

Métodos Bootstrap aplicados as distribuições Exponenciais Bivariada

André F. B. Menezes¹

¹Departamento de Estatística, Universidade Estadual de Maringá, DEs, PR, Brazil

1 Introdução

Seguramente a distribuição exponencial é um dos principais modelos escolhidos para analisar dados de sobrevivência e confiabilidade. Além disso, em diversas aplicações usualmente dois tempos de sobrevida são conjuntamente medidos. Neste contexto, os modelos mais comuns a serem adotados são as distribuições exponenciais bivariadas. No presente trabalho, a estimação dos parâmetros das distribuições exponenciais bivariada proposta por Gumbel (1960) e Hougaard (1986) será abordada sob o ponto de vista clássico e os métodos Bootstrap serão aplicados para alcançar inferências mais refinadas a cerca dos parâmetros.

Dentro do contexto de confiabilidade, o modelo de estresse-resistência descreve a vida de um componente que tem uma força aleatória Y e que é submetido a um estresse X . O componente falha no instante em que o estresse aplicado excede a resistência, e o componente irá funcionar sempre que satisfazer a relação $Y > X$. Assim, $R = \Pr(X < Y)$ é uma medida de confiabilidade do componente (Nadarajah and Kotz, 2006). A quantidade R é de interesse em diversas aplicações principalmente em engenharia (Kotz et al., 2003), logo é importante obter estimativas pontuais e intervalares confiáveis. Usualmente, R é calculado considerando que X e Y são variáveis aleatórias independentes pertencentes a mesma família univariada de distribuições.

Nadarajah and Kotz (2006) derivaram expressões para R considerando 6 distribuições exponenciais bivariadas. Ressalta-se que a confiabilidade para as distribuições exponenciais bivariadas é calculada por

$$R = \Pr(X < Y) = \int_0^\infty \int_x^\infty f(x, y) dy dx, \quad (1)$$

em que $f(x, y)$ denota a função densidade conjunta de (X, Y) .

Em geral, inferências sob os parâmetros das distribuições exponenciais bivariadas são realizadas sob o enfoque clássico ou Bayesiano. Ao nosso melhor conhecimento, nenhum trabalho explorou os métodos Bootstrap aplicados para melhorar a estimação dos parâmetros e funções dos parâmetros para as distribuições exponenciais bivariadas estudadas neste trabalho.

2 Distribuições Exponenciais Bivariada

2.1 Distribuição Gumbel Exponencial Bivariada (GBE)

Introduzida por Gumbel (1960) essa versão bivariada da distribuição exponencial tem sido bastante explorada na literatura. Dentre as principais referências podemos citar Damien and Müller (1998), Nair and Nair (1988), Pal and Murthy (2003), Nadarajah and Kotz (2006), Nadarajah (2006) e Balakrishnan and Lai (2009).

Sua função de sobrevivência conjunta é definida por:

$$S(x, y) = \exp [- (\lambda_1 x + \lambda_2 y + \theta \lambda_1 \lambda_2 xy)] . \quad (2)$$

Enquanto que a função de densidade conjunta de X e Y pode ser expressa por:

$$f(x, y) = [(1 - \theta) \lambda_1 \lambda_2 + \theta \lambda_1^2 \lambda_2 x + \theta \lambda_1 \lambda_2^2 y + \theta^2 \lambda_1^2 \lambda_2^2 x y] S(x, y) \quad (3)$$

para $x, y > 0$, $\lambda_1, \lambda_2 \geq 0$ e $0 < \theta < 1$. As distribuições marginais de X e Y são exponenciais com parâmetros de taxa λ_1 e λ_2 , respectivamente. Além disso, o coeficiente de correlação $\rho = \text{Corr}(X, Y)$ é definido por

$$\rho = \frac{1}{\theta} \exp \left(\frac{1}{\theta} \right) E_1 \left(\frac{1}{\theta} \right) - 1, \quad (4)$$

em que $E_1(\cdot)$ representa a função exponencial integral (Abramowitz and Stegun, 1974), definida como

$$E_1(x) = \int_x^\infty \frac{e^{-t}}{t} dt. \quad (5)$$

Nadarajah and Kotz (2006) deduziram uma expressão analítica para confiabilidade, $R = P(X < Y)$, quando o vetor aleatório (X, Y) tem distribuição Gumbel Exponencial bivariada. A expressão de R é dada por

$$R = \frac{1}{2} + \frac{\sqrt{\pi} (\lambda_1 - \lambda_2)}{2\sqrt{\theta \lambda_1 \lambda_2}} \exp \left[\frac{(\lambda_1 + \lambda_2)}{4\theta \lambda_1 \lambda_2} \right] \left[1 - \Phi \left(\frac{\lambda_1 + \lambda_2}{\sqrt{2\theta \lambda_1 \lambda_2}} \right) \right] \quad (6)$$

em que $\Phi(\cdot)$ denota a função distribuição acumulada da distribuição normal padrão. Percebe-se que se $\lambda_1 < \lambda_2$, então $R < 1/2$ e se $\lambda_1 > \lambda_2$, então $R > 1/2$. Além disso, se $\lambda_1 = \lambda_2$, então $R = 1/2$.

