Relatório (escolher o nome depois)

Universidade Federal da Paraíba - CCEN

Gabriel de Jesus Pereira

17 de agosto de 2024

Questão 1: Considere as seguintes funções $S(t)$ apresentadas abaixo e responda o que se pede.

1. $S_1(t) = e^{-t/5}$, em que $t \geq 0$
2. $S_2(t) = \frac{1}{1+t}$, em que $t \geq 0$.
3. $S_3(t) = 1 - \frac{t}{2}$, em que $t \geq 0$
4. $S_4(t) = 2e^{-t/2}$, em que $t \geq 0$

(a) Considerando as condições, vistas em sala de aula, para que $S(t)$ seja uma função de sobrevivência, identifique quais das funções apresentadas são realmente funções de sobrevivência.

Resolução (a)

Para que $S(t)$ seja uma função de sobrevivência, deve satisfazer as seguintes condições: $S(t) \to 0$ para $t \to \infty$ e $S(t) \to 1$ para $t \to 0$. Assim, temos os seguintes resultados:

1. $\lim_{t\to\infty} e^{-t/5} = 0$ e $\lim_{t\to 0} e^{-t/5} = 1$
2. $\lim_{t\to\infty} \frac{1}{1+t} = 0$ e $\lim_{t\to 0} \frac{1}{1+t} = 1$
3. $\lim_{t\to\infty} (1 - \frac{t}{2}) = -\infty$ e $\lim_{t\to 0} (1 - \frac{t}{2}) = 1$
4. $\lim_{t\to\infty} 2e^{-t/2} = 0$ e $\lim_{t\to 0} 2e^{-t/2} = 2$

Dessa forma, pelas condições necessárias, vemos que apenas $S_1(t)$ e $S_2(t)$ validam a condição para ser uma função de sobrevivência.

Questão 2:

Escolha um dos bancos de dados disponíveis no seguinte endereço eletrônico: **http://sobrevida.fiocruz.br/dados.html**. A partir do banco de dados escolhido por você, faça o que se pede a seguir

2b)

Considerando a variável de tempo até ocorrência do evento de interesse na base de dados escolhida, forneça as seguintes informações:

i) É possível montar uma tabela para descrever os dados de acordo com o número de ocorrências do evento de interesse registradas em intervalos de tempo da pesquisa? Se sim, apresente-a.

tempo	intervalo	falhas	censura	amplitude	risco
1	[1,2)	47	66	1	2453
2	[2,3)	41	47	1	2340
3	[3,4)	32	33	1	2252
4	[4,5)	30	33	1	2187
5	[5,6)	29	37	1	2124
6	[6,7)	22	25	1	2058
7	[7,8)	25	40	1	2011
8	[8,9)	18	39	1	1946
9	[9,10)	32	43	1	1889
10	[10,11)	19	36	1	1814
11	[11,12)	31	35	1	1759
12	[12,13)	27	47	1	1693
13	[13,14)	22	36	1	1619
14	[14,15)	13	27	1	1561
15	[15,16)	25	23	1	1521
16	[16,17)	10	24	1	1473
17	[17,18)	11	38	1	1439
18	[18,19)	11	28	1	1390
19	[19,20)	12	24	1	1351
20	[20,21)	15	14	1	1315
21	[21,22)	15	31	1	1286
22	[22,23)	14	31	1	1240
23	[23,24)	11	30	1	1195
24	[24,25)	7	33	1	1154
25	[25,26)	13	28	1	1114
26	[26,27)	7	16	1	1073
27	[27,28)	14	14	1	1050
28	[28,29)	12	17	1	1022
29	[29,30)	10	19	1	993
30	[30,31)	14	9	1	964
31	[31,32)	10	12	1	941
32	[32,33)	11	29	1	919
33	[33,34)	7	18	1	879
34	[34,35)	4	18	1	854
35	[35,36)	8	20	1	832

tempo	intervalo	falhas	censura	amplitude	risco
36	[36,37)	15	25	1	804
37	[37,38)	8	22	1	764
38	[38,39)	11	18	1	734
39	[39,40)	5	13	1	705
40	[40,41)	8	19	1	687
41	[41,41)	2	16	0	660
42	[41,42)	3	34	1	642
43	[42,43)	5	600	1	605

ii. Apresente o cálculo de sobrevivência empírica (pela definição apresentada na aula 1). Apresente também as estimativas empíricas das seguintes quantidades: função densidade, função de risco, função de risco acumulada.

