



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Vitor Macedo Rocha

**Gráficos de Controle Bootstrap para
percentis skew Birnbaum-Saunders**

Brasília - DF, 2021

Vitor Macêdo Rocha

Gráficos de Controle Bootstrap para percentis skew Birnbaum-Saunders

Monografia apresentada ao Departamento de Estatística da Universidade de Brasília como requisito para a obtenção do grau de Bacharel em Estatística.

Orientador: Prof. **Dr. Helton Saulo Bezerra dos Santos**

Brasília - DF, 2021

Agradecimentos

Aos meus amados pais por todo apoio, amor, carinho e inspiração. Ao meu ilustre irmão que sempre esteve ao meu lado prestando apoio de seu jeito impar, por todo o companheirismo, carinho, lembranças memoráveis e noites em claro jogando.

Aos professores Dr. Guilherme Souza Rodrigues, Dr. Eduardo Yoshio Nakano, Dr. Juliana Betini Fachini Gomes e Dr. Roberto Vila Gabriel por todo o conhecimento e oportunidades de aprendizado que me proporcionaram.

Ao professor Dr. Helton Saulo Bezerra dos Santos por servir de inspiração, pela troca de conhecimento e toda ajuda prestada ao longo da minha graduação que proporcionaram ferramentas para minha formação tanto acadêmica quanto pessoal.

“O bootstrap raramente é a estrela da estatística, mas é o melhor ator coadjuvante.”

-Bradley Efron

Conteúdo

| | | |
|----------|---|-----------|
| 1 | Introdução | 9 |
| 2 | Revisão de literatura | 10 |
| 2.1 | Gráfico de controle | 10 |
| 2.1.1 | Comprimento médio da sequência | 11 |
| 2.2 | Estimador de máxima verossimilhança e suas propriedades | 11 |
| 2.3 | Bootstrap paramétrico | 12 |
| 2.3.1 | Intervalo de confiança bootstrap percentil | 13 |
| 2.3.2 | Intervalo de confiança bootstrap BCa | 14 |
| 2.4 | Modelo Birnbaum-Saunders | 16 |
| 3 | Modelo skew Birnbaum-Saunders | 16 |
| 3.1 | Estimação via Algoritmo ECM | 18 |
| 4 | Metodologia | 20 |
| 4.1 | Materiais | 20 |
| 4.2 | Métodos | 20 |
| 5 | Estudo de simulação | 21 |
| 6 | Exemplo ilustrativo | 26 |
| 7 | Conclusões | 28 |
| | Referências | 29 |
| A | Densidade SNBS | 31 |
| B | Código R | 31 |

Resumo

O problema de detectar uma mudança no percentil de uma população skew Birnbaum-Saunders em uma situação de monitoramento de processo é considerado. A técnica bootstrap paramétrico é empregada na construção dos limites de controle no monitoramento de percentis quando o processo segue o modelo skew BS, que é útil para descrever o comportamento de dados de tempo de vida com uma assimetria que não é bem capitada pelos modelos convencionais como Birnbaum-Saunders. Esse trabalho generaliza a metodologia de Lio e Park (2008). A performance dos gráficos de controle para percentis propostos foram investigados via simulação Monte Carlo. Um exemplo ilustrativo é apresentado.

Palavras-chave: Cartas de controle, Bootstrap, percentis skew Birnbaum-Saunders, algoritmo ECM, simulação Monte Carlo.

Abstract

The problem of detecting a shift in the percentiles of a skewed Birnbaum-Saunders population in a process monitoring situation is considered. The parametric bootstrap method is used in the construction of control limits in monitoring percentiles when the process follows the BS skew model, useful to describe the behavior of lifetime data with an asymmetry that is not well accepted by usual models like Birnbaum-Saunders. This work generalizes the methodology of Lio e Park (2008). The performance of the control charts for proposed percentiles was investigated through Monte Carlo simulation. An illustrative example is presented.

Keywords: Control charts, Bootstrap, skew Birnbaum-Saunders percentiles, ECM algorithm, Monte Carlo simulation.

1 Introdução

Na literatura, existe uma enorme variedade de métodos para o monitoramento de processos, com o intuito de manter somente a variabilidade inerente ao processo, ver Montgomery (2000). Para tal, é importante compreender a característica de interesse à monitorar e encontrar o melhor método para realizar. Em um cenário geral de engenharia, faz-se o uso frequente de componentes que são submetidos a repetitivos ciclos de choque e usualmente são demandados uma resistência mínima desses componentes. Por esse motivo, é de grande interesse do engenheiro o monitoramento de percentis inferiores da resistência desses componentes, como visto em Lio e Park (2008) e Nichols e Padgett (2006) nos quais são feitos estudos de monitoramento de percentis provindos das distribuições Birnbaum-Saunders e Weibull, respectivamente.

Na construção de cartas de controle para o monitoramento de percentis extremos é empregada a técnica bootstrap paramétrica, discutido em Efron (1979), que possibilita a distribuição amostral de um estimador através da geração de amostras parametricamente. Essa abordagem é bastante robusta pois não há restrições em relação a distribuição dos dados e proporciona um ótimo amparo para a construção de cartas de controle apropriados para percentis, como visto em Lio e Park (2008) e Nichols e Padgett (2006) que empregaram essa metodologia para construção de gráficos de controle para percentis provindos da distribuição Birnbaum-Saunders e Weibull respectivamente. O uso de gráfico de controle do tipo Shewhart não são convenientes para processos que são modelados por distribuições assimétricas e com suporte positivo, uma vez que ele assume a normalidade do processo e com isso tais gráficos podem cometer falhas em detectar importantes mudanças em percentis inferiores como ressaltam Padgett e Spurrier (1990).

Nesse trabalho o modelo probabilístico de interesse é o skew Birnbaum-Saunders, proposto em Vilca et al. (2011), esse modelo é capaz de captar uma assimetria de maneira mais eficaz que o modelo usual Birnbaum-Saunders, o que vem a ser bastante útil no caso de dados com uma acentuada assimetria positiva. Esse modelo também é estudado em Santana, Vilca e Leiva (2011), no qual propõe análise de diagnóstico para modelos de regressão SNBS e é motivado pela detecção de influência na estimação de percentis extremos inferiores.

O principal objetivo no presente trabalho é propor um gráfico de controle para dados de tempo de falha provindos do modelo skew Birnbaum-Saunders através do uso da técnica bootstrap. Será feito a comparação dos gráficos de controle utilizando diferentes métodos de intervalo bootstrap para a obtenção dos limites, sendo eles o intervalo percentil e BCa, além de avaliar o desempenho dos gráficos de controle proposta através de uma extensa simulação de Monte Carlo.

2 Revisão de literatura

Na presente seção será aludido conceitos e definições com o objetivo de facilitar a compreensão deste trabalho. Primeiramente, será discutido o contexto de gráficos de controle e o método de máxima verossimilhança, posteriormente será tratado a uma aplicação do bootstrap e um modelo de probabilidade subjacente ao principal que será trabalhado. A principal finalidade dessa seção é introduzir o modelo BS e uma técnica bastante útil na construção de gráficos de controle.

2.1 Gráfico de controle

No cenário de controle estatísticos de qualidades podemos classificar a variação do processo em causas aleatórias, as quais são inerentes ao processo, e as causa atribuíveis que são identificáveis, de maior impacto, e devem ser eliminadas. Um processo é dito sob controle quando a única variabilidade é devido a causas aleatórias, caso contrário o processo é denominado fora de controle. Segundo (MONTGOMERY, 2000) os principais objetivos do gráfico de controle são estimar algum parâmetro de interesse do processo, monitorar esse parâmetro, sinalizar a presença de causa de variações atribuíveis e reduzir a variabilidade eliminando essas causas.

Um gráfico de controle representa os resultados de alguma característica de interesse versus a ordem da coleta ou tempo, aditivamente são plotados 3 linhas, a linha central, limite superior de controle e o limite inferior de controle. Duas etapas são consideradas no CEQ, a fase 1 e fase 2 que possuem diferentes intuítos principais, os quais são melhor descritos a seguir.

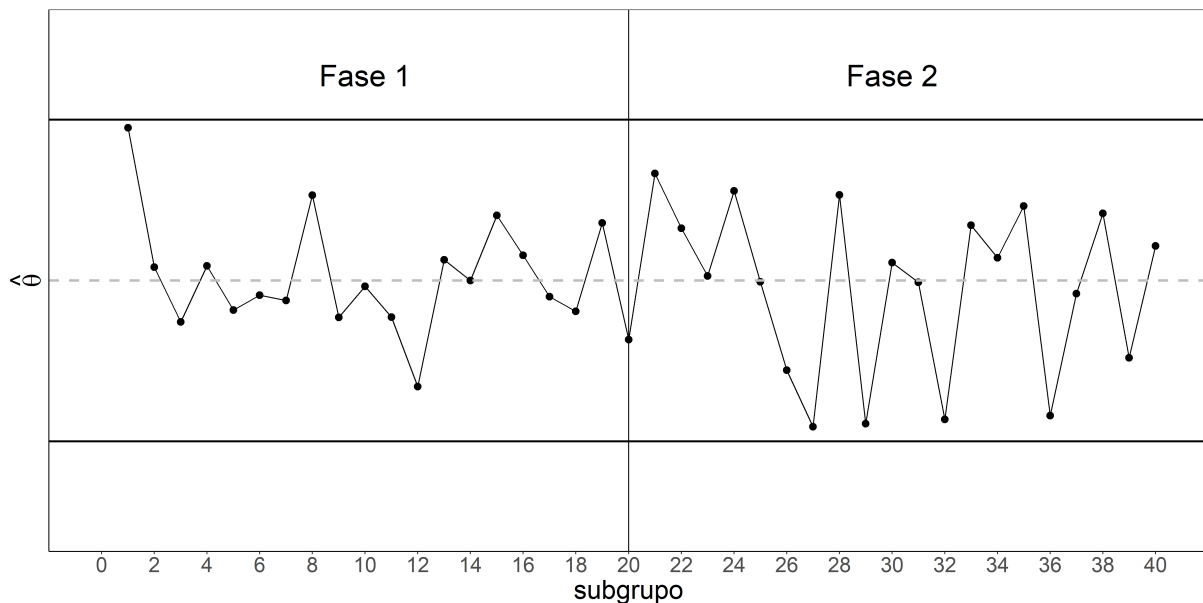


Figura 1: Exemplo de um gráfico de controle para um parâmetro genérico θ .

Fase I: Um conjunto de dados do processo é coletado, construindo os limites de controle de teste para determinar se o processo estava sob controle no decorrer da coleta, o obje-

tivo principal dessa etapa é obter limites de controle confiáveis para ser estabelecidos no monitoramento de produção futura.

Fase II: Nessa fase o objetivo é simplesmente o monitoramento, visto que do fase I o objetivo foi garantir que os limites provinham de um processo sob "controle", então o monitoramento é realizado comparando o valor da estatística amostral para cada amostra que é obtida do processo, com os limites de controle.

2.1.1 Comprimento médio da sequência

Um fator importante do uso de carta de controle é a avaliação do desempenho do mesmo, segundo Montgomery (2000) o comprimento médio da sequência (CMS) é uma base válida para a avaliação da fase II. O comprimento da sequência é o número de pontos até que um ponto seja computado fora dos limites de controle e segue uma distribuição geométrica de parâmetro $p = P(\text{Ponto fora dos limites de controle})$. Essencialmente o CMS é o número de amostras não correlacionadas em média que devem ser coletadas antes que o gráfico indique uma situação fora de controle, logo

$$\text{CMS} = \frac{1}{p} \quad (1)$$

A probabilidade de um ponto ser observado fora dos limites de controle pode variar em duas situações, sob e fora de controle. Desse modo o gráfico de controle pode ser estudando quanto ao desempenho do processo sob ou fora de controle.

2.2 Estimador de máxima verossimilhança e suas propriedades

O método de estimação por máxima verossimilhança, segundo Mood (1950), procura encontrar de qual densidade os valores observados são os mais prováveis de ocorrerem. Seja uma amostra aleatória $\mathbf{X} = [X_1, X_2, \dots, X_n]^T$ independente e identicamente distribuída com função de densidade $f(x|\boldsymbol{\theta})$, com $\boldsymbol{\theta} = [\theta_1, \dots, \theta_k]^T \in \Theta$. A função de verossimilhança é definida por

$$L(\boldsymbol{\theta}|\mathbf{X}) = \prod_{i=1}^n f(x_i|\boldsymbol{\theta}).$$

A partir da definição dessa função define-se formalmente por CASELLA e BERGER (2010) o estimador de máxima verossimilhança, como sendo

Definição 1. *A estimativa de máxima verossimilhança do vetor de parâmetros $\boldsymbol{\theta}$ é o vetor de valores $\hat{\boldsymbol{\theta}} \in \Theta$, que maximiza a função de verossimilhança $L(\boldsymbol{\theta}|\mathbf{X})$.*

Na prática é muito mais simples maximizar a função de log-verossimilhança $l(\boldsymbol{\theta}|\mathbf{X}) = \ln(L(\boldsymbol{\theta}|\mathbf{X}))$, visto que como é uma transformação bijetora monótona, isto é, de correspondência

biunívoca, o valor que maximiza a $l(\theta|\mathbf{X})$ é o mesmo valor que maximiza a função de verossimilhança $L(\theta|\mathbf{X})$. Porém nem sempre é possível obter o máximo de forma analítica, logo, recorreremos a métodos numéricos de maximização. Os estimadores de máxima verossimilhança possuem ótimas propriedades, ver Bolfarine e Sandoval (2001) que traz ótimas exemplificações e formalizações das propriedades a seguir.

