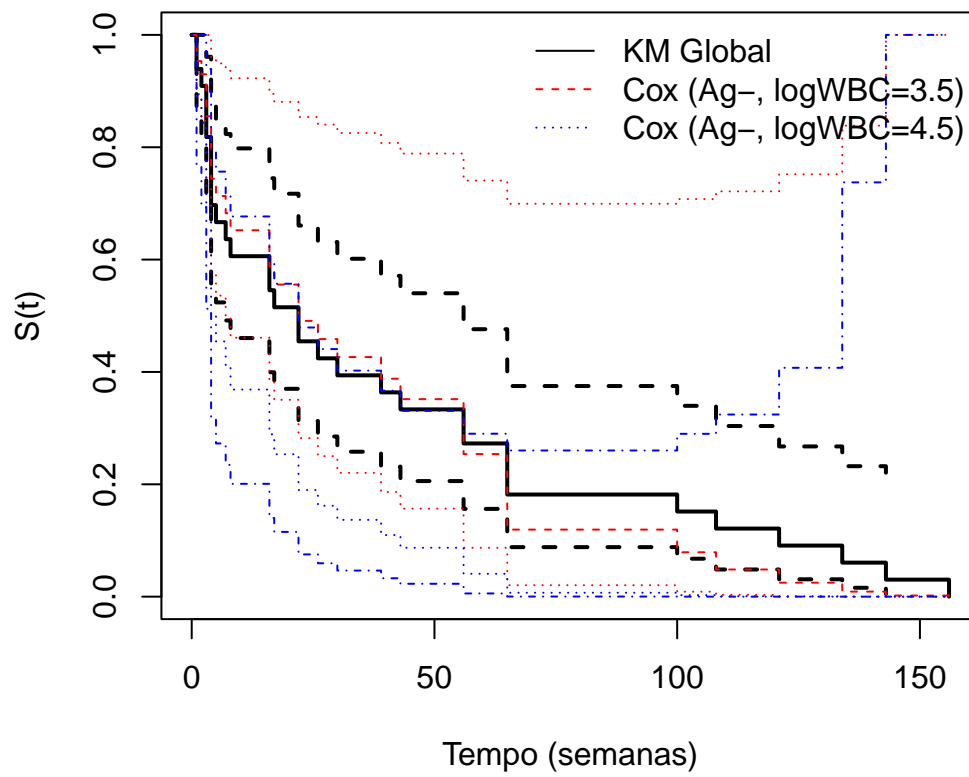
Tabela 6: Estimativa da vida média residual em $T=50$, dado que já se passaram 50 semanas

Tempo	VMR
50	5.512

(v) Verificação gráfica da qualidade do ajuste



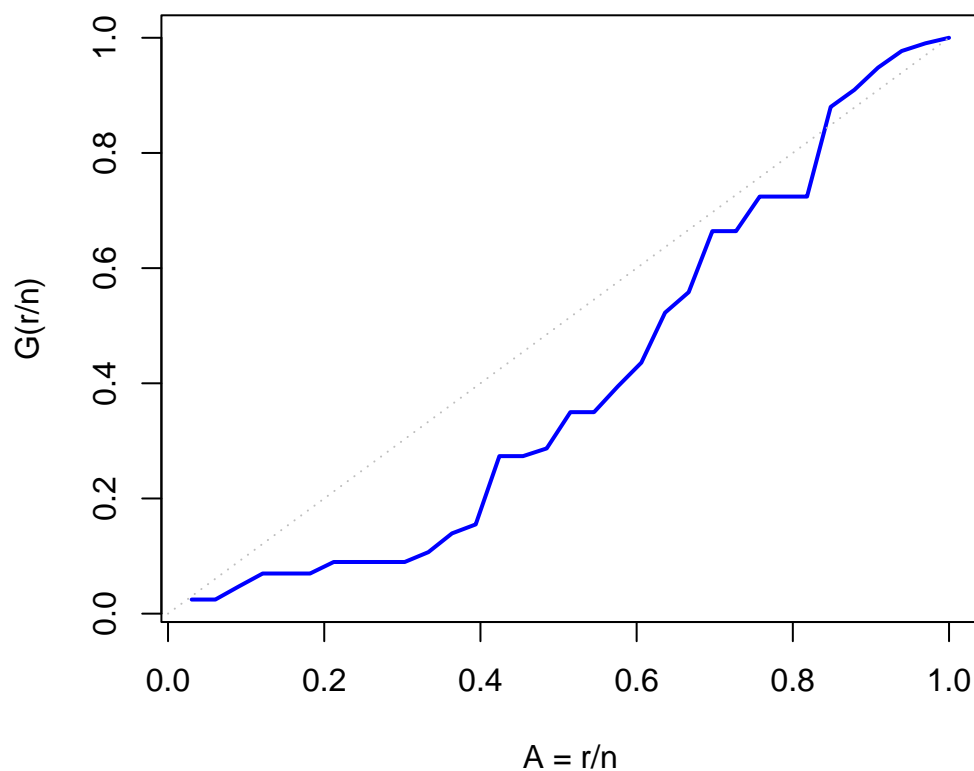
S(t) Empírica (Kaplan–Meier) vs. Cox (Grupo Ag–)



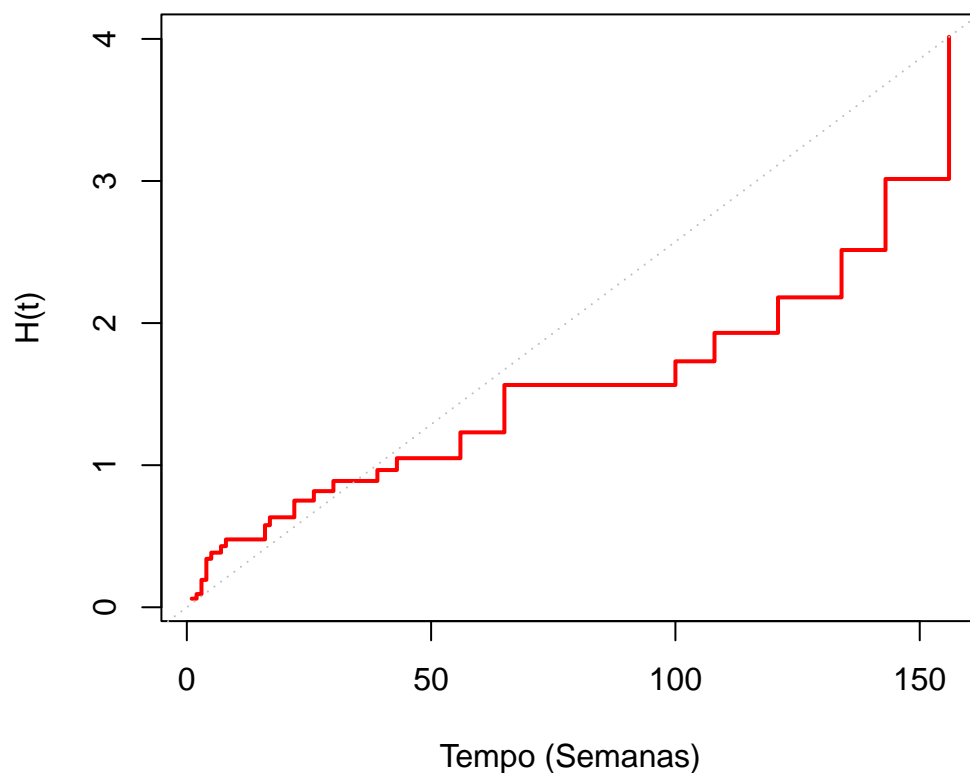
Exercício 2: Ajuste de modelo paramétrico

Análise Preliminar

TTT plot para os dados



Estimativa de $H(t)$ (Nelson–Aalen)



Modelos paramétricos (Weibull e Log-Normal)

(i.1) Ajuste do modelo

--- Sumário do Modelo Weibull de Regressão ---

grupoAg=0.484848484848485, logWBC=4.13546583270838

	time	est	lcl	ucl
1	1	0.96341429	0.9112257156	0.98985468
2	2	0.93000571	0.8557361115	0.97571528
3	3	0.89839352	0.8107354434	0.95953581
4	4	0.86824939	0.7720056513	0.94166063
5	5	0.83939632	0.7350504792	0.92429699
6	7	0.78511837	0.6700097332	0.88751924
7	8	0.75953301	0.6393800590	0.87011839
8	16	0.58537705	0.4525590849	0.72557120
9	17	0.56686708	0.4342295317	0.70725204
10	22	0.48322866	0.3545591053	0.63237367
11	26	0.42573427	0.3033597808	0.57869772

```

12  30 0.37536470 0.2566063594 0.52725701
13  39 0.28339681 0.1705753558 0.42471104
14  43 0.25033084 0.1444932449 0.38892219
15  56 0.16774936 0.0799963374 0.28627334
16  65 0.12742192 0.0511170853 0.23146604
17 100 0.04428735 0.0081111045 0.11325673
18 108 0.03486127 0.0050544849 0.09737294
19 121 0.02366323 0.0021966758 0.07714399
20 134 0.01608816 0.0009025704 0.06080796
21 143 0.01232739 0.0005198245 0.05303768
22 156 0.00840111 0.0002140757 0.04337401

```

--- Sumário do Modelo Log-Normal de Regressão ---

grupoAg=-0.484848484848485,logWBC=4.13546583270838

	time	est	lcl	ucl
1	1	0.99130693	0.963799788	0.99895659
2	2	0.96373963	0.907636080	0.99133877
3	3	0.92717009	0.850276752	0.97517426
4	4	0.88747942	0.792395216	0.95365190
5	5	0.84747773	0.744330453	0.92750785
6	7	0.77121810	0.657171433	0.86961230
7	8	0.73585834	0.618086201	0.84025878
8	16	0.51915205	0.386436146	0.63658725
9	17	0.49883085	0.366733383	0.61851888
10	22	0.41307545	0.284449346	0.53289212
11	26	0.35940378	0.235165361	0.48054538
12	30	0.31549572	0.194540333	0.43853256
13	39	0.24169709	0.134041832	0.36034207
14	43	0.21683678	0.116020024	0.33330400
15	56	0.15746161	0.073308696	0.26559717
16	65	0.12919336	0.054948764	0.22826165
17	100	0.06781110	0.021045430	0.14391478
18	108	0.05973753	0.017712111	0.13229270
19	121	0.04921495	0.013381723	0.11724430
20	134	0.04108005	0.010069749	0.10443585
21	143	0.03649241	0.008559711	0.09700297
22	156	0.03101646	0.006780149	0.08517358

Tabela 7: Log-Verossimilhança dos Modelos Ajustados

Modelo	LogVerossimilhança
Weibull	-146.4855
Log-Normal	-145.9392

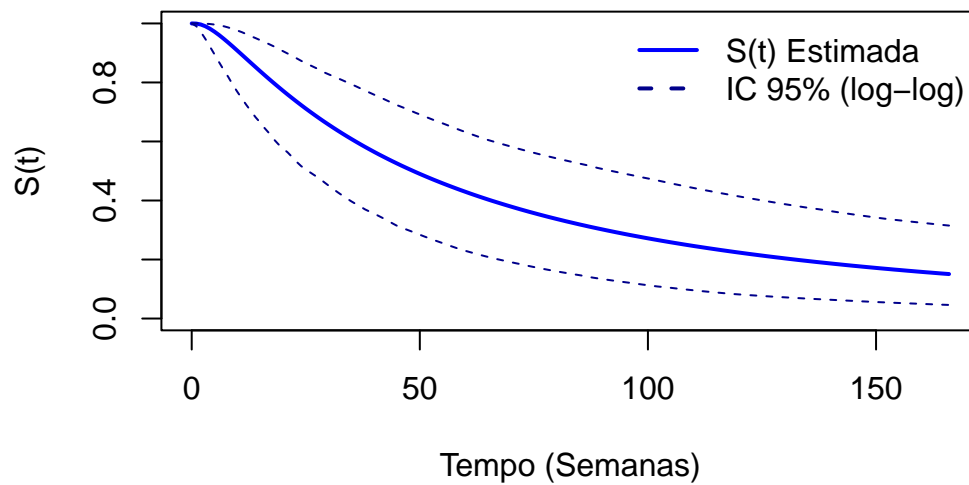
Tabela 8: AIC e BIC dos Modelos Ajustados

Modelo	AIC	BIC
Weibull	300.9709	306.9570
Log-Normal	299.8785	305.8645

Assim, o modelo Log-Normal apresentou melhor ajuste, pois possui menor AIC e BIC.

(ii) Estimativa da função de sobrevivência

Função de Sobrevivência Estimada (Modelo Log-Norma Indivíduo: Ag+, WBC=5.000



Estimativa de $S(t=50)$ para um indivíduo Ag+ com WBC=5000: 0.4902

Intervalo de Confiança 95% (log-log) em $t=50$: [0.2841; 0.6916]