tempo	intervalo	falhas	censura	risco	S(t)	f(t)	h(t)	H(t)
1	[1,2)	47	66	2453	1.0000	0.0192	0.0192	0.0000
2	[2,3)	41	47	2340	0.9539	0.0167	0.0175	0.0472
3	[3,4)	32	33	2252	0.9181	0.0130	0.0142	0.0855
4	[4,5)	30	33	2187	0.8916	0.0122	0.0137	0.1148
5	[5,6)	29	37	2124	0.8659	0.0118	0.0137	0.1440
6	[6,7)	22	25	2058	0.8390	0.0090	0.0107	0.1756
7	[7,8)	25	40	2011	0.8198	0.0102	0.0124	0.1987
8	[8,9)	18	39	1946	0.7933	0.0073	0.0092	0.2315
9	[9,10)	32	43	1889	0.7701	0.0130	0.0169	0.2613
10	[10,11)	19	36	1814	0.7395	0.0077	0.0105	0.3018
11	[11,12)	31	35	1759	0.7171	0.0126	0.0176	0.3326
12	[12,13)	27	47	1693	0.6902	0.0110	0.0159	0.3708
13	[13,14)	22	36	1619	0.6600	0.0090	0.0136	0.4155
14	[14,15)	13	27	1561	0.6364	0.0053	0.0083	0.4520
15	[15,16)	25	23	1521	0.6201	0.0102	0.0164	0.4779
16	[16,17)	10	24	1473	0.6005	0.0041	0.0068	0.5100
17	[17,18)	11	38	1439	0.5866	0.0045	0.0076	0.5334
18	[18,19)	11	28	1390	0.5667	0.0045	0.0079	0.5680
19	[19,20)	12	24	1351	0.5508	0.0049	0.0089	0.5965
20	[20,21)	15	14	1315	0.5361	0.0061	0.0114	0.6235
21	[21,22)	15	31	1286	0.5243	0.0061	0.0117	0.6458
22	[22,23)	14	31	1240	0.5055	0.0057	0.0113	0.6822
23	[23,24)	11	30	1195	0.4872	0.0045	0.0092	0.7192
24	[24,25)	7	33	1154	0.4704	0.0029	0.0061	0.7541

tempo	intervalo	falhas	censura	risco	S(t)	f(t)	h(t)	H(t)
25	[25,26)	13	28	1114	0.4541	0.0053	0.0117	0.7894
26	[26,27)	7	16	1073	0.4374	0.0029	0.0065	0.8269
27	[27,28)	14	14	1050	0.4280	0.0057	0.0133	0.8485
28	[28,29)	12	17	1022	0.4166	0.0049	0.0117	0.8756
29	[29,30)	10	19	993	0.4048	0.0041	0.0101	0.9043
30	[30,31)	14	9	964	0.3930	0.0057	0.0145	0.9340
31	[31,32)	10	12	941	0.3836	0.0041	0.0106	0.9581
32	[32,33)	11	29	919	0.3746	0.0045	0.0120	0.9818
33	[33,34)	7	18	879	0.3583	0.0029	0.0080	1.0263
34	[34,35)	4	18	854	0.3481	0.0016	0.0047	1.0551
35	[35,36)	8	20	832	0.3392	0.0033	0.0096	1.0812
36	[36,37)	15	25	804	0.3278	0.0061	0.0187	1.1155
37	[37,38)	8	22	764	0.3115	0.0033	0.0105	1.1665
38	[38,39)	11	18	734	0.2992	0.0045	0.0150	1.2066
39	[39,40)	5	13	705	0.2874	0.0020	0.0071	1.2469
40	[40,41)	8	19	687	0.2801	0.0033	0.0116	1.2727
41	[41,41)	2	16	660	0.2691	Inf	Inf	1.3128
42	[41,42)	3	34	642	0.2617	0.0012	0.0047	1.3405
43	[42,43)	5	600	605	0.2466	0.0020	0.0083	1.3998