Propriedades dos estimadores de máxima verossimilhança

P1. Seja $T = T(X_1, \dots, X_n)$ uma estatística suficiente para θ , então o EMV $\hat{\theta}$ (se existir) é uma função de T .

P2. Consistência: Se temos um número suficientemente grande de observações na nossa amostra, nosso estimador de máxima verossimilhança é consistente, isto é

$$\hat{\theta}_{emv} \xrightarrow{P} \theta$$

P3. Invariância: Se $\hat{\theta}$ é um estimador de máxima verossimilhança para θ , então um estimador de máxima verossimilhança para $g(\theta)$ é dado por $g(\hat{\theta})$

P4. Distribuição assintótica: Sob condições de regularidade temos que um vetor de $\hat{\theta}_{emv}$ que estima o vetor de p parâmetros θ

$$\hat{\theta}_{emv} \sim N_p(\theta, [I(\theta)]^{-1})$$

Onde $I(\theta)$ é a matriz de informação de Fisher esperada, que é definido como

$$I_{ij}(\theta) = -E \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln(L(\theta|\mathbf{X})) \right]$$

Há casos em que calcular essa expectância é bastante árduo, nessas circunstâncias usualmente é utilizado a informação de Fisher observada, definida como

$$I_{ij}(\theta) = - \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln(L(\theta|\mathbf{X})) \right]$$

É possível notar que os EMV são assintoticamente os melhores estimadores lineares não viciados, visto que a variância assintótica coincide com o limite inferior de Cramér-Rao.

2.3 Bootstrap paramétrico

A técnica de bootstrap proposta em Efron (1979) é bastante utilizada para obter boas estimativas. Existem dois diferentes tipos de bootstrap, o paramétrico e não paramétrico, ver Efron e Tibshirani (1994). Diferente do bootstrap não paramétrico que as amostras bootstrap

são geradas através da re-amostragem de uma amostra principal, o bootstrap paramétrico assume que a amostra original provem de uma distribuição conhecida $F_X(x|\boldsymbol{\theta})$ e após obter uma estimativa $\hat{\boldsymbol{\theta}}$ são gerados B amostras bootstrap $\mathbf{X}^{*1}, \mathbf{X}^{*2}, \dots, \mathbf{X}^{*B}$, e a partir dessas amostras bootstrap serão obtidos estimativas para diversos parâmetros de interesse.

O bootstrap possui diversas aplicações, dentre elas obter estimativas para parâmetros, construção de intervalos de confiança aproximada e correção de viés, para mais informações teóricas ver DiCiccio e Efron (1996) e Efron (1982). Vários algoritmos utilizando essa técnica são apresentados em Rizzo (2019). A seguir serão tratados dois diferentes tipos de intervalo de confiança bootstrap e suas propriedades.

2.3.1 Intervalo de confiança bootstrap percentil

O intervalo de confiança bootstrap percentil, em Rizzo (2019), é definida da seguinte maneira, suponha que geramos um número B finito de amostras bootstrap $\mathbf{X}^{*1}, \mathbf{X}^{*2}, \dots, \mathbf{X}^{*B}$ e computamos as réplicas bootstrap $\hat{\theta}^*(b) = S(\mathbf{X}^{*b})$, $b = 1, 2, \dots, B$. Seja \hat{G} a função de densidade acumulada empírica de θ . Podemos então definir o intervalo de confiança percentil aproximado para θ com $(1 - \alpha) \cdot 100\%$ de confiança como sendo.

$$\left[\hat{\theta}_{inf}; \hat{\theta}_{sup} \right] = \left[\hat{G}^{-1}(\alpha/2); \hat{G}^{-1}(1 - \alpha/2) \right] \quad (2)$$

Em Efron e Tibshirani (1994) tem-se uma breve discussão de como o intervalo de confiança bootstrap percentil incorpora transformações que deveriam ser achadas de modo que o intervalo de confiança normal fosse acurado. O lema a seguir formaliza o fato dessa incorporação inerente do intervalo percentil de transformações normalizadoras.

Lema 2.1 (lema do intervalo percentil). *Suponha a transformação $\hat{\phi} = m(\hat{\theta})$ que normaliza perfeitamente a distribuição de $\hat{\theta}$:*

$$\hat{\phi} \sim N(\phi, c^2)$$

para algum desvio padrão c . Então o intervalo de confiança de percentil baseado direto em $\hat{\theta}$ é dado por $\left[m^{-1} \left(\hat{\phi} - z_{1-\alpha/2}c \right); m^{-1} \left(\hat{\phi} - z_{\alpha/2}c \right) \right]$.

Propriedades do intervalo percentil

P1. O intervalo percentil para qualquer transformação de parâmetro monótona $\phi = m(\theta)$ é simplesmente o intervalo percentil para θ mapeado por $m(\theta)$:

$$\left[\hat{\phi}_{inf}; \hat{\phi}_{sup} \right] = \left[m \left(\hat{\theta}_{inf} \right); m \left(\hat{\theta}_{sup} \right) \right]$$

P2. Suponha $\theta \in [a, b]$ seja o parâmetro de interesse, então o procedimento irá sempre gerar

intervalos que irão cair dentro desse intervalo , isto é

$$a \leq \hat{\theta}_{inf} < \hat{\theta}_{sup} \leq b$$

P3. O intervalo percentil é de acurácia de primeira ordem, isto é , o erro de aproximação tende a zero em velocidade $1/\sqrt{n}$.

2.3.2 Intervalo de confiança bootstrap BCa

O Intervalo de confiança bootstrap *bias-corrected and accelerated* (BCa) é uma versão modificada do intervalo percentil que possui melhores propriedades teóricas e performance prática, ver Rizzo (2019). Para um intervalo de confiança com $\alpha \cdot 100\%$ de significância, os quantis $\alpha/2$ e $1 - \alpha/2$ são ajustados por dois fatores, são eles correção de viés e de assimetria. Seja z_0 a correção do viés, a o ajuste de assimetria ou "aceleração". Fazendo o ajuste

$$\alpha_1 = \Phi \left(\hat{z}_0 + \frac{\hat{z}_0 + z_{\alpha/2}}{1 - \hat{a} (\hat{z}_0 + z_{\alpha/2})} \right), \quad (3)$$

$$\alpha_2 = \Phi \left(\hat{z}_0 + \frac{\hat{z}_0 + z_{1-\alpha/2}}{1 - \hat{a} (\hat{z}_0 + z_{1-\alpha/2})} \right), \quad (4)$$

Onde $\Phi(\cdot)$ é a função de densidade acumulada normal padrão, $z_\alpha = \Phi^{-1}(\alpha)$ e \hat{G} a função de densidade acumulada empírica de θ . Podemos então definir o intervalo de confiança BCa aproximado para θ com $(1 - \alpha) \cdot 100\%$ de confiança como sendo.

$$\left[\hat{\theta}_{inf}; \hat{\theta}_{sup} \right] = \left[\hat{G}^{-1}(\alpha_1); \hat{G}^{-1}(\alpha_2) \right] \quad (5)$$

Então o intervalo BCa é o mesmo do percentil exceto para valores não nulos de \hat{z}_0 e \hat{a} , que alteram os percentis usados para os limites no intervalo BCa. A correção de viés z_0 e obtido a partir da proporção de réplicas bootstrap menores que a estimativa da amostra observada $\hat{\theta}$

$$\hat{z}_0 = \Phi^{-1} \left(\frac{1}{B} \sum_{b=1}^B I_{(\hat{\theta}^{(b)} < \hat{\theta})} \right) \quad (6)$$

Já a estimação do fator de assimetria (ou aceleração) é realizada a partir das replicas jackknife de $\hat{\theta} = s(\mathbf{X})$. Seja $\mathbf{X}_{(i)}$ a amostra original com a i -ésima observação excluída, $\hat{\theta}_{(i)} = s(\mathbf{X}_{(i)})$, e defina $\hat{\theta}_{(\cdot)} = \sum_{i=1}^n \hat{\theta}_{(i)}/n$, então definimos a estimativa do fator de aceleração como sendo

$$\hat{a} = \frac{\sum_{i=1}^n (\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^3}{6 \sum_{i=1}^n \left\{ (\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^2 \right\}^{3/2}} \quad (7)$$

O fator \hat{a} é chamado de aceleração pois estima a razão de mudança do erro padrão de $\hat{\theta}$ em relação ao real valor do parâmetro θ . Apesar de parecer estranho essa adaptação do intervalo percentil DiCiccio e Efron (1996) e Efron (1987) trazem melhores explicações do intervalo BCa. Como o intervalo BCa é uma modificação do intervalo percentil, é razoável ele herdar as propriedades inerentes do mesmo, exceto que o intervalo BCa é de acurácia de segunda ordem, isto é, o erro de aproximação tende a zero em velocidade $1/n$.

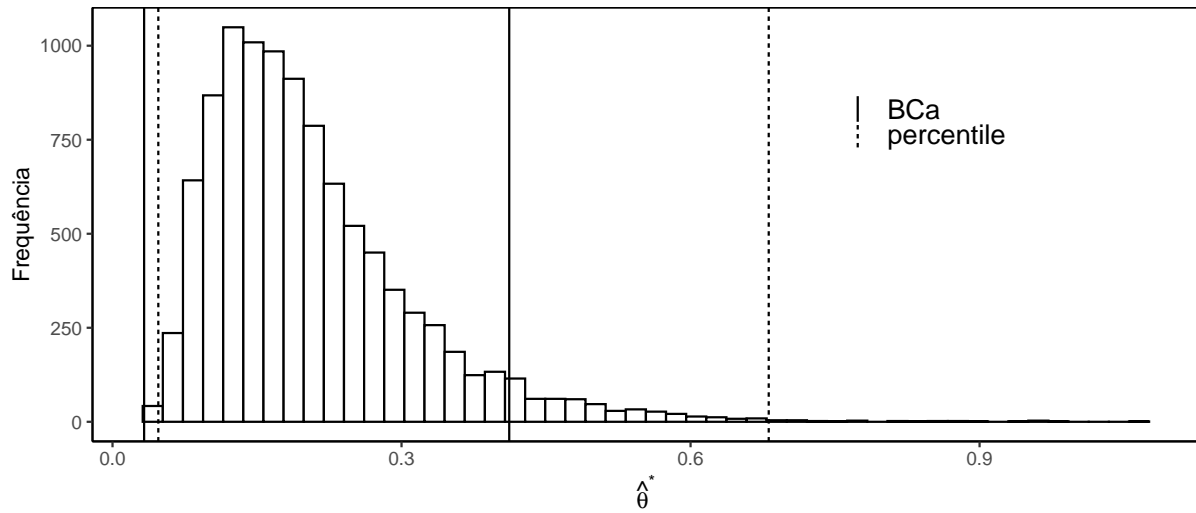


Figura 2: Histograma das réplicas bootstrap de uma estatística genérica $\hat{\theta}$, tendo os respectivos intervalos de confiança bootstrap com 99.5% de confiança.

2.4 Modelo Birnbaum-Saunders

A distribuição Birnbaum Saunders proposta em Birnbaum e Saunders (1969) descreve o tempo de falha de um material submetido a um ciclo padrão de estresse repetido. Esse modelo pode ser obtida a partir de uma variável aleatória T que admite a seguinte representação estocástica

$$T \stackrel{d}{=} \frac{\beta}{4} \left[\alpha Z + \sqrt{(\alpha Z)^2 + 4} \right]^2 \quad (8)$$

Sendo $Z \sim N(0, 1)$, $\alpha > 0$ e $\beta > 0$ os parâmetros de forma e escala do modelo BS, respectivamente. Convencionalmente denotado por $T \sim BS(\alpha, \beta)$. A v.a Z em 8 também assume a seguinte representação estocástica.

$$Z \stackrel{d}{=} \frac{1}{\alpha} \left[\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right] \quad (9)$$

consequentemente, sua função de densidade é dado por .

$$f_T(t) = \phi \left(\frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right] \right) \frac{t + b}{2\alpha\sqrt{\beta t^3}}, \quad t > 0$$

Sendo $\phi(\cdot)$ a função densidade de uma v.a normal padrão. Johnson, Kotz e Balakrishnan (1995) traz algumas transformações de $T \sim BS(\alpha, \beta)$ como $cT \sim BS(\alpha, c\beta)$ ($c > 0$) e $T^{-1} \sim BS(\alpha, \beta^{-1})$. A estimação de percentis inferiores extremos é difícil de se empregar usando os modelos usuais quando a distribuição real dos dados é mais assimétrica, a seguir então será tratado uma extensão do modelo BS que melhora a estimação de percentis nesses casos.

