UNIVERSIDADE FEDERAL DE MINAS GERAIS

ESPECIALIZAÇÃO EM ESTATÍSTICA

Confiabilidade

Lista 4

Professor: Enrico Colosimo

Aluno: Augusto Filho

Belo Horizonte - MG

Exercício 1. *Os dados abaixo foram apresentados por Nelson & Schmee (1979) e mostram o número de milhas necessárias para que um dispositivo de controle de locomotivas falhasse em um teste de vida. Foram usados 96 dispositivos de controle e o teste terminava após 135.000 milhas (censura do tipo I). Foram observadas 37 falhas que estão listadas abaixo. Conseqüentemente ocorreram 59 censuras em 135.000 milhas. O número de mil milhas até a falha para os 37 dispositivos não censurados foram:*

22,5	*51,5*	*66,5*	*77,0*	*82,0*	*93,5*	*112,5*	*118,5*	*123,0*	*134,0*
37,5	*53,0*	*68,0*	*78,5*	*83,0*	*102,5*	*113,5*	*119,0*	*127,5*	
46,0	*54,5*	*69,5*	*80,0*	*84,0*	*107,0*	*116,0*	*120,0*	*131,0*	
48,5	*57,5*	*76,5*	*81,5*	*91,5*	*108,5*	*117,0*	*122,5*	*132,5*	

O responsável pelo estudo tem interesse em estimar o tempo médio (MTTF) e mediano de vida destes dispositivos e o percentual de falhas após 100.000 milhas de uso.

(a) *Construa a estimativa de Kaplan-Meier para a função de confiabilidade ($R(t)$) para os resultados do teste. Obtenha uma estimativa para o MTTF, o tempo mediano e para o percentual de falhas após 100.000 milha de uso. Reporte estes resultados na Tabela Resumo em anexo.*

Resolução*:*

A seguir, temos os resultados obtidos pelo MINITAB. Chamamos a atenção para a versão do MINITAB em uso na resolução deste exercício. **MINITAB 14***.*

```
Distribution Analysis: Observ


Variable: Observ

Censoring Information  Count
Uncensored value          37
Right censored value      59


Censoring value: censur = 0


Nonparametric Estimates


Characteristics of Variable


            Standard   95,0% Normal CI
Mean(MTTF)     Error     Lower    Upper
   116,802   2,95724  111,006  122,598
Median = *
IQR = *  Q1 = 108,5  Q3 = *
```

Kaplan-Meier Estimates

Time	at Risk	Number Failed	Survival Probability	Standard Error	95,0% Normal CI Lower	Upper
22,5	96	1	0,989583	0,0103623	0,969274	1,00000
37,5	95	1	0,979167	0,0145771	0,950596	1,00000
46,0	94	1	0,968750	0,0177580	0,933945	1,00000
48,5	93	1	0,958333	0,0203947	0,918360	0,99831
51,5	92	1	0,947917	0,0226777	0,903469	0,99236
53,0	91	1	0,937500	0,0247053	0,889079	0,98592
54,5	90	1	0,927083	0,0265361	0,875074	0,97909
57,5	89	1	0,916667	0,0282085	0,861379	0,97195
66,5	88	1	0,906250	0,0297491	0,847943	0,96456
68,0	87	1	0,895833	0,0311776	0,834726	0,95694
69,5	86	1	0,885417	0,0325087	0,821701	0,94913
76,5	85	1	0,875000	0,0337539	0,808844	0,94116
77,0	84	1	0,864583	0,0349224	0,796137	0,93303
78,5	83	1	0,854167	0,0360217	0,783565	0,92477
80,0	82	1	0,843750	0,0370579	0,771118	0,91638
81,5	81	1	0,833333	0,0380363	0,758784	0,90788
82,0	80	1	0,822917	0,0389611	0,746554	0,89928
83,0	79	1	0,812500	0,0398361	0,734423	0,89058
84,0	78	1	0,802083	0,0406645	0,722382	0,88178
91,5	77	1	0,791667	0,0414491	0,710428	0,87291
93,5	76	1	0,781250	0,0421923	0,698555	0,86395
102,5	75	1	0,770833	0,0428964	0,686758	0,85491
107,0	74	1	0,760417	0,0435631	0,675035	0,84580
108,5	73	1	0,750000	0,0441942	0,663381	0,83662
112,5	72	1	0,739583	0,0447912	0,651794	0,82737
113,5	71	1	0,729167	0,0453554	0,640272	0,81806
116,0	70	1	0,718750	0,0458880	0,628811	0,80869
117,0	69	1	0,708333	0,0463902	0,617410	0,79926
118,5	68	1	0,697917	0,0468629	0,606067	0,78977
119,0	67	1	0,687500	0,0473070	0,594780	0,78022
120,0	66	1	0,677083	0,0477233	0,583547	0,77062
122,5	65	1	0,666667	0,0481125	0,572368	0,76097
123,0	64	1	0,656250	0,0484753	0,561240	0,75126
127,0	63	1	0,645833	0,0488122	0,550163	0,74150
131,0	62	1	0,635417	0,0491238	0,539136	0,73170
132,5	61	1	0,625000	0,0494106	0,528157	0,72184
134,0	60	1	0,614583	0,0496730	0,517226	0,71194

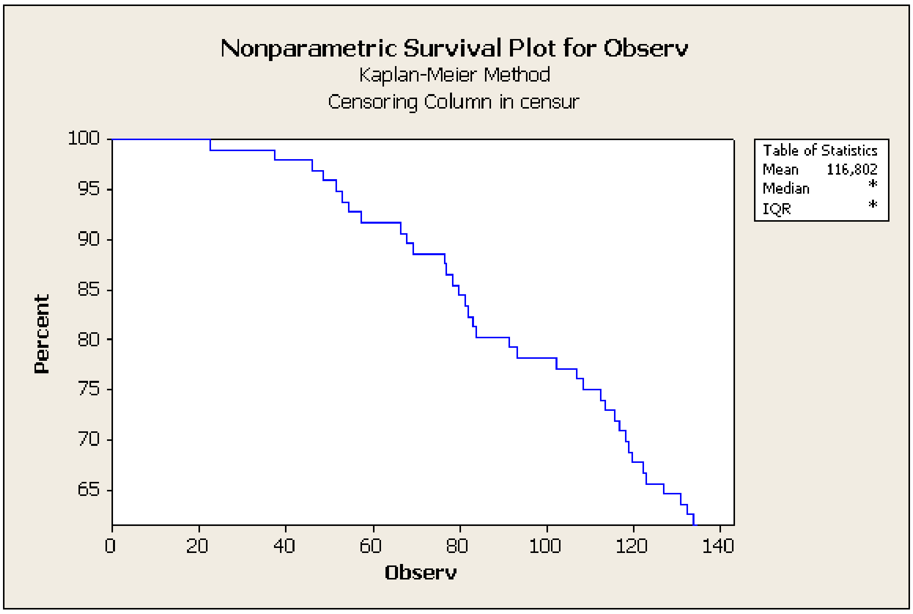

Figura 1: Survival plot

A seguir, descreveremos uma estimativa para o MTTF, o tempo mediano e para o percentual de falhas após 100.000 milhas de uso.

- *Muito embora a estimativa para o MTTF (tempo médio), pelo método de Kaplan-Meier não tenha uma boa aproximação, o valor encontrado foi:* $\overline{X}$ *=116,802.* ***(Utilizou-se o MINITAB 14.)***
- *Não foi possível encontrar a mediana* $(\widetilde{X})$ *pelo método de Kaplan-Meier.*
- *A seguir, temos via interpolação linear o percentual de falhas após 100.00 milhas de uso:*

$$\begin{aligned}
\frac{102,5-93,5}{0,770833-0,781250} &= \frac{100-93,5}{x-0,781250} \\
\frac{9}{-0,010417} &= \frac{6,5}{x-0,781250} \\
-863,9723529 &= \frac{6,5}{x-0,781250} \\
-863,9723529x+674,9784007 &= 6,5 \\
-863,9723529x &= 6,5-674,9784007 \\
-863,9723529x &= -668,4784007 \\
x &= \frac{-668,4784007}{-863,9723529} \qquad \textit{Portanto, o valor para } x \textit{ é:} \\
x &= 0,773726611.
\end{aligned}$$

Desta forma, o percentual de falhas após 100.000 milhas é de aproximadamente $77,4\%$.