iii. Apresente o cálculo da função de sobrevivência $S(t)$ considerando os seguintes estimadores: Kaplan-Meier, Nelson-Aalen e Tabela de Vida. Para cada versão desses estimadores, apresente também as estimativas das seguintes quantidades: função densidade, função de risco, função de risco acumulada. Interprete os resultados.

n	events	dropouts	hazard	S	lower	upper	cumHazard	f(t)
2453	47	66	0.0194215	0.9805785	0.9759643	0.9851928	0.0194215	0.0190443
2340	41	47	0.0176991	0.9632231	0.9568931	0.9695532	0.0371206	0.0170482
2252	32	33	0.0143145	0.9494351	0.9420342	0.9568360	0.0514351	0.0135907
2187	30	33	0.0138217	0.9363123	0.9280307	0.9445940	0.0652568	0.0129414
2124	29	37	0.0137735	0.9234161	0.9143600	0.9324721	0.0790302	0.0127186
2058	22	25	0.0107553	0.9134844	0.9038794	0.9230895	0.0897855	0.0098248
2011	25	40	0.0125565	0.9020142	0.8918155	0.9122130	0.1023420	0.0113261
1946	18	39	0.0093434	0.8935864	0.8829724	0.9042003	0.1116854	0.0083491
1889	32	43	0.0171352	0.8782746	0.8669472	0.8896021	0.1288206	0.0150494
1814	19	36	0.0105791	0.8689833	0.8572456	0.8807210	0.1393997	0.0091930
1759	31	35	0.0178007	0.8535147	0.8411283	0.8659012	0.1572004	0.0151932
1693	27	47	0.0161725	0.8397113	0.8267774	0.8526451	0.1733729	0.0135802
1619	22	36	0.0137414	0.8281724	0.8147983	0.8415466	0.1871144	0.0113803
1561	13	27	0.0084006	0.8212153	0.8075821	0.8348485	0.1955150	0.0068987
1521	25	23	0.0165618	0.8076145	0.7934919	0.8217370	0.2120768	0.0133755
1473	10	24	0.0068446	0.8020867	0.7877711	0.8164022	0.2189214	0.0054900
1439	11	38	0.0077465	0.7958733	0.7813408	0.8104059	0.2266679	0.0061652
1390	11	28	0.0079942	0.7895110	0.7747560	0.8042659	0.2346621	0.0063115
1351	12	24	0.0089619	0.7824354	0.7674351	0.7974358	0.2436240	0.0070121
1315	15	14	0.0114679	0.7734625	0.7581578	0.7887673	0.2550919	0.0088700
1286	15	31	0.0118064	0.7643308	0.7487231	0.7799384	0.2668982	0.0090240
1240	14	31	0.0114332	0.7555920	0.7396970	0.7714870	0.2783315	0.0086389
1195	11	30	0.0093220	0.7485483	0.7324222	0.7646744	0.2876535	0.0069780
1154	7	33	0.0061538	0.7439419	0.7276627	0.7602210	0.2938074	0.0045781
1114	13	28	0.0118182	0.7351498	0.7185763	0.7517234	0.3056255	0.0086881
1073	7	16	0.0065728	0.7303179	0.7135832	0.7470525	0.3121983	0.0048002
1050	14	14	0.0134228	0.7205149	0.7034591	0.7375708	0.3256211	0.0096713
1022	12	17	0.0118402	0.7119839	0.6946557	0.7293122	0.3374613	0.0084300
993	10	19	0.0101678	0.7047446	0.6871882	0.7223011	0.3476291	0.0071657
964	14	9	0.0145909	0.6944618	0.6765890	0.7123345	0.3622200	0.0101328
941	10	12	0.0106952	0.6870344	0.6689400	0.7051287	0.3729152	0.0073480
919	11	29	0.0121614	0.6786790	0.6603364	0.6970217	0.3850766	0.0082537
879	7	18	0.0080460	0.6732184	0.6547119	0.6917249	0.3931226	0.0054167
854	4	18	0.0047337	0.6700316	0.6514280	0.6886352	0.3978563	0.0031717