3 Modelo skew Birnbaum-Saunders

A distribuição skew Birnbaum Saunders, uma extensão da distribuição BS desenvolvida em Vilca et al. (2011), apresenta um melhor desempenho na estimação de percentis inferiores extremos do que o modelo usual BS ou um modelo BS baseado em uma distribuição simétrica, esse modelo é obtido utilizando uma v.a skew normal com parâmetro de assimetria λ , ver Azzalini (2013), denotado por $Z \sim SN(\lambda)$ na equação (8). Consequentemente, agora obtivemos uma v.a $T \sim SNBS(\alpha, \beta, \lambda)$ a qual possui a seguinte função de densidade

$$f_T(t) = 2\phi \left(\frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right] \right) \Phi \left(\lambda \frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right] \right) \frac{t + b}{2\alpha\sqrt{\beta t^3}}, \quad t > 0 \quad (10)$$

Sendo $\phi(\cdot)$ e $\Phi(\cdot)$ a função densidade e a função de distribuição acumulada, respectivamente, de uma v.a normal padrão.

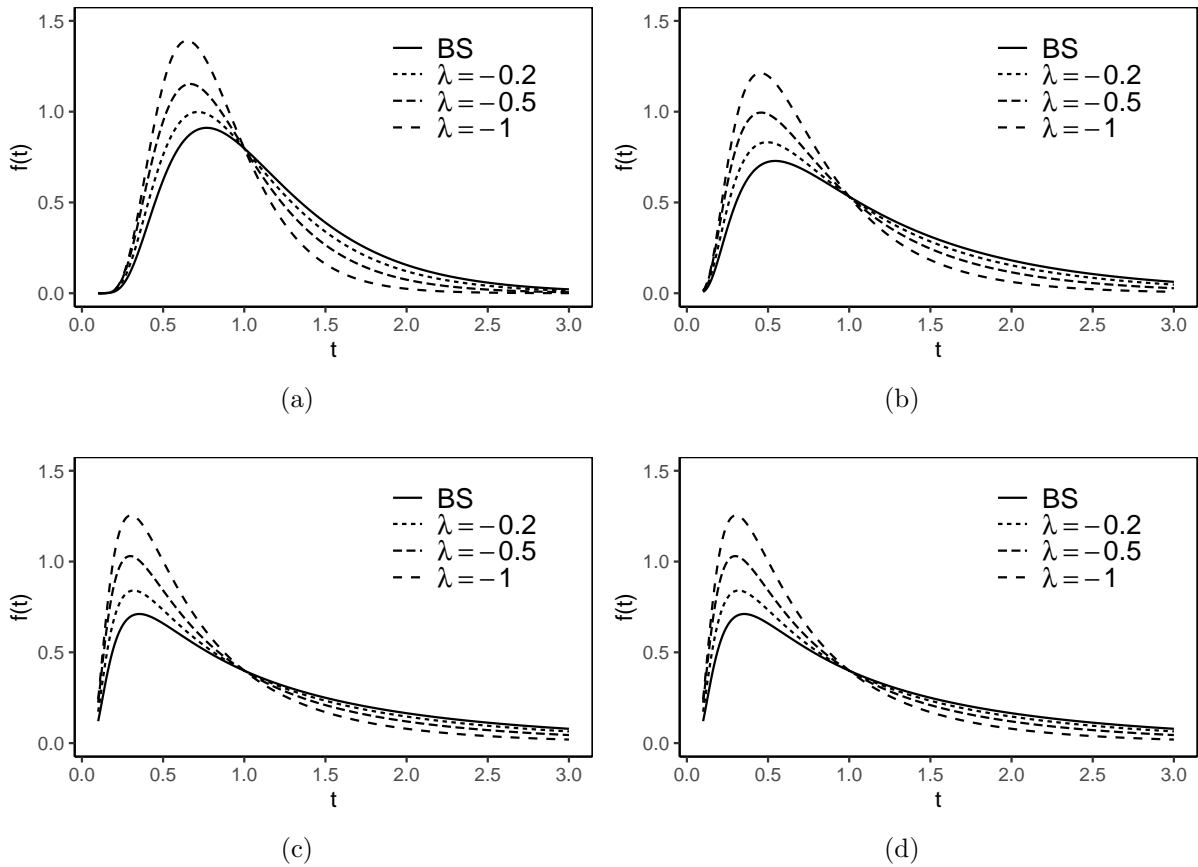


Figura 3: Densidades $SNBS(\alpha, \beta = 1, \lambda)$ variando λ com $\alpha = 0.5$ no painel (a), $\alpha = 0.75$ no painel (b), $\alpha = 1$ no painel (c) e $\alpha = 2$ no painel (d).

Usando a representação estocástica conhecida para o modelo skew-normal, ver Genton (2004), é dado por

$$Z \stackrel{d}{=} \delta|Z_0| + Z_1\sqrt{1 - \delta^2} \quad (11)$$

Onde $\delta = \lambda/\sqrt{1 + \lambda^2}$, $Z_0, Z_1 \stackrel{iid}{\sim} N(0, 1)$. Então das equações 8 e 11, obtemos a seguinte representação para $T \sim SNBS(\alpha, \beta, \lambda)$

$$T \stackrel{d}{=} \frac{\beta}{4} \left[\alpha \left\{ \delta|Z_0| + Z_1\sqrt{1 - \delta^2} \right\} + \sqrt{\left(\alpha \left\{ \delta|Z_0| + Z_1\sqrt{1 - \delta^2} \right\} \right)^2 + 4} \right]^2 \quad (12)$$

A seguir serão apresentado teoremas que serão úteis para a estimação via algoritmo ECM, ver Vilca et al. (2011).

Teorema 3.1. *Seja $T \sim SNBS(\alpha, \beta, \lambda)$ e $H \sim HN(0, 1)$ então $T|(H = h)$ tem densidade dada por*

$$f_{T|H}(t|h) = \phi \left(\lambda_h + \frac{1}{\alpha_\delta} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right] \right) \frac{t + b}{2\alpha_\delta \sqrt{\beta t^3}}, \quad t > 0$$

Onde $\alpha_\delta = \alpha\sqrt{1-\delta^2}$, $\lambda_h = -\delta h/\sqrt{1-\delta^2}$

O modelo apresentado no teorema 3.1 é conhecido como Birnbaum-Saunders estendido e é denotado por $EBS(\alpha_h, \beta, \sigma = 2, \lambda_h)$; ver Vilca et al. (2010).

Teorema 3.2. *Seja $T \sim SNBS(\alpha, \beta, \lambda)$ e $H \sim HN(0, 1)$ então $H|(T = t)$ também segue uma half-normal com densidade*

$$f_{H|T}(h|t) = \frac{\phi(h|\delta a(t; \alpha, \beta), 1 - \delta^2)}{\Phi(\lambda a(t; \alpha, \beta))}, \quad t > 0 \quad (13)$$

Onde $a(t; \alpha, \beta) = (1/\alpha)[\sqrt{t/\beta} - \sqrt{\beta/t}]$ e $\phi(\cdot|\mu, \sigma)$ denota o valor da densidade de uma variável normal com média μ e desvio padrão σ . logo

$$E[H|(T = t)] = \delta a(t; \alpha, \beta) + \sqrt{1 - \delta^2} a(t; \alpha, \beta) W_\Phi \left(\frac{\delta a(t; \alpha, \beta)}{\sqrt{1 - \delta^2}} \right)$$

$$E[H^2|(T = t)] = \delta^2 a^2(t; \alpha, \beta) + [1 - \delta^2] + \delta \sqrt{1 - \delta^2} a(t; \alpha, \beta) W_\Phi \left(\frac{\delta a(t; \alpha, \beta)}{\sqrt{1 - \delta^2}} \right)$$

Onde $W_\Phi(u) = \phi(u)/\Phi(u)$

3.1 Estimação via Algoritmo ECM

O algoritmo EM Dempster, Laird e Rubin (1977) é um famoso método iterativo para encontrar estimativas de máxima verossimilhança para dados incompletos, que consiste em uma etapa E de obter uma expectância e uma etapa M de maximização. Apesar da lenta convergência, esse algoritmo é numericamente estável, isto é, cada iteração aumenta o valor da verossimilhança. Uma generalização desse método é o algoritmo ECM Meng e Rubin (1993), que substitui a etapa de maximização por uma sequência de maximizações condicionais chamado de etapa CM, essas etapas são realizadas iterativamente até atingir um critério de parada que é definida inicialmente.

Seja uma amostra aleatória $T_i \sim SNBS(\alpha, \beta, \lambda)$, $i = 1, \dots, n$. A função de log-verossimilhança para $\boldsymbol{\theta} = [\alpha, \beta, \lambda]^T$ para a amostra observada t_1, \dots, t_n é dado por.

$$l(\boldsymbol{\theta}|\mathbf{t}) = \ln \left\{ 2\phi \left(\frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right] \right) \right\} + \ln \left\{ \Phi \left(\frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right] \right) \right\} + \ln \left\{ \frac{t + b}{2\alpha\sqrt{\beta t^3}} \right\} \quad (14)$$

Seja $\mathbf{t}^{(c)} = [\mathbf{t}^T, \mathbf{h}^T]^T$ os dados completos, que corresponde aos dados observados \mathbf{t} e os dados não observados $\mathbf{h} = [h_1, \dots, h_n]^T$, do teorema 3.1 obtemos a seguinte representação hierárquica.

$$T_i|(H_i = h_i) \stackrel{ind}{\sim} EBS(\alpha_\delta, \beta, \sigma = 2, \lambda_{h_i})$$

$$H_i \stackrel{iid}{\sim} HN(0, 1)$$

Essa representação nos permite encontrar a densidade conjunta $f_{T,H}(t, h)$ e assim construir a log-verossimilhança associada a $\mathbf{t}^{(c)}$ como $l^{(c)}(\boldsymbol{\theta}|\mathbf{t}^{(c)})$, dada por

$$l^{(c)}(\boldsymbol{\theta}|t_i, h_i) = \sum_{i=1}^n \left[k - \ln\{\alpha_\delta\} + \ln \left\{ \frac{t_i + \beta}{\sqrt{\beta}} \right\} - \frac{[a(t_i; \alpha, \beta) - \delta h_i]^2}{2[1 - \delta^2]} \right] \quad (15)$$

Sendo k uma constante que não depende de $\boldsymbol{\theta}$. Como descrito anteriormente, o algoritmo ECM utiliza a esperança condicional da log-verossimilhança completa dado $T_i = t_i$, logo se fará o uso da função $Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) = E \left[l^{(c)}(\boldsymbol{\theta}) \mid (T_i = t_i); \hat{\boldsymbol{\theta}} \right]$, onde

$$Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n k - \ln\{\alpha_\delta\} + \ln \left\{ \frac{t_i + \beta}{\sqrt{\beta}} \right\} - \frac{\xi^2(t_i; \beta)}{2\alpha_\delta^2} + \frac{\lambda}{\alpha_\delta} \xi(t_i; \beta) \hat{h}_i - \frac{\lambda^2}{2} \hat{h}_i^2 \quad (16)$$

onde $\xi(t_i; \beta) = \left[\sqrt{t_i/\beta} - \sqrt{\beta/t_i} \right]$, $\lambda = \delta/\sqrt{1 - \delta^2}$, $\hat{h}_i = E \left[H_i \mid (T_i = t_i); \hat{\boldsymbol{\theta}} \right]$ e $\hat{h}_i^2 = E \left[H_i^2 \mid (T_i = t_i); \hat{\boldsymbol{\theta}} \right]$ dados no teorema 3.2 avaliados em $\hat{\boldsymbol{\theta}}$. Temos então as etapas do algoritmo ECM no caso da estimação dos parâmetros do modelos skew Birnbaum-Saunders.

Etapa E: Dado $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^{(r)}$, calcular $\hat{h}_i^{(r)}$ e $\hat{h}_i^{2(r)}$ para $i = 1, 2, \dots, n$.

Etapa CM-1: Fixando $\hat{\beta}^{(r)}$, atualize $\hat{\alpha}^{(r)}$ e $\hat{\delta}^{(r)}$ com

$$\begin{aligned} \hat{\alpha}^{2(r+1)} &= \frac{1}{n} \sum_{i=1}^n \xi^2(t_i; \hat{\beta}^{(r)}) + \left[1 - \frac{\sum_{i=1}^n \hat{h}_i^{2(r)}}{n} \right] \left[\frac{\sum_{i=1}^n \hat{h}_i^{(r)} \xi(t_i; \hat{\beta}^{(r)})}{\sum_{i=1}^n \hat{h}_i^{2(r)}} \right]^2 \\ \hat{\delta}^{(r+1)} &= \frac{1}{\alpha_\delta^{(r)}} \frac{\sum_{i=1}^n \hat{h}_i^{(r)} b(t_i; \hat{\beta}^{(r)})}{\sum_{i=1}^n \hat{h}_i^{2(r)}} \end{aligned}$$

onde

$$b(t_i; \hat{\beta}^{(r)}) = 2 \sinh \left(\frac{t_i - \hat{\beta}^{(r)}}{2} \right)$$

Etapa CM-2: Fixando $\hat{\alpha}^{(r+1)}$ e $\hat{\delta}^{(r+1)}$, atualize $\hat{\beta}^{(r)}$ utilizando

$$\hat{\beta}^{(r+1)} = \underset{\beta}{\operatorname{argmax}} Q \left(\hat{\alpha}^{(r+1)}, \beta, \hat{\delta}^{(r+1)} \mid \hat{\boldsymbol{\theta}}^{(r)} \right)$$

Vale ressaltar que é sempre ter bons valores iniciais para usar num procedimento deste tipo, no caso da distribuição $SNBS(\alpha, \beta, \lambda)$ podemos usar para os parâmetros de forma e escala os valores iniciais sendo a estimativa de momentos modificados. Vilca et al. (2011) também traz a matriz de informação de fisher observada

4 Metodologia

4.1 Materiais

Este trabalho consiste em propor o gráfico de controle bootstrap para percentis skew Birnbaum-Saunders e verificar seu desempenho por meio de simulações. O estudo simulado foi feito utilizando o software gratuito Rstudio (R Core Team, 2020), o estudo se tornou possível devido ao pacote `bssn` (MAEHARA; SANCHEZ, 2016) o qual já havia implementado diversas funções do modelo skew Birnbaum-Saunders como também a estimação dos parâmetros por máxima verossimilhança via algoritmo ECM.

4.2 Métodos

Os seguintes passos baseados em (LIO; PARK, 2008) são seguidos para a construção do gráfico de controle bootstrap para percentis da distribuição skew Birnbaum-Saunders, estimando os parâmetros via máxima verossimilhança.

1. A partir de um processo sob controle, observe k amostras aleatórias independentes de tamanho n_j ($j = 1, \dots, k$), usualmente igual a 4,5 ou 6.
2. Obtenha a estimativa de máxima verossimilhança dos parâmetros α , β e λ do modelo SNBS usando a amostra combinada de tamanho $N = \sum_{j=1}^k n_j$, denotando por $\hat{\alpha}$, $\hat{\beta}$ e $\hat{\lambda}$.
3. Gerar amostras bootstrap de tamanho m $x_1^*, x_2^*, \dots, x_m^*$ da skew Birnbaum-Saunders usando as estimativas MV obtidas no passo 2.
4. Obter as estimativas MV para os parâmetros do modelo SNBS na amostra bootstrap gerada no passo 3, denotando por $\hat{\alpha}^*$, $\hat{\beta}^*$ e $\hat{\lambda}^*$.
5. Para a amostra bootstrap obtida no passo 3 e as estimativas MV obtidas no passo 4, calcular a estimativa bootstrap do p -ésimo percentil \hat{W}_p^* .
6. Repetir os passos 3-6 um elevado número de vezes, B , obtendo B amostra bootstrap de \hat{W}_p^* , denotado por $\hat{W}_{p1}^*, \dots, \hat{W}_{pB}^*$.
7. A partir da amostra *bootstrap* de W_{bp}^* serão calculados os limites de controle usando intervalo percentil e BCa como descrito em 2.3.1 e 2.3.2, sendo γ a probabilidade de uma observação indicar uma situação fora de controle quando o processo esta sob controle estatístico, isto é, a taxa de falso alarme FAR (*false alarm rate*). A linha central corresponde à estimativa \hat{W}_p a partir das estimativas de máxima verossimilhança obtidas no passo 2.

Uma vez que os limites de controle foram computados, subgrupos de tamanho n , são coletados do processo em intervalos regulares de tempo, e \hat{W}_p é obtido via máxima verossimilhança e comparado com os limites para avaliar o estado processo. Enquanto $LIC < \hat{W}_p < LSC$ o

processo é considerado sob controle, caso $\hat{W}_p < LIC$ ou $\hat{W}_p > LSC$ o processo é considerado fora de controle e deve ser interrompido a fim de eliminar as possíveis causas.

5 Estudo de simulação

Para avaliar a performance do gráfico de controle bootstrap para percentis skew Birnbaum-Saunders, será conduzido uma simulação computacional. O comportamento do gráfico de controle em questão é analisado pelo uso dos limites de controle médio, CMS médio e seus respectivos desvios padrões. A simulação foi gerida com diferentes valores do parâmetro de forma, percentis de interesse e taxas de falso alarme.

Os limites de controle médio foram computados do seguinte modo: para dados valores de α, β e λ foi reinicido 10000 o algoritmo descrito em 4.2, os limites de controle médio foram calculados a partir da média aritmética desses 10000 valores gerados de LICs e LSCs assim como seus respectivos desvios padrão. Os resultados dos valores médios dos limites de controle são apresentados em 1 e 4 e seus respectivos desvios padrão são apresentados em 2 e 5

Os CMSs médio quando o processo esta sob controle foram investigados da seguinte maneira: uma vez que os limites de controle são computados, geram-se amostras do mesmo processo e estima-se \widehat{W}_p . O número de amostras geradas até que a primeira estimativa de \widehat{W}_p se apresente fora dos limites de controle é o comprimento da sequência associado a esses limites, repete-se esse processo 10000, a partir desses valores de CMSs são calculados a média e seu erro padrão, esses resultados são apresentados em 3 e 6. O comprimento médio da sequência quando o processo está fora de controle também é considerado no estudo realizado da seguinte: uma vez que os limites de controle são computados, geram-se amostras processo com o valor do parâmetro α defasado e estima-se \widehat{W}_p . O número de amostras geradas até que a primeira estimativa de \widehat{W}_p se apresente fora dos limites de controle é o comprimento da sequência associado a esses limites, repete-se esse processo 10000, a partir desses valores de CMSs são calculados a média e seu respectivo desvio padrão, esses resultados são apresentados em 7.

Baseado na teoria padrão discutido em 2.1.1, o comprimento médio da sequência esperado é o inverso da taxa de falso alarme, CMSs menores do que isso é uma indicação que os limites estão tipicamente estreitos e CMSs maiores indicam que os limites estão tipicamente largos.

Tabela 1: EMV para LIC e LSC sob controle utilizando o método de intervalo bootstrap percentil, para os percentis $p=0.01, 0.1, 0.5$, $FAR=0.1, 0.01, 0.0027, 0.002$ ($n_j = 5$ e $k = 20$).

| Parâmetro de forma | p=0.01 | | p=0.1 | | p=0.5 | |
|--------------------|----------|----------|------------|----------|----------|----------|
| | LIC | LSC | LIC | LSC | LIC | LSC |
| | | | FAR=0.1 | | | |
| $\alpha = 0.5$ | 0.208263 | 0.590436 | 0.335264 | 0.724898 | 0.567701 | 1.018142 |
| $\alpha = 1$ | 0.062353 | 0.370976 | 0.133730 | 0.543338 | 0.346534 | 1.036138 |
| $\alpha = 2$ | 0.014829 | 0.166786 | 0.039745 | 0.318117 | 0.160986 | 1.110915 |
| | | | FAR=0.01 | | | |
| $\alpha = 0.5$ | 0.153852 | 0.760563 | 0.269261 | 0.889654 | 0.482437 | 1.198210 |
| $\alpha = 1$ | 0.039459 | 0.604097 | 0.094516 | 0.811079 | 0.258061 | 1.403554 |
| $\alpha = 2$ | 0.008766 | 0.401285 | 0.026305 | 0.680652 | 0.102431 | 1.934855 |
| | | | FAR=0.0027 | | | |
| $\alpha = 0.5$ | 0.134606 | 0.846124 | 0.244374 | 0.972839 | 0.448888 | 1.288253 |
| $\alpha = 1$ | 0.032312 | 0.743594 | 0.081689 | 0.965595 | 0.227066 | 1.602816 |
| $\alpha = 2$ | 0.006987 | 0.597538 | 0.022263 | 0.959131 | 0.084993 | 2.448898 |
| | | | FAR=0.002 | | | |
| $\alpha = 0.5$ | 0.130972 | 0.864534 | 0.239564 | 0.990850 | 0.442204 | 1.307769 |
| $\alpha = 1$ | 0.031007 | 0.775311 | 0.079328 | 1.000305 | 0.221265 | 1.647064 |
| $\alpha = 2$ | 0.006669 | 0.648552 | 0.021533 | 1.029333 | 0.081849 | 2.569828 |

Tabela 2: Desvio padrão dos limites de controle na tabela 1 ($n_j = 5$ e $k = 20$).

| Parâmetro de forma | p=0.01 | | p=0.1 | | p=0.5 | |
|--------------------|----------|----------|------------|----------|----------|----------|
| | LIC | LSC | LIC | LSC | LIC | LSC |
| | | | FAR=0.1 | | | |
| $\alpha = 0.5$ | 0.019815 | 0.026165 | 0.023623 | 0.029297 | 0.025838 | 0.045335 |
| $\alpha = 1$ | 0.009002 | 0.030764 | 0.015440 | 0.041850 | 0.027612 | 0.089082 |
| $\alpha = 2$ | 0.002285 | 0.025009 | 0.005522 | 0.045193 | 0.019520 | 0.168186 |
| | | | FAR=0.01 | | | |
| $\alpha = 0.5$ | 0.017098 | 0.031960 | 0.022214 | 0.038035 | 0.025147 | 0.061036 |
| $\alpha = 1$ | 0.006541 | 0.049551 | 0.012323 | 0.067549 | 0.022941 | 0.138817 |
| $\alpha = 2$ | 0.001429 | 0.068046 | 0.003823 | 0.109842 | 0.012893 | 0.340149 |
| | | | FAR=0.0027 | | | |
| $\alpha = 0.5$ | 0.016082 | 0.038492 | 0.021618 | 0.045633 | 0.025363 | 0.070935 |
| $\alpha = 1$ | 0.005736 | 0.066866 | 0.011220 | 0.088702 | 0.021579 | 0.170964 |
| $\alpha = 2$ | 0.001179 | 0.109494 | 0.003322 | 0.167918 | 0.011154 | 0.464175 |
| | | | FAR=0.002 | | | |
| $\alpha = 0.5$ | 0.015903 | 0.040664 | 0.021574 | 0.047848 | 0.025450 | 0.073733 |
| $\alpha = 1$ | 0.005607 | 0.071725 | 0.011027 | 0.094259 | 0.021422 | 0.179442 |
| $\alpha = 2$ | 0.001140 | 0.122140 | 0.003239 | 0.185297 | 0.010865 | 0.497504 |

Tabela 3: CMS e erro padrão utilizando o método de intervalo bootstrap percentil, para os percentis $p=0.01, 0.1, 0.5$, FAR=0.1, 0.01, 0.0027, 0.002 ($n_j = 5$ e $k = 20$).

| Parâmetro de forma | p=0.01 | | p=0.1 | | p=0.5 | |
|--------------------|------------|----------|------------|-----------|------------|----------|
| | CMS | EP | CMS | EP | CMS | EP |
| | | | FAR=0.1 | | | |
| $\alpha = 0.5$ | 8.825400 | 0.086328 | 8.778300 | 0.085891 | 9.363400 | 0.096754 |
| $\alpha = 1$ | 10.276000 | 0.102246 | 10.168800 | 0.103147 | 9.428500 | 0.095302 |
| $\alpha = 2$ | 11.017000 | 0.114057 | 10.778000 | 0.110971 | 9.603700 | 0.096501 |
| | | | FAR=0.01 | | | |
| $\alpha = 0.5$ | 80.542400 | 0.950346 | 80.600400 | 0.989332 | 93.697000 | 1.259010 |
| $\alpha = 1$ | 111.108100 | 1.350899 | 110.333300 | 1.459769 | 90.454800 | 1.169375 |
| $\alpha = 2$ | 123.490400 | 1.584037 | 115.237400 | 1.512275 | 94.996800 | 1.220188 |
| | | | FAR=0.0027 | | | |
| $\alpha = 0.5$ | 287.371900 | 3.961586 | 302.945100 | 4.645574 | 350.379800 | 5.612581 |
| $\alpha = 1$ | 437.414900 | 6.818166 | 436.104900 | 7.100006 | 340.155300 | 5.153968 |
| $\alpha = 2$ | 472.510100 | 6.928664 | 429.268400 | 6.576070 | 347.059000 | 5.095428 |
| | | | FAR=0.002 | | | |
| $\alpha = 0.5$ | 384.914100 | 5.564490 | 405.168400 | 6.370340 | 476.176800 | 7.933619 |
| $\alpha = 1$ | 586.892700 | 9.354750 | 589.494300 | 10.199445 | 453.985200 | 7.179456 |
| $\alpha = 2$ | 640.731000 | 9.730546 | 587.396400 | 9.673676 | 464.649300 | 7.143877 |

Tabela 4: EMV para LIC e LSC sob controle utilizando o método de intervalo bootstrap BCa, para os percentis $p=0.01, 0.1, 0.5$, FAR=0.1, 0.01, 0.0027, 0.002 ($n_j = 5$ e $k = 20$).

| Parâmetro de forma | p=0.01 | | p=0.1 | | p=0.5 | |
|--------------------|----------|----------|------------|----------|----------|----------|
| | LIC | LSC | LIC | LSC | LIC | LSC |
| | | | FAR=0.1 | | | |
| $\alpha = 0.5$ | 0.158476 | 0.450265 | 0.287461 | 0.619423 | 0.561874 | 1.008416 |
| $\alpha = 1$ | 0.042740 | 0.233546 | 0.105958 | 0.404997 | 0.336436 | 1.006151 |
| $\alpha = 2$ | 0.011150 | 0.102106 | 0.030893 | 0.199450 | 0.137341 | 0.920161 |
| | | | FAR=0.01 | | | |
| $\alpha = 0.5$ | 0.122957 | 0.607873 | 0.236630 | 0.775733 | 0.470226 | 1.170722 |
| $\alpha = 1$ | 0.029247 | 0.409790 | 0.079414 | 0.633420 | 0.244168 | 1.330542 |
| $\alpha = 2$ | 0.007146 | 0.262524 | 0.022163 | 0.448895 | 0.085563 | 1.549731 |
| | | | FAR=0.0027 | | | |
| $\alpha = 0.5$ | 0.111917 | 0.693717 | 0.219115 | 0.859085 | 0.434024 | 1.249439 |
| $\alpha = 1$ | 0.025350 | 0.528992 | 0.071070 | 0.778238 | 0.212423 | 1.500759 |
| $\alpha = 2$ | 0.005973 | 0.420195 | 0.019527 | 0.672054 | 0.070976 | 1.933667 |
| | | | FAR=0.002 | | | |
| $\alpha = 0.5$ | 0.110103 | 0.713228 | 0.215847 | 0.877883 | 0.426974 | 1.266356 |
| $\alpha = 1$ | 0.024705 | 0.558160 | 0.069631 | 0.812975 | 0.206570 | 1.538522 |
| $\alpha = 2$ | 0.005769 | 0.464866 | 0.019061 | 0.733268 | 0.068623 | 2.023246 |

Tabela 5: Desvio padrão dos limites de controle na tabela 4 ($n_j = 5$ e $k = 20$).

| Parâmetro de forma | p=0.01 | | p=0.1 | | p=0.5 | |
|--------------------|------------|----------|----------|----------|----------|----------|
| | LIC | LSC | LIC | LSC | LIC | LSC |
| | FAR=0.1 | | | | | |
| $\alpha = 0.5$ | 0.017560 | 0.025241 | 0.022628 | 0.026777 | 0.025813 | 0.044872 |
| $\alpha = 1$ | 0.007602 | 0.023282 | 0.013391 | 0.031461 | 0.028620 | 0.087563 |
| $\alpha = 2$ | 0.001877 | 0.012897 | 0.004472 | 0.024588 | 0.020309 | 0.153754 |
| | FAR=0.01 | | | | | |
| $\alpha = 0.5$ | 0.015732 | 0.027185 | 0.021366 | 0.031583 | 0.025141 | 0.058654 |
| $\alpha = 1$ | 0.005824 | 0.034947 | 0.011094 | 0.048798 | 0.022967 | 0.127243 |
| $\alpha = 2$ | 0.001263 | 0.036177 | 0.003349 | 0.061549 | 0.012530 | 0.264689 |
| | FAR=0.0027 | | | | | |
| $\alpha = 0.5$ | 0.015725 | 0.029541 | 0.021507 | 0.036383 | 0.025556 | 0.066743 |
| $\alpha = 1$ | 0.005333 | 0.045315 | 0.010515 | 0.063883 | 0.021490 | 0.152050 |
| $\alpha = 2$ | 0.001096 | 0.066537 | 0.003046 | 0.101167 | 0.010719 | 0.340182 |
| | FAR=0.002 | | | | | |
| $\alpha = 0.5$ | 0.015801 | 0.030320 | 0.021553 | 0.037712 | 0.025775 | 0.068644 |
| $\alpha = 1$ | 0.005249 | 0.048237 | 0.010445 | 0.068056 | 0.021223 | 0.157893 |
| $\alpha = 2$ | 0.001070 | 0.076313 | 0.003008 | 0.112927 | 0.010497 | 0.358030 |

Tabela 6: CMS e erro padrão utilizando o método de intervalo bootstrap BCa, para os percentis p=0.01, 0.1, 0.5, FAR=0.1, 0.01, 0.0027, 0.002 ($n_j = 5$ e $k = 20$).

| Parâmetro de forma | p=0.01 | | p=0.1 | | p=0.5 | |
|--------------------|------------|----------|------------|------------|------------|----------|
| | CMS | EP | CMS | EP | CMS | EP |
| | FAR=0.1 | | | | | |
| $\alpha = 0.5$ | 4.162800 | 0.037940 | 5.366500 | 0.051580 | 9.357000 | 0.095750 |
| $\alpha = 1$ | 4.740600 | 0.043259 | 5.915500 | 0.056513 | 9.345700 | 0.093841 |
| $\alpha = 2$ | 7.022100 | 0.067185 | 6.565700 | 0.062735 | 8.478000 | 0.084127 |
| | FAR=0.01 | | | | | |
| $\alpha = 0.5$ | 24.382500 | 0.264058 | 36.293000 | 0.411048 | 89.057400 | 1.139375 |
| $\alpha = 1$ | 32.579400 | 0.352323 | 46.530900 | 0.512738 | 86.262800 | 1.132142 |
| $\alpha = 2$ | 65.338900 | 0.714357 | 56.094000 | 0.634668 | 69.358200 | 0.919530 |
| | FAR=0.0027 | | | | | |
| $\alpha = 0.5$ | 74.640300 | 0.841885 | 119.416300 | 1.499368 | 328.636200 | 4.989957 |
| $\alpha = 1$ | 110.456800 | 1.258208 | 165.229900 | 2.038226 | 305.532700 | 4.902920 |
| $\alpha = 2$ | 250.757600 | 2.935138 | 207.111500 | 2.568755 | 224.813800 | 3.597842 |
| | FAR=0.002 | | | | | |
| $\alpha = 0.5$ | 97.414500 | 1.120009 | 160.157300 | 2.085820 | 439.540700 | 7.044231 |
| $\alpha = 1$ | 147.593300 | 1.744478 | 223.966100 | 3.063862 | 403.256100 | 6.611553 |
| $\alpha = 2$ | 347.563000 | 4.158005 | 281.858300 | 3.50649468 | 287.887100 | 4.817728 |

Tabela 7: CMS fora de controle e seu respectivo erro padrão com α_1 deslocado para α_2 ($n_j = 5$ e $k = 20$).

| percentile | α_1 | α_2 | CMS_{perc} | $CMS_{perc}(EP)$ | CMS_{BCa} | $CMS_{BCA}(EP)$ |
|------------|------------|------------|--------------|------------------|-------------|-----------------|
| FAR=0.1 | | | | | | |
| p=0.01 | 0.5 | 1.0 | 1.351400 | 0.006965 | 1.743600 | 0.011895 |
| p=0.1 | 0.5 | 1.0 | 1.375400 | 0.007433 | 1.627200 | 0.010573 |
| p=0.5 | 0.5 | 1.0 | 2.004000 | 0.014695 | 2.032200 | 0.014870 |
| p=0.01 | 1.0 | 1.5 | 2.506200 | 0.020425 | 4.332200 | 0.041568 |
| p=0.1 | 1.0 | 1.5 | 2.446800 | 0.019886 | 3.594400 | 0.034009 |
| p=0.5 | 1.0 | 1.5 | 3.300000 | 0.028241 | 3.467000 | 0.030455 |
| p=0.01 | 2.0 | 2.5 | 5.275000 | 0.052488 | 6.977000 | 0.067816 |
| p=0.1 | 2.0 | 2.5 | 4.785600 | 0.045968 | 6.027800 | 0.055948 |
| FAR=0.01 | | | | | | |
| p=0.01 | 0.5 | 1.0 | 1.896000 | 0.013849 | 2.793600 | 0.024253 |
| p=0.1 | 0.5 | 1.0 | 1.906800 | 0.014083 | 2.494800 | 0.021205 |
| p=0.5 | 0.5 | 1.0 | 3.552000 | 0.031306 | 3.844800 | 0.034732 |
| p=0.01 | 1.0 | 1.5 | 7.256000 | 0.088487 | 16.808400 | 0.218197 |
| p=0.1 | 1.0 | 1.5 | 6.293200 | 0.075481 | 10.972600 | 0.135844 |
| p=0.5 | 1.0 | 1.5 | 10.011800 | 0.103968 | 11.656600 | 0.121693 |
| p=0.01 | 2.0 | 2.5 | 33.880400 | 0.522369 | 47.123600 | 0.556147 |
| p=0.1 | 2.0 | 2.5 | 23.060800 | 0.319285 | 33.281200 | 0.404244 |
| FAR=0.0027 | | | | | | |
| p=0.01 | 0.5 | 1.0 | 2.356400 | 0.019124 | 3.523200 | 0.034913 |
| p=0.1 | 0.5 | 1.0 | 2.336400 | 0.019098 | 3.108600 | 0.030828 |
| p=0.5 | 0.5 | 1.0 | 4.895600 | 0.046825 | 5.571200 | 0.054170 |
| p=0.01 | 1.0 | 1.5 | 14.544000 | 0.230446 | 34.818800 | 0.558149 |
| p=0.1 | 1.0 | 1.5 | 11.121400 | 0.150275 | 19.443400 | 0.294827 |
| p=0.5 | 1.0 | 1.5 | 18.103000 | 0.199508 | 23.159000 | 0.264949 |
| p=0.01 | 2.0 | 2.5 | 96.553600 | 1.612005 | 130.057400 | 1.726174 |
| p=0.1 | 2.0 | 2.5 | 60.779600 | 1.097303 | 85.138600 | 1.207928 |
| FAR=0.002 | | | | | | |
| p=0.01 | 0.5 | 1.0 | 2.492000 | 0.020795 | 3.675000 | 0.037162 |
| p=0.1 | 0.5 | 1.0 | 2.443000 | 0.020370 | 3.266000 | 0.032843 |
| p=0.5 | 0.5 | 1.0 | 5.241200 | 0.050798 | 6.068600 | 0.060692 |
| p=0.01 | 1.0 | 1.5 | 17.214000 | 0.280593 | 40.318600 | 0.664671 |
| p=0.1 | 1.0 | 1.5 | 12.695400 | 0.191945 | 22.037400 | 0.342464 |
| p=0.5 | 1.0 | 1.5 | 21.274600 | 0.241752 | 26.854800 | 0.311091 |
| p=0.01 | 2.0 | 2.5 | 121.932200 | 2.133754 | 162.675800 | 2.249715 |
| p=0.1 | 2.0 | 2.5 | 75.597600 | 1.465827 | 104.238400 | 1.476556 |

6 Exemplo ilustrativo

A presente seção tem por objetivo ilustrar a aplicação dos gráficos de controle Bootstrap propostos nesse trabalho. Os dados utilizados foram gerados artificialmente no software R, e como descrito anteriormente, pode descrever a resistência de algum componente utilizado em projetos de engenharia que se tenha o interesse de monitorar a qualidade.

Primeiramente foi gerado 20 amostras $T \sim SNBS(\alpha = 0.5, \beta = 1, \lambda = -1)$ de tamanho 5, esses dados foram os considerados sob-controle e foram utilizados para obtenção dos limites de controle, como descrito anteriormente. O interesse de se monitorar o percentil $W_{0.01}$ é considerado, e para os parâmetros escolhidos temos $W_{0.01} = 0.2976$. A tabela a seguir apresenta amostras sob controle

Tabela 8: 20 subgrupos do processo sob controle.

| Amostra | Observações | | | | |
|---------|-------------|--------|--------|--------|--------|
| 1 | 0.5077 | 0.6633 | 0.6892 | 0.5504 | 0.8894 |
| 2 | 0.6608 | 0.3407 | 1.3813 | 1.1186 | 0.7519 |
| 3 | 1.3179 | 0.6534 | 0.7955 | 0.6029 | 0.8168 |
| 4 | 1.1002 | 0.5593 | 0.4093 | 0.5604 | 0.4093 |
| 5 | 0.7915 | 0.7205 | 0.9716 | 1.8620 | 0.8174 |
| 6 | 2.1323 | 0.6627 | 1.2814 | 1.4174 | 1.2050 |
| 7 | 1.0587 | 0.5375 | 0.4659 | 1.1832 | 0.4395 |
| 8 | 0.6199 | 0.6307 | 0.7426 | 0.9967 | 1.0166 |
| 9 | 0.6139 | 0.7767 | 0.5723 | 0.6222 | 0.5273 |
| 10 | 0.5723 | 0.7686 | 0.7043 | 0.3455 | 0.6187 |
| 11 | 0.6585 | 0.9595 | 1.1361 | 0.6459 | 0.9558 |
| 12 | 0.5301 | 0.9390 | 0.4657 | 0.5230 | 0.5432 |
| 13 | 0.7215 | 1.3306 | 0.9351 | 1.4821 | 0.6011 |
| 14 | 0.4484 | 0.6166 | 1.2070 | 1.2821 | 0.3347 |
| 15 | 0.6606 | 0.8468 | 1.1483 | 0.5463 | 1.1953 |
| 16 | 0.4883 | 0.6262 | 0.6782 | 0.9394 | 1.3164 |
| 17 | 0.5524 | 0.8185 | 0.8144 | 0.9947 | 0.4770 |
| 18 | 0.8261 | 0.6457 | 1.3970 | 0.8420 | 0.5176 |
| 19 | 1.8544 | 1.4822 | 0.8261 | 1.1876 | 0.5012 |
| 20 | 0.4697 | 0.5644 | 1.8327 | 0.4963 | 0.5659 |

Após a obtenção dos limites de controle, foi gerado artificialmente mais 20 amostras de $T \sim SNBS(\alpha = 1, \beta = 1, \lambda = -1)$, considerados fora de controle, o percentil de interesse se deslocou para $W_{0.01} = 0.1174$. A tabela a seguir apresenta amostras fora controle.

Tabela 9: 20 subgrupos do processo fora de controle.

| Amostra | Observações | | | | |
|---------|-------------|----------|----------|----------|----------|
| 21 | 0.902438 | 0.141468 | 0.308701 | 0.735370 | 0.541203 |
| 22 | 0.468219 | 0.631039 | 0.414408 | 0.604209 | 1.047210 |
| 23 | 0.304980 | 0.254401 | 1.960308 | 0.471464 | 0.805628 |
| 24 | 0.217720 | 0.439465 | 0.671292 | 0.132100 | 0.891714 |
| 25 | 0.669691 | 0.565652 | 0.343406 | 0.536471 | 0.407239 |
| 26 | 0.580438 | 0.398907 | 0.234454 | 0.872837 | 0.645091 |
| 27 | 0.248572 | 0.494120 | 0.365167 | 1.003820 | 2.680040 |
| 28 | 0.326631 | 0.927139 | 0.451735 | 0.584455 | 0.433797 |
| 29 | 1.007431 | 0.757501 | 1.467377 | 0.310743 | 0.231010 |
| 30 | 0.411667 | 0.808563 | 0.523147 | 0.422900 | 0.565318 |
| 31 | 0.759139 | 0.520551 | 0.893608 | 0.755516 | 0.648034 |
| 32 | 3.092582 | 0.309092 | 0.132928 | 0.466010 | 0.874684 |
| 33 | 0.423863 | 0.261264 | 0.430302 | 0.388483 | 0.794531 |
| 34 | 1.917422 | 0.552687 | 0.909764 | 0.962089 | 0.550445 |
| 35 | 0.192949 | 0.533974 | 1.154707 | 0.382135 | 1.801227 |
| 36 | 1.888733 | 0.544255 | 0.592602 | 0.623935 | 0.387754 |
| 37 | 0.636466 | 1.947942 | 0.691112 | 0.672561 | 0.424623 |
| 38 | 0.659927 | 0.342604 | 1.721541 | 0.762863 | 0.325597 |
| 39 | 0.262650 | 0.385634 | 0.302290 | 3.368886 | 0.848429 |
| 40 | 0.206989 | 0.488725 | 0.228894 | 0.406216 | 0.507850 |

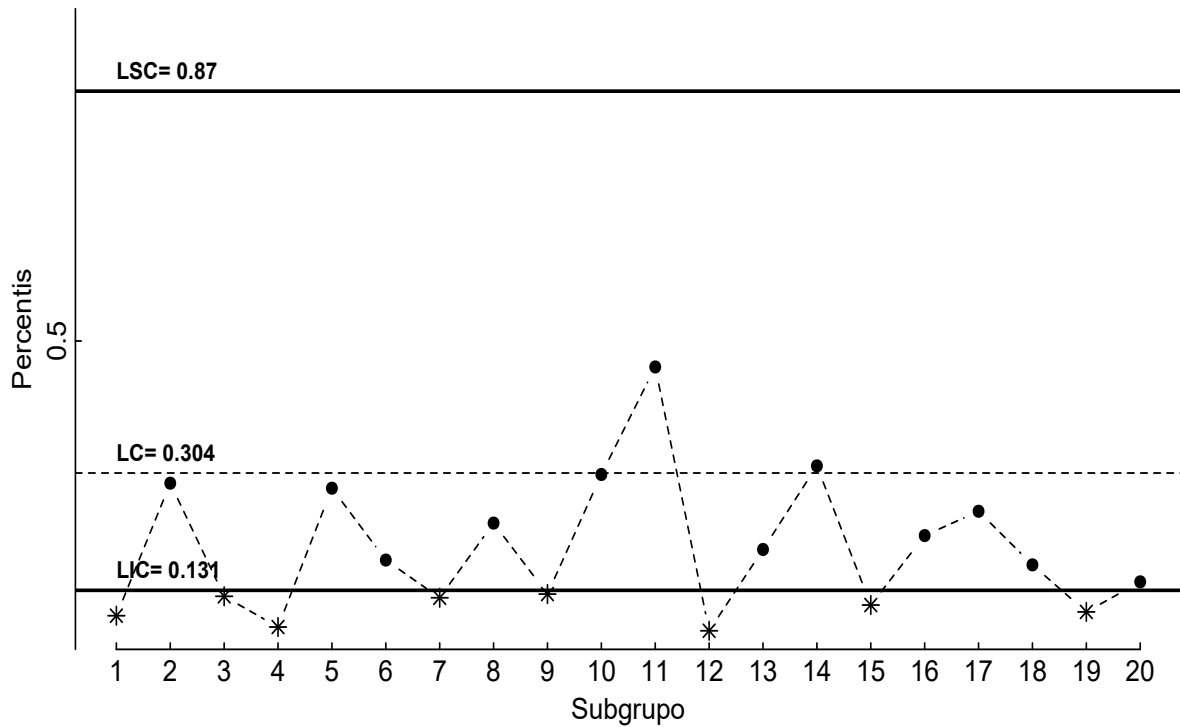


Figura 4: Carta de controle EMV bootstrap percentile com FAR=0.002

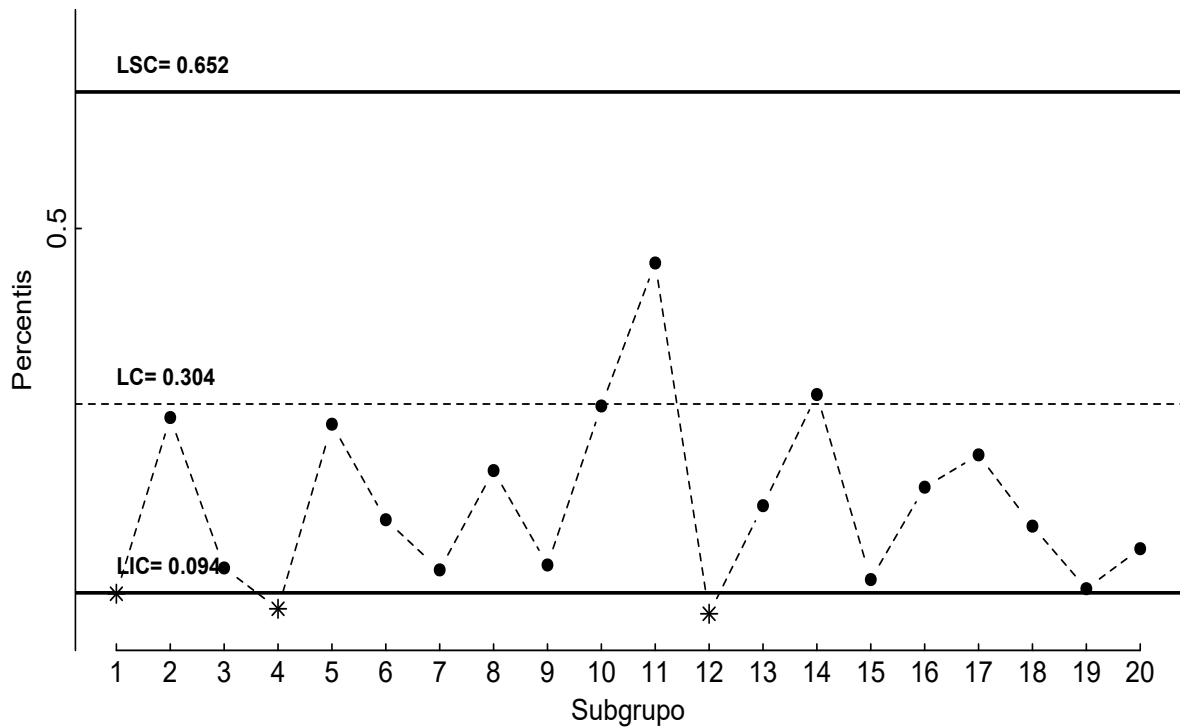


Figura 5: Carta de controle EMV bootstrap BCa com FAR=0.002

7 Conclusões

No presente trabalho o gráfico de controle bootstrap baseado no estimador de máxima verossimilhança para monitorar percentis do modelo skew Birnbaum-Saunders é proposto. Através de uma exaustiva simulação de Monte Carlo foi demonstrado que o uso dos limites de controle usando a metodologia do intervalo bootstrap percentil é adequado para percentis da distribuição skew Birnbaum-Saunders. Já a metodologia do intervalo bootstrap BCa apresentou CMSs sob controle usualmente inferiores ao esperado. O uso de ambos os gráficos de controle em situações práticas se apresentaram eficiente em detectar deslocamentos para baixo dos percentis inferiores, sendo neste caso mais apropriados que os gráficos de controle convencionais de Shewhart.

Referências

- AZZALINI, A. *The skew-normal and related families*. [S.l.]: Cambridge University Press, 2013. v. 3.
- BIRNBAUM, Z. W.; SAUNDERS, S. C. A new family of life distributions. *Journal of applied probability*, Cambridge University Press, v. 6, n. 2, p. 319–327, 1969.
- BOLFARINE, H.; SANDOVAL, M. C. *Introdução à inferência estatística*. [S.l.]: SBM, 2001. v. 2.
- BOURGUIGNON, M.; HO, L. L.; FERNANDES, F. H. Control charts for monitoring the median parameter of birnbaum-saunders distribution. *Quality and Reliability Engineering International*, Wiley Online Library, v. 36, n. 4, p. 1333–1363, 2020.
- CASELLA, G.; BERGER, R. L. *Inferência estatística*. São Paulo: Cengage Learning, 2010.
- CHENG, S. W.; XIE, H. Control charts for lognormal data. *Tamkang Journal of Science and Engineering*, TAMKANG UNIVERSITY, v. 3, n. 3, p. 131–138, 2000.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 39, n. 1, p. 1–22, 1977.
- DICICCIO, T. J.; EFRON, B. Bootstrap confidence intervals. *Statistical science*, JSTOR, p. 189–212, 1996.
- EFRON, B. Bootstrap methods: another look at the jackknife. In: *Breakthroughs in statistics*. [S.l.]: Springer, 1979. p. 569–593.
- EFRON, B. *The jackknife, the bootstrap, and other resampling plans*. [S.l.]: Siam, 1982. v. 38.
- EFRON, B. Better bootstrap confidence intervals. *Journal of the American statistical Association*, Taylor & Francis Group, v. 82, n. 397, p. 171–185, 1987.
- EFRON, B.; TIBSHIRANI, R. J. *An introduction to the bootstrap*. [S.l.]: CRC press, 1994.
- GENTON, M. G. *Skew-elliptical distributions and their applications: a journey beyond normality*. [S.l.]: CRC Press, 2004.
- JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. *Continuous univariate distributions*, wiley. 1995.
- LIO, Y. L.; PARK, C. A bootstrap control chart for birnbaum-saunders percentiles. *Quality and Reliability Engineering International*, Wiley Online Library, v. 24, n. 5, p. 585–600, 2008.
- MAEHARA, R. P.; SANCHEZ, L. B. *bssn: Birnbaum-Saunders Model Based on Skew-Normal Distribution*. [S.l.], 2016. R package version 0.7. Disponível em: (<https://CRAN.R-project.org/package=bssn>).
- MENG, X.-L.; RUBIN, D. B. Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika*, Oxford University Press, v. 80, n. 2, p. 267–278, 1993.

- MONTGOMERY, D. C. *Introdução Ao Controle Estatístico Da Qualidade*. [S.l.]: Grupo Gen-LTC, 2000.
- MOOD, A. M. *Introduction to the Theory of Statistics*. [S.l.: s.n.], 1950.
- NICHOLS, M. D.; PADGETT, W. A bootstrap control chart for weibull percentiles. *Quality and reliability engineering international*, Wiley Online Library, v. 22, n. 2, p. 141–151, 2006.
- PADGETT, W.; SPURRIER, J. D. Shewhart-type charts for percentiles of strength distributions. *Journal of Quality Technology*, Taylor & Francis, v. 22, n. 4, p. 283–288, 1990.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2020. Disponível em: <http://www.R-project.org/>.
- RIZZO, M. L. *Statistical computing with R*. [S.l.]: CRC Press, 2019.
- SANTANA, L.; VILCA, F.; LEIVA, V. Influence analysis in skew-birnbaum–saunders regression models and applications. *Journal of Applied Statistics*, Taylor & Francis, v. 38, n. 8, p. 1633–1649, 2011.
- VILCA, F. et al. An extended birnbaum–saunders model and its application in the study of environmental quality in santiago, chile. *Stochastic Environmental Research and Risk Assessment*, Springer, v. 24, n. 5, p. 771–782, 2010.
- VILCA, F. et al. Estimation of extreme percentiles in birnbaum–saunders distributions. *Computational Statistics & Data Analysis*, Elsevier, v. 55, n. 4, p. 1665–1678, 2011.

A Densidade SNBS

Seja $Z \sim SN(\lambda)$, então a função de distribuição acumulada de T em é dado por

$$\begin{aligned} P(T \leq t) &= P\left(\frac{\beta}{4} \left[\alpha Z + \sqrt{\{\alpha Z\}^2 + 4}\right]^2 \leq t\right) \\ &= P\left(Z \leq \frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right]\right) \end{aligned}$$

portanto, derivando a expressão obtida acima em relação a t , temos que a densidade de T é dada por

$$f_T(t) = 2\phi\left(\frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right]\right) \Phi\left(\lambda \frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right]\right) \frac{t+b}{2\alpha\sqrt{\beta t^3}}, \quad t > 0$$

B Código R

```
##### Função Auxiliares #####
#Desvio padrão das colunas
colSD <- function(data) {
  apply(data, 2, sd,na.rm=T)
}
#Função criada para ordenar as colunas ta tabela ARL
ord <- function(x) {
  z <- 1:(2*x)
  aux <- matrix(z,nrow = x,ncol=2)
  return(c(t(aux)))
}
#####Funções referentes ao IC Bootstrap BCa #####
#coeficiente de correção de viés pro BCa
z0 <- function(th,th0){
  th <- th[>%na.omit]
  z0 <- qnorm(sum(th < th0) / length(th))
  return(z0)
}
# IC bootstrap BCa
boot.BCa.qbssn <- function(tt,th,th0,conf,perc){
  tt <- as.matrix(tt)
  n <- nrow(tt)
  alpha <- (1+c(-conf,conf))/2
  zalpha <- qnorm(alpha)
  #correção
  z <- vector(mode = "list",length=length(perc))
  names(z) <- paste0("z0",perc)
  for (i in 1:length(perc)) {
    z[[i]] <- z0(th[,i],th0[i])
  }
  #jackknife a estimation
  w <- matrix(NA,nrow = n,ncol=length(perc))
  for (i in 1:n) {
    tryCatch({
```

```

    ti <- tt[-i]
    alpha.init <-sqrt(2*sqrt(s(ti)/r(ti)) - 1)
    beta.init <- sqrt(s(ti)*r(ti))
    EMV <- EMbssn(ti,alpha = alpha.init,beta = beta.init,delta =0.5,loglik=F,accuracy = 10e-5,iter.max=5000)
    beta.jack <- EMV[[1]]$beta
    alpha.jack <- EMV[[1]]$alpha
    lambda.jack <- EMV[[1]]$lambda
    w[i,]<- qbssn_rip(p =perc,alpha = alpha.jack,beta = beta.jack,lambda =lambda.jack)
  },error=function(e){cat(conditionMessage(e), "\n")})
}
#fator de aceleração a
L <- colMeans(w)-w
a <- vector(mode = "list",length=length(perc))
names(a) <- paste0("z",perc)
for (j in 1:length(perc)) {
  a[[j]] <- sum(L[,j]^3)/(6 * sum(L[,j]^2)^1.5)
}
# BCa conf. limits
adj.alpha <- vector(mode = "list",length=length(perc))
names(adj.alpha) <- paste0("adj.alpha",perc)
for (j in 1:length(perc)) {
  adj.alpha[[j]] <- pnorm(z[[j]] + ((z[[j]]+zalpha)/(1-a[[j]]*(z[[j]]+zalpha))))
}
#BCa limits
limits <- vector(mode = "list",length=length(perc))
names(limits) <- paste0("w",perc)
for (k in 1:length(perc)) {
  limits[[k]] <- quantile(th[,k], adj.alpha[[k]], na.rm=T)
}
BCa=data.frame(limits)
return(list("BCa"=BCa))
}
#===== Função da simulação =====#
simu <- function(perc,k,n,alpha,beta,lambda,conf,B,MC){
  N <- k*n
  FAR <- (1-conf)
  quantis_limites <- c(FAR/2,1-FAR/2)
  #Data frames auxiliares
  df1 <- matrix(NA,nrow = MC,ncol = 2*length(conf))
  colnames(df1) <- paste0("p",round(quantis_limites,5))
  df2 <- matrix(0,nrow = MC,ncol = length(conf))
  colnames(df2) <- paste0("FAR",round(FAR,5))
  #Percentile
  percentile <- rep(list(df1),length(perc))
  names(percentile) <- paste0("W",perc)
  ARL.p <- rep(list(df2),length(perc))
  names(ARL.p) <- paste0("CMS",perc)
  #BCa
  BCA <- rep(list(df1),length(perc))
  names(BCA) <- paste0("W",perc)
  ARL.BCA <- rep(list(df2),length(perc))
  names(ARL.BCA) <- paste0("CMS",perc)
  for (j in 1:MC) {
    tt <- rbssn(N,beta = beta,alpha = alpha,lambda = lambda)
    alpha.init <-sqrt(2*sqrt(s(tt)/r(tt)) - 1)
    beta.init <- sqrt(s(tt)*r(tt))
    EMV <- EMbssn(tt,alpha = alpha.init,beta = beta.init,
      delta =0,loglik=F,accuracy = 10e-4,iter.max=5000)
  }
}

```



```

print(paste0("j=",j))
beta.hat <- EMV[[1]]$beta
alpha.hat <- EMV[[1]]$alpha
lambda.hat <- EMV[[1]]$lambda
th0=qbssn_rip(perc,alpha = alpha.hat,beta = beta.hat,lambda = lambda.hat)
w <- matrix(NA,nrow = B,ncol = length(perc));colnames(w) <- paste0("W",perc)
for (i in 1:B) {# Bootstrap para obter as réplicas W
  capture.output(tryCatch({
    ti <- rbssn(n,beta = beta.hat,alpha = alpha.hat,lambda = lambda.hat)
    alpha.init <-sqrt(2*sqrt(s(ti)/r(ti)) - 1)
    beta.init <- sqrt(s(ti)*r(ti))
    EMV <- EMbssn(ti,alpha = alpha.init,beta = beta.init,
      delta =0.5,loglik=F,accuracy = 10e-4,iter.max=5000)
    beta.boot <- EMV[[1]]$beta
    alpha.boot <- EMV[[1]]$alpha
    lambda.boot <- EMV[[1]]$lambda
    w[i,]<- qbssn_rip(p =perc,alpha = alpha.boot,beta = beta.boot,lambda =lambda.boot)
  },error=function(e){cat(conditionMessage(e), "\n")}),file = NULL)
}
limites.p <- apply(w, 2, quantile,probs=quantis_limites,na.rm=T)
capture.output(limites.BCa <- boot.BCa.qbssn(tt = tt,th0 = th0,th = w,
  conf = conf,perc = perc)$BCa,file=NULL)
cat("\f")
for (k in 1:length(perc)) {
  percentile[[k]][j,] <- limites.p[,k]
  BCA[[k]][j,] <- limites.BCa[,k]
}
cs <- 1
while (sum(c(first(ARL.BCa,j),first(ARL.p,j)) %in% 0)!=0){
  capture.output(tryCatch({
    ti <- rbssn(n,beta = beta,alpha = alpha,lambda = lambda)
    alpha.init <-sqrt(2*sqrt(s(ti)/r(ti)) - 1)
    beta.init <- sqrt(s(ti)*r(ti))
    EMV <- EMbssn(ti,alpha = alpha.init,beta = beta.init,
      delta =0.5,loglik=F,accuracy = 10e-4,iter.max=5000)
    beta.boot <- EMV[[1]]$beta
    alpha.boot <- EMV[[1]]$alpha
    lambda.boot <- EMV[[1]]$lambda
    w<- qbssn_rip(p =c(.01,.1,.5),alpha = alpha.boot,beta = beta.boot,lambda =lambda.boot)
  },error=function(e){cat(conditionMessage(e), "\n")}),file = NULL)
  for (m in 1:length(perc)) {
    for (z in 1:length(FAR)) {
      #CMS BCa
      if(ARL.BCa[[m]][j,z]==0){
        ARL.BCa[[m]][j,z] <- ifelse(w[m] > BCA[[m]][j,z] &
          w[m] < BCA[[m]][j,z+length(FAR)] | is.na(w[m]),0,cs)
      }
      #CMS Percentile
      if(ARL.p[[m]][j,z]==0){
        ARL.p[[m]][j,z] <- ifelse(w[m] > percentile[[m]][j,z] &
          w[m] < percentile[[m]][j,z+length(FAR)] | is.na(w[m]),0,cs)
      }
    }
  }
  cs <- cs+1
}
}
#resultados percentile tabela

```

```

p.m <- cbind(sapply(percentile, colMeans),alpha=alpha,FAR=rep(FAR,2))
p.sd <- cbind(sapply(percentile, colSD),alpha=alpha,FAR=rep(FAR,2))
aux1 <- sapply(ARL.p, colMeans)
aux2 <- sapply(ARL.p, colSD)
colnames(aux2) <- paste0("sd.",colnames(aux2))
p.arl <- cbind(cbind(aux1,aux2)[,ord(length(perc))],alpha=alpha)
#resultador BCa tabelas
bca.m <- cbind(sapply(BCA, colMeans),alpha=alpha,FAR=rep(FAR,2))
bca.sd <- cbind(sapply(BCA, colSD),alpha=alpha,FAR=rep(FAR,2))
aux1 <- sapply(ARL.BCA, colMeans)
aux2 <- sapply(ARL.BCA, colSD)
colnames(aux2) <- paste0("sd.",colnames(aux2))
bca.arl <- cbind(cbind(aux1,aux2)[,ord(length(perc))],alpha=alpha)
return(list("percentile"=list(limites=p.m,
                             limistes.sd=p.sd,
                             ARL=p.arl),
          "BCA"=list(limites=bca.m,
                    limistes.sd=bca.sd,
                    ARL=bca.arl)))
}

simu_ARL_out <- function(perc,k,n,alpha,beta,lambda,conf,B,MC,shift){
  N <- k*n
  FAR <- (1-conf)
  quantis_limites <- c(FAR/2,1-FAR/2)
  #Data frames auxiliares
  df1 <- matrix(NA,nrow = MC,ncol = 2*length(conf))
  colnames(df1) <- paste0("p",round(quantis_limites,5))
  df2 <- matrix(0,nrow = MC,ncol = length(conf))
  colnames(df2) <- paste0("FAR",round(FAR,5))
  #Percentile
  percentile <- rep(list(df1),length(perc))
  names(percentile) <- paste0("W",perc)
  ARL.p <- rep(list(df2),length(perc))
  names(ARL.p) <- paste0("CMS",perc)
  #BCa
  BCA <- rep(list(df1),length(perc))
  names(BCA) <- paste0("W",perc)
  ARL.BCA <- rep(list(df2),length(perc))
  names(ARL.BCA) <- paste0("CMS",perc)
  for (j in 1:MC) {
    tt <- rbssn(N,beta = beta,alpha = alpha,lambda = lambda)
    alpha.init <-sqrt(2*sqrt(s(tt)/r(tt)) - 1)
    beta.init <- sqrt(s(tt)*r(tt))
    EMV <- EMbssn(tt,alpha = alpha.init,beta = beta.init,
                 delta =0,loglik=F,accuracy = 10e-4,iter.max=5000)
    print(paste0("j=",j))
    beta.hat <- EMV[[1]]$beta
    alpha.hat <- EMV[[1]]$alpha
    lambda.hat <- EMV[[1]]$lambda
    th0=qbssn_rip(perc,alpha = alpha.hat,beta = beta.hat,lambda = lambda.hat)
    w <- matrix(NA,nrow = B,ncol = length(perc));colnames(w) <- paste0("W",perc)
    for (i in 1:B) {# Bootstrap para obter as réplicas W
      capture.output(tryCatch({
        ti <- rbssn(n,beta = beta.hat,alpha = alpha.hat,lambda = lambda.hat)
        alpha.init <-sqrt(2*sqrt(s(ti)/r(ti)) - 1)
        beta.init <- sqrt(s(ti)*r(ti))
        EMV <- EMbssn(ti,alpha = alpha.init,beta = beta.init,

```

```

        delta =0.5,loglik=F,accuracy = 10e-4,iter.max=5000)
    beta.boot <- EMV[[1]]$beta
    alpha.boot <- EMV[[1]]$alpha
    lambda.boot <- EMV[[1]]$lambda
    w[i,]<- qbssn_rip(p =perc,alpha = alpha.boot,beta = beta.boot,lambda =lambda.boot)
  },error=function(e){cat(conditionMessage(e), "\n")},file = NULL)
}
limites.p <- apply(w, 2, quantile,probs=quantis_limites,na.rm=T)
capture.output(limites.BCa <- boot.BCa.qbssn(tt = tt,th0 = th0,th = w,
                                           conf = conf,perc = perc)$BCa,file=NULL)

cat("\f")
for (k in 1:length(perc)) {
  percentile[[k]][j,] <- limites.p[,k]
  BCA[[k]][j,] <- limites.BCa[,k]
}
cs <- 1
while (sum(c(first(ARL.BCa,j),first(ARL.p,j)) %in% 0)!=0){
  capture.output(tryCatch({
    ti <- rbssn(n,beta = beta,alpha = alpha+shift,lambda = lambda)
    alpha.init <-sqrt(2*sqrt(s(ti)/r(ti)) - 1)
    beta.init <- sqrt(s(ti)*r(ti))
    EMV <- EMbssn(ti,alpha = alpha.init,beta = beta.init,
                 delta =0.5,loglik=F,accuracy = 10e-4,iter.max=5000)
    beta.boot <- EMV[[1]]$beta
    alpha.boot <- EMV[[1]]$alpha
    lambda.boot <- EMV[[1]]$lambda
    w<- qbssn_rip(p =c(.01,.1,.5),alpha = alpha.boot,beta = beta.boot,lambda =lambda.boot)
  },error=function(e){cat(conditionMessage(e), "\n")},file = NULL)
  for (m in 1:length(perc)) {
    for (z in 1:length(FAR)) {
      #CMS BCa
      if(ARL.BCa[[m]][j,z]==0){
        ARL.BCa[[m]][j,z] <- ifelse(w[m] > BCA[[m]][j,z] &
                                   w[m] < BCA[[m]][j,z+length(FAR)] | is.na(w[m]),0,cs)
      }
      #CMS Percentile
      if(ARL.p[[m]][j,z]==0){
        ARL.p[[m]][j,z] <- ifelse(w[m] > percentile[[m]][j,z] &
                                   w[m] < percentile[[m]][j,z+length(FAR)] | is.na(w[m]),0,cs)
      }
    }
  }
  cs <- cs+1
}
}

#resultados percentile tabela
p.m <- cbind(sapply(percentile, colMeans),alpha=alpha,FAR=rep(FAR,2))
p.sd <- cbind(sapply(percentile, colSD),alpha=alpha,FAR=rep(FAR,2))
aux1 <- sapply(ARL.p, colMeans)
aux2 <- sapply(ARL.p, colSD)
colnames(aux2) <- paste0("sd.",colnames(aux2))
p.arl <- cbind(cbind(aux1,aux2)[,ord(length(perc))],alpha1=alpha,alpha2=alpha+shift)
#resultador BCa tabelas
bca.m <- cbind(sapply(BCA, colMeans),alpha=alpha,FAR=rep(FAR,2))
bca.sd <- cbind(sapply(BCA, colSD),alpha=alpha,FAR=rep(FAR,2))
aux1 <- sapply(ARL.BCa, colMeans)
aux2 <- sapply(ARL.BCa, colSD)
colnames(aux2) <- paste0("sd.",colnames(aux2))

```

```

bca.ar1 <- cbind(cbind(aux1,aux2)[,ord(length(perc))],alpha1=alpha,alpha2=alpha+shift)
return(list("percentile"=list(limites=p.m,
                             limistes.sd=p.sd,
                             ARL=p.ar1),
          "BCA"=list(limites=bca.m,
                    limistes.sd=bca.sd,
                    ARL=bca.ar1)))
}
#===== Estudo de simulação =====#
r1<- simu(perc=c(.01,.1,.5),
         k=20,
         n=5,
         alpha=.5,
         beta=1,
         lambda=-1,
         conf=c(0.9980,0.9973,0.9900,0.9000),
         B=1000,
         MC=1)
r2<- simu(perc=c(.01,.1,.5),
         k=20,
         n=5,
         alpha=1,
         beta=1,
         lambda=-1,
         conf=c(0.9980,0.9973,0.9900,0.9000),
         B=1000,
         MC=1)
r3<- simu(perc=c(.01,.1,.5),
         k=20,
         n=5,
         alpha=2,
         beta=1,
         lambda=-1,
         conf=c(0.9980,0.9973,0.9900,0.9000),
         B=1000,
         MC=1)
out1<- simu_ARL_out(perc=c(.01,.1,.5),
                  k=20,
                  n=5,
                  alpha=.5,
                  beta=1,
                  lambda=-1,
                  conf=c(0.9980,0.9973,0.9900,0.9000),
                  B=10000,
                  MC=10000,shift = .5)
out2<- simu_ARL_out(perc=c(.01,.1,.5),
                  k=20,
                  n=5,
                  alpha=1,
                  beta=1,
                  lambda=-1,
                  conf=c(0.9980,0.9973,0.9900,0.9000),
                  B=10000,
                  MC=10000,shift = .5)
out3<- simu_ARL_out(perc=c(.01,.1,.5),
                  k=20,
                  n=5,
                  alpha=2,

```

```

        beta=1,
        lambda=-1,
        conf=c(0.9980,0.9973,0.9900,0.9000),
        B=10000,
        MC=10000,shift = .5)
#####
# Resultados percentile
#####
lim_perc <- rbind(r1[["percentile"]][["limites"]],
                 r2[["percentile"]][["limites"]],
                 r3[["percentile"]][["limites"]],
                 r4[["percentile"]][["limites"]])%>%as.data.frame()%>%
  tibble::rownames_to_column("lim")%>%
  gather(key="w",value = "valor",-alpha,-FAR,-lim)%>%
  mutate(lim=ifelse(as.numeric(str_sub(string = lim,start = 2,end = 5))>.5,"LSC","LIC"))%>%
  transmute(lim.perc=paste(lim,w),alpha,FAR,valor)%>%
  dcast(formula = FAR+alpha~lim.perc)%>%
  group_by(FAR)%>%
  arrange(desc(FAR),alpha)%>%
  select(FAR,alpha,"LIC W0.01","LSC W0.01","LIC W0.1","LSC W0.1","LIC W0.5","LSC W0.5")

sd_perc <- rbind(r1[["percentile"]][["limistes.sd"]],
                 r2[["percentile"]][["limistes.sd"]],
                 r3[["percentile"]][["limistes.sd"]],
                 r4[["percentile"]][["limistes.sd"]])%>%as.data.frame()%>%
  tibble::rownames_to_column("lim")%>%
  gather(key="w",value = "valor",-alpha,-FAR,-lim)%>%
  mutate(lim=ifelse(as.numeric(str_sub(string = lim,start = 2,end = 5))>.5,"LSC.sd","LIC.sd"))%>%
  transmute(lim.perc=paste(lim,w),alpha,FAR,valor)%>%
  dcast(formula = FAR+alpha~lim.perc)%>%
  group_by(FAR)%>%
  arrange(desc(FAR),alpha)%>%
  select(FAR,alpha,"LIC.sd W0.01","LSC.sd W0.01","LIC.sd W0.1","LSC.sd W0.1","LIC.sd W0.5","LSC.sd W0.5")

ARL_perc <- rbind(r1[["percentile"]][["ARL"]],
                 r3[["percentile"]][["ARL"]],
                 r4[["percentile"]][["ARL"]])%>%as.data.frame()%>%
  tibble::rownames_to_column("FAR")%>%
  mutate(FAR=as.numeric(str_extract(FAR,pattern = c("0.002","0.0027","0.01","0.1"))))%>%
  group_by(FAR)%>%
  arrange(desc(FAR),alpha)

#####
# Resultados BCa
#####
lim_BCa <- rbind(r1[["BCA"]][["limites"]],
                 r3[["BCA"]][["limites"]],
                 r4[["BCA"]][["limites"]])%>%as.data.frame()%>%
  tibble::rownames_to_column("lim")%>%
  gather(key="w",value = "valor",-alpha,-FAR,-lim)%>%
  mutate(lim=ifelse(as.numeric(str_sub(string = lim,start = 2,end = 5))>.5,"LSC","LIC"))%>%
  transmute(lim.perc=paste(lim,w),alpha,FAR,valor)%>%
  dcast(formula = FAR+alpha~lim.perc)%>%
  group_by(FAR)%>%
  arrange(desc(FAR),alpha)%>%
  select(FAR,alpha,"LIC W0.01","LSC W0.01","LIC W0.1","LSC W0.1","LIC W0.5","LSC W0.5")

sd_BCa <- rbind(r1[["BCA"]][["limistes.sd"]],

```

```

      r2[["BCA"]][["limistes.sd"]],
      r3[["BCA"]][["limistes.sd"]],
      r4[["BCA"]][["limistes.sd"]])%>%as.data.frame()%>%
  tibble::rownames_to_column("lim")%>%
  gather(key="w", value = "valor", -alpha, -FAR, -lim)%>%
  mutate(lim=ifelse(as.numeric(str_sub(string = lim, start = 2, end = 5))>.5, "LSC.sd", "LIC.sd"))%>%
  transmute(lim.perc=paste(lim,w), alpha,FAR,valor)%>%
  dcast(formula = FAR+alpha~lim.perc)%>%
  group_by(FAR)%>%
  arrange(desc(FAR), alpha)%>%
  select(FAR,alpha,"LIC.sd W0.01", "LSC.sd W0.01", "LIC.sd W0.1", "LSC.sd W0.1", "LIC.sd W0.5", "LSC.sd W0.5")

ARL_BCa <- rbind(r1[["BCA"]][["ARL"]],
               r3[["BCA"]][["ARL"]],
               r4[["BCA"]][["ARL"]])%>%as.data.frame()%>%
  tibble::rownames_to_column("FAR")%>%
  mutate(FAR=as.numeric(str_extract(FAR,pattern = c("0.002", "0.0027", "0.01", "0.1"))))%>%
  group_by(FAR)%>%
  arrange(desc(FAR), alpha)
ARL_BCa <- ARL_BCa[c(1,8,2:7)]
#=====
# Resultados ARL out
#=====
ARL_out <- rbind(data.frame(out1$percentile[["ARL"]], type="percentile")%>%tibble::rownames_to_column("FAR"),
                data.frame(out1$BCA[["ARL"]], type="BCa")%>%tibble::rownames_to_column("FAR"),
                data.frame(out2$percentile[["ARL"]], type="percentile")%>%tibble::rownames_to_column("FAR"),
                data.frame(out2$BCA[["ARL"]], type="BCa")%>%tibble::rownames_to_column("FAR"),
                data.frame(out3$percentile[["ARL"]], type="percentile")%>%tibble::rownames_to_column("FAR"),
                data.frame(out3$BCA[["ARL"]], type="BCa")%>%tibble::rownames_to_column("FAR"))%>%
  gather(key="w", value = "valor", -alpha1, -alpha2, -FAR, -type)%>%rowwise()%>%
  mutate(n=str_length(w)-3,
         med=ifelse(str_detect(w,pattern = "sd"), "sd", "ARL"),
         w=as.numeric(str_split(w,pattern =paste0(".",n,"}"),n = 2,simplify = T)[,2]),
         FAR=as.numeric(str_remove(FAR,pattern = "FAR")),
         med1=paste(med,type))%>%
  select(FAR,alpha1,alpha2,w,valor,med1)
tab<- spread(data = ARL_out,key = "med1",value = "valor")%>%group_by(FAR)
tab_final=tab%>%select(w,alpha1,alpha2,'ARL percentile','sd percentile','ARL BCa','sd BCa')%>%
  arrange(desc(FAR))

```

Referências

- AZZALINI, A. *The skew-normal and related families*. [S.l.]: Cambridge University Press, 2013. v. 3.
- BIRNBAUM, Z. W.; SAUNDERS, S. C. A new family of life distributions. *Journal of applied probability*, Cambridge University Press, v. 6, n. 2, p. 319–327, 1969.
- BOLFARINE, H.; SANDOVAL, M. C. *Introdução à inferência estatística*. [S.l.]: SBM, 2001. v. 2.
- BOURGUIGNON, M.; HO, L. L.; FERNANDES, F. H. Control charts for monitoring the median parameter of birnbaum-saunders distribution. *Quality and Reliability Engineering International*, Wiley Online Library, v. 36, n. 4, p. 1333–1363, 2020.

- CASELLA, G.; BERGER, R. L. Inferência estatística. *São Paulo: Cengage Learning*, 2010.
- CHENG, S. W.; XIE, H. Control charts for lognormal data. *Tamkang Journal of Science and Engineering*, TAMKANG UNIVERSITY, v. 3, n. 3, p. 131–138, 2000.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 39, n. 1, p. 1–22, 1977.
- DICICCIO, T. J.; EFRON, B. Bootstrap confidence intervals. *Statistical science*, JSTOR, p. 189–212, 1996.
- EFRON, B. Bootstrap methods: another look at the jackknife. In: *Breakthroughs in statistics*. [S.l.]: Springer, 1979. p. 569–593.
- EFRON, B. *The jackknife, the bootstrap, and other resampling plans*. [S.l.]: Siam, 1982. v. 38.
- EFRON, B. Better bootstrap confidence intervals. *Journal of the American statistical Association*, Taylor & Francis Group, v. 82, n. 397, p. 171–185, 1987.
- EFRON, B.; TIBSHIRANI, R. J. *An introduction to the bootstrap*. [S.l.]: CRC press, 1994.
- GENTON, M. G. *Skew-elliptical distributions and their applications: a journey beyond normality*. [S.l.]: CRC Press, 2004.
- JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. Continuous univariate distributions, wiley. 1995.
- LIO, Y. L.; PARK, C. A bootstrap control chart for birnbaum–saunders percentiles. *Quality and Reliability Engineering International*, Wiley Online Library, v. 24, n. 5, p. 585–600, 2008.
- MAEHARA, R. P.; SANCHEZ, L. B. *bssn: Birnbaum-Saunders Model Based on Skew-Normal Distribution*. [S.l.], 2016. R package version 0.7. Disponível em: <https://CRAN.R-project.org/package=bssn>.
- MENG, X.-L.; RUBIN, D. B. Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika*, Oxford University Press, v. 80, n. 2, p. 267–278, 1993.
- MONTGOMERY, D. C. *Introdução Ao Controle Estatístico Da Qualidade*. [S.l.]: Grupo Gen-LTC, 2000.
- MOOD, A. M. *Introduction to the Theory of Statistics*. [S.l.: s.n.], 1950.
- NICHOLS, M. D.; PADGETT, W. A bootstrap control chart for weibull percentiles. *Quality and reliability engineering international*, Wiley Online Library, v. 22, n. 2, p. 141–151, 2006.
- PADGETT, W.; SPURRIER, J. D. Shewhart-type charts for percentiles of strength distributions. *Journal of Quality Technology*, Taylor & Francis, v. 22, n. 4, p. