

# Tópicos em Estatística: Modelagem com Apoio Computacional

Notas de Aula

[Prof. Helton Saulo](#), Universidade de Brasília (UnB)

21 agosto, 2025

## Sumário

<b>1</b>	<b>Distribuição Birnbaum-Saunders</b>	<b>2</b>
1.1	Introdução . . . . .	2
1.2	Histórico . . . . .	3
1.3	Fadiga, gênese e derivação . . . . .	4
1.4	Funções de probabilidade e propriedades . . . . .	7
1.5	Função característica e momentos . . . . .	13
1.6	Geração de números aleatórios . . . . .	14
1.7	Método de Estimação por Máxima Verossimilhança . . . . .	18
1.8	Método de Estimação por Momentos . . . . .	21
1.9	A distribuição log-Birnbaum-Saunders . . . . .	22
1.10	Modelo de regressão Birnbaum-Saunders . . . . .	25
1.11	Modelo de regressão Birnbaum-Saunders para dados censurados . . . . .	33
1.12	Análise de resíduos . . . . .	34
	Referências . . . . .	37

# 1 Distribuição Birnbaum-Saunders

## 1.1 Introdução

- De modo geral, a literatura relacionada a métodos estatísticos e modelos para dados de tempo de vida reconhece o conceito de “tempo de vida” (lifetime) como uma variável aleatória contínua positiva que representa o tempo até a ocorrência de algum evento de interesse.
- No entanto, em algumas ocasiões, o envelhecimento de itens (por exemplo, componentes, sistemas, subsistemas, espécimes, estruturas, órgãos, unidades, etc.) não é medido em termos cronológicos. Nessas situações, o tempo de vida é medido por meio de outras variáveis aleatórias. Por exemplo: a quantidade de quilômetros percorridos, a resistência de espécimes de material até sua ruptura, o nível de degradação, a flexibilidade de um adesivo e o número de ciclos até que um espécime de material falhe devido à fadiga (tempo de vida por fadiga).
- Além disso, a terminologia “variável aleatória de tempo de vida”, que doravante denotamos por  $T$ , é amplamente utilizada para se referir a qualquer variável aleatória contínua positiva (por exemplo, quantidade de água da chuva ou concentração de contaminantes).
- Qualquer modelo probabilístico associado a uma variável aleatória de tempo de vida é frequentemente chamado de distribuição de tempo de vida.
- A teoria da confiabilidade refere-se a problemas probabilísticos e estatísticos relacionados às distribuições de vida de itens sujeitos a falha. Os métodos probabilísticos desta teoria descrevem o desempenho e a {degradação} dos itens por meio de variáveis aleatórias de tempo de vida.
- As distribuições de tempo de vida são usadas para determinar a função de confiabilidade (ou sobrevivência), a taxa de falha (ou risco) e o tempo de vida médio desses itens. Os aspectos estatísticos desta teoria resolvem problemas de estimativa e inferência dos parâmetros da distribuição de tempo de vida.
- Tecnicamente, a confiabilidade de um item é definida como a probabilidade de que uma variável aleatória de tempo de vida  $T > 0$  ultrapasse um tempo fixo  $t$ . Assim, a função de confiabilidade de  $T$  ao longo do tempo  $t$  é dada por:

$$R_T(t) = P(T > t) = 1 - P(T \leq t) = 1 - F_T(t) = 1 - \int_0^t f_T(u) du, \quad t > 0, \quad (1)$$

a qual, como observado, depende da distribuição contínua de tempo de vida quando uma análise paramétrica é realizada. Essa distribuição de vida para  $T$  pode ser caracterizada por sua função densidade de probabilidade  $f_T(\cdot)$  ou por sua função de distribuição acumulada  $F_T(\cdot)$ , como mostrado na Equação (1), ou ainda por sua taxa de falha ou risco, dada por:

$$h_T(t) = \frac{f_T(t)}{R_T(t)}, \quad t > 0,$$

ou pela sua taxa de falha acumulada definida como:

$$H_T(t) = \int_0^t h_T(u) du.$$

## 1.2 Histórico

- O modelo probabilístico mais utilizado é a distribuição normal (ou Gaussiana). No entanto, as distribuições de vida frequentemente são modelos probabilísticos com dois parâmetros que apresentam assimetria à direita (assimetria positiva), unimodalidade e suporte positivo ( $T > 0$ ).
- Com a fama da distribuição exponencial, surgiu o uso da distribuição gama para modelar a soma de tempos de vida com distribuição exponencial. {A vantagem} da distribuição gama é que ela surge naturalmente como a convolução de distribuições exponenciais, mas sua desvantagem é que o modelo gama não é tratável algebricamente.
- Vários estudos começaram a encontrar dados de tempo de vida que são melhor modelados por uma distribuição de vida diferente do modelo exponencial. Por exemplo, Lieblein and Zelen (1956) e Kao (1959) consideraram a distribuição de Weibull, porque este modelo {se ajusta bem aos dados de falha}. Essa distribuição foi desenvolvida por Weibull (1951) para descrever a resistência à ruptura de materiais.
- Motivado pela fadiga de materiais em aeronaves comerciais causadas por vibrações constantes e pelos problemas decorrentes, Birnbaum and Saunders (1968) apresentou um modelo probabilístico engenhoso que descreve tempos de vida associados a espécimes de material expostos à fadiga, resultado de tensão e estresse cíclicos.
- No ano seguinte, Birnbaum and Saunders (1969a); Birnbaum and Saunders (1969b) formalizaram a distribuição de tempo de vida por fadiga, que mais tarde recebeu seus nomes, a caracterizaram como uma distribuição de vida e apresentaram um método de estimativa para os parâmetros dessa distribuição de dois parâmetros
- Birnbaum and Saunders (1958) já se interessavam em encontrar um modelo que descrevesse o tempo de falha associado a espécimes de material expostos à fadiga. Simultaneamente, Birnbaum e Saunders também trabalharam com {um grupo composto por} Barlow, Esary, Marshall e Proschan em processos de desgaste causados por danos cumulativos. Os principais resultados dessas investigações foram estabelecidos nos seguintes três trabalhos:
  - (i) Birnbaum, Esary, and Marshall (1966) desenvolveu um modelo estocástico de desgaste;
  - (ii) Birnbaum and Saunders (1969a) desenvolveu a distribuição de tempo de vida por fadiga; e
  - (iii) Esary, Marshall, and Proschan (1973) propôs o modelo de choque.
- A distribuição Birnbaum-Saunders é baseada em danos cumulativos que produzem fadiga nos materiais. Este argumento é a regra de Miner; veja Miner (1945).

- Birnbaum and Saunders (1968) obtiveram uma interpretação probabilística dessa regra. Assim, a distribuição de tempo de vida por fadiga de Birnbaum and Saunders (1969a) foi derivada de um modelo que mostra que o tempo total decorrido {até o dano cumulativo}, produzido pelo desenvolvimento e crescimento de uma fissura dominante, supera um valor limite e causa a falha do espécime de material.
- Desmond (1985) reforçou a justificativa para o uso dessa distribuição ao relaxar algumas suposições feitas inicialmente por Birnbaum and Saunders (1969a). Além disso, uma derivação mais geral da distribuição Birnbaum-Saunders foi fornecida por Desmond (1985) com base em um modelo biológico introduzido por Cramér (1946).
- Embora a distribuição de tempo de vida por fadiga seja conhecida como o modelo Birnbaum-Saunders, um trabalho anterior atribuído a Freudenthal and Shinozuka (1961) introduz um modelo semelhante, embora com uma parametrização diferente. Além disso, o trabalho de Fréchet (1927) apresentou algumas ideias semelhantes relacionadas à distribuição Birnbaum-Saunders, mas também com uma parametrização diferente da considerada por Birnbaum and Saunders (1969a). Este fato foi mencionado por Birnbaum and Saunders (1969a) e também é descrito em Desmond (1985), Johnson, Kotz, and Balakrishnan (1995), Marshall and Olkin (2007) e S. Saunders (2007).

### 1.3 Fadiga, gênese e derivação

- A fadiga é um processo de dano cumulativo presente em materiais sujeitos a níveis periódicos e flutuantes de estresse durante cada ciclo de operação (ou ciclos de operação).
- Uma teoria para a fadiga de materiais baseia-se na suposição de que a propagação de uma fissura dominante é a principal causa da falha devido à fadiga. O processo de fadiga pode ser dividido em três etapas:

[(A1)] O início de uma fissura imperceptível;

[(A2)] O crescimento e a propagação da fissura, que provoca uma rachadura no espécime de material devido ao estresse e à tensão cíclicos; e

[(A3)] A ruptura ou falha do espécime de material devido à fadiga.

- O modelo Birnbaum-Saunders considera um espécime de material exposto a uma sequência de  $m$  cargas cíclicas, denotadas por  $\{l_i, i = 1, 2, \dots, m, m \in \mathbb{N}\}$ . O esquema de carregamento pode ser representado da seguinte forma:

$$\underbrace{l_1, l_2, \dots, l_m}_{1^\circ \text{ ciclo}} \quad \underbrace{l_{m+1}, l_{m+2}, \dots, l_{2m}}_{2^\circ \text{ ciclo}} \quad \dots \quad \underbrace{l_{jm+1}, l_{jm+2}, \dots, l_{jm+m}}_{(j+1)^\circ \text{ ciclo}}$$

em  $l_{jm+i} = l_{km+i}$ , para  $j \neq k$ . Birnbaum and Saunders (1969a) consideraram que a carga é contínua. Assim, a cada imposição de carga,  $l_i$ , a fissura é estendida por uma quantidade aleatória. Na prática, não é possível observar as extensões da fissura por ciclo; apenas o momento da falha é conhecido.

- Birnbaum and Saunders (1969a) utilizaram o conhecimento de certos {tipos} de falhas em materiais devido à fadiga para derivar seu modelo. O processo de fadiga que eles usaram foi baseado no seguinte:

[(B1)] Um material é sujeito a cargas cíclicas ou choques repetitivos, que produzem uma fissura ou desgaste no material;

[(B2)] A falha no material ocorre quando o tamanho da fissura excede um certo nível de resistência (limite), denotado por  $\omega$ ;

[(B3)] A sequência de cargas impostas ao material é a mesma de um ciclo para outro;

[(B4)] Durante um ciclo de carga, a extensão da fissura causada pela carga  $l_i$ ,  $X_i$ , é uma variável aleatória cuja distribuição é governada por todas as cargas  $l_j$ , para  $j < i, i = 1, \dots, m$ , e pela extensão da fissura que a precede;

[(B5)] O tamanho total da fissura causado pelo  $k$ -ésimo ciclo,  $Y_k$ , é uma variável aleatória que segue uma distribuição com média  $\mu$  e variância finita  $\sigma^2$ ; e

[(B6)] O tamanho das fissuras ( $Y_k$ ) em diferentes ciclos são independentes.

- Note que a extensão da fissura no  $(k + 1)$ -ésimo ciclo de carga é

$$Y_{k+1} = X_{km+1} + X_{km+2} + \dots + X_{km+m}, \quad k = 0, 1, 2, \dots, m \in \mathbb{N}.$$

Como mencionado por Mann, Schafer, and Singpurwalla (1974), a suposição (B4) é bastante restritiva e pode não ser válida para certas aplicações. Essa suposição garante que, independentemente da dependência entre as extensões aleatórias sucessivas causadas pelas cargas em um ciclo específico, as extensões totais de fissuras são independentes de ciclo para ciclo. A plausibilidade dessa suposição em estudos de fadiga aeronáutica é brevemente afirmada por Birnbaum and Saunders (1969a).

- O modelo Birnbaum-Saunders busca a distribuição do menor número de ciclos, denotado por  $N^*$ , tal que a soma

$$S_n = \sum_{j=1}^n Y_j \tag{2}$$

de  $n$  variáveis aleatórias positivas exceda o limite  $\omega$ , ou seja,

$$N^* = \inf \{n \in \mathbb{N} : S_n > \omega\},$$

em que  $S_n$  é dado em (2). No caso mais simples, a distribuição Birnbaum-Saunders é derivada supondo que as variáveis  $Y_k$  são independentes e identicamente distribuídas, e então aplica-se a versão clássica do teorema central do limite. Além disso,  $N^*$  é tratado como uma variável aleatória contínua  $T$ .

- Especificamente, com base no teorema central do limite, (2) e as suposições (B1)-(B6), à medida que  $n \rightarrow \infty$ , é possível estabelecer que  $S_n$  segue aproximadamente uma distribuição normal ( $\sim^a$ ) com média  $n\mu$  e variância  $n\sigma^2$ :

$$S_n \sim^a N(n\mu, n\sigma^2). \quad (3)$$

- Seja  $N$  o número de ciclos necessários até a falha de um material. Como  $Y_k > 0$  para todo  $k \geq 1$ , o dano é irreversível e, por complementaridade, temos  $\{N > n\} \equiv \{S_n \leq \omega\}$  e  $\{N \leq n\} \equiv \{S_n > \omega\}$ . Assim, pela suposição (B5) e pela Equação (2), temos que  $E(S_n) = n\mu$  e  $V(S_n) = n\sigma^2$ . Portanto, da Equação (3), obtemos:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \sim^a N(0, 1). \quad (4)$$

- Assim, a partir de (4), temos:

$$\begin{aligned} P(N \leq n) &= P(S_n > \omega) \\ &= P\left(Z_n < \sqrt{\frac{\mu\omega}{\sigma}} \left(\sqrt{n}\sqrt{\frac{\mu}{\omega}} - \frac{1}{\sqrt{n}}\sqrt{\frac{\omega}{\mu}}\right)\right) \\ &\approx \Phi\left(\frac{\sqrt{\mu\omega}}{\sigma} \left(\sqrt{n}\sqrt{\frac{\mu}{\omega}} - \frac{1}{\sqrt{n}}\sqrt{\frac{\omega}{\mu}}\right)\right), \end{aligned} \quad (5)$$

em que  $\Phi(\cdot)$  é a função de distribuição acumulada da normal padrão. Birnbaum and Saunders (1969a) usaram (5) para definir uma distribuição contínua de vida, a partir da variável aleatória discreta  $N$  por meio de uma variável aleatória contínua  $T$ , e o argumento discreto  $n$  pelo argumento contínuo  $t$ .