2.2 Distribuição Hougaard Exponencial Bivariada (HBE)

Essa distribuição foi proposta por Hougaard (1986), ela tem função de sobrevivência conjunta e densidade conjunta especificada por:

$$S(x, y) = \exp \left\{ - \left[\left(\frac{x}{\mu_1} \right)^r + \left(\frac{y}{\mu_2} \right)^r \right]^{1/r} \right\}, \quad (7)$$

$$f(x, y) = \frac{(x y)^{r-1}}{(\mu_1 \mu_2)^r} \left[\left(\frac{x}{\mu_1} \right)^r + \left(\frac{y}{\mu_2} \right)^r \right]^{1/r-2} \left\{ r - 1 + \left[\left(\frac{x}{\mu_1} \right)^r + \left(\frac{y}{\mu_2} \right)^r \right]^{1/r} \right\} S(x, y), \quad (8)$$

respectivamente, para $x, y > 0$, $\mu_1, \mu_2 \geq 0$ e $r > 0$. As distribuições marginais de X e Y são exponenciais com médias μ_1 e μ_2 , respectivamente. Além disso, o coeficiente de correlação $\rho = \text{Corr}(X, Y)$ é dado por

$$\rho = \frac{\Gamma(1/r)^2}{r \Gamma(2/r)}. \quad (9)$$

Nadarajah and Kotz (2006) mostraram que confiabilidade, $R = P(X < Y)$, quando o vetor aleatório (X, Y) tem distribuição Hougaard Exponencial bivariada é definida por

$$R = \frac{\mu_2^r}{\mu_1^r + \mu_2^r}. \quad (10)$$

Note que, se $\mu_1 = \mu_2$ então $R = 1/2$

Por fim, apresentamos na Figura 1 o comportamento do coeficiente de correlação de ambas as distribuições para diferentes valores dos parâmetros θ e r . Note que para a distribuição GBE a correlação é zero quando $\theta = 0$ e decresce até -0.40365 a medida que θ tende a 1, isto é, $-0.40365 < \rho < 0$. Em contrapartida, na distribuição HBE a medida que r tende para o infinito a correlação se aproxima de 1, ou seja, $-1 < \rho < 1$. Tendo em vista o comportamento teórico e empírico da correlação entre as variáveis aleatórias X e Y , podemos propor distribuições apropriadas para os dados em estudo.

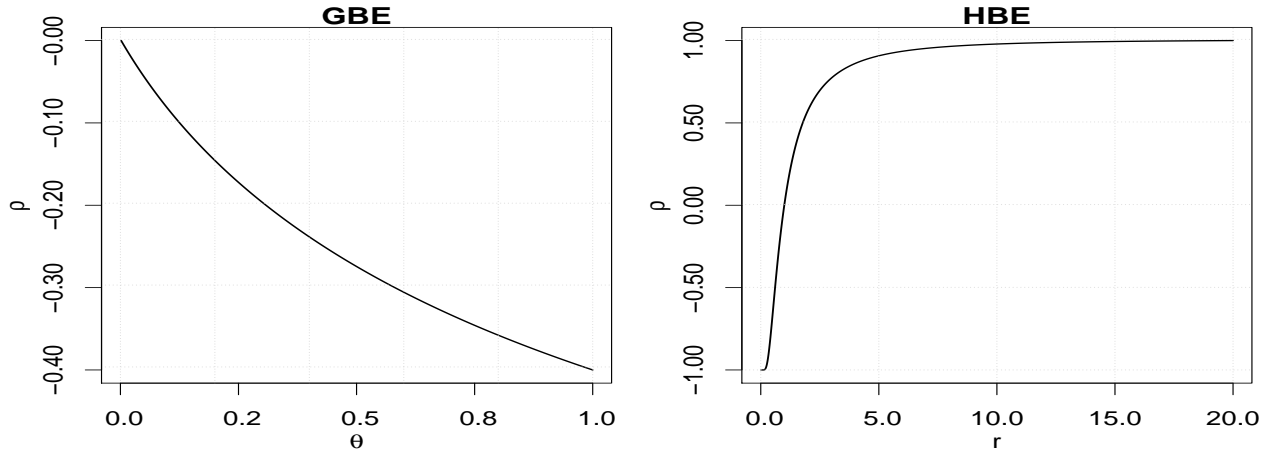


Figura 1: Coeficiente de correlação teórico das distribuições GBE e HBE em função dos parâmetros θ e r .

3 Métodos Bootstrap

Independente do modelo ou metodologia estatística empregada para resolução de problemas, em geral o interesse envolve a estimação de uma quantidade θ desconhecida. Esta quantidade, denominada de parâmetro, é tratada como fixa na abordagem clássica. Em geral, ela está associada ao modelo empregado na análise. Como aponta Efron and Tibshirani (1986), duas principais questões surgem ao propor uma estimativa para θ , sendo elas (i) qual estimador $\hat{\theta}$ deve ser utilizado? (ii) Tendo escolhido usar um determinado estimador $\hat{\theta}$, o quão preciso é ele como um estimador de θ ? O Bootstrap é um método computacionalmente intensivo para responder a segunda pergunta, sendo uma alternativa a análise teórica dos estimadores.

Para introduzir o conceito de Bootstrap considere que $\mathbf{x} = (x_1, \dots, x_n)$ seja uma amostra aleatória de tamanho n , onde cada observação vem aleatoriamente da variável aleatória Y , a qual tem função de distribuição denotada por $F = F_\theta(x)$. Deve-se ressaltar que θ é o parâmetro que indexa a distribuição e é visto como uma função de F , isto é, $\theta = t(F)$. Finalmente, seja $\hat{\theta}$ um estimador de θ baseado em \mathbf{x} , ou seja, $\hat{\theta} = s(\mathbf{x})$.

A aplicação dos métodos Bootstrap proposto por Efron (1979) consiste em obter da amostra observada \mathbf{x} um número grande (B) de amostras artificiais $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ e então extrair informações dessas amostras para realizar e aperfeiçoar a inferência sob θ . Podemos classificar os métodos Bootstrap conforme a amostragem realizada: paramétrico ou não paramétrico. No Bootstrap paramétrico, as amostras são obtidas a partir do modelo adotado, e considerando as estimativas obtidas dos dados observados, isto é, baseado em $F(\hat{\theta})$. Enquanto que na versão não paramétrica, as amostras são obtidas a partir da função distribuição empírica \hat{F} , através da amostragem com reposição dos dados observados. Utilizando cada uma das amostras artificiais x^* nós estimamos θ , e então obtemos $\hat{\theta}^* = \hat{\theta}_1^*, \dots, \hat{\theta}_B^*$ a distribuição Bootstrap do parâmetro θ .

Afim de mensurar a precisão da estimativa pontual $\hat{\theta}$ obtida utilizando os dados observados podemos calcular o erro padrão e o vício Bootstrap. A estimativa Bootstrap do erro padrão é

obtida calculando o desvio padrão da distribuição Bootstrap do parâmetro, isto é,

$$\hat{\sigma}_B(\hat{\theta}) = \sqrt{\frac{1}{B-1} \sum_{i=1}^B \left(\hat{\theta}_{(i)}^* - \hat{\theta}_{(\cdot)}^* \right)^2}, \quad (11)$$

em que B é o número de amostras Bootstrap, $\hat{\theta}_{(i)}^*$ a i -ésima estimativa Bootstrap e $\hat{\theta}_{(\cdot)}^*$ é o valor médio das estimativas Bootstrap.

Além da variação em torno de um valor médio, a acurácia de uma estimativa pontual também é determinada pelo seu viés, isto é, o desvio que o valor esperado apresenta do estimador do verdadeiro valor do parâmetro. Um estimativa Bootstrap para o viés é definida como

$$\hat{\mathcal{B}}(\hat{\theta}) = \hat{\theta}_{(\cdot)}^* - \hat{\theta}, \quad (12)$$

isto é, a diferença entre a média da distribuição Bootstrap do parâmetro e a estimativa obtida pela amostra observada. Assim, uma estimativa para θ corrigida por vício Bootstrap é definida como

$$\hat{\theta}_{BCE} = 2\hat{\theta} - \hat{\theta}_{(\cdot)}^*. \quad (13)$$

Em geral, estimativas intervalares para θ são mais informativas do que uma estimativa pontual. Na literatura diversos métodos Bootstrap foram propostos para obtenção de intervalos de confiança (Carpenter and Bithell, 2000). Neste trabalho, dois intervalos de confiança Bootstrap foram obtidos. Na sequência, dois esquemas apresentados em Mazucheli et al. (2006) para obter os intervalos de confiança p -Bootstrap e t -Bootstrap será exposto.

Intervalos p -Bootstrap

1. Obter uma amostra Bootstrap $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$, seja paramétrica ou não paramétrica.
2. Da amostra Bootstrap \mathbf{x}^* obter o estimador de máxima verossimilhança de θ , denotado por $\hat{\theta}^*$.
3. Repetir os passos [1] e [2] um número B grande de vezes.
4. A partir de $\hat{\theta}^* = (\hat{\theta}_{(1)}^*, \dots, \hat{\theta}_{(B)}^*)$, para algum valor de α , ($0 \leq \alpha \leq 1$), o intervalo de $100 \times (1 - \alpha)\%$ de confiança p -Bootstrap é dado por

$$\left(\hat{\theta}_{(q_1)}^*; \hat{\theta}_{(q_2)}^* \right), \quad (14)$$

em que $q_1 = [B \times (\alpha/2)]$, $q_2 = B - q_1$ e $[\cdot]$ indica o menor número inteiro maior ou igual ao argumento.

Por outro lado, para a obtenção do intervalo de confiança t -Bootstrap, redefinimos o passo 4. na seguinte forma:

Intervalos t -Bootstrap

4. A partir de $\hat{\theta}^* = *(\hat{\theta}_{(1)}^*, \leq \dots, \leq \hat{\theta}_{(B)}^*)$, obter $\mathbf{T}^* = (T_{(1)}^*, \dots, T_{(2)}^*)$, com $T_{(i)} < T_{(j)}$ para $i, j = 1, \dots, B$ e $i \neq j$, em que

$$T = \frac{(\hat{\theta}_i^* - \hat{\theta})}{\sigma_i^*}, \quad (15)$$

$\hat{\theta}$ é o estimador de máxima verossimilhança de θ e σ_i^* é o erro padrão de $\hat{\theta}_i^*$. Note que para alguns casos σ_i^* , $i = 1, \dots, B$, pode ser calculado diretamente pela inversa da matriz de informação, caso contrário, será necessário obter suas estimativas via *Double Bootstrap*.

5. Usando \mathbf{T}^* , para algum valor de α , ($0 < \alpha < 1$), o intervalo de $100 \times (1 - \alpha)\%$ de confiança t -Bootstrap é definido por

$$(\hat{\theta} - \hat{\sigma} \mathbf{T}_{(q_2)}^*; \hat{\theta} - \hat{\sigma} \mathbf{T}_{(q_1)}^*), \quad (16)$$

em que $\hat{\sigma} = \sqrt{\widehat{\text{Var}}(\hat{\theta})}$.

Os métodos Bootstrap descritos foram empregados sob os parâmetros e as respectivas funções R e ρ da distribuições Gumbel e Hougaard Exponencial Bivariada. Importante destacar que os estimadores de máxima verossimilhança foram calculados e a partir deles $B = 20000$ amostras das distribuições consideradas foram geradas, ou seja, o Bootstrap paramétrico foi adotado.

4 Aplicações

Nesta seção duas aplicações com dados simulados são apresentadas para propósitos de ilustração dos métodos Bootstrap discutidos anteriormente. Os parâmetros dos modelos foram estimados tanto para a amostra observada quanto para as amostras Bootstrap pela *procedure* NL MIXED (SAS, 2010b) do software SAS. A geração das amostras Bootstrap foi feita no R. Os resumos a posteriori para os parâmetros das distribuições foram obtidos via *procedure* MCMC (SAS, 2010a).

4.1 Dados simulados da distribuição GEB

Nesta subseção apresentamos um exemplo considerando os dados simulados de uma distribuição Gumbel Exponencial bivariada com parâmetros $\lambda_1 = 2$, $\lambda_2 = 3$ e $\theta = 0.8$, com este cenário têm-se que $R = 0.4207$ e $\rho = -0.3612$. As $n = 100$ observações (ver Tabela 1) foram simuladas utilizando o método da condicional, com auxílio do método da transformação inversa para obter amostras aleatórias da distribuição condicional de Y dado X .

Tabela 1: Dados simulados da distribuição GEB com parâmetros $\lambda_1 = 2$, $\lambda_2 = 3$ e $\theta = 0.8$.

$x :$	0.274, 0.982, 0.070, 0.005, 0.276, 1.757, 1.306, 0.762, 1.038, 0.311, 1.471, 0.162, 0.013, 0.306, 0.380, 1.570, 0.112, 0.053, 0.138, 0.305, 0.009, 0.135, 0.444, 0.059, 0.734, 0.131, 0.106, 0.012, 0.024, 0.246, 1.508, 0.186, 0.769, 0.013, 0.275, 0.312, 0.368, 0.152, 0.092, 0.205, 0.104, 0.503, 0.166, 0.024, 0.071, 1.669, 0.265, 0.472, 0.342, 0.475, 0.257, 0.059, 0.723, 1.168, 0.603, 0.080, 0.036, 1.196, 0.146, 0.249, 0.078, 0.585, 0.084, 0.078, 0.236, 1.017, 0.770, 0.075, 0.056, 0.067, 0.923, 0.032, 0.003, 0.534, 0.091, 0.074, 0.131, 0.165, 0.356, 0.598, 1.731, 0.588, 1.074, 1.071, 0.559, 0.694, 0.218, 0.377, 0.808, 0.427, 1.078, 0.046, 0.259, 0.589, 0.035, 0.571, 0.157, 0.861, 0.055, 0.031
$y :$	0.164, 0.020, 0.157, 0.220, 0.376, 0.061, 0.136, 0.185, 0.028, 0.317, 0.132, 0.469, 0.875, 0.039, 0.126, 0.031, 0.444, 0.222, 0.374, 0.507, 0.180, 0.237, 0.013, 0.317, 0.730, 0.486, 0.740, 0.559, 0.160, 0.210, 0.152, 0.453, 0.055, 0.985, 0.506, 0.044, 0.063, 0.151, 0.188, 0.300, 0.133, 0.254, 0.282, 0.845, 1.037, 0.018, 0.046, 0.113, 0.179, 0.117, 0.693, 0.163, 0.011, 0.310, 0.105, 0.517, 0.367, 0.149, 0.015, 0.337, 0.180, 0.315, 1.015, 0.649, 0.520, 0.228, 0.378, 0.635, 0.349, 0.606, 0.362, 1.491, 0.602, 0.745, 0.240, 0.524, 0.504, 1.023, 0.055, 0.034, 0.151, 0.152, 0.242, 0.014, 0.101, 0.240, 0.076, 0.455, 0.079, 0.590, 0.165, 0.297, 0.238, 0.451, 0.001, 0.412, 0.061, 0.069, 0.943

Na Tabela 2 apresenta-se as estimativas pontuais e o erro padrão para os parâmetros de acordo com o método de estimação. Observa-se que o erro padrão Bootstrap ficou menor do que o assintótico para os parâmetros θ , e ρ . Para os demais parâmetros os erros padrões se comportaram de forma bastante similar.

Tabela 2: Estimativas pontuais e erro padrão Clássico, Bootstrap e Bayesiano.

Parâmetro	Estimativas			Erro padrão		
	MLE	BCE [†]	Bayes	MLE	Bootstrap	Bayes
λ_1	2.2785	2.2410	2.2996	0.2244	0.2284	0.2247
λ_2	3.1237	3.0812	3.1272	0.2959	0.3103	0.2934
θ	0.8255	0.8744	0.7729	0.1626	0.1372	0.1354
R	0.4385	0.4384	0.4395	0.0310	0.0320	0.0309
ρ	-0.3752	-0.3893	-0.3603	0.0366	0.0349	0.0347

[†]: estimativa corrigida via Bootstrap paramétrico.

Os intervalo de 95% de confiança (credibilidade) estão expostos na Tabela 3. Os intervalos

de confiança p -Bootstrap e t -Bootstrap foram obtidos considerando $B = 20000$ amostras geradas da distribuição GBE com parâmetros dados pelas estimativas de verossimilhança. Para comparar os intervalos de confiança duas medidas foram analisada, a amplitude (A) e a forma (F), a qual mede a assimetria do intervalo em torno de sua estimativa. Uma vez que os intervalos assintóticos são simétricos então $F = 1$.

Tabela 3: Intervalos de 95% de confiança (credibilidade) assintótico, p -Bootstrap, t -Bootstrap e Bayesiano.

Parâmetro	Assintótico	p -Bootstrap	t -Bootstrap	HPD
λ_1	(1.8333, 2.7236)	(1.9088, 2.8044)	(1.8510, 2.7222)	(1.8660, 2.7300)
λ_2	(2.5366, 3.7107)	(2.6155, 3.8340)	(2.5663, 3.7133)	(2.5786, 3.7179)
θ	(0.5029, 1.1480)	(0.4790, 0.9874)	(0.6744, 1.1607)	(0.5289, 0.9996)
R	(0.3769, 0.5000)	(0.3763, 0.4998)	(0.3777, 0.5010)	(0.3803, 0.5005)
ρ	(-0.4478, -0.3026)	(-0.4087, -0.2798)	(-0.4468, -0.3015)	(-0.4110, -0.2951)

Das medidas apresentadas na Tabela 4 percebe-se que o intervalo t -Bootstrap exibiu menores amplitudes para todos os parâmetros estimado, ao ser comparado com os intervalos assintóticos e p -Bootstrap.

Tabela 4: Amplitude (A) e forma (F) para os intervalos de 95% de confiança (credibilidade).

Parâmetro	Assintótico		p -Bootstrap		t -Bootstrap		HPD	
	A	F	A	F	A	F	A	F
λ_1	0.8903	1.0000	0.8955	1.4226	0.8712	1.0381	0.8640	1.0946
λ_2	1.1741	1.0000	1.2185	1.3979	1.1470	1.0580	1.1393	1.0902
θ	0.6451	1.0000	0.5084	0.4675	0.4863	2.2200	0.4707	0.5872
R	0.1231	1.0000	0.1235	0.9871	0.1233	1.0295	0.1202	1.0664
ρ	0.1452	1.0000	0.1289	2.8478	0.1453	1.0295	0.1159	2.2374

A propriedade de assimetria dos intervalos de confiança assintótico parece não ser razoável para o parâmetro θ e ρ da distribuição GBE, conforme indicam os gráficos da Figura 2. Nesse sentido, os intervalos t -Bootstrap e p -Bootstrap são preferíveis ao intervalo assintótico.

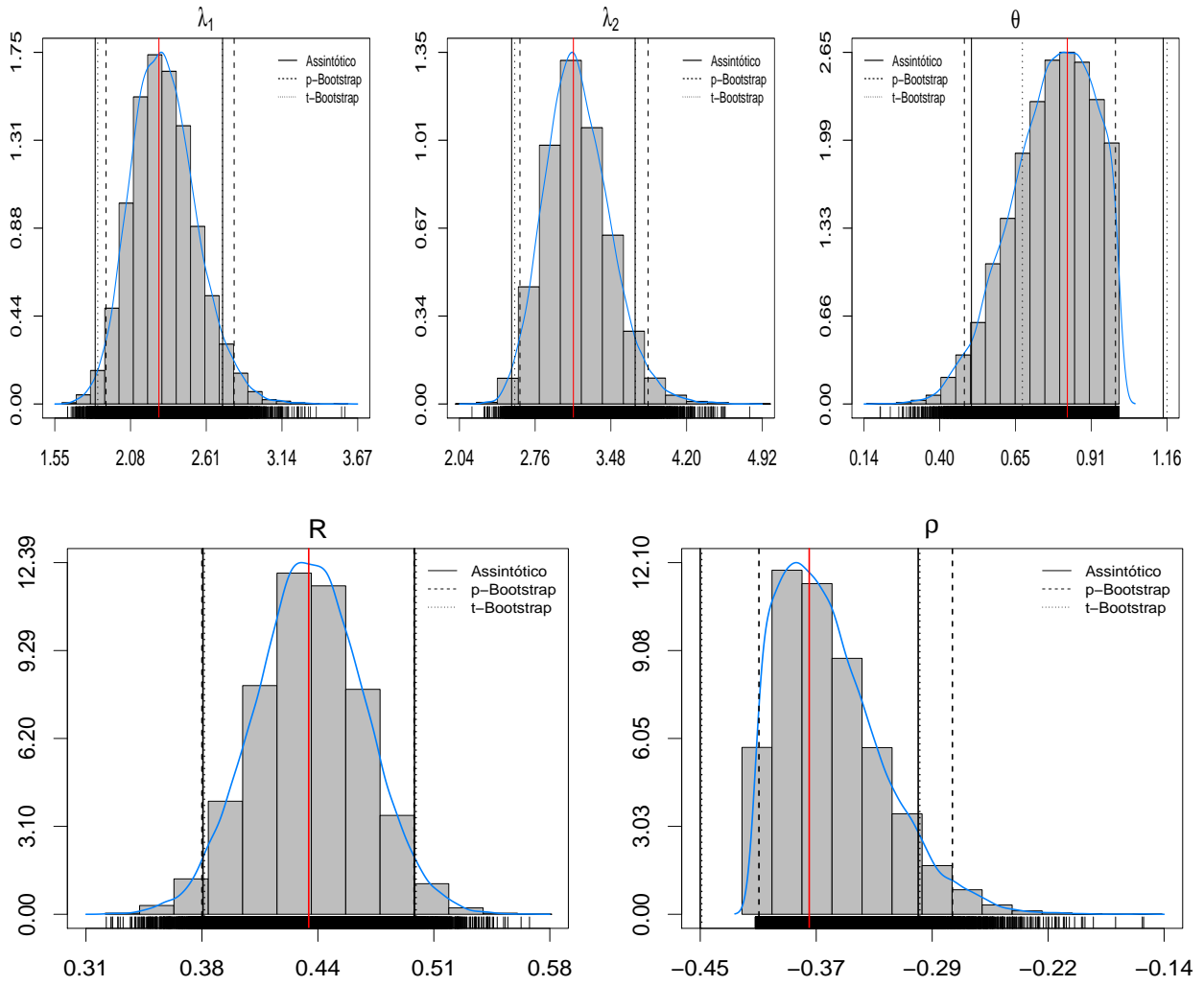


Figura 2: Distribuição Bootstrap dos parâmetros e intervalos de 95% de confiança.

Por fim, a Figura 3 ilustra o ajuste da marginal distribuição GBE para os dados em questão. As curvas de sobrevivência marginais são comparadas com o estimador não paramétrico de Kaplan-Meier. Nota-se que as estimativas originais levam a ajustes bastante parecidos das estimativas corrigidas por Bootstrap.

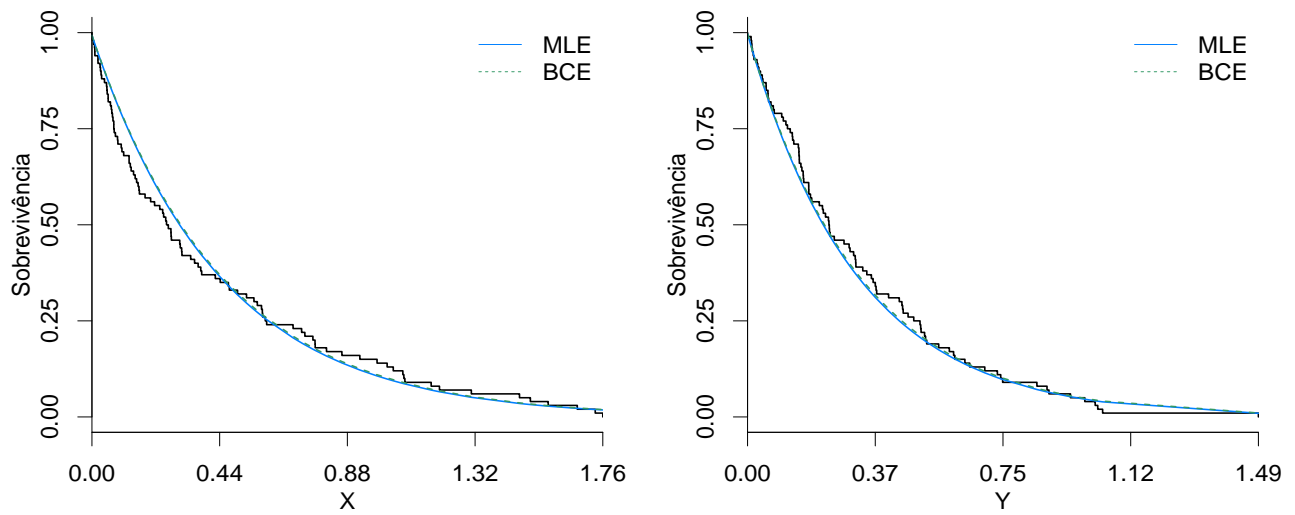


Figura 3: Kaplan-Meier versus as curvas de sobrevivência marginais de X e Y considerando as estimativas MLE e BCE.

4.2 Dados simulados da distribuição HEB

Nesta subseção apresentamos um exemplo considerando os dados simulados de uma distribuição Hougaard Exponencial bivariada com parâmetros $\mu_1 = 5$, $\mu_2 = 10$ e $r = 8$, neste cenário $R = 0.9961$ e $\rho = 0.9714$. As $n = 100$ observações (ver Tabela 5) foram simuladas utilizando o método da condicional, com auxílio do método da transformação inversa para obter amostras aleatórias da distribuição condicional de Y dado X .

Tabela 5: Dados simulados da distribuição HBE com parâmetros $\mu_1 = 5$, $\mu_2 = 10$ e $r = 8$.

$x :$	3.191, 20.446, 1.205, 8.054, 0.204, 23.439, 1.801, 7.470, 1.456, 1.497, 2.246, 0.963, 5.845, 2.269, 2.489, 1.927, 1.569, 10.259, 0.077, 1.095, 13.918, 0.165, 8.911, 9.197, 9.389, 7.008, 1.782, 2.397, 5.536, 2.905, 1.929, 2.463, 2.222, 4.791, 5.164, 0.089, 17.252, 3.169, 0.680, 0.939, 0.024, 1.056, 13.068, 2.367, 3.865, 9.168, 1.142, 1.337, 3.731, 4.402, 2.052, 8.813, 20.865, 0.419, 0.491, 2.992, 3.624, 5.215, 0.243, 2.673, 3.493, 0.084, 12.675, 7.031, 4.875, 15.600, 1.932, 11.515, 6.852, 6.080, 1.488, 5.731, 0.610, 1.839, 0.210, 0.433, 3.211, 2.470, 5.209, 2.784, 5.466, 9.606, 0.610, 19.552, 0.181, 15.159, 20.961, 17.413, 1.348, 5.358, 2.973, 11.284, 14.277, 7.494, 1.277, 3.269, 0.314, 1.636, 7.424, 1.926
$y :$	5.424, 31.232, 2.721, 11.796, 0.321, 57.900, 4.073, 9.937, 3.167, 2.726, 3.767, 1.946, 6.495, 3.931, 7.078, 4.243, 3.706, 19.774, 0.161, 1.712, 24.071, 0.145, 19.015, 23.219, 16.551, 13.864, 2.883, 4.852, 10.823, 4.530, 4.470, 4.361, 3.263, 12.497, 12.187, 0.227, 39.609, 8.505, 0.571, 3.191, 0.077, 2.333, 27.622, 5.310, 7.838, 17.248, 1.963, 1.863, 5.416, 8.432, 3.049, 17.195, 37.650, 1.126, 0.812, 6.505, 6.023, 14.669, 0.469, 8.250, 5.470, 0.175, 27.177, 10.893, 13.202, 27.299, 5.791, 21.304, 13.634, 9.755, 2.906, 10.192, 1.443, 3.930, 0.507, 1.262, 9.683, 4.218, 9.959, 6.237, 8.724, 17.895, 1.172, 47.333, 0.449, 31.392, 49.003, 29.384, 2.648, 15.114, 7.323, 24.800, 23.643, 16.653, 2.795, 6.531, 0.675, 4.010, 13.046, 4.082

A Tabela 6 mostra as estimativas pontuais e o erro padrão para os parâmetros conforme com o método de estimação. Verifica-se que o erro padrão Bootstrap ficou menor do que o assintótico para os parâmetros μ_1 , e μ_2 . Nos demais casos os erros padrões se comportaram de forma bastante equivalente. Em relação, a estimativa corrigida verifica-se que não há notáveis diferença da estimativa pontual obtida por máxima verossimilhança.

Tabela 6: Estimativas pontuais e erro padrão Clássico, Bootstrap e Bayesiano.

Parâmetro	Estimativas			Erro padrão		
	MLE	BCE [†]	Bayes	MLE	Bootstrap	Bayes
μ_1	5.2379	5.2428	5.2953	0.5201	0.5187	0.5378
μ_2	10.5160	10.5227	10.6353	1.0442	1.0419	1.0776
r	7.5571	7.4559	7.5097	0.6253	0.6503	0.6335
R	0.9949	0.9951	0.9941	0.0024	0.0025	0.0028
ρ	0.9522	0.9519	0.9508	0.0072	0.0073	0.0076

[†]: estimativa corrigida via Bootstrap paramétrico.

Os intervalo de 95% de confiança assintótico, p -Bootstrap, t -Bootstrap e a região HPD para os parâmetros do modelo e as funções R e ρ estão exibidos na Tabela 7. Além disso, na Tabela 8 têm-se os valores das amplitude (A) e forma (F) dos intervalos. Observa-se que os três intervalos assintótico, p -Bootstrap, t -Bootstrap estão em concordância, com exceção do parâmetro R , no qual sua distribuição Bootstrap possui uma forte assimetria a esquerda (ver Figura 4). Logo, neste caso os intervalos de confiança Bootstrap são preferíveis.

Tabela 7: Intervalos de 95% de confiança (credibilidade) assintótico, p -Bootstrap, t -Bootstrap e Bayesiano.

Parâmetro	Assintótico	p -Bootstrap	t -Bootstrap	HPD
μ_1	(4.2185, 6.2573)	(4.2667, 6.2982)	(4.3570, 6.4289)	(4.2923, 6.3727)
μ_2	(8.4694, 12.5625)	(8.5697, 12.6402)	(8.7500, 12.9026)	(8.5209, 12.6988)
r	(6.3317, 8.7826)	(6.4970, 9.0499)	(6.3266, 8.7784)	(6.2896, 8.7592)
R	(0.9902, 0.9996)	(0.9887, 0.9982)	(0.9866, 0.9981)	(0.9884, 0.9985)
ρ	(0.9381, 0.9663)	(0.9372, 0.9654)	(0.9272, 0.9619)	(0.9350, 0.9645)

Tabela 8: Amplitude (A) e forma (F) para os intervalos de 95% de confiança (credibilidade).

Parâmetro	Assintótico		p-Bootstrap		t-Bootstrap		HPD	
	A	F	A	F	A	F	A	F
μ_1	2.0388	1.0000	2.0316	1.0918	2.0719	1.3520	2.0804	1.2001
μ_2	4.0931	1.0000	4.0706	1.0914	4.1526	1.3515	4.1779	1.0941
r	2.4509	1.0000	2.5529	1.4081	2.4518	0.9926	2.4696	0.9484
R	0.0094	1.0000	0.0095	0.5323	0.0116	0.3872	0.0101	0.5538
ρ	0.0282	1.0000	0.0282	0.8800	0.0347	0.3872	0.0295	0.7151

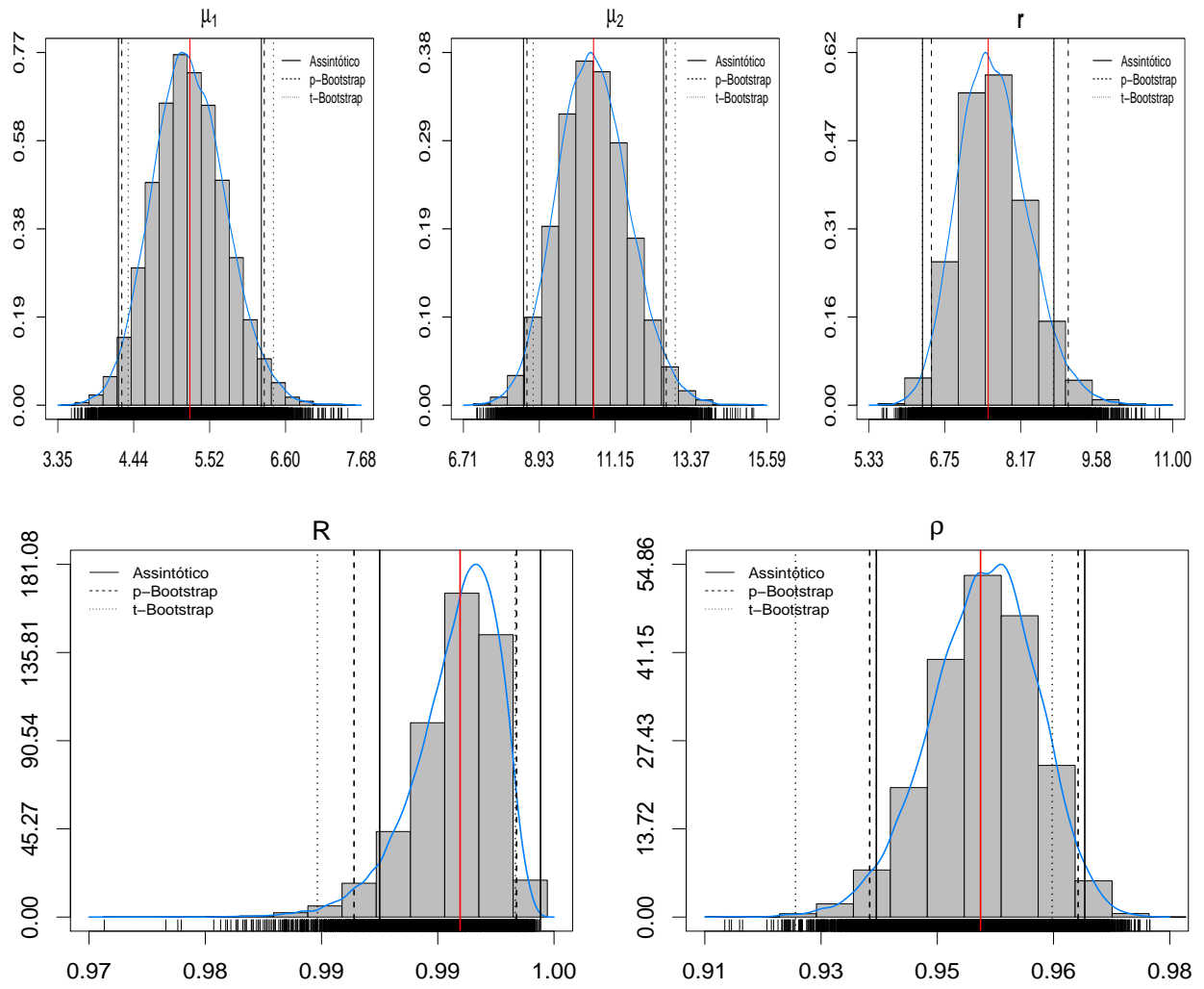


Figura 4: Distribuição Bootstrap dos parâmetros e intervalos de 95% de confiança.

Finalmente, a título de ilustração temos na Figura 5 os ajustes marginais considerando as estimativa de máxima verossimilhança (MLE) e sua versão corrigida por meio do vício Bootstrap (BCE). É possível notar que as estimativas via MLE proporcionam ajuste muito próximos das estimativas corrigidas.

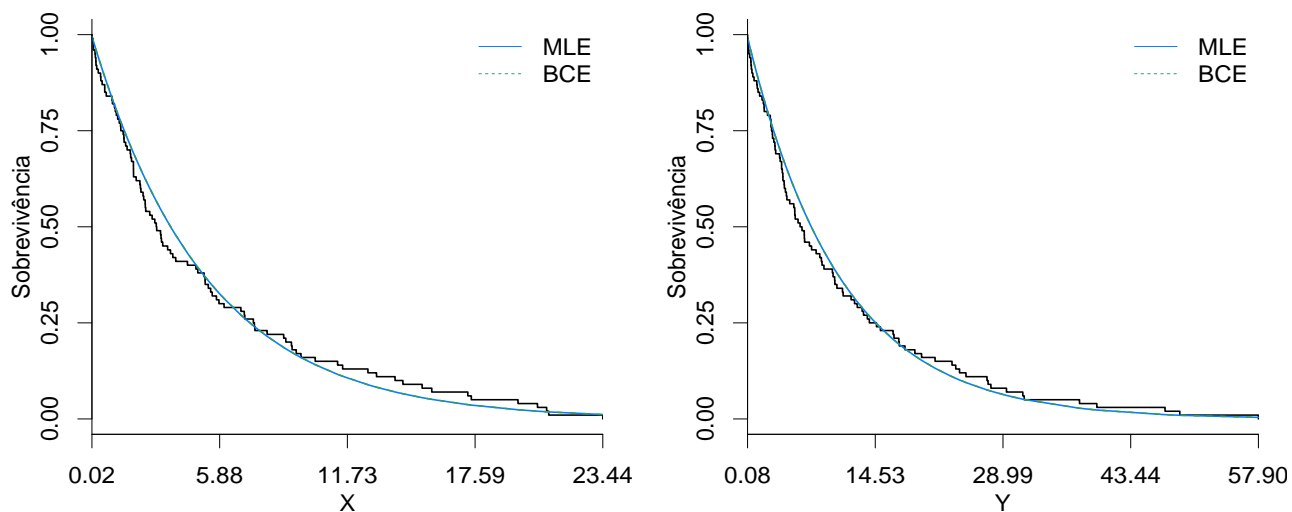


Figura 5: Kaplan-Meier versus as curvas de sobrevivência marginais de X e Y considerando as estimativas MLE e BCE.

Referências

- Abramowitz, M., Stegun, I. A., 1974. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series. Dover Publications, Incorporated, New York.
- Balakrishnan, N., Lai, C.-D., 2009. Continuous Bivariate Distributions, 2nd Edition. Springer.
- Carpenter, J., Bithell, J., 2000. Bootstrap confidence intervals: When, which, what? A practical guide for medical statisticians. *Statistics in Medicine* 19 (9), 1141–1164.
- Damien, P., Müller, P., 1998. A Bayesian bivariate failure time regression model. *Computational Statistics & Data Analysis* 28 (1), 77–85.
- Efron, B., 1979. Bootstrap methods: Another look at the jackknife. *Annals of Statistics* 7 (1), 1–26.
- Efron, B., Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1 (1), 54–75.
- Gumbel, E. J., 1960. Bivariate Exponential distributions. *Journal of the American Statistical Association* 55 (292), 698–707.
- Hougaard, P., 1986. A class of multivariate failure time distributions. *Biometrika* 73, 671–678.
- Kotz, S., Lumelskii, Y., Pensky, M., 2003. The Stress-Strength Model and its Generalizations Theory and Applications. World Scientific.
- Mazucheli, J., Coelho-Barros, E. A., Achcar, J. A., 2006. Inference for the mode of the Log-Logistic hazard function. *Rev. Mat. Estat., São Paulo* 24 (4), 73–87.
- Nadarajah, S., 2006. Information matrix for the bivariate Gumbel distribution. *Applied Mathematics and Computation* 172 (1), 394–405.
- Nadarajah, S., Kotz, S., 2006. Reliability for some bivariate Exponential distributions. *Mathematical Problems in Engineering*, 1–14.
- Nair, N. U., Nair, V. K. R., 1988. A characterization of the bivariate Exponential distribution. *Biometrical Journal* 30 (1), 107–112.
- Pal, S., Murthy, G. S. R., 2003. An application of Gumbel’s bivariate exponential distribution in estimation of warranty cost of motor cycles. *International Journal of Quality & Reliability Management* 20, 488–502.
- SAS, 2010a. The MCMC Procedure, SAS/STAT® User’s Guide, Version 9.4. Cary, NC: SAS Institute Inc.

SAS, 2010b. The NLMIXED Procedure, SAS/STAT® User's Guide, Version 9.4. Cary, NC: SAS Institute Inc.