(b) *Ajuste o modelo log-normal aos dados. Anote o valor das estimativas dos parâmetros:*

i) $B(2) = \hat{\mu}$

ii) $B(1) = \hat{\sigma}$

iii) *com os valores das estimativas dos parâmetros, escreva a forma funcional da estimativa da função de confiabilidade da log-normal:* $\hat{R}_{logn}(t) =$

Resolução*:*

A seguir, veremos os resultados computacionais do Software ***Minitab 14*** *para as perguntas acima.*

```
Distribution Analysis: Observ
Variable: Observ

Censoring Information  Count
Uncensored value          37
Right censored value      59

Censoring value: censur = 0
Estimation Method: Maximum Likelihood
Distribution:   Lognormal

Parameter Estimates
                      Standard    95,0% Normal CI
Parameter  Estimate      Error      Lower     Upper
Location    5,11687   0,104159    4,91273   5,32102
Scale      0,705501  0,0932023   0,544562  0,914004

Log-Likelihood = -237,092
Goodness-of-Fit
Anderson-Darling (adjusted) = 367,588

Characteristics of Distribution
                                        Standard   95,0% Normal CI
                             Estimate      Error    Lower     Upper
Mean(MTTF)                    213,949    32,6279  158,672   288,484
Standard Deviation            171,825    52,6957  94,1973   313,426
Median                        166,813    17,3750  136,010   204,593
First Quartile(Q1)            103,650    8,71630  87,9000   122,222
Third Quartile(Q3)            268,467    40,2969  200,044   360,293
Interquartile Range(IQR)      164,817    36,2763  107,066   253,718
```

Table of Percentiles

Percent	Percentile	Standard Error	95,0% Normal CI Lower	Upper
0,1	18,8537	4,55585	11,7412	30,2749
1	32,3182	5,70873	22,8608	45,6883
2	39,1716	6,06260	28,9221	53,0532
3	44,2551	6,26246	33,5360	58,4004
4	48,5096	6,40092	37,4551	62,8268
5	52,2702	6,50808	40,9518	66,7170
6	55,6999	6,59777	44,1599	70,2556
7	58,8918	6,67759	47,1562	73,5479
8	61,9046	6,75224	49,9894	76,6598
9	64,7782	6,82490	52,6924	79,6361
10	67,5412	6,89784	55,2888	82,5088
30	115,228	9,79675	97,5408	136,121
40	139,510	12,8648	116,443	167,147
50	166,813	17,3750	136,010	204,593
60	199,459	23,8600	157,772	252,161
70	241,493	33,5032	183,998	316,953
80	302,064	49,3080	219,357	415,955
90	411,995	82,1941	278,660	609,128
91	429,567	87,8551	287,702	641,385
92	449,508	94,3962	297,842	678,405
93	472,504	102,087	309,385	721,625
94	499,581	111,334	322,784	773,213
95	532,360	122,790	338,747	836,635
96	573,630	137,593	358,476	917,918
97	628,777	157,984	384,262	1028,88
98	710,378	189,316	421,350	1197,66
99	861,018	250,316	487,024	1522,21

Desta forma, utilizando-se o método de ***máxima verossimilhança*** *aplicado à distribuição lognormal, encontrou-se as seguintes estimativas para μ e σ:*

i) Encontrou-se $\hat{\mu} = 5,11687$;

ii) Encontrou-se $\hat{\sigma} = 0,705501$

iii) A seguir, a forma funcional da estimativa da função de confiabilidade da lognormal:

$$\hat{R}_{logn}(t) = \Phi\{-\left[log(t) - \mu\right]/\sigma\} \therefore \Phi\{-\left[log(t) - 5,11687\right]/0,705501\}$$

(c) *Utilize as técnicas gráficas apresentadas em sala (gráfico de linearização da função de confiabilidade e gráfico KM vs $\hat{R}_{logn}$) e verifique a adequação do modelo log-normal aos dados. O modelo log-normal lhe parece adequado para representar a distribuição do tempo de falha do dispositivo de controle? Justifique.*

Resolução*:*

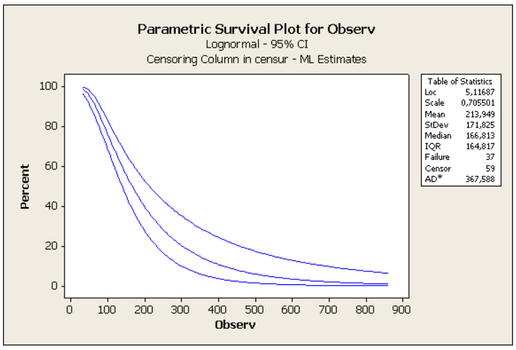

Figura 2: Survival plot for Lognormal

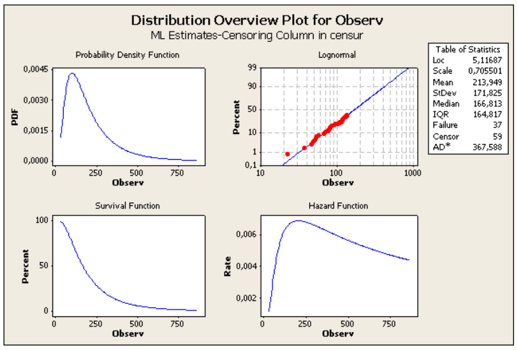

Figura 3: Distribution plot Lognormal

Uma rápida comparação entre os gráficos obtidos pelos métodos de KM e $\hat{R}_{logn}$:

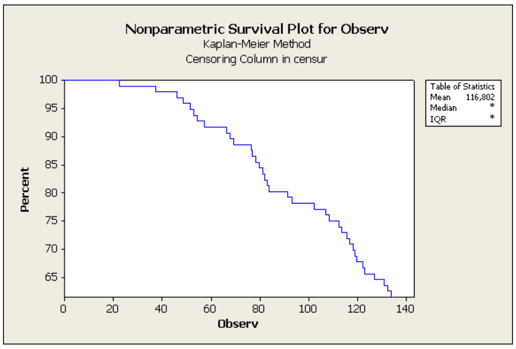

Figura 4: Survival plot:KM

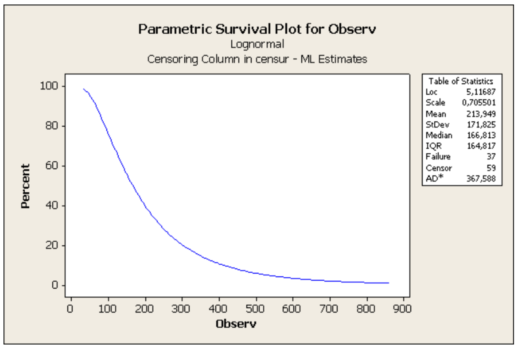

Figura 5: Survival plot: $\hat{R}_{logn}(t)$

Pelos gráficos acima, o modelo Lognormal pareceu mais adequado para representar a distribuição de falhas do dispositivo de controle por se adequar melhor aos dados originais com um coeficiente de correlação linear de 0,991.

(d) *Ajuste o modelo exponencial aos dados. Anote o valor das estimativas dos parâmetros:*

i) $B(2) = \hat{\mu} = log(\hat{\alpha})$
(no caso da exponencial, $\sigma = \frac{1}{\delta} = 1$*)*

ii) *E também encontre,* $\hat{R}(t)$
(escreva a forma funcional)

Resolução*:*

Para o modelo Exponencial, obteve-se as seguintes estimativas para μ e σ.

i) *Temos* $B(2) = \hat{\mu} = log(\hat{\alpha}) = 304,649$ *e* $\hat{\sigma} = 1$*, pelo método de máxima verossimilhança.*

A forma funcional da função exponencial é dada a seguir:

ii) $\hat{R}(t) = exp\{-t/\alpha\} = exp\{-t/304,649\}$

Ajuste o modelo Weibull aos dados e anote o valor das estimativas dos parâmetros:

i) $B(2) = \hat{\mu} = log(\hat{\alpha})$

ii) $B(1) = \hat{\sigma} = \frac{1}{\delta}$

iii) *E também encontre,* $\hat{R}(t)$
(escreva a forma funcional do modelo Weibull).

Resolução*:*

A seguir, obtemos os resultados para μ e σ pelo método de máxima verossimilhança.

i) *Logo, temos* $\hat{\mu} = 183,399$*;*

ii) *O valor do desvio padrão foi de* $\hat{\sigma} = 2,33096$

A forma funcional é dada a seguir:

iii) $R(t) = exp\left[-\left(\frac{t}{\alpha}\right)^{\delta}\right] = exp\left[-\left(\frac{t}{183,40}\right)^{0,43}\right].$

(e) *Utilize as técnicas gráficas e verifique a adequação dos modelos exponencial e Weibull aos dados. Compare os gráficos com aqueles obtidos para o modelo log-normal. Dentre os 3 modelos escolha o modelo mais adequado para representar a distribuição do tempo de falha do dispositivo de controle (justifique sua escolha). Para este modelo, estime as 3 figuras de mérito de interesse (MTBF, tempo mediano e percentual de falhas após 100.000 milhas). Anote este valores na Tabela Resumo.*

Resolução*:*

A seguir, apresentamos a saída computacional do programa ***Minitab***[1] ***14.***

[1] Para editar as saídas computacionais do Minitab utilizou-se o package Sweave.

Distribution ID Plot: Observ

Goodness-of-Fit

Distribution	Anderson-Darling (adj)
Weibull	367,591
Exponential	368,213
Lognormal	367,588
Normal	367,621

Table of Percentiles

Distribution	Percent	Percentile	Standard Error	95% Normal CI Lower	Upper
Weibull	1	25,4868	6,61528	15,3243	42,3888
Exponential	1	3,06182	0,503360	2,21842	4,22587
Lognormal	1	32,3182	5,70873	22,8608	45,6883
Normal	1	9,39255	15,4310	-20,8517	39,6368
Weibull	5	51,2864	8,10958	37,6190	69,9193
Exponential	5	15,6264	2,56897	11,3220	21,5673
Lognormal	5	52,2702	6,50808	40,9518	66,7170
Normal	5	50,9498	10,8383	29,7071	72,1924
Weibull	10	69,8421	8,19140	55,4989	87,8922
Exponential	10	32,0979	5,27687	23,2563	44,3010
Lognormal	10	67,5412	6,89784	55,2888	82,5088
Normal	10	73,1038	8,85666	55,7450	90,4625
Weibull	50	156,715	12,1003	134,706	182,319
Exponential	50	211,166	34,7155	152,999	291,448
Lognormal	50	166,813	17,3750	136,010	204,593
Normal	50	151,252	9,01149	133,590	168,914

Table of MTTF

Distribution	Mean	Standard Error	95% Normal CI Lower	Upper
Weibull	162,503	14,3993	136,596	193,324
Exponential	304,649	50,0839	220,731	420,471
Lognormal	213,949	32,6279	158,672	288,484
Normal	151,252	9,0115	133,590	168,914

A seguir, veremos os gráficos para as distribuições Weibull, Exponencial e Lognormal.

Começaremos com a distribuição ***Weibull****:*

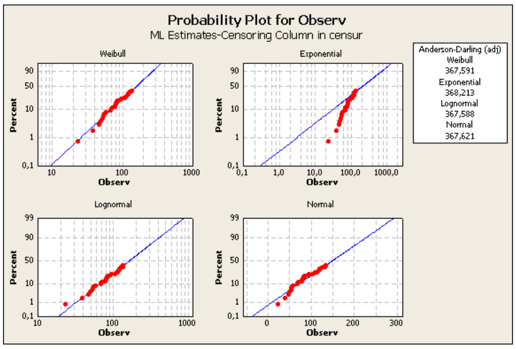

Figura 6: Probability Plot

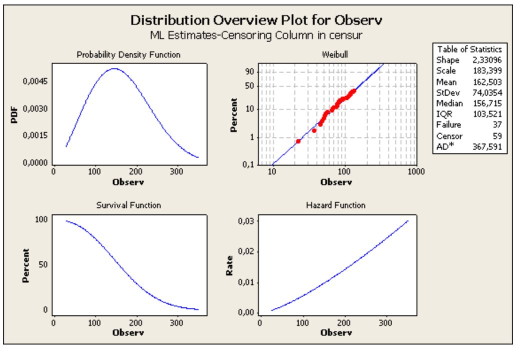

Figura 7: Distribuição Weibull

A seguir, a distribuição ***Exponencial****:*

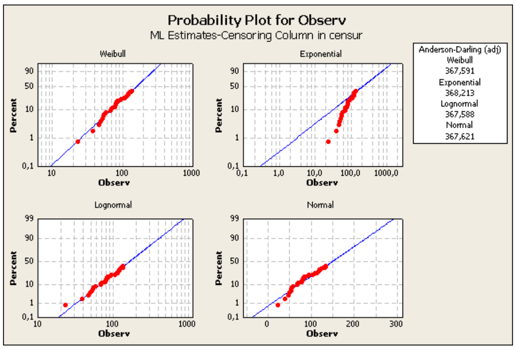

Figura 8: Probability Plot

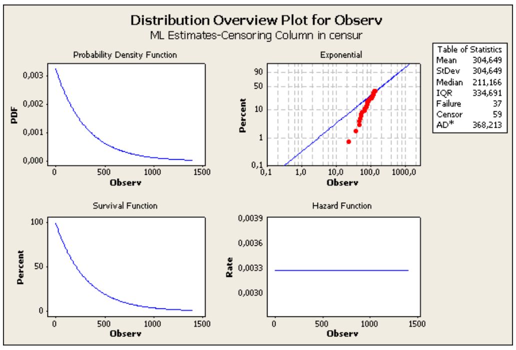

Figura 9: Distribuição Exponencial

E finalmente a distribuição ***Lognormal****:*

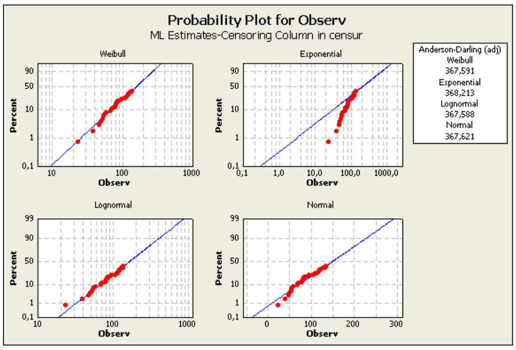

Figura 10: Probability Plot

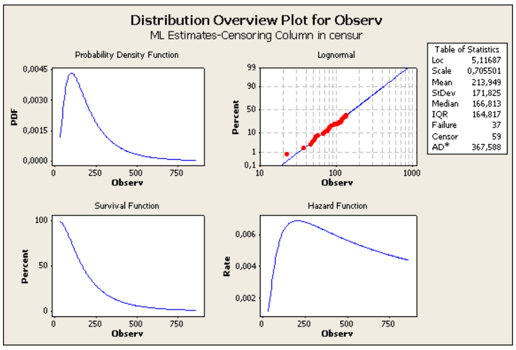

Figura 11: Distribuição Lognormal

A tabela a seguir resume os valores de "Mean(MTTF)", os valores medianos e os percentuais de falhas após 100.000 milhas para cada distribuição mencionada acima.

Tabela 1: Resumo das estimativas

	Weibull	Exponencial	Lognormal
$Mean(MTTF)$	162,503	304,649	213,949
I.C. para $MTTF$	(136,596 ; 193,324)	(220,731 ; 420,471)	(158,672 ; 288,484)
$\hat{t}_{0,50}$	156,715	211,166	166,813
I.C. para $\hat{t}_{0,50}$	(134,706 ; 182,319)	(152,999 ; 291,448)	(136,010 ; 204,593)
$1 - \hat{R}(100)$	21,59%	27,98%	23,41%
I.C. para $1 - \hat{R}(100)$	(15,5 ; 30)	(21,16 ; 36,43)	(16,86 ; 31,16)

*Os valores acima são medidos em x 100 milhas.

Os valores de Anderson-Darling obtidos foram:

```
        Anderson-Darling

Distribution           (adj)
Weibull              367,591
Exponential          368,213
Lognormal            367,588
Normal               367,621
```

*Desta forma, baseados nas medidas das Tabelas 1 e nos valores de Anderson-Darling, é possível concluir que o melhor ou melhores ajustamento são obtidos utilizando-se os modelos de **Weibull** e **Lognormal**. Muito embora os dois modelos sejam praticamente iguais, para efeito didático escolheremos o modelo Lognormal por apresentar menor valor para o coeficiente ajustado de Anderson-Darling.*

(f) *Analise os valores estimados que estão na Tabela Resumo. Se você necessitar representar a performance deste dispositivo em termos de uma grandeza, você utilizaria o MTTF ou o tempo mediano? Qual valor você reportaria: o valor baseado na estimativa de KM ou no modelo escolhido?*

Tabela 2: Tabela Resumo

	MTTF (mil milhas)	$t_{0,50}$ (mil milhas)	$\hat{R}(100)$ (%)
KM	116,802	*	77,37%
Lognormal	213,9	166,8	76,6%

Resolução*:*

A principal desvantagem de usarmos o método de Kaplan-Meier neste exemplo é que ele subestima o tempo médio de vida. Logo, o valor encontrado para MTTF não é um bom valor e por haver mais censura do que falha não foi possível encontrar o valor de $t_{0,5}$ pelo método de Kaplan-Meier.

Já o método paramétrico, em particular com a utilização do modelo Lognormal, foi possível obtermos uma maior precisão nas estimativas, por apresentar um menor intervalo de confiança, ou seja, quando se acrescenta uma informação a mais se ganha em precisão, pois o I.C. é menor. O ganho nos valores estimados com os modelos paramétricos é sem dúvida um diferencial, no entanto, deve-se levar em conta o cuidado com os pressupostos das distribuições paramétricas para o seu uso.

Portanto se houver necessidade de reportarmos a performance deste dispositivo em termos de uma grandeza, deve-se escolher o valor de MTTF, levando-se em conta a distribuição Lognormal e o Método de Kaplan-Meier.