n	events	dropouts	hazard	S	lower	upper	cumHazard	f(t)
832	8	20	0.0097324	0.6635106	0.6447055	0.6823157	0.4075887	0.0064575
804	15	25	0.0189514	0.6509362	0.6317442	0.6701282	0.4265400	0.0123361
764	8	22	0.0106242	0.6440205	0.6246156	0.6634254	0.4371642	0.0068422
734	11	18	0.0151724	0.6342492	0.6145429	0.6539555	0.4523366	0.0096231
705	5	13	0.0071582	0.6297091	0.6098629	0.6495553	0.4594948	0.0045076
687	8	19	0.0118081	0.6222734	0.6021960	0.6423508	0.4713029	0.0073479
660	2	16	0.0030675	0.6203646	0.6002264	0.6405028	0.4743704	0.0019030
642	3	34	0.0048000	0.6173868	0.5971477	0.6376260	0.4791704	0.0029635
605	5	600	0.0163934	0.6072657	0.5860332	0.6284983	0.4955638	0.0099552

tempo	risco	falha	censura	S(t)	upper	lower	H(t)	h(t)	f(t)
1	2453	47	66	0.9808398	0.9862648	0.9754148	0.0191602	0.0191602	0.0187931
2	2340	41	47	0.9636541	0.9711104	0.9561979	0.0366816	0.0175214	0.0168845
3	2252	32	33	0.9499610	0.9586912	0.9412308	0.0508912	0.0142096	0.0134986
4	2187	30	33	0.9369300	0.9467067	0.9271533	0.0646086	0.0137174	0.0128523
5	2124	29	37	0.9241376	0.9348322	0.9134431	0.0782621	0.0136535	0.0126177
6	2058	22	25	0.9142586	0.9256076	0.9029096	0.0889521	0.0106900	0.0097734
7	2011	25	40	0.9028929	0.9149436	0.8908421	0.1013837	0.0124316	0.0112244
8	1946	18	39	0.8945414	0.9070830	0.8819997	0.1106334	0.0092497	0.0082743
9	1889	32	43	0.8793877	0.8927708	0.8660045	0.1275736	0.0169402	0.0148970
10	1814	19	36	0.8701769	0.8840459	0.8563078	0.1380477	0.0104741	0.0091143
11	1759	31	35	0.8548412	0.8694788	0.8402036	0.1556714	0.0176236	0.0150654
12	1693	27	47	0.8412082	0.8564889	0.8259274	0.1716194	0.0159480	0.0134156
13	1619	22	36	0.8297773	0.8455793	0.8139752	0.1852080	0.0135886	0.0112755
14	1561	13	27	0.8228669	0.8389777	0.8067562	0.1935360	0.0083280	0.0068528
15	1521	25	23	0.8093418	0.8260373	0.7926463	0.2099726	0.0164366	0.0133028
16	1473	10	24	0.8038473	0.8207732	0.7869214	0.2167614	0.0067889	0.0054572
17	1439	11	38	0.7977025	0.8148842	0.7805209	0.2244056	0.0076442	0.0060978
18	1390	11	28	0.7913898	0.8088358	0.7739438	0.2323193	0.0079137	0.0062628
19	1351	12	24	0.7843604	0.8020990	0.7666218	0.2412016	0.0088823	0.0069669
20	1315	15	14	0.7754133	0.7935182	0.7573085	0.2526084	0.0114068	0.0088450
21	1286	15	31	0.7663689	0.7848320	0.7479057	0.2642725	0.0116641	0.0089390
22	1240	14	31	0.7577163	0.7765191	0.7389135	0.2755628	0.0112903	0.0085549
23	1195	11	30	0.7507415	0.7698177	0.7316654	0.2847679	0.0092050	0.0069106
24	1154	7	33	0.7461876	0.7654440	0.7269312	0.2908337	0.0060659	0.0045263
25	1114	13	28	0.7374799	0.7570847	0.7178750	0.3025034	0.0116697	0.0086061
26	1073	7	16	0.7326687	0.7524670	0.7128704	0.3090271	0.0065238	0.0047798

tempo	risco	falha	censura	S(t)	upper	lower	H(t)	h(t)	f(t)
27	1050	14	14	0.7228998	0.7430846	0.7027150	0.3223605	0.0133333	0.0096387
28	1022	12	17	0.7144117	0.7349229	0.6939006	0.3341022	0.0117417	0.0083884
29	993	10	19	0.7072173	0.7280009	0.6864336	0.3441727	0.0100705	0.0071220
30	964	14	9	0.6969465	0.7181132	0.6757798	0.3586955	0.0145228	0.0101216
31	941	10	12	0.6895400	0.7109738	0.6681063	0.3693225	0.0106270	0.0073277
32	919	11	29	0.6812866	0.7030116	0.6595615	0.3812920	0.0119695	0.0081547
33	879	7	18	0.6758611	0.6977817	0.6539404	0.3892556	0.0079636	0.0053823
34	854	4	18	0.6726954	0.6947319	0.6506590	0.3939394	0.0046838	0.0031508
35	832	8	20	0.6662272	0.6885029	0.6439515	0.4035548	0.0096154	0.0064060
36	804	15	25	0.6537976	0.6765285	0.6310667	0.4222115	0.0186567	0.0121977
37	764	8	22	0.6469516	0.6699341	0.6239690	0.4326827	0.0104712	0.0067744
38	734	11	18	0.6372561	0.6605975	0.6139147	0.4476691	0.0149864	0.0095502
39	705	5	13	0.6327365	0.6562462	0.6092269	0.4547613	0.0070922	0.0044875
40	687	8	19	0.6253684	0.6491522	0.6015846	0.4664062	0.0116448	0.0072823
41	660	2	16	0.6234734	0.6473297	0.5996171	0.4694365	Inf	Inf
42	642	3	34	0.6205600	0.6445315	0.5965884	0.4741094	0.0046729	0.0028998
43	605	5	600	0.6154314	0.6396226	0.5912401	0.4823738	0.0082645	0.0050862

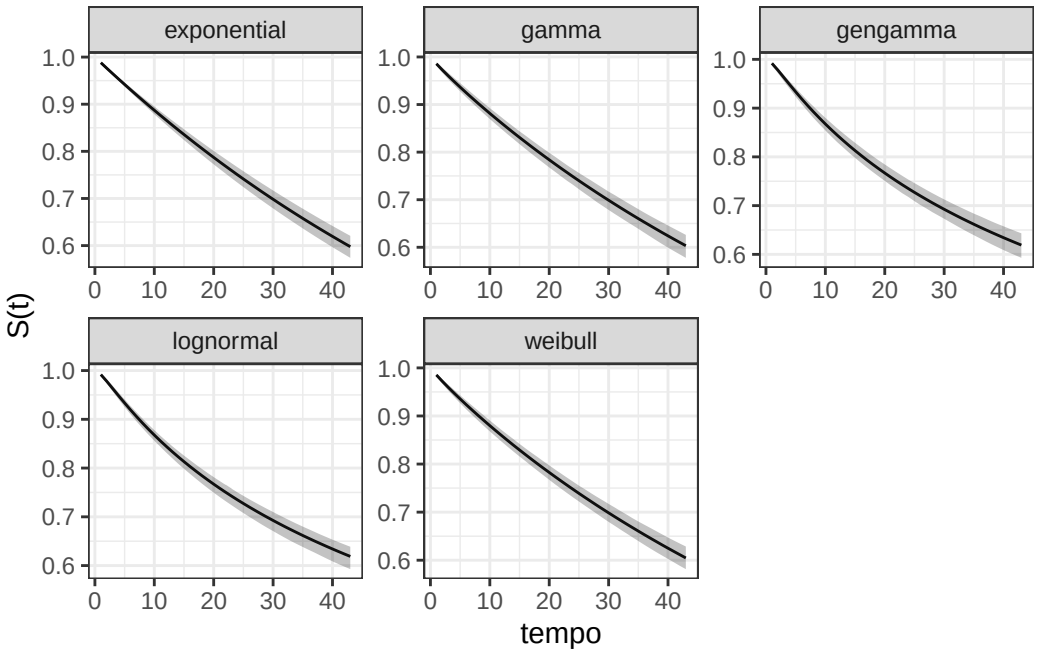
iv. Explique como o teste de LogRank deve ser aplicado. Escolha uma variável qualitativa de sua base e realize o teste de comparação de curvas de sobrevivência. Interprete adequadamente os resultados.

comparação	estatística	p-valor
outr vs hiper	0.6326518	0.4263843
outr vs diab	20.9698863	0.0000047
outr vs rim	0.5012909	0.4789335
outr vs cong	1.8960397	0.1685223
hiper vs diab	26.9880266	0.0000002
hiper vs rim	4.2081902	0.0402292
hiper vs cong	1.0446274	0.3067477
diab vs rim	39.1219659	0.0000000
diab vs cong	1.9015330	0.1679068
rim vs cong	3.1526686	0.0758029

Modelo	logvero	AIC	BIC	TRV	p-valor
Gama Generalizado	-3708.408	7422.817	7440.232	NA	NA
Exponencial	-3721.816	7445.632	7451.437	-26.8150245	1

Modelo	logvero	AIC	BIC	TRV	p-valor
Weibull	-3719.989	7443.979	7455.589	-23.1616870	1
Log-Normal	-3708.411	7420.822	7432.432	-0.0052486	1
Gama	-3720.558	7445.115	7456.725	-24.2983908	1

```
# A tibble: 43 x 16
   tempo `S(t)_exponential` `S(t)_weibull` `S(t)_gengamma` `S(t)_lognormal`
   <dbl>              <dbl>          <dbl>           <dbl>             <dbl>
 1     1              0.988          0.985           0.992             0.992
 2     2              0.976          0.972           0.978             0.978
 3     3              0.965          0.960           0.963             0.963
 4     4              0.953          0.947           0.948             0.948
 5     5              0.942          0.935           0.933             0.933
 6     6              0.931          0.924           0.919             0.919
 7     7              0.920          0.913           0.906             0.906
 8     8              0.909          0.901           0.892             0.893
 9     9              0.898          0.891           0.880             0.880
10    10              0.887          0.880           0.868             0.868
# i 33 more rows
# i 11 more variables: `S(t)_gamma` <dbl>, lower_exponential <dbl>,
#   lower_weibull <dbl>, lower_gengamma <dbl>, lower_lognormal <dbl>,
#   lower_gamma <dbl>, upper_exponential <dbl>, upper_weibull <dbl>,
#   upper_gengamma <dbl>, upper_lognormal <dbl>, upper_gamma <dbl>
```


exponential
gamma
gengamma
lognormal
weibull
S(t)
tempo
1.0
0.9
0.8
0.7
0.6
0
10
20
30
40