- Especificamente, o número de ciclos até a falha  $N$  é convertido no tempo total até a ocorrência da falha  $T$ , enquanto o  $n$ -ésimo ciclo é convertido no tempo  $t$ . Assim, a partir de (5), definindo:

$$\alpha = \frac{\sigma}{\sqrt{\omega\mu}}, \quad \beta = \frac{\omega}{\mu}, \quad \xi(x) = \sqrt{x} - \frac{1}{\sqrt{x}} = 2 \sinh(\log(\sqrt{x})), \quad x > 0,$$

obtemos a função de distribuição acumulada da distribuição Birnbaum-Saunders com parâmetro de forma  $\alpha > 0$  e parâmetro de escala  $\beta > 0$  como:

$$F_T(t) = \Phi\left(\frac{1}{\alpha}\xi\left(\frac{t}{\beta}\right)\right), \quad t > 0, \alpha > 0, \beta > 0. \quad (6)$$

- Note que aqui e adiante usamos a notação  $\log$  para se referir ao logaritmo natural. Em S. Saunders (1976),  $\alpha$  também é chamado de “parâmetro de flexão”, que está diretamente associado à fadiga do material. Se uma variável aleatória  $T$  tem uma função de distribuição acumulada como em (6), então a notação  $T \sim \text{BS}(\alpha, \beta)$  é usada. Isso significa que, ao admitirmos  $T \sim \text{BS}(\alpha, \beta)$ , podemos escrever:

$$T = \beta \left( \frac{\alpha}{2} Z + \sqrt{\left(\frac{\alpha}{2} Z\right)^2 + 1} \right)^2, \quad (7)$$

em que  $Z$  é uma variável aleatória que segue uma distribuição normal com média  $\mu = 0$  e variância  $\sigma^2 = 1$ , ou seja, a distribuição normal padrão, denotada por  $Z \sim N(0, 1)$ . Assim,

$$Z = \frac{1}{\alpha} \left( \sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim N(0, 1). \quad (8)$$

- Portanto, temos duas abordagens que permitem obter a distribuição Birnbaum-Saunders.
- A primeira abordagem tem origem na física dos materiais, o que possibilita interpretar o modelo Birnbaum-Saunders como uma distribuição de vida, assumindo que  $S_n$  e  $Y_k$  representam a extensão total de uma fissura e sua extensão parcial causada pelo  $k$ -ésimo ciclo de carga, respectivamente. Nessa abordagem, a distribuição descreve o tempo decorrido até que a extensão da fissura exceda um limite crítico, levando à falha, permitindo que o modelo Birnbaum-Saunders seja considerado como uma distribuição de vida.
- A segunda abordagem baseia-se nas Equações (7) e (8), que são assumidas como a definição da distribuição Birnbaum-Saunders. Nessa abordagem, podemos assumir que qualquer variável aleatória que siga uma distribuição Birnbaum-Saunders é uma transformação de uma variável aleatória com distribuição normal padrão.

## 1.4 Funções de probabilidade e propriedades

- Seja  $T \sim \text{BS}(\alpha, \beta)$ . Então, a função densidade de probabilidade de  $T$  é dada por:

$$\begin{aligned} f_T(t; \alpha, \beta) &= \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2\alpha^2} \left( \frac{t}{\beta} + \frac{\beta}{t} - 2 \right) \right) \frac{1}{2\alpha} \left( \left( \frac{t}{\beta} \right)^{-1/2} + \left( \frac{t}{\beta} \right)^{-3/2} \right) \\ &= \frac{1}{\sqrt{8\pi}} \exp \left( \frac{1}{\alpha^2} \right) \exp \left( -\frac{1}{2\alpha^2} \left( \frac{t}{\beta} + \frac{\beta}{t} \right) \right) \frac{t^{-3/2}}{\alpha\beta^{1/2}} (t + \beta), \end{aligned} \quad (9)$$

para  $t > 0$ ,  $\alpha > 0$  e  $\beta > 0$ .

- A expressão dada em (9) pode ser obtida facilmente usando o teorema da transformação de variáveis aleatórias, considerando

$$Z = \frac{1}{\alpha} \left( \sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim N(0, 1) \quad (10)$$

e sua função densidade de probabilidade é:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} z^2 \right), \quad z \in \mathbb{R}. \quad (11)$$

Assim,

$$\begin{aligned} f_T(t; \alpha, \beta) &= \phi\left(\frac{1}{\alpha}\xi\left(\frac{t}{\beta}\right)\right) \left| \frac{d}{dt} \left(\frac{1}{\alpha}\xi\left(\frac{t}{\beta}\right)\right) \right| \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{1}{\alpha}\xi\left(\frac{t}{\beta}\right)\right)^2\right) \frac{1}{\alpha\beta} \xi'\left(\frac{t}{\beta}\right), \quad t > 0, \end{aligned} \quad (12)$$

em que

$$\xi(u) = u^{1/2} - u^{-1/2} = \sinh(\log(u)), \quad u > 0, \quad (13)$$

e a derivada da função  $\xi(\cdot)$  dada em (13) é:

$$\xi'(u) = \frac{d}{dy} \xi(u) = \frac{1}{2} (y^{-1/2} + y^{-3/2}), \quad y > 0.$$

A expressão obtida em (12) comprova a Equação definida em (9).

A figura abaixo apresenta gráficos da função densidade de probabilidade da distribuição Birnbaum–Saunders para diferentes valores de seu parâmetro de forma  $\alpha$ , considerando o parâmetro de escala  $\beta = 1$ , sem perda de generalidade. Com base nessa figura, note que a distribuição Birnbaum–Saunders é contínua, unimodal e assimetricamente inclinada para a direita (assimetria positiva). Também é possível observar que, à medida que  $\alpha$  tende a zero, a distribuição Birnbaum–Saunders se aproxima de uma distribuição simétrica em torno de  $\beta$  (a mediana da distribuição) e sua variabilidade é reduzida. Além disso, observe que, à medida que o parâmetro de forma aumenta, a distribuição Birnbaum–Saunders apresenta caudas mais pesadas. Isso significa que  $\alpha$  não apenas modifica a forma da distribuição, mas também sua assimetria e curtose.

```
#### Densidade Birnbaum-Saunders f(t; alpha, beta)
bs_pdf <- function(t, alpha, beta = 1) {
  u <- t / beta
  xi <- sqrt(u) - 1 / sqrt(u) # xi(u)
  xi_prime <- 0.5 * (u^(-1/2) + u^(-3/2)) # xi'(u)
  (1 / sqrt(2*pi)) * exp(-(xi^2) / (2*alpha^2)) * (1 / (alpha*beta)) * xi_prime
}

#### Curvas para diferentes alphas
alphas <- c(0.3, 0.5, 1, 2)
t <- seq(1e-4, 4, length.out = 2000)

dens <- sapply(alphas, function(a) bs_pdf(t, a, beta = 1))
ylim <- c(0, max(dens, na.rm = TRUE))

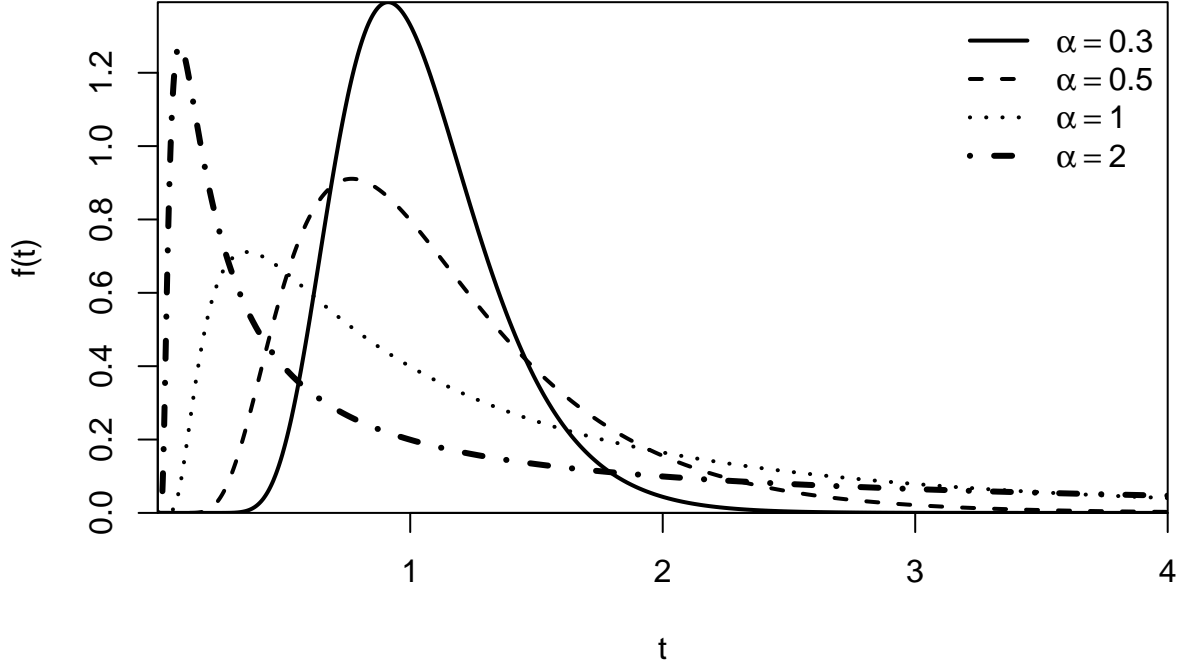
plot(t, dens[,1], type = "l", lwd = 2, lty = 1,
      xlab = "t", ylab = "f(t)", ylim = ylim, xaxs = "i", yaxs = "i")
```



```

lines(t, dens[,2], lwd = 2, lty = 2)
lines(t, dens[,3], lwd = 2, lty = 3)
lines(t, dens[,4], lwd = 3, lty = 4)
legend("topright",
      legend = parse(text = paste0("alpha==", alphas)),
      lty = c(1,2,3,4), lwd = c(2,2,2,3), bty = "n")

```



- Podemos obter a função de distribuição acumulada de  $T \sim \text{BS}(\alpha, \beta)$  como:

$$\begin{aligned}
 F_T(t; \alpha, \beta) &= P(T \leq t) = P\left(\beta \left(\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1}\right)^2 \leq t\right) \\
 &= P\left(Z \leq \frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right).
 \end{aligned}$$

Portanto,

$$F_T(t; \alpha, \beta) = \Phi\left(\frac{1}{\alpha} \xi\left(\frac{t}{\beta}\right)\right), \quad t > 0, \alpha > 0, \beta > 0, \quad (14)$$

em que  $\xi(\cdot)$  é dada em (13) e

$$\Phi(z) = \int_{-\infty}^z \phi(u) du, \quad z \in \mathbb{R}, \quad (15)$$

é a função de distribuição acumulada normal padrão, com  $\phi(\cdot)$  dada em (11). Note que a função definida em (14) também pode ser obtida integrando a função densidade de probabilidade definida em (12).

- Outro indicador útil para análises estatísticas é a função quantílica, ou o  $q \times 100$ -ésimo quantil da distribuição. Com base na Equação (14), a função quantílica de  $T \sim \text{BS}(\alpha, \beta)$  é dada por:

$$t(q; \alpha, \beta) = F_T^{-1}(q; \alpha, \beta) = \beta \left( \frac{\alpha z(q)}{2} + \sqrt{\left( \frac{\alpha z(q)}{2} \right)^2 + 1} \right)^2, \quad (16)$$

para  $0 < q \leq 1$ , em que  $F_T^{-1}(\cdot)$  é a função inversa de  $F_T(\cdot)$  expressa em (14), e  $z(q) = \Phi^{-1}(q)$  é o  $q \times 100$ -ésimo quantil da distribuição normal padrão, com  $\Phi^{-1}(\cdot)$  sendo a função inversa de  $\Phi(\cdot)$  dada em (15). Note que, se  $q = 0,5$ , então  $z(0,5) = 0$ , o que corresponde à média, mediana e moda de  $Z \sim N(0,1)$ . Assim, a partir da função quantílica definida em (16), temos que  $t(0,5; \alpha, \beta) = \beta$ , o que confirma que o parâmetro de escala  $\beta > 0$  é também a mediana da distribuição Birnbaum-Saunders. A Equação dada em (16) é utilizada para gerar números aleatórios em processos de simulação da distribuição Birnbaum-Saunders e também para derivar ferramentas de aderência associadas a ela.

- Dois indicadores úteis na análise de tempos de vida são a função de confiabilidade (ou sobrevivência) e a taxa de falha (ou risco). A partir de (12) e (9), a função de confiabilidade e a taxa de falha de  $T \sim \text{BS}(\alpha, \beta)$  são, respectivamente, dadas por:

$$R_T(t; \alpha, \beta) = 1 - F_T(t; \alpha, \beta) = \Phi \left( -\frac{1}{\alpha} \xi \left( \frac{t}{\beta} \right) \right), \quad t > 0, \quad (17)$$

e

$$h_T(t; \alpha, \beta) = \frac{f_T(t; \alpha, \beta)}{R_T(t; \alpha, \beta)} = \frac{\phi \left( \frac{1}{\alpha} \xi \left( \frac{t}{\beta} \right) \right) \frac{1}{\alpha} \xi' \left( \frac{t}{\beta} \right)}{\Phi \left( -\frac{1}{\alpha} \xi \left( \frac{t}{\beta} \right) \right)}, \quad t > 0, \quad (18)$$

para  $0 < R_T(t; \alpha, \beta) < 1$ .

- Assim, as seguintes propriedades são válidas para a taxa de falha da distribuição Birnbaum-Saunders definida em (18):

(C1)  $h_T(t)$  é unimodal para qualquer  $\alpha$ , crescente para  $t < t_c$  e decrescente para  $t > t_c$ , em que  $t_c$  é o ponto de mudança de  $h_T(t)$ ;

(C2)  $h_T(t)$  se aproxima de  $1/(2\alpha^2\beta)$  à medida que  $t \rightarrow \infty$ ; e

(C3)  $h_T(t)$  tende a ser crescente à medida que  $\alpha \rightarrow 0$ .

Para mais detalhes sobre (C1)-(C3), ver Kundu, Kannan, and Balakrishnan (2008). Estudos numéricos indicam que a distribuição Birnbaum-Saunders apresenta aproximadamente uma taxa de falha crescente quando  $\alpha < 0,41$  e  $0 < t < 8\beta$ , o que implica uma taxa de falha crescente em média. Birnbaum and Saunders (1969a) mostraram, por meio de cálculos numéricos, que a taxa média de falha de  $T$  diminui lentamente para  $t < 1,64$ .

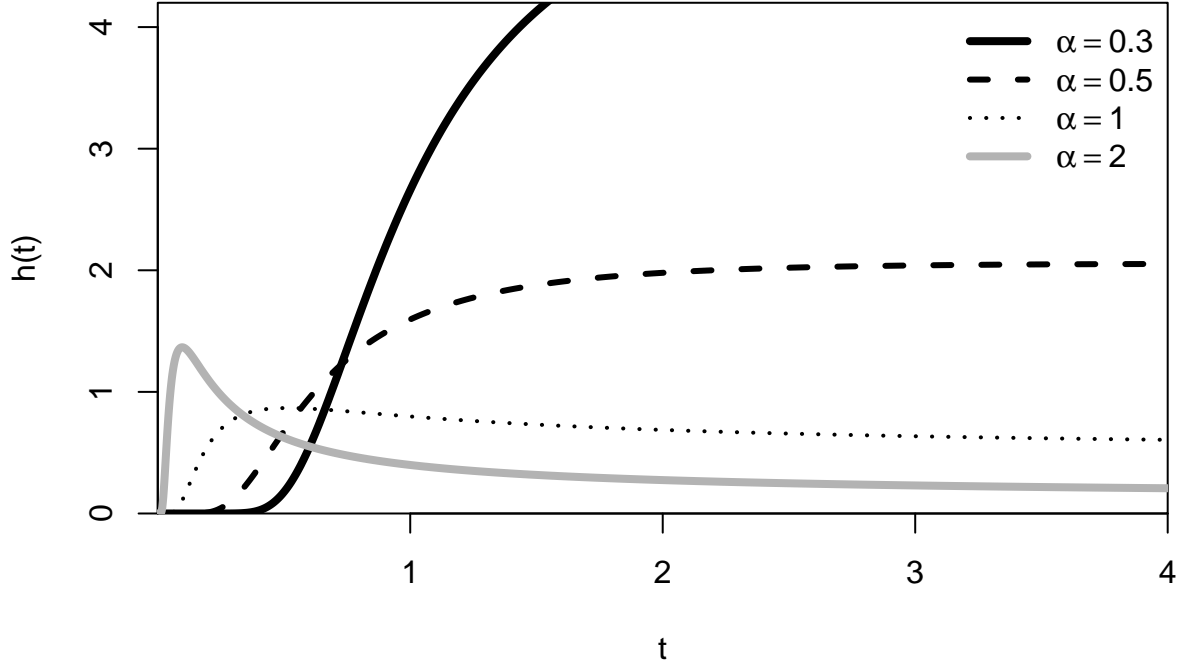
- A figura abaixo apresenta gráficos da taxa de falha da distribuição Birnbaum-Saunders para diferentes valores de  $\alpha$  e  $\beta = 1$ . As considerações teóricas descritas em (C1)-(C3) são verificadas por meio dessa análise gráfica.

```
#### Birnbaum-Saunders: PDF, CDF e Hazard
bs_pdf <- function(t, alpha, beta = 1) {
  stopifnot(all(t > 0), alpha > 0, beta > 0)
  u <- t / beta
  xi <- sqrt(u) - 1/sqrt(u)                # xi(u)
  xi_p <- 0.5 * (u^(-1/2) + u^(-3/2))      # xi'(u)
  z <- xi / alpha
  dnorm(z) * xi_p / (alpha * beta)
}

bs_cdf <- function(t, alpha, beta = 1) {
  stopifnot(all(t > 0), alpha > 0, beta > 0)
  z <- (sqrt(t/beta) - sqrt(beta/t)) / alpha
  pnorm(z)
}

bs_hazard <- function(t, alpha, beta = 1) {
  f <- bs_pdf(t, alpha, beta)
  S <- 1 - bs_cdf(t, alpha, beta)
  f / S
}

#### Curvas para alphas e domínio t
alphas <- c(0.3, 0.5, 1.0, 2.0)
t <- seq(1e-4, 4, length.out = 2000)
H <- sapply(alphas, function(a) bs_hazard(t, a, beta = 1))
ylim <- c(0, min(4.2, max(H[is.finite(H)])))
plot(t, H[,1], type = "l", lwd = 4, lty = 1, col = "black",
      xlab = "t", ylab = "h(t)", ylim = ylim, xaxs = "i", yaxs = "i")
lines(t, H[,2], lwd = 3, lty = 2, col = "black") # tracejado
lines(t, H[,3], lwd = 2, lty = 3, col = "black") # pontilhado
lines(t, H[,4], lwd = 4, lty = 1, col = "grey70") # cinza espesso
legend("topright",
      legend = parse(text = c("alpha==0.3", "alpha==0.5", "alpha==1.0", "alpha==2.0")),
      lty = c(1,2,3,1), lwd = c(4,3,2,4), col = c("black","black","black","grey70"),
      bty = "n")
```



- A distribuição Birnbaum-Saunders possui propriedades de proporcionalidade e reciprocidade, ou seja, a distribuição pertence à família de escalas e é fechada sob reciprocidade; veja S. C. Saunders (1974) e Marshall and Olkin (2007). A partir da função densidade de probabilidade da distribuição Birnbaum-Saunders dada em (9), é claro que, se  $T \sim \text{BS}(\alpha, \beta)$ , para todo  $a > 0$ , a variável aleatória  $Y = bT$  segue uma distribuição Birnbaum-Saunders com parâmetros  $\alpha$  e  $b\beta$ . Além disso, a variável aleatória  $Y = 1/T$  tem a mesma distribuição de  $T$ , com o parâmetro  $\beta$  substituído por  $1/\beta$ . Especificamente, seja  $T \sim \text{BS}(\alpha, \beta)$ . Então:

[(D1)]  $bT \sim \text{BS}(\alpha, b\beta)$  para  $b > 0$ ; e

[(D2)]  $\frac{1}{T} \sim \text{BS}\left(\alpha, \frac{1}{\beta}\right)$ .

- Além disso, a partir da representação dada em (10), note que qualquer variável aleatória que segue uma distribuição Birnbaum-Saunders está relacionada à distribuição qui-quadrado com um grau de liberdade e, portanto, tem média igual a um e variância igual a dois. Assim, a partir de (10), uma outra propriedade da distribuição Birnbaum-Saunders é:

[(D3)]  $V = Z^2 = \frac{1}{\alpha^2} \left( \frac{T}{\beta} + \frac{\beta}{T} - 2 \right) \sim \chi^2(1)$ , com  $E(V) = 1$  e  $V(V) = 2$ .

- Além disso, uma representação importante de uma variável aleatória que segue a distribuição Birnbaum-Saunders está relacionada à sua transformação logarítmica. Assim, se  $T \sim \text{BS}(\alpha, \beta)$ , então:

[(D4)]  $Y = \log(T) \sim \text{log-BS}(\alpha, \log(\beta))$ , em que  $\text{log-BS}(\alpha, \log(\beta))$  representa uma distribuição conhecida como Birnbaum-Saunders logarítmica. Nós a chamamos de distribuição log-Birnbaum-Saunders.

- As propriedades (D1)-(D4) da distribuição Birnbaum-Saunders são úteis para diversos

propósitos estatísticos, como obtenção de momentos, geração de números aleatórios, estimativa de parâmetros e modelagem baseada em regressão; veja mais detalhes em S. C. Saunders (1974), Rieck and Nedelman (1991), e Rieck (1999).

## 1.5 Função característica e momentos

- Seja  $T \sim \text{BS}(\alpha, \beta)$ . Então, a função característica de  $T$  é dada por

$$\begin{aligned}\varphi_T(s) &= \int_0^\infty \frac{\exp(i s t)}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right) \frac{1}{2\alpha} \left(\frac{t}{\beta^{-1/2}} + \frac{t}{\beta^{-3/2}}\right) dt \\ &= \frac{1}{2} \left( \left(1 + \frac{1}{\sqrt{1 - 2i s \alpha^2 \beta}}\right) \exp\left(-\frac{(1 + \sqrt{1 - 2i s \alpha^2 \beta})}{\alpha^2}\right) \right). \end{aligned} \quad (19)$$

- O *r*-ésimo momento da variável aleatória contínua  $T$  é dado em geral por

$$E(T^r) = \int_0^\infty t^r f_T(t) dt, \quad (20)$$

em que  $f_T(\cdot)$  é a função de densidade de probabilidade de  $T$ .

- Note que

$$E(T) = \beta \left(1 + \frac{\alpha^2}{2}\right) \quad (21)$$

e

$$E(T^2) = \beta^2 \left(\frac{3}{2}\alpha^4 + 2\alpha^2 + 1\right). \quad (22)$$

Logo,

$$\begin{aligned}V(T) &= \beta^2 \left(\frac{3}{2}\alpha^4 + 2\alpha^2 + 1\right) - \left(\beta \left(1 + \left(\frac{\alpha^2}{2}\right)\right)\right)^2 \\ &= \beta^2 \alpha^2 \left(1 + \frac{5}{4}\alpha^2\right). \end{aligned} \quad (23)$$

- Os coeficientes de assimetria e curtose de  $T \sim \text{BS}(\alpha, \beta)$  são, respectivamente, dados por

$$\text{CS}(T) = \alpha_3(T) = \sqrt{\beta_1(T)} = \frac{16\alpha^2(11\alpha^2 + 6)^2}{(5\alpha^2 + 4)^3} \quad (24)$$

$$\text{CK}(T) = \alpha_4(T) = \beta_2(T) = \beta_2(T) = 3 + \frac{6\alpha^2(93\alpha^2 + 40)}{(5\alpha^2 + 4)^2}. \quad (25)$$

- Momentos negativos da distribuição Birnbaum-Saunders podem ser obtidos da seguinte maneira. Se  $T \sim \text{BS}(\alpha, \beta)$ , então  $T/\beta$  e  $\beta/T$  têm a mesma distribuição, isto é,

$$\frac{T}{\beta} \stackrel{d}{=} \frac{\beta}{T} \sim \text{BS}(\alpha, 1). \quad (26)$$

- Assim, com base no resultado dado em (26), obtemos

$$E(T^{-r}) = \frac{E(T^r)}{\beta^{2r}}.$$

- Portanto, podemos obter a média e a variância de  $T^{-1}$  por

$$E\left(\frac{1}{T}\right) = \frac{1}{\beta} \left(1 + \frac{\alpha^2}{2}\right), \quad (27)$$

$$V\left(\frac{1}{T}\right) = \frac{\alpha^2}{\beta^2} \left(\frac{5}{4}\alpha^2 + 1\right). \quad (28)$$

## 1.6 Geração de números aleatórios

- Chang and Tang (1994) apresentou o seguinte método para gerar números aleatórios para a distribuição Birnbaum-Saunders. Seja  $T \sim \text{BS}(\alpha, \beta)$ . Então, lembramos que sua função de distribuição acumulada dada em (14) é:

$$F_T(t; \alpha, \beta) = \Phi\left(\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right), \quad t > 0, \alpha > 0, \beta > 0, \quad (29)$$

o que implica que, conforme mencionado em (D3),

$$V = \frac{1}{\alpha^2} \left(\frac{T}{\beta} + \frac{\beta}{T} - 2\right), \quad (30)$$

que segue uma distribuição  $\chi^2(1)$ . Então, a partir de (30), temos:

$$T^2 - \beta(2 + \alpha V)T + \beta^2 = 0. \quad (31)$$

Observe que a Equação (31) tem duas soluções, as raízes  $t_1$  e  $t_2$ , em que  $t_1 t_2 = \beta^2$  e  $t_2 = \beta^2 / t_1$ .

- Michael, Schucany, and Haas (1976) apresentou um método para gerar variáveis aleatórias por meio de transformações com múltiplas raízes. Esse método pode ser adaptado à distribuição Birnbaum-Saunders, pois consiste em duas variáveis aleatórias contínuas, digamos  $V$  e  $T$ , relacionadas por  $V = g(T)$ , cuja equação tem  $k$  raízes diferentes, denotadas por  $t_1, t_2, \dots, t_k$ , uma das quais corresponde a um valor  $v_0$  com probabilidade:

$$p_i(v_0) = \left(1 + \sum_{j=1 \neq i}^k \left| \frac{g'(t_i)}{g'(t_j)} \right| \frac{f_T(t_j)}{f_T(t_i)} \right)^{-1}, \quad i = 1, \dots, k, \quad (32)$$

em que  $g'(\cdot)$  é a derivada de  $g(\cdot)$ ,  $|x|$  é o valor absoluto de  $x$ , e  $f_T(\cdot)$  é a função densidade de probabilidade de  $T$ . Para a distribuição Birnbaum-Saunders, temos  $k = 2$  e a probabilidade  $p_1(v_0) = 1 - p_2(v_0)$  dada em (32), na qual devemos escolher a menor raiz, digamos  $t_1$ , é igual a  $1/2$ . Isso é obtido a partir da Equação (32), em que, considerando  $t_1 t_2 = \beta^2$  e  $t_2 = \beta^2/t_1$ , temos:

$$\frac{g'(t_1)}{g'(t_2)} = \frac{\frac{1}{\beta} - \frac{\beta}{t_1^2}}{\frac{1}{\beta} - \frac{\beta}{t_2^2}} = \frac{\frac{1}{\beta} - \frac{\beta}{t_1^2}}{\frac{1}{\beta} - \frac{t_1^2}{\beta^3}} = - \left( \frac{\beta}{t_1} \right)^2, \quad (33)$$

e

$$\begin{aligned} \frac{f_T(t_1)}{f_T(t_2)} &= \frac{(t_2 + \beta)}{t_2^{3/2}} \frac{t_1^{3/2}}{(t_1 + \beta)} \\ &\times \exp \left( -\frac{1}{2\alpha^2} \left( \left( \sqrt{\frac{t_2}{\beta}} - \sqrt{\frac{\beta}{t_2}} \right)^2 + \left( \sqrt{\frac{t_1}{\beta}} - \sqrt{\frac{\beta}{t_1}} \right)^2 \right) \right) = \left( \frac{t_1}{\beta} \right)^2. \end{aligned} \quad (34)$$

Portanto, a partir das Equações (33), (34) e com base em (32), é comprovado que  $p_1(v_0) = 1 - p_2(v_0) = 1/2$ . Assim, uma variável aleatória com distribuição Birnbaum-Saunders de parâmetro de forma  $\alpha$  e parâmetro de escala  $\beta$  pode ser gerada por:

$$t_1 1_{]-\infty, 0.5]}(u) + t_2 1_{]0.5, \infty]}(u), \quad (35)$$

em que  $1_A(\cdot)$  é a função indicadora do conjunto  $A$ . Observe que a expressão em (35), de acordo com Michael, Schucany, and Haas (1976), assegura que a probabilidade de escolher uma das duas raízes  $t_1$  ou  $t_2$ , dado que  $V = v$ , é igual a  $1/2$ . O procedimento de Michael, Schucany, and Haas (1976) para gerar números aleatórios é conhecido como o método das múltiplas raízes e é descrito para a distribuição Birnbaum-Saunders no Algoritmo 1.

- Além do método das múltiplas raízes para gerar números aleatórios a partir da distribuição Birnbaum-Saunders proposto por Chang and Tang (1994), Rieck (2003) introduziu um método baseado na relação entre as distribuições Birnbaum-Saunders e log-Birnbaum-Saunders; veja Rieck and Nedelman (1991) e a Propriedade (D4). Se  $T \sim \text{BS}(\alpha, \beta)$ , então  $Y = \log(T)$  tem uma função de distribuição acumulada dada por:

$$F_Y(y) = \Phi \left( \frac{2}{\alpha} \sinh \left( \frac{y - \log(\beta)}{2} \right) \right), \quad y \in \mathbb{R}, \alpha > 0, \log(\beta) \in \mathbb{R}. \quad (36)$$

Assim,

$$Z = \left( \frac{2}{\alpha} \sinh \left( \frac{Y - \log(\beta)}{2} \right) \right) \sim N(0, 1). \quad (37)$$

A solução da Equação (37) para  $Y$  é:

**Algorithm 1** Gerador 1 de números aleatórios de uma distribuição Birnbaum-Saunders

- 1: Gere um número aleatório  $u$  de  $U \sim U(0, 1)$ .
- 2: Gere um número aleatório  $z$  de  $Z \sim N(0, 1)$ .
- 3: Defina valores para  $\alpha$  e  $\beta$  de  $T \sim BS(\alpha, \beta)$ .
- 4: Calcule um número aleatório  $t = t_1$  ou  $t = t_2$  de  $T \sim BS(\alpha, \beta)$  com o seguinte critério:
  - 4.1: Se  $u \leq 0.5$ , então

$$t_1 = \frac{(\beta(2 + \alpha^2 z^2)) - \sqrt{(-\beta(2 + \alpha^2 z^2)^2 - 4\beta^2)}}{2};$$

- 4.2: Senão, isto é, se  $u > 0.5$ ,

$$t_2 = \frac{(\beta(2 + \alpha^2 z^2)) + \sqrt{(-\beta(2 + \alpha^2 z^2)^2 - 4\beta^2)}}{2}.$$

- 5: Repita as etapas de 1 a 4 até que a quantidade necessária de números aleatórios seja concluída.

$$Y = \log(\beta) + 2 \log \left( \frac{\alpha Z}{2} + \left( \frac{\alpha^2 Z^2}{4} + 1 \right)^{1/2} \right). \quad (38)$$

Então, usando a relação entre as distribuições Birnbaum-Saunders e log-Birnbaum-Saunders, um gerador de números aleatórios para a distribuição Birnbaum-Saunders é obtido tomando  $\exp(Y)$  na Equação (39), o que é resumido no Algoritmo 2

**Algorithm 2** Gerador 2 de números aleatórios a partir de uma distribuição Birnbaum-Saunders

- 1: Gerar um número aleatório  $z$  a partir de  $Z \sim N(0, 1)$ .
- 2: Definir valores para  $\alpha$  e  $\beta$  de  $T \sim BS(\alpha, \beta)$ .
- 3: Calcular um número aleatório  $y$  de  $Y = \log(T) \sim \log\text{-BS}(\alpha, \log(\beta))$  usando (39), levando a:

$$y = \log(\beta) + 2 \log \left( \frac{\alpha z}{2} + \left( \frac{\alpha^2 z^2}{4} + 1 \right)^{1/2} \right).$$

- 4: Calcular um número aleatório  $t = \exp(y)$  de  $T \sim BS(\alpha, \beta)$  usando o valor  $y$  obtido no Passo 3.
- 5: Repetir os Passos 1 a 4 até que a quantidade necessária de números aleatórios seja gerada.

- Um terceiro método para gerar números aleatórios a partir de um modelo Birnbaum-Saunders é baseado na relação entre as distribuições Birnbaum-Saunders e a inversa gaussiana; veja Bhattacharyya and Fries (1982), Desmond (1986) e Balakrishnan et al. (2009). Para mais detalhes sobre a distribuição inversa gaussiana, consulte os livros de Chhikara and Folks (1989) e Seshadri (1999a); Seshadri (1999b). Seja  $X_1 \sim \text{IG}(\beta, \beta/\alpha)$ ,  $X_2 \sim \text{IG}(1/\beta, 1/(\alpha^2\beta))$  e  $Y = 1/X_2$ . Então, a variável aleatória



$$T = WX_1 + (1 - W)Y \quad (39)$$

segue uma distribuição Birnbaum-Saunders com parâmetros  $\alpha$  e  $\beta$ , em que

$$P(W = 0) = P(W = 1) = \frac{1}{2}.$$

Assim, utilizando a relação entre as distribuições Birnbaum-Saunders e a inversa gaussiana, um gerador de números aleatórios para a distribuição Birnbaum-Saunders é resumido no Algoritmo 3. Em Chhikara and Folks (1989), é possível encontrar um método para gerar números aleatórios a partir de uma distribuição inversa gaussiana.

---

**Algorithm 3** Gerador 3 de números aleatórios a partir de uma distribuição Birnbaum-Saunders

---

- 1: Definir valores para  $\alpha$  e  $\beta$  de  $T \sim \text{BS}(\alpha, \beta)$ .
- 2: Gerar um número aleatório  $x_1$  de  $X_1 \sim \text{IG}(\beta, \beta/\alpha)$ .
- 3: Simular um número aleatório  $x_2$  de  $X_2 \sim \text{IG}(1/\beta, 1/(\alpha^2\beta))$ .
- 4: Calcular um número aleatório  $y = 1/x_2$ .
- 5: Obter um número aleatório  $w$  de acordo com a distribuição de probabilidade de  $W$  dada por

$$P(W = 0) = P(W = 1) = 1/2.$$

- 6: Calcular um número aleatório  $t$  de  $T \sim \text{BS}(\alpha, \beta)$  usando

$$t = wx_1 + (1 - w)y.$$

- 7: Repetir os Passos 1 a 6 até que a quantidade necessária de números aleatórios seja gerada.
- 

- Os métodos baseados em (i) múltiplas raízes, (ii) a distribuição log-Birnbaum-Saunders e (iii) as distribuições inversa gaussiana têm algumas vantagens uns sobre os outros. No entanto, o procedimento mais utilizado para gerar números aleatórios a partir de um modelo probabilístico consiste em inverter a função de distribuição acumulada, conhecido como o método da transformação inversa (ou, analogamente, a transformação integral de probabilidade). Como é sabido, se uma variável aleatória  $T$  tem uma função de distribuição acumulada  $F_T(\cdot)$ , então a transformação da variável aleatória  $T$  é dada por

$$U = F_T(T), \quad (40)$$

seguinto uma distribuição uniforme no intervalo  $[0, 1]$ , ou seja,  $U = F_T(T) \sim U(0, 1)$ . A transformação  $U = F_T(T)$  definida em (40) é conhecida como a transformação integral de probabilidade. Assim, a variável aleatória  $T = F_T^{-1}(U)$  segue um modelo probabilístico com função de distribuição acumulada  $F_T(\cdot)$ , em que  $F_T^{-1}(\cdot)$  é a inversa da função  $F_T(\cdot)$ . Tendo isso em mente, e se a função de distribuição acumulada tem uma forma analítica fechada, é possível obter números aleatórios da distribuição de interesse com um gerador adequado de números aleatórios seguindo uma distribuição uniforme no intervalo  $[0, 1]$ .

Seja  $T \sim \text{BS}(\alpha, \beta)$ . Então, sua função quantílica (ou inversa da função de distribuição acumulada) é dada em (16). Note que a função quantílica de Birnbaum-Saunders depende da função quantílica da distribuição normal padrão, a qual não possui uma forma analítica fechada. De qualquer forma, é possível assumir que a função quantílica de Birnbaum-Saunders tem uma forma analítica fechada, que recordamos ser dada por

$$t(q) = \beta \left( \frac{\alpha z(q)}{2} + \sqrt{\frac{\alpha^2 z(q)^2}{4} + 1} \right)^2, \quad 0 < q < 1, \quad (41)$$

em que  $z(q)$  é o quantil  $q \times 100$  da distribuição normal padrão. Assim, utilizando a função quantílica de Birnbaum-Saunders, ou seja, o método de transformação inversa, um gerador de números aleatórios para a distribuição Birnbaum-Saunders é resumido no Algoritmo 4.

---

**Algorithm 4** Gerador 4 de números aleatórios a partir de uma distribuição Birnbaum-Saunders

---

- 1: Gerar um número aleatório  $z$  de  $Z \sim N(0, 1)$ .
- 2: Definir valores para  $\alpha$  e  $\beta$  de  $T \sim \text{BS}(\alpha, \beta)$ .
- 3: Calcular um número aleatório  $t$  de  $T \sim \text{BS}(\alpha, \beta)$  usando (41), o que leva a

$$t = \beta \left( \frac{\alpha z}{2} + \sqrt{\frac{\alpha^2 z^2}{4} + 1} \right)^2.$$

- 4: Repetir os Passos 1 a 3 até que a quantidade necessária de números aleatórios seja gerada.
- 

## 1.7 Método de Estimação por Máxima Verossimilhança

- Sejam  $T_1, T_2, \dots, T_n$  uma amostra aleatória de tamanho  $n$  de uma variável aleatória  $T$  com função densidade de probabilidade  $f_T(\cdot)$  e  $t_1, t_2, \dots, t_n$  suas observações (dados). Então, devido à independência entre  $T_1, T_2, \dots, T_n$ , a função de verossimilhança para um parâmetro  $\theta$  é definida em geral como

$$L(\theta) = \prod_{i=1}^n L_i(\theta), \quad (42)$$

em que  $L_i(\theta) = L(\theta; t_i) = f_T(t_i; \theta)$  é a contribuição individual de cada observação para a função de verossimilhança.

- No entanto, como é bem conhecido, porque a função de verossimilhança para  $\theta$  dada em (42) e seu logaritmo natural (doravante denominado log-verossimilhança) atingem seus valores máximos nos mesmos pontos, frequentemente prefere-se trabalhar com a função de log-verossimilhança devido à sua tratabilidade matemática. Assim, a função de log-verossimilhança para  $\theta$  é dada em geral por

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}), \quad (43)$$

em que  $\ell_i(\boldsymbol{\theta}) = \log(f_T(t_i; \boldsymbol{\theta}))$  é a contribuição individual de cada observação para a função de log-verossimilhança.

- Sejam  $T_1, T_2, \dots, T_n$  uma amostra aleatória de tamanho  $n$  de  $T \sim \text{BS}(\alpha, \beta)$  e  $t_1, t_2, \dots, t_n$  suas observações. Então, a função de log-verossimilhança para  $\boldsymbol{\theta} = (\alpha, \beta)^\top$  baseada em  $t_1, t_2, \dots, t_n$  é dada por

$$\ell(\boldsymbol{\theta}) = c_1 + \frac{n}{\alpha^2} - \frac{1}{2\alpha^2} \sum_{i=1}^n \left( \frac{t_i}{\beta} + \frac{\beta}{t_i} \right) - n \log(\alpha) - \frac{n}{2} \log(\beta) + \sum_{i=1}^n \log(t_i + \beta), \quad (44)$$

em que  $c_1$  é uma constante que não depende de  $\boldsymbol{\theta}$ .

- Para maximizar a função de log-verossimilhança  $\ell(\boldsymbol{\theta}) = \ell(\alpha, \beta)$  expressa em (44), precisamos das suas derivadas primeiras em relação a  $\alpha$  e  $\beta$ , formando o vetor escore definido por  $\dot{\boldsymbol{\ell}} = (\dot{\ell}_\alpha, \dot{\ell}_\beta)^\top$ , cujos elementos são dados por

$$\dot{\ell}_\alpha = \frac{\partial \ell(\alpha, \beta)}{\partial \alpha} = -\frac{2n}{\alpha^3} + \frac{1}{\alpha^3} \sum_{i=1}^n \left( \frac{t_i}{\beta} + \frac{\beta}{t_i} \right) - \frac{n}{\alpha}, \quad (45)$$

$$\dot{\ell}_\beta = \frac{\partial \ell(\alpha, \beta)}{\partial \beta} = \frac{1}{2\alpha^2} \sum_{i=1}^n \left( \frac{t_i}{\beta^2} - \frac{1}{t_i} \right) - \frac{n}{2\beta} + \sum_{i=1}^n \frac{1}{t_i + \beta}.$$

- Assim, para obter as equações de log-verossimilhança, fazemos  $\dot{\ell}_\alpha = 0$  e  $\dot{\ell}_\beta = 0$  e avaliamos em  $\alpha = \hat{\alpha}$  e  $\beta = \hat{\beta}$ , obtendo

$$\hat{\alpha} = \left( \frac{1}{n} \sum_{i=1}^n \left( \frac{t_i}{\hat{\beta}} + \frac{\hat{\beta}}{t_i} - 2 \right) \right)^{1/2} \quad (46)$$

$$\hat{\beta} = \sqrt{\frac{\frac{1}{2\hat{\alpha}^2} \sum_{i=1}^n t_i}{\frac{1}{2\hat{\alpha}^2} \sum_{i=1}^n \frac{1}{t_i} - \sum_{i=1}^n \frac{1}{t_i + \hat{\beta}} + \frac{n}{2\hat{\beta}}}}. \quad (47)$$

- Portanto, usando um algoritmo numérico, podemos resolver as equações (46) e (47). Esse algoritmo nos permite ter o procedimento iterativo conjunto

$$\hat{\alpha}^{(m+1)} = \left( \frac{1}{n} \sum_{i=1}^n \left( \frac{t_i}{\hat{\beta}^{(m)}} + \frac{\hat{\beta}^{(m)}}{t_i} - 2 \right) \right)^{1/2}, \quad m = 0, 1, 2, \dots, \quad (48)$$

$$\hat{\beta}^{(m+1)} = \sqrt{\frac{\frac{1}{2(\hat{\alpha}^{(m)})^2} \sum_{i=1}^n t_i}{\frac{1}{2(\hat{\alpha}^{(m)})^2} \sum_{i=1}^n \frac{1}{t_i} - \sum_{i=1}^n \frac{1}{t_i + \hat{\beta}^{(m)}} + \frac{n}{2\hat{\beta}^{(m)}}}},$$

que necessita de valores iniciais  $\hat{\alpha}^{(0)}$  e  $\hat{\beta}^{(0)}$ .

- Por exemplo, de acordo com Birnbaum and Saunders (1969b), podemos usar a estimativa média-média de  $\beta$  como valor inicial  $\hat{\beta}^{(0)}$  dada por

$$\tilde{\beta} = (sr)^{1/2}, \quad (49)$$

em que

$$s = \frac{\sum_{i=1}^n t_i}{n} \quad \text{e} \quad r = \frac{1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{t_i}} \quad (50)$$

são as médias aritmética e harmônica dos dados  $t_1, t_2, \dots, t_n$ , respectivamente. Assim, com base na expressão para  $\hat{\alpha}$  dada em (46) e no valor inicial para  $\hat{\beta}$  dado em (49), temos o valor inicial  $\hat{\alpha}^{(0)}$ . Alternativamente, como  $\beta$  é a mediana da distribuição Birnbaum-Saunders, a mediana amostral dos dados  $t_1, t_2, \dots, t_n$  pode ser usada como valor inicial  $\hat{\beta}^{(0)}$ , e novamente procedendo como antes, obtemos o valor inicial  $\hat{\alpha}^{(0)}$  para o procedimento iterativo apresentado em (48).

- Uma vez estabelecidas as estimativas pontuais de  $\alpha$  e  $\beta$ , podemos derivar inferências para grandes amostras com base nas propriedades dos estimadores de máxima verossimilhança. Um aspecto que facilita essa inferência é a matriz Hessiana (ou das derivadas segundas). A matriz Hessiana corresponde à derivada segunda da Equação (44) ou às derivadas primeiras dos elementos do vetor escore expressos na Equação (45) com respeito aos parâmetros desconhecidos, dada por

$$\ddot{\ell} = \left( \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right) = \begin{pmatrix} \ddot{\ell}_{\alpha\alpha} & \ddot{\ell}_{\alpha\beta} \\ \ddot{\ell}_{\beta\alpha} & \ddot{\ell}_{\beta\beta} \end{pmatrix}, \quad i, j = 1, 2, \quad (51)$$

em que

$$\ddot{\ell}_{\alpha\alpha} = \frac{\partial^2 \ell(\alpha, \beta)}{\partial \alpha^2} = \frac{6n}{\alpha^4} - \frac{3}{\alpha^4} \sum_{i=1}^n \left( \frac{t_i}{\beta} + \frac{\beta}{t_i} \right) + \frac{n}{\alpha^2}, \quad (52)$$

$$\ddot{\ell}_{\alpha\beta} = \frac{\partial^2 \ell(\alpha, \beta)}{\partial \alpha \partial \beta} = \ddot{\ell}_{\beta\alpha} = \frac{1}{\alpha^3} \sum_{i=1}^n \left( \frac{1}{t_i} - \frac{t_i}{\beta^2} \right), \quad (53)$$

$$\ddot{\ell}_{\beta\beta} = \frac{\partial^2 \ell(\alpha, \beta)}{\partial \beta^2} = -\frac{1}{\alpha^2 \beta^3} \sum_{i=1}^n t_i + \frac{n}{2\beta^2} - \sum_{i=1}^n \frac{1}{(t_i + \beta)^2}. \quad (54)$$

- A matriz de informação esperada de Fisher correspondente pode ser obtida da matriz Hessiana dada em (51) por

$$\mathcal{I}(\boldsymbol{\theta}) = -E \left( \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right) = \begin{pmatrix} \mathcal{I}_{\alpha\alpha} & \mathcal{I}_{\alpha\beta} \\ \mathcal{I}_{\beta\alpha} & \mathcal{I}_{\beta\beta} \end{pmatrix}, \quad i, j = 1, 2, \quad (55)$$

em que  $\mathcal{I}_{\alpha\alpha}$  e  $\mathcal{I}_{\beta\beta}$  são definidos como:

$$E(T) = \frac{\beta}{2} (2 + \alpha^2), \quad E\left(\frac{1}{T}\right) = \frac{1}{2\beta} (2 + \alpha^2). \quad (56)$$

- Portanto, os parâmetros  $\alpha$  e  $\beta$  são ortogonais, ou seja, os estimadores  $\hat{\alpha}$  e  $\hat{\beta}$  são independentes. Assim, a matriz de informação esperada de Fisher é estabelecida como:

$$\mathcal{I}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{2n}{\alpha^2} & 0 \\ 0 & \frac{n}{\beta^2} \left( \frac{1}{4} + \frac{1}{\alpha^2} + I(\alpha) \right) \end{pmatrix}, \quad (57)$$

com

$$I(\alpha) = 2 \int_0^\infty \left( \frac{1}{1 + \frac{1}{\xi(\alpha z)}} - \frac{1}{2} \right)^2 \phi(z) dz,$$

em que  $\phi(\cdot)$  e  $\xi(\cdot)$  são dados, respectivamente, em (13) e (11).

- Com base na normalidade assintótica dos estimadores de máxima verossimilhança, temos:

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \sim^a N_2(\mathbf{0}_{2 \times 1}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}} = \mathcal{J}(\boldsymbol{\theta})^{-1}), \quad (58)$$

em que  $\mathcal{J}(\boldsymbol{\theta}) = \lim_{n \rightarrow \infty} \frac{1}{n} \mathcal{I}(\boldsymbol{\theta})$ . Em prática, a matriz de informação observada pode ser usada como uma aproximação da matriz de variância-covariância dos estimadores.

- Com essa normalidade assintótica, podemos construir testes de hipóteses e intervalos de confiança para  $\boldsymbol{\theta}$ . A região de confiança aproximada  $(1 - \zeta) \times 100\%$  para  $\boldsymbol{\theta}$  é dada por:

$$\mathcal{R} \equiv \left\{ \boldsymbol{\theta} \in \mathbb{R}^2 : (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^\top \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq \chi_{1-\zeta}^2(2) \right\},$$

em que  $\chi_{1-\zeta}^2(2)$  é o quantil  $(1 - \zeta) \times 100$  da distribuição qui-quadrado com dois graus de liberdade.

## 1.8 Método de Estimação por Momentos

- Como é amplamente conhecido, os estimadores por momentos podem não ser únicos e, além disso, podem não existir em todos os casos. Isso ocorre ao estimar os parâmetros da distribuição Birnbaum-Saunders utilizando o método dos momentos. Uma alternativa é empregar o método dos momentos modificados, que fornece estimadores que sempre existem e são únicos; veja Ng, Kundu, and Balakrishnan (2003).
- A seguir, apresentamos os estimadores por momentos modificados para os parâmetros  $\alpha$  e  $\beta$  da distribuição Birnbaum-Saunders e suas distribuições assintóticas, as quais podem ser utilizadas para construir intervalos de confiança e testes de hipóteses para esses parâmetros.

- O método dos momentos modificado é um caso particular do método geral de estimação por momentos. Para mais detalhes sobre o método geral, consulte Mátyás (1999) e Santos-Neto et al. (2014), que aplicam o método a uma versão reparametrizada da distribuição Birnbaum-Saunders proposta por Santos-Neto et al. (2012); veja também Leiva et al. (2014).
- Seja  $T_1, T_2, \dots, T_n$  uma amostra de tamanho  $n$  de  $T \sim \text{BS}(\alpha, \beta)$ . As estimativas por momentos modificados de  $\alpha$  e  $\beta$ , denotadas por  $\tilde{\alpha}$  e  $\tilde{\beta}$ , são dadas por:

$$\tilde{\alpha} = \left( 2 \left( \sqrt{\frac{s}{r}} - 1 \right) \right)^{1/2} \quad \text{e} \quad \tilde{\beta} = \sqrt{s r}. \quad (59)$$

- Lembramos que  $s$  e  $r$  são as médias aritmética e harmônica, respectivamente, definidas em (50). Note que a estimativa por momentos modificados de  $\beta$  coincide com a estimativa média-média proposta por Birnbaum and Saunders (1969b), apresentada em (49).
- A distribuição conjunta assintótica dos estimadores por momentos modificados  $\tilde{\alpha}$  e  $\tilde{\beta}$  é bivariada normal e dada por:

$$\sqrt{n} \left( \begin{pmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{pmatrix} - \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right) \sim^a N_2(\mathbf{0}_{2 \times 1}, \boldsymbol{\Sigma}_{\tilde{\theta}}),$$

à medida que  $n \rightarrow \infty$ , em que  $\mathbf{0}_{2 \times 1}$  é um vetor  $2 \times 1$  de zeros e

$$\boldsymbol{\Sigma}_{\tilde{\theta}} = \begin{pmatrix} \frac{\alpha^2}{2} & 0 \\ 0 & (\alpha \beta)^2 \left( \frac{4+3\alpha^2}{(2+\alpha^2)^2} \right) \end{pmatrix}$$

é a matriz de variância-covariância assintótica de  $\tilde{\boldsymbol{\theta}}$ . Assim, as distribuições marginais assintóticas de  $\tilde{\alpha}$  e  $\tilde{\beta}$ , à medida que  $n \rightarrow \infty$ , são dadas por:

$$\begin{aligned} \sqrt{n}(\tilde{\alpha} - \alpha) &\sim^a N\left(0, \frac{\alpha^2}{2}\right) \\ \sqrt{n}(\tilde{\beta} - \beta) &\sim^a N\left(0, (\alpha \beta)^2 \frac{(4 + 3\alpha^2)}{(2 + \alpha^2)^2}\right), \end{aligned} \quad (60)$$

respectivamente. A partir do resultado assintótico dado em (60), podemos construir intervalos de confiança e testes de hipóteses para os parâmetros  $\alpha$  e  $\beta$  da distribuição Birnbaum-Saunders.

## 1.9 A distribuição log-Birnbaum-Saunders

- A distribuição sinh-normal é obtida a partir da transformação

$$Y = \mu + \sigma \operatorname{arcsinh}\left(\frac{\alpha Z}{2}\right),$$

em que  $Z \sim N(0, 1)$ ,  $\alpha > 0$  é um parâmetro de forma,  $\mu \in \mathbb{R}$  é um parâmetro de localização, e  $\sigma > 0$  é um parâmetro de escala. Nesse caso, a notação  $Y \sim \text{SHN}(\alpha, \mu, \sigma)$  é utilizada. A função densidade de probabilidade de  $Y$  é dada por

$$f_Y(y; \alpha, \mu, \sigma) = \phi\left(\frac{2}{\alpha} \sinh\left(\frac{y - \mu}{\sigma}\right)\right) \frac{2 \cosh((y - \mu)/\sigma)}{\alpha \sigma}, \quad y \in \mathbb{R}. \quad (61)$$

- A função de distribuição acumulada de  $Y$  é expressa por

$$F_Y(y; \alpha, \mu, \sigma) = \Phi\left(\frac{2}{\alpha} \sinh\left(\frac{y - \mu}{\sigma}\right)\right), \quad y \in \mathbb{R}. \quad (62)$$

A função quantílica de  $Y$  é definida como

$$y(q; \alpha, \mu, \sigma) = F_Y^{-1}(q; \alpha, \mu, \sigma) = \mu + \sigma \operatorname{arcsinh}\left(\frac{\alpha z(q)}{2}\right), \quad 0 < q \leq 1,$$

em que  $z(q)$  é o quantil de ordem  $q \times 100$  de  $Z \sim N(0, 1)$  e  $F_Y^{-1}(\cdot)$  é a função inversa de  $F_Y(\cdot)$ . A média e a variância de  $Y$  podem ser obtidas usando a função geradora de momentos dada por

$$m(r) = \exp(\mu r) \left( \frac{K_{\frac{\sigma r + 1}{2}}(1/\delta^2) + K_{\frac{\sigma r - 1}{2}}(1/\delta^2)}{2K_{1/2}(1/\delta^2)} \right), \quad (63)$$

em que  $E(Y) = \mu$  e  $K_\lambda(\cdot)$  é a função de Bessel modificada de terceira ordem.

- A função de confiabilidade e a taxa de falha de  $Y \sim \text{SHN}(\alpha, \mu, \sigma)$  são dadas por

$$R_Y(y; \alpha, \mu, \sigma) = \Phi\left(-\frac{2}{\alpha} \sinh\left(\frac{y - \mu}{\sigma}\right)\right), \quad y \in \mathbb{R}, \quad (64)$$

$$h_Y(y; \alpha, \mu, \sigma) = \frac{2\phi\left(\frac{2}{\alpha} \sinh\left(\frac{y - \mu}{\sigma}\right)\right) \cosh\left(\frac{y - \mu}{\sigma}\right)}{\alpha \sigma \Phi\left(-\frac{2}{\alpha} \sinh\left(\frac{y - \mu}{\sigma}\right)\right)}, \quad y \in \mathbb{R}, \quad (65)$$

respectivamente.

- Rieck (1989) observaram que uma distribuição sinh-normal com parâmetros  $\alpha, \mu$  e  $\sigma$  apresenta as seguintes propriedades:

[(i)] É simétrica em torno de sua média  $\mu$ ; [(ii)] É unimodal para  $\alpha < 2$  e sua curtose é menor que a da distribuição normal; [(iii)] É platicúrtica para  $\alpha = 2$ ; e [(iv)] É bimodal para  $\alpha > 2$ , e à medida que  $\alpha$  aumenta, a bimodalidade é enfatizada, com modos mais separados e curtose maior que a da distribuição normal.

- Além disso, se  $Y \sim \text{SHN}(\alpha, \mu, \sigma)$ , então, à medida que  $\alpha$  se aproxima de zero,

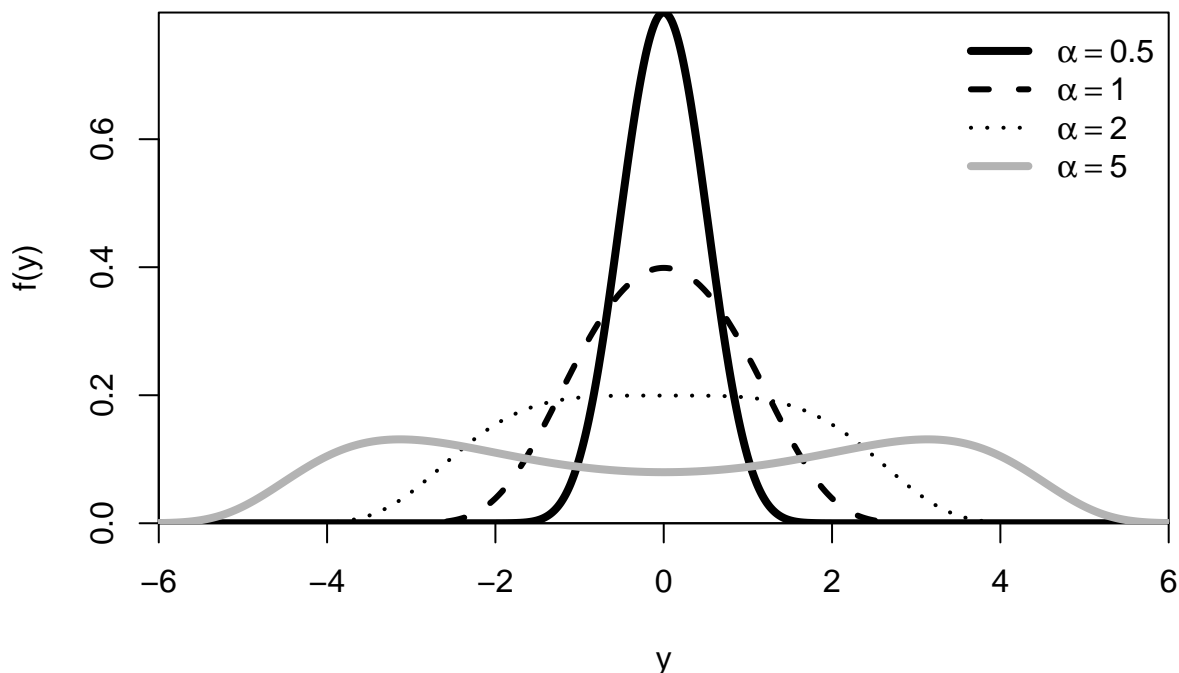
$$U = \frac{2(Y - \mu)}{\alpha \sigma} \sim N(0, 1).$$

- Assim, como mencionado, o parâmetro  $\alpha$  modifica a forma da distribuição sinh-normal, pois, à medida que  $\alpha$  aumenta, a curtose também aumenta. O parâmetro  $\mu$  altera a localização, enquanto o parâmetro  $\sigma$  altera a escala dessa distribuição, conforme pode ser verificado na figura abaixo.

```
#### Log-BS: PDF de  $Y = \mu + \sigma * \log(T)$ 
logbs_pdf <- function(y, alpha, mu = 0, sigma = 2) {
  stopifnot(alpha > 0, sigma > 0)
  z <- (y - mu) / sigma
  dnorm((2/alpha) * sinh(z)) * (2 * cosh(z)) / (alpha * sigma)
}

#### Curvas para alphas e domínio y
alphas <- c(0.5, 1.0, 2.0, 5.0)
y <- seq(-6, 6, length.out = 2000)

dens <- sapply(alphas, function(a) logbs_pdf(y, a, mu = 0, sigma = 2))
ylim <- c(0, max(dens, na.rm = TRUE))
plot(y, dens[,1], type = "l", lwd = 4, lty = 1, col = "black",
      xlab = "y", ylab = "f(y)", ylim = ylim, xaxs = "i", yaxs = "i")
lines(y, dens[,2], lwd = 3, lty = 2, col = "black") # tracejado
lines(y, dens[,3], lwd = 2, lty = 3, col = "black") # pontilhado
lines(y, dens[,4], lwd = 4, lty = 1, col = "grey70") # cinza espesso
legend("topright",
      legend = parse(text = c("alpha==0.5", "alpha==1.0", "alpha==2.0", "alpha==5.0")),
      lty = c(1, 2, 3, 1), lwd = c(4, 3, 2, 4), col = c("black", "black", "black", "grey70"),
      bty = "n")
```





- Rieck and Nedelman (1991) provaram que, se  $T \sim \text{BS}(\alpha, \beta)$ , então  $Y = \log(T) \sim \text{SHN}(\alpha, \mu, \sigma = 2)$ , em que  $\mu = \log(\beta)$ . Por essa razão, a distribuição sinh-normal é também conhecida como distribuição log-Birnbaum-Saunders. Assim, com base na Equação (63), seja  $Y = \log(T)$ , e a relação

$$M_Y(r) = E(\exp(rY)) = E(T^r), \quad (66)$$

os momentos de qualquer ordem da distribuição Birnbaum-Saunders podem ser calculados; ver Rieck (1999). Além disso, a estimação de parâmetros e a geração de números aleatórios dessa distribuição podem ser obtidas de forma mais eficiente a partir da distribuição log-Birnbaum-Saunders.

## 1.10 Modelo de regressão Birnbaum-Saunders

- Considere o modelo de regressão Birnbaum-Saunders

$$T_i = \beta_i \varphi_i = \exp(\mu_i) \varphi_i = \exp(\mathbf{x}_i^\top \boldsymbol{\eta}) \varphi_i, \quad i = 1, 2, \dots, n, \quad (67)$$

em que  $T_i$  e  $\beta_i = \exp(\mu_i)$  são a variável resposta e a mediana para o caso  $i$ , respectivamente,  $\boldsymbol{\eta} = (\eta_0, \eta_1, \dots, \eta_p)^\top$  é um vetor  $(p+1) \times 1$  de parâmetros desconhecidos a serem estimados,  $\mathbf{x}_i^\top = (1, x_{i1}, \dots, x_{ip})$  contém valores de  $p$  variáveis explicativas, e  $\varphi_i \sim \text{BS}(\alpha, 1)$  é o erro do modelo, para  $i = 1, \dots, n$ . Note que, usando a Propriedade (D1) da distribuição Birnbaum-Saunders, temos  $T_i \sim \text{BS}(\alpha, \beta_i)$ .

- Ao aplicar o logaritmo na Equação (67), obtemos

$$Y_i = \mu_i + \varepsilon_i = \mathbf{x}_i^\top \boldsymbol{\eta} + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (68)$$

em que  $Y_i = \log(T_i)$  é a variável resposta logarítmica para o caso  $i$ ,  $\boldsymbol{\eta}$  e  $\mathbf{x}_i$  são os mesmos da Equação (67), e  $\varepsilon_i = \log(\varphi_i) \sim \log\text{-BS}(\alpha, 0)$  é o termo de erro do modelo, para  $i = 1, \dots, n$ .

- Agora, considere o modelo de regressão Birnbaum-Saunders dado em (67) tal que  $T_1, T_2, \dots, T_n$  é uma amostra de tamanho  $n$  de  $T \sim \text{BS}(\alpha, \beta_i)$ . Então,  $Y_1 = \log(T_1), Y_2 = \log(T_2), \dots, Y_n = \log(T_n)$  pode ser considerada uma amostra de tamanho  $n$  de  $Y \sim \log\text{-BS}(\alpha, \mu_i = \log(\beta_i))$ , com  $y_1, \dots, y_n$  sendo suas observações.
- Assim, a função de log-verossimilhança para  $\boldsymbol{\theta} = (\alpha, \boldsymbol{\eta}^\top)^\top$  com base em  $y_1, \dots, y_n$  é dada por

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= \sum_{i=1}^n \log \left( \frac{1}{\alpha \sqrt{2\pi}} \cosh \left( \frac{y_i - \mu_i}{2} \right) \exp \left( -\frac{2}{\alpha^2} \left( \sinh \left( \frac{y_i - \mu_i}{2} \right) \right)^2 \right) \right) \\ &= c_3 + \sum_{i=1}^n \log(\xi_{i1}) - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2, \end{aligned} \quad (69)$$

em que  $c_3$  é uma constante que não depende de  $\boldsymbol{\theta}$ , e

$$\xi_{i1} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mu_i}{2}\right), \quad \xi_{i2} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mu_i}{2}\right), \quad (70)$$

com  $\mu_i = \mathbf{x}_i^\top \boldsymbol{\eta}$  para  $i = 1, 2, \dots, n$ .

- As estimativas de máxima verossimilhança do parâmetro de forma e dos coeficientes de regressão são soluções das equações

$$\dot{\ell}_\alpha = 0, \quad \dot{\ell}_{\eta_j} = 0, \quad j = 0, 1, \dots, p, \quad (71)$$

em que  $\dot{\ell}_\alpha$  e  $\dot{\ell}_{\eta_j}$  são as derivadas primeiras da função de verossimilhança dada na Equação (69), constituindo o vetor gradiente correspondente, dado por

$$\dot{\ell} = \begin{pmatrix} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\eta}} \end{pmatrix} = \begin{pmatrix} \dot{\ell}_\alpha \\ \dot{\ell}_\eta \end{pmatrix},$$

em que

$$\dot{\ell}_\alpha = \frac{1}{\alpha} \sum_{i=1}^n (\xi_{i2}^2 - 1), \quad \dot{\ell}_\eta = (\dot{\ell}_{\eta_0}, \dot{\ell}_{\eta_1}, \dots, \dot{\ell}_{\eta_p})^\top,$$

com

$$\dot{\ell}_{\eta_j} = \sum_{i=1}^n \left( \frac{x_{ij}}{\alpha^2} \sinh(y_i - \mu_i) - \frac{x_{ij}}{2} \tanh\left(\frac{y_i - \mu_i}{2}\right) \right), \quad j = 0, 1, \dots, p.$$

- Ao resolver a equação de log-verossimilhança  $\dot{\ell}_\alpha = 0$ , dada na Equação (71) em relação a  $\alpha$ , e avaliá-la em  $\hat{\alpha}$ , obtém-se a estimativa de máxima verossimilhança do parâmetro de forma:

$$\hat{\alpha} = \sqrt{\frac{4}{n} \sum_{i=1}^n \sinh^2\left(\frac{y_i - \mathbf{x}_i^\top \hat{\boldsymbol{\eta}}}{2}\right)},$$

em que  $\hat{\boldsymbol{\eta}}$  é a estimativa de máxima verossimilhança de  $\boldsymbol{\eta}$ . No entanto, as equações  $\dot{\ell}_{\eta_j} = 0$ , para  $j = 0, 1, \dots, p$ , dadas também em (71), não apresentam soluções analíticas. Portanto, é necessário utilizar métodos iterativos para resolver problemas de otimização não linear. Por exemplo, pode-se utilizar o método quase-Newton de Broyden-Fletcher-Goldfarb-Shanno (BFGS); ver Nocedal and Wright (1999), Lange (2001) e Press et al. (1992). O método BFGS está disponível no software R por meio das funções `optim` e `optimx`; ver [www.R-project.org](http://www.R-project.org) e R Core Team (2014).

- Note que a bimodalidade da distribuição log-Birnbaum-Saunders quando  $\alpha > 2$  pode causar múltiplos máximos na função de log-verossimilhança. No entanto, a experiência com dados de fadiga indica que o caso de  $\alpha > 2$  é incomum na prática. Além disso, foi demonstrado que, se  $\alpha \leq 2$ , a estimativa de máxima verossimilhança de  $\boldsymbol{\eta}$  é única se  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$  tem posto  $p$ .

- Rieck and Nedelman (1991) observaram que o estimador de mínimos quadrados de  $\boldsymbol{\eta}$  é menos eficiente do que o estimador de máxima verossimilhança. Como é bem conhecido, a estimativa de mínimos quadrados de  $\boldsymbol{\eta}$  é dada por

$$\bar{\boldsymbol{\eta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y},$$

em que  $\mathbf{y}$  é um vetor  $n \times 1$  de observações e  $\mathbf{X}$  é a matriz de desenho do modelo.

- A inferência em grandes amostras para  $\boldsymbol{\theta}$  pode ser novamente baseada na normalidade assintótica do estimador  $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\boldsymbol{\eta}}^\top)^\top$ , à medida que  $n \rightarrow \infty$ , o que é dado por

$$\hat{\boldsymbol{\theta}} \sim^a N_{p+2}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}), \quad (72)$$

em que  $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}$  é a matriz assintótica de variância-covariância  $(p+2) \times (p+2)$  de  $\hat{\boldsymbol{\theta}}$ .

- Note que  $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}$  pode ser aproximada por  $-\ddot{\boldsymbol{\ell}}^{-1}$ , em que  $-\ddot{\boldsymbol{\ell}}$  é a matriz de informação observada  $(p+2) \times (p+2)$ , obtida a partir da matriz Hessiana com as segundas derivadas dadas por

$$\ddot{\boldsymbol{\ell}} = \begin{pmatrix} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha^2} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \boldsymbol{\eta}} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\eta} \partial \alpha} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^\top} \end{pmatrix} = \begin{pmatrix} \ddot{\ell}_{\alpha\alpha} & \ddot{\ell}_{\alpha\boldsymbol{\eta}} \\ \ddot{\ell}_{\boldsymbol{\eta}\alpha} & \ddot{\ell}_{\boldsymbol{\eta}\boldsymbol{\eta}} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n \frac{(1-3\xi_{i2}^2)}{\alpha^2} & \mathbf{k}^\top \mathbf{X} \\ \mathbf{X}^\top \mathbf{k} & \mathbf{X}^\top \mathbf{V} \mathbf{X} \end{pmatrix}, \quad (73)$$

em que  $\mathbf{V} = \text{diag}\{v_1(\boldsymbol{\theta}), \dots, v_n(\boldsymbol{\theta})\}$  e  $\mathbf{k} = (k_1, \dots, k_n)^\top$ , com

$$\begin{aligned} v_i(\boldsymbol{\theta}) &= \frac{1}{4} \left( \text{sech} \left( \frac{y_i - \mu_i}{2} \right) \right)^2 - \frac{1}{\alpha^2} \cosh(y_i - \mu_i), \\ k_i &= -\frac{2}{\alpha^3} \sinh(y_i - \mu_i), \quad i = 1, 2, \dots, n. \end{aligned} \quad (74)$$

- Assim, como é bem conhecido, com base no resultado dado na Equação (72), uma região de confiança aproximada de  $(1 - \varrho) \times 100\%$  para o parâmetro  $\boldsymbol{\theta}$  pode ser obtida por

$$\mathcal{R} \equiv \left\{ \boldsymbol{\theta} \in \mathbb{R}^{p+2} : (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^\top \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq \chi_{1-\varrho}^2(p+2) \right\},$$

em que  $\chi_{1-\varrho}^2(p+2)$  denota o quantil de ordem  $(1 - \varrho) \times 100$  da distribuição qui-quadrado central com  $p+2$  graus de liberdade. Nesse caso, uma estimativa de  $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}$  pode ser obtida aproximadamente a partir da matriz de informação observada, com base na Equação (73), avaliada em  $\hat{\boldsymbol{\theta}}$ .

```
#####
## Regression by Rieck and Nedelman (1991)
## Rieck, J.R., Nedelman, J.R., 1991. A log-linear model for the
## Birnbaum-Saunders distribution. Technometrics 33, 51-60.
#####
```

```
mle_unc <- function(x, t, v, kernel = "normal"){
```

```

x      <- as.matrix(x)
t      <- as.matrix(t)
y      <- as.matrix(log(t))
fit     <- lm.fit(x, y)
beta    <- c(fit$coef)
k       <- length(beta)
n       <- length(y)
mu      <- x %*% beta
alpha   <- sqrt((4 / n) * sum((sinh((y - mu) / 2)) ^ 2))
thetaStar <- c(beta,alpha)

loglik <- function(par){

  z1     <- par[1:k]
  z2     <- par[k + 1]
  mu     <- (x %*% z1)
  xi1    <- (2 / z2) * cosh((y - mu) / 2)
  xi2    <- (2 / z2) * sinh((y - mu) / 2)
  g      <- switch(kernel,
                    "normal" = exp(-(xi2 ^ 2) / 2),
                    "t"      = (v + (xi2 ^ 2)) ^ (-(v + 1) / 2)
  )

  const  <- switch(kernel,
                    "normal" = 1 / (sqrt(2 * pi)),
                    "t"      = (gamma((v + 1) / 2) * v ^ (v / 2)) /
                      (sqrt(pi) * gamma(v / 2))
  )

  result <- switch(kernel,
                    "normal" = sum( log(const / 2) + log(xi1) + log(g) ) ,
                    "t"      = sum(log(const / 2) + log(xi1) + log(g) )
  )
  return(-result)
}

loglikplus <- function(par){

  z1     <- par[1:k]
  z2     <- par[k + 1]
  mu     <- (x %*% z1)
  xi1    <- (2 / z2) * cosh((y - mu) / 2)

```

```

xi2    <- (2 / z2) * sinh((y - mu) / 2)
result  <- log(dnorm(xi2)) + log(xi1/2)
return(sum(result))

}

score <- function(par){

  z1      <-par[1:k]
  z2      <-par[k+1]
  mu      <-(x %*% z1)
  xi1     <- (2 / z2) * cosh((y - mu) / 2)
  xi2     <- (2 / z2) * sinh((y - mu) / 2)
  status  <- 1
  const   <- switch(kernel,
    "normal" = 1 / (sqrt(2 * pi)),
    "t"      = (gamma((v + 1.0) / 2) * v ^ (v / 2)) /
      (sqrt(pi) * gamma(v / 2))
  )

  derivativeF <- switch(kernel,
    "normal" = -(const) * xi2 * exp(-(xi2 ^ 2) / 2),
    "t"      = -(const) * (v + 1) * xi2 * ((v + (xi2 ^ 2)) ^
      (-(v + 3) / 2))
  )

  h          <- switch(kernel,
    "normal" = dnorm(xi2) / (1 - pnorm(xi2)),
    "t"      = dt(xi2, df = v) / (1 - pt(xi2, df = v))
  )

  wg         <- switch(kernel,
    "normal" = -(1 / 2),
    "t"      = -(v + 1) / (2 * (v + (xi2 ^ 2)))
  )

  Ubeta      <- switch(kernel,
    "normal" = -status * ((2 / (z2 ^ 2)) * sinh(y - mu) * wg +
      (1 / 2) * tanh((y - mu) / 2)) +
      ((1 - status) / 2) * xi1 * h,
    "t"      = -status * ((2 / (z2 ^ 2)) * sinh(y - mu) * wg +
      (1 / 2) * tanh((y - mu) / 2)) +
      ((1 - status) / 2) * xi1 * h
  )

```

```

)

Ualpha      <- switch(kernel,
                        "normal" = -sum((status / z2) *
                                         (((xi2 ^ 2) * 2 * wg) + 1) -
                                         ((1 - status) / z2) * h * xi2),
                        "t"       = -sum((status / z2) * (((xi2 ^ 2) * 2 * wg) + 1) -
                                         ((1 - status) / z2) * h * xi2)
)

result      <- c(t(x) %*% Ubeta, Ualpha)

return(result)
}

# est      <- optim(thetaStar, loglik, method = "BFGS", hessian = TRUE)
est        <- optim(thetaStar, loglik, method = "BFGS", hessian = TRUE)
# control = list(fnscale = -1, maxit = 2000, reltol = 1e-12))

if(est$conv != 0)
  warning("FUNCTION DID NOT CONVERGE!")

coef        <- (est$par)[1:k]
alphaest    <- est$par[k+1]
muhat       <- x %*% coef
muhat       <- as.vector(muhat)
etahat      <- exp(muhat)

SHess       = solve(est$hessian)
SE          = sqrt(diag(SHess))
se.coef     = SE
tval        = c(coef,alphaest)/se.coef
matcoef     = cbind(c(coef,alphaest), se.coef, tval, 2*(1-pnorm(abs(tval))))

AIC <- 2 * loglik(c(coef,alphaest)) + 2 * (length(coef) + 1)

BIC = 2 * loglik(c(coef,alphaest)) + (length(coef) + 1) * log(length(y))

result <- list(alphaHat = alphaest,
               betaHat  = coef,
               muHat    = muhat,

```

```

        AIC      = AIC,
        BIC      = BIC,
        matcoef  = matcoef
    )
    return(result)
}

## Teste com dados do artigo
## Rieck, J., and J. Nedelman. 1991. "A Log-Linear Model for the Birnbaum-Saunders
## Distribution." Technometrics 3: 51-60.

fatigue_df <- data.frame(
  work_MJ_m3 = c(
    11.5, 13.0, 14.3, 15.6, 16.0, 17.3, 19.3, 21.1, 21.5, 22.6, 22.6, 24.0,
    24.0, 24.6, 25.2, 25.5, 26.3, 27.9, 28.3, 28.4, 28.6, 30.9, 31.9, 34.5,
    40.1, 40.1, 43.0, 44.1, 46.5, 47.3, 48.7, 52.9, 56.6, 59.9, 60.2,
    60.3, 60.5, 62.1, 62.8, 66.5, 67.0, 67.1, 67.9, 68.8, 75.4, 100.5
  ),
  life_cycles = c(
    3280, 5046, 1563, 4707, 977, 2834, 2266, 2208, 1040, 700, 1583, 482,
    804, 1093, 1125, 884, 1300, 852, 580, 1066, 1114, 386, 745, 736,
    750, 316, 456, 552, 355, 242, 190, 127, 185, 255, 195,
    283, 212, 327, 373, 125, 187, 135, 245, 137, 200, 190
  )
)

# quick check
nrow(fatigue_df)

## [1] 46

head(fatigue_df)

##   work_MJ_m3 life_cycles
## 1      11.5      3280
## 2      13.0      5046
## 3      14.3      1563
## 4      15.6      4707
## 5      16.0       977
## 6      17.3      2834

Zx <- model.matrix(~ log(fatigue_df$work_MJ_m3) )
Zy <- fatigue_df$life_cycles

```

```

fit_logbs1 <- mle_unc (x = Zx , t = Zy , kernel = "normal")

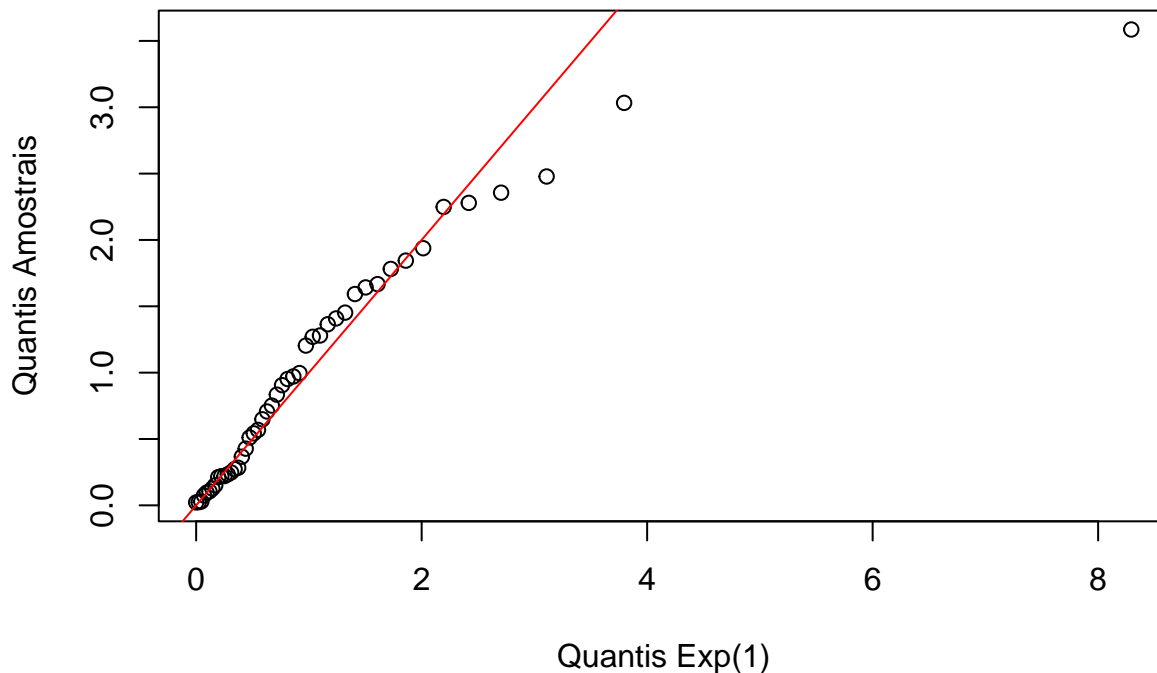
fit_logbs1$matcoef

##                                se.coef      tval
## (Intercept)                   12.2797340 0.3893978 31.535187 0
## log(fatigue_df$work_MJ_m3)    -1.6707690 0.1084439 -15.406766 0
##                               0.4103574 0.0427819  9.591846 0

xi2hat <- (2 / fit_logbs1$alphaHat) * sinh((log(Zy) - fit_logbs1$muHat) / 2)
SurvivalBS_fit_logbs1 <- 1 - pnorm(xi2hat)
CSresidualModel_fit_logbs1 <- -log(SurvivalBS_fit_logbs1)
DSresidualModel_fit_logbs1 <- qnorm(SurvivalBS_fit_logbs1)

## QQ plot
a <- ppoints(2000)
QGG <- qexp(a) #quantis da Exp(1)
qqplot(QGG, CSresidualModel_fit_logbs1, main="",
xlab="Quantis Exp(1)", ylab="Quantis Amostras")
abline(0, 1,col="red")

```

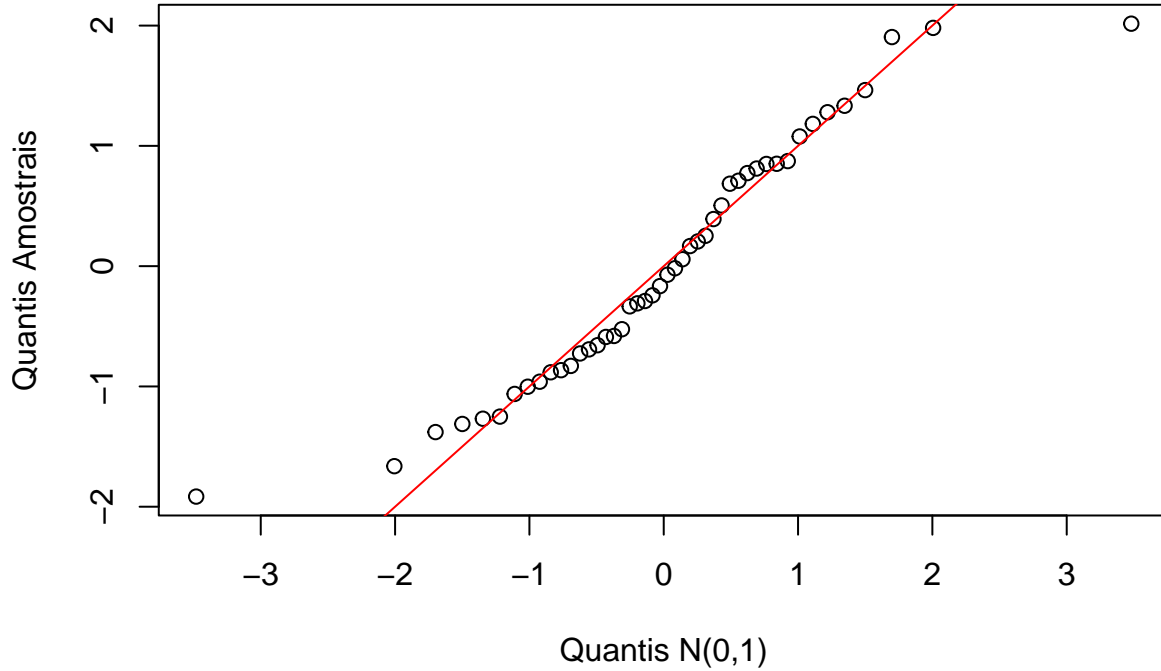


```

a <- ppoints(2000)
QGG <- qnorm(a) #quantis da Exp(1)
qqplot(QGG, DSresidualModel_fit_logbs1, main="",
xlab="Quantis N(0,1)", ylab="Quantis Amostras")
abline(0, 1,col="red")

```





### 1.11 Modelo de regressão Birnbaum-Saunders para dados censurados

- Assim como na Equação (68), considere novamente os modelos de regressão log-linear Birnbaum-Saunders dados por

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\eta} + \varepsilon_i, \quad i = 1, \dots, n, \quad (75)$$

em que  $Y_i$  é a variável resposta logarítmica censurada ou não censurada para o caso  $i$ ,  $\boldsymbol{\eta} = (\eta_0, \eta_1, \dots, \eta_p)^\top$  é um vetor de parâmetros desconhecidos a serem estimados,  $\mathbf{x}_i^\top = (1, x_{i1}, \dots, x_{ip})^\top$  contém os valores das variáveis explicativas, e  $\varepsilon_i \sim \text{log-BS}(\alpha, 0)$  é o termo de erro do modelo, para  $i = 1, \dots, n$ .

- Assume-se censura não-informativa e que as variáveis de resposta logarítmicas censuradas e não censuradas são independentes.
- A função de log-verossimilhança total do modelo definido na Equação (75) para  $\boldsymbol{\theta} = (\alpha, \boldsymbol{\eta}^\top)^\top$  é dada por

$$\ell(\boldsymbol{\theta}) = c_4 + \sum_{i \in D} \ell_i(\boldsymbol{\theta}) + \sum_{i \in C} \ell_i^{(c)}(\boldsymbol{\theta}),$$

em que  $D$  e  $C$  denotam os conjuntos de casos em que  $y_i$  é a resposta logarítmica não censurada e censurada, respectivamente,  $c_4$  é uma constante que não depende de  $\boldsymbol{\theta}$ , e  $\ell_i(\boldsymbol{\theta}) = \log(f_Y(y_i; \boldsymbol{\theta}))$ ,  $\ell_i^{(c)}(\boldsymbol{\theta}) = \log(R_Y(y_i; \boldsymbol{\theta}))$ , com  $f_Y(\cdot)$  e  $R_Y(\cdot)$  sendo as funções densidade de probabilidade e confiabilidade da distribuição log-Birnbaum-Saunders, dadas nas Equações (61) e (64), respectivamente.

- Assim, a função de log-verossimilhança para  $\boldsymbol{\theta}$  é expressa por

$$\ell(\boldsymbol{\theta}) = c_4 + \sum_{i \in D} \left( \log(\xi_{i1}) - \frac{\xi_{i2}^2}{2} \right) + \sum_{i \in C} \log(\Phi(-\xi_{i2})), \quad (76)$$

em que  $\xi_{i1}$ ,  $\xi_{i2}$  e  $\mu_i$ , para  $i = 1, 2, \dots, n$ , são dados na Equação (70).

- Para obter a estimativa de máxima verossimilhança de  $\boldsymbol{\theta}$ , as equações de log-verossimilhança  $\dot{\ell}_\alpha = 0$  e  $\dot{\ell}_{\eta_j}(\boldsymbol{\theta}) = 0$ , para  $j = 0, 1, \dots, p$ , devem ser resolvidas, em que

$$\begin{aligned} \dot{\ell}_\alpha &= \frac{1}{\alpha} \left( \sum_{i \in D} (\xi_{i2}^2 - 1) + \sum_{i \in C} \xi_{i2} h_Y(\xi_{i2}; \boldsymbol{\theta}) \right) \\ \dot{\ell}_{\eta_j} &= \sum_{i \in D} \left( \frac{x_{ij}}{\alpha^2} \sinh(y_i - \mu_i) - \frac{x_{ij}}{2} \tanh\left(\frac{y_i - \mu_i}{2}\right) \right) + \sum_{i \in C} \frac{x_{ij} \xi_{i1} h_Y(\xi_{i2}; \boldsymbol{\theta})}{2}, \end{aligned}$$

com  $h_Y(\cdot)$  sendo a taxa de falha da distribuição log-Birnbaum-Saunders dada na Equação (65).

- Como novamente as equações de log-verossimilhança não apresentam soluções analíticas, pode-se utilizar o método BFGS e as estimativas de máxima verossimilhança de  $\boldsymbol{\theta}$  no caso não censurado como valores iniciais; ver Leiva et al. (2007).
- Assim como no caso não censurado, a inferência assintótica para o vetor de parâmetros  $\boldsymbol{\theta}$  pode ser utilizada. A matriz assintótica de variância-covariância de  $\hat{\boldsymbol{\theta}}$  pode ser aproximada por  $-\ddot{\ell}^{-1}$ , obtida neste caso como

$$\ddot{\ell} = \begin{pmatrix} \ddot{\ell}_{\alpha\alpha} & \ddot{\ell}_{\alpha\eta} \\ \ddot{\ell}_{\eta\alpha} & \ddot{\ell}_{\eta\eta} \end{pmatrix} = \begin{pmatrix} \text{tr}(\mathbf{G}) & \mathbf{K}^\top \mathbf{X} \\ \mathbf{X}^\top \mathbf{K} & \mathbf{X}^\top \mathbf{V} \mathbf{X} \end{pmatrix},$$

em que  $\mathbf{V} = \text{diag}\{v_1(\boldsymbol{\theta}), v_2(\boldsymbol{\theta}), \dots, v_n(\boldsymbol{\theta})\}$ ,  $\mathbf{K} = (k_1(\boldsymbol{\theta}), \dots, k_n(\boldsymbol{\theta}))^\top$ , e  $\mathbf{G} = \text{diag}\{g_1(\boldsymbol{\theta}), \dots, g_n(\boldsymbol{\theta})\}$ , com

$$v_i(\boldsymbol{\theta}) = \begin{cases} \frac{1}{4} \left( \text{sech}\left(\frac{y_i - \mu_i}{2}\right) \right)^2 - \frac{1}{\alpha^2} \cosh(y_i - \mu_i), & i \in D; \\ -\frac{1}{4} \xi_{i2} h_Y(\xi_{i2}; \boldsymbol{\theta}) - \frac{1}{4} \xi_{i1}^2 h'_Y(\xi_{i2}; \boldsymbol{\theta}), & i \in C; \end{cases} \quad (77)$$

$$k_i(\boldsymbol{\theta}) = \begin{cases} -\frac{2}{\alpha^3} \sinh(y_i - \mu_i), & i \in D; \\ -\frac{1}{2\alpha} \xi_{i1} h_Y(\xi_{i2}; \boldsymbol{\theta}) - \frac{1}{\alpha^3} \sinh(y_i - \mu_i) h'_Y(\xi_{i2}; \boldsymbol{\theta}), & i \in C; \end{cases} \quad (78)$$

$$g_i(\boldsymbol{\theta}) = \begin{cases} \frac{1}{\alpha^2} - \frac{3\xi_{i2}^2}{\alpha^2}, & i \in D; \\ -\frac{2}{\alpha^2} \xi_{i2} h_Y(\xi_{i2}; \boldsymbol{\theta}) - \frac{1}{\alpha^2} \xi_{i2}^2 h'_Y(\xi_{i2}; \boldsymbol{\theta}), & i \in C; \end{cases} \quad (79)$$

em que  $h'_Y(\cdot)$  é a derivada da taxa de falha  $h_Y(\cdot)$ , e  $\xi_{i1}$  e  $\xi_{i2}$  são dados na Equação (76).

## 1.12 Análise de resíduos

- Para estudar desvios das suposições de erro e a presença de observações atípicas, consideramos dois tipos de resíduos: componente de deviance e do tipo martingale. O

leitor interessado pode consultar, por exemplo, McCullagh and Nelder (1989) para o resíduo do componente de deviance, Barlow and Prentice (1988) e Therneau, Grambsch, and Fleming (1990) para o resíduo do tipo martingale, e ambos em Ortega, Bolfarine, and Paula (2003).

- Em geral, o resíduo do componente de deviance é definido como

$$r_{\text{DC}_i} = \text{sign}(y_i - \hat{\mu}_i) \sqrt{2(\ell_i(\hat{\boldsymbol{\theta}}_s) - \ell_i(\hat{\boldsymbol{\theta}}))}, \quad i = 1, 2, \dots, n, \quad (80)$$

em que  $\hat{\boldsymbol{\theta}}_s$  é a estimativa de máxima verossimilhança de  $\boldsymbol{\theta}$  sob o modelo saturado (com  $n$  parâmetros),  $\hat{\boldsymbol{\theta}}$  é a estimativa de máxima verossimilhança de  $\boldsymbol{\theta}$  sob o modelo de interesse (com  $p + 2$  parâmetros),  $\hat{\mu}_i = E(Y_i)$ , e  $\text{sign}(z)$  denota o sinal de  $z$ .

- Davison and Gigli (1989) definiram a partir de (80) o resíduo do componente de deviance para dados censurados como

$$r_{\text{DC}_i} = \text{sign}(y_i - \hat{\mu}_i) \sqrt{-2 \log(\hat{R}_Y(y_i; \boldsymbol{\theta}))}, \quad i = 1, 2, \dots, n, \quad (81)$$

em que  $\hat{R}_Y(\cdot)$  é a estimativa de máxima verossimilhança da função de confiabilidade correspondente; ver, por exemplo, Ortega (2001).

- Therneau, Grambsch, and Fleming (1990) introduziram o resíduo do componente de deviance em processos de contagem usando basicamente resíduos martingale, que são enviesados e têm valor máximo em  $+1$  e mínimo em  $-\infty$ . Para dados censurados, o resíduo martingale pode ser expresso por

$$r_{\text{M}_i} = \delta_i + \log(\hat{R}_Y(y_i; \boldsymbol{\theta})), \quad i = 1, 2, \dots, n, \quad (82)$$

em que  $\delta_i = 0$  indica se a observação é censurada,  $\delta_i = 1$  se a observação não é censurada e, novamente,  $\hat{R}(\cdot)$  é a estimativa de máxima verossimilhança da função de confiabilidade correspondente com dados censurados; ver Klein and Moeschberger (1997) e Ortega, Bolfarine, and Paula (2003).

- O resíduo do componente de deviance proposto por Therneau, Grambsch, and Fleming (1990) é uma transformação do resíduo martingale para atenuar a assimetria. Essa transformação foi motivada pelos resíduos do componente de deviance encontrados em modelos lineares generalizados. Em particular, o resíduo do componente de deviance para o modelo de riscos proporcionais de Cox sem variáveis explicativas dependentes do tempo é chamado de resíduo do tipo martingale e é descrito como

$$r_{\text{MT}_i} = \text{sign}(r_{\text{M}_i}) \sqrt{-2(r_{\text{M}_i} + \delta_i \log(\delta_i - r_{\text{M}_i}))}, \quad i = 1, 2, \dots, n, \quad (83)$$

em que  $r_{M_i}$  é o resíduo martingale definido em (82). Embora o resíduo do tipo martingale expresso em (83) não seja um componente de deviance dos modelos de regressão log-Birnbaum-Saunders, usamos essa transformação do resíduo martingale para obter resíduos distribuídos simetricamente em torno de zero.

- Ortega (2001) sugeriram padronizar os resíduos do componente de deviance e do tipo martingale para dados censurados como

$$r_{DC_i}^* = \frac{r_{DC_i}}{\sqrt{1 - GL_{ii}}}, \quad r_{MT_i}^* = \frac{r_{MT_i}}{\sqrt{1 - GL_{ii}}}, \quad i = 1, 2, \dots, n,$$

respectivamente, com  $GL_{ii}$  sendo o  $i$ -ésimo elemento da diagonal principal da matriz de leverage generalizada, avaliada em  $\hat{\theta}$ .

- Considerando  $\alpha < 2$  fixo ou conhecido, o resíduo do componente de deviance para os modelos de regressão Birnbaum-Saunders é dado por

$$r_{DC_i} = \begin{cases} \text{sign}(y_i - \hat{\mu}_i) \sqrt{-2 \log \left( \cosh \left( \frac{y_i - \hat{\mu}_i}{2} \right) + \frac{2}{\alpha^2} \left( \sinh \left( \frac{y_i - \hat{\mu}_i}{2} \right) \right)^2 \right)}, & i \in D; \\ \text{sign}(y_i - \hat{\mu}_i) \sqrt{-2 \log \left( \Phi \left( -\frac{2}{\alpha} \sinh \left( \frac{y_i - \hat{\mu}_i}{2} \right) \right) \right)}, & i \in C. \end{cases}$$

- Além disso, o resíduo martingale para os modelos de regressão Birnbaum-Saunders assume a forma

$$r_{M_i} = \delta_i + \log \left( \Phi \left( -\frac{2}{\alpha} \sinh \left( \frac{y_i - \hat{\mu}_i}{2} \right) \right) \right), \quad i = 1, 2, \dots, n.$$

## Referências

- Balakrishnan, N., V. Leiva, A. Sanhueza, and E. Cabrera. 2009. “Mixture Inverse Gaussian Distribution and Its Transformations, Moments and Applications.” *Statistics* 43: 91–104.
- Barlow, W., and R. Prentice. 1988. “Residuals for Relative Risk Regression.” *Biometrika* 75: 65–74.
- Bhattacharyya, G., and A. Fries. 1982. “Fatigue Failure Models: Birnbaum-Saunders Versus Inverse Gaussian.” *IEEE Transactions on Device and Materials Reliability* 31: 439–40.
- Birnbaum, Z., J. Esary, and A. Marshall. 1966. “Stochastic Characterization of Wear-Out for Components and Systems.” *The Annals of Mathematical Statistics* 37: 816–25.
- Birnbaum, Z., and S. Saunders. 1958. “A Statistical Model for Life-Length of Materials.” *Journal of the American Statistical Association* 53: 151–60.
- . 1968. “A Probabilistic Interpretation of Miner’s Rule.” *Journal of Applied Mechanics* 16: 637–52.
- . 1969a. “A New Family of Life Distributions.” *Journal of Applied Probability* 6: 319–27.
- . 1969b. “Estimation for a Family of Life Distributions with Applications to Fatigue.” *Journal of Applied Probability* 6: 328–47.
- Chang, D., and L. Tang. 1994. “Percentile Bounds and Tolerance Limits of the Birnbaum-Saunders Distribution.” *Communications in Statistics: Theory and Methods* 23: 2853–63.
- Chhikara, R., and J. Folks. 1989. *The Inverse Gaussian Distribution – Theory, Methodology, and Applications*. New York, US: Marcel Dekker.
- Cramér, H. 1946. *Mathematical Methods of Statistics*. Princeton, US: Princeton University Press.
- Davison, A., and A. Gigli. 1989. “Deviance Residuals and Normal Scores Plots.” *Biometrika* 76: 211–21.
- Desmond, A. 1985. “Stochastic Models of Failure in Random Environments.” *Canadian Journal of Statistics* 13: 171–83.
- . 1986. “On the Relationship Between Two Fatigue Life Models.” *IEEE Transactions on Reliability* 35: 167–69.
- Esary, J., A. W. Marshall, and F. Proschan. 1973. “Shock Models and Wear Processes.” *The Annals of Probability* 1: 627–49.
- Fréchet, M. 1927. “Sur La Loi de Probabilité de l’ecart Maximum.” *Annales de La Société Polonaise de Mathématique* 6: 93.
- Freudenthal, A., and M. Shinozuka. 1961. “Structural Safety Under Conditions of Ultimate Load Failure and Fatigue.” WADD-TR-61-177. Wright Air Force Base, Ohio, US: Wright Air Development Division.
- Johnson, N., S. Kotz, and N. Balakrishnan. 1995. *Continuous Univariate Distributions*. Vol. 2. New York, US: Wiley.
- Kao, J. 1959. “A Graphical Estimation of Mixed Weibull Parameters in Life-Testing of Electron Tubes.” *Technometrics* 1: 389–407.
- Klein, J., and M. Moeschberger. 1997. *Survival Analysis: Techniques for Censored and Truncated Data*. New York, US: Springer.
- Kundu, D., N. Kannan, and N. Balakrishnan. 2008. “On the Hazard Function of Birnbaum-Saunders Distribution and Associated Inference.” *Computational Statistics and Data*

- Analysis* 52: 2692–702.
- Lange, K. 2001. *Numerical Analysis for Statisticians*. New York, US: Springer.
- Leiva, V., M. Barros, G. Paula, and M. Galea. 2007. “Influence Diagnostics in Log-Birnbaum-Saunders Regression Models with Censored Data.” *Computational Statistics and Data Analysis* 51: 5694–5707.
- Leiva, V., M. Santos-Neto, F. Cysneiros, and M. Barros. 2014. “Birnbbaum-Saunders Statistical Modelling: A New Approach.” *Statistical Modelling* 14: 21–48.
- Lieblein, J., and M. Zelen. 1956. “Statistical Investigation of the Fatigue Life of Deep Ball Bearing.” *Journal of Research of the National Bureau of Standards* 57: 273–316.
- Mann, N., R. Schafer, and N. Singpurwalla. 1974. *Methods for Statistical Analysis of Reliability and Life Data*. New York, US: Wiley.
- Marshall, A., and I. Olkin. 2007. *Life Distributions*. New York, US: Springer.
- Mátyás, L. 1999. *Generalized Method of Moments Estimation*. New York, US: Cambridge University Press.
- McCullagh, P., and J. Nelder. 1989. *Generalized Linear Models*. London, UK: Chapman; Hall.
- Michael, J., W. Schucany, and R. Haas. 1976. “Generating Random Variates Using Transformations with Multiple Roots.” *The American Statistician* 30: 88–90.
- Miner, M. 1945. “Cumulative Damage in Fatigue.” *Journal of Applied Mechanics* 12: A159–64.
- Ng, H., D. Kundu, and N. Balakrishnan. 2003. “Modified Moment Estimation for the Two-Parameter Birnbbaum-Saunders Distribution.” *Computational Statistics and Data Analysis* 43: 283–98.
- Nocedal, J., and S. Wright. 1999. *Numerical Optimization*. New York, US: Springer.
- Ortega, E. 2001. “Influence Analysis in Generalized Log-Gamma Regression Models.” PhD thesis, Brazil: University of São Paulo.
- Ortega, E., H. Bolfarine, and G. Paula. 2003. “Influence Diagnostics in Generalized Log-Gamma Regression Models.” *Computational Statistics and Data Analysis* 42: 165–86.
- Press, W., S. Teukolsky, W. Vetterling, and B. Flannery. 1992. *Numerical Recipes in c: The Art of Scientific Computing*. London, UK: Prentice-Hall.
- R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rieck, J. 1989. “Statistical Analysis for the Birnbbaum-Saunders Fatigue Life Distribution.” PhD thesis, Clemson, US: Clemson University.
- . 1999. “A Moment-Generating Function with Application to the Birnbbaum-Saunders Distribution.” *Communications in Statistics: Theory and Methods* 28: 2213–22.
- . 2003. “A Comparison of Two Random Number Generators for the Birnbbaum-Saunders Distribution.” *Communications in Statistics: Theory and Methods* 32: 929–34.
- Rieck, J., and J. Nedelman. 1991. “A Log-Linear Model for the Birnbbaum-Saunders Distribution.” *Technometrics* 3: 51–60.
- Santos-Neto, M., F. Cysneiros, V. Leiva, and S. Ahmed. 2012. “On New Parameterizations of the Birnbbaum-Saunders Distribution.” *Pakistan Journal of Statistics* 28: 1–26.
- Santos-Neto, M., F. Cysneiros, V. Leiva, and M. Barros. 2014. “On New Parameterizations of the Birnbbaum-Saunders Distribution and Its Moments, Estimation and Application.” *Revstat Statistical Journal* 12: 247–72.
- Saunders, S. 1976. “The Problems of Estimating a Fatigue Service Life with a Low Probability

- of Failure.” *Engineering Fracture Mechanics* 8: 205–15.
- . 2007. *Reliability, Life Testing and Prediction of Services Lives*. New York, US: Springer.
- Saunders, S. C. 1974. “A Family of Random Variables Closed Under Reciprocation.” *Journal of the American Statistical Association* 69: 533–39.
- Seshadri, V. 1999a. *The Inverse Gaussian Distribution: A Case Study in Exponential Families*. New York, US: Claredon.
- . 1999b. *The Inverse Gaussian Distribution: Statistical Theory and Applications*. New York, US: Springer.
- Therneau, T., P. Grambsch, and T. Fleming. 1990. “Martingale-Based Residuals for Survival Models.” *Biometrika* 77: 147–60.
- Weibull, W. 1951. “A Statistical Distribution Function of Wide Applicability.” *Journal of Applied Mechanics* 1: 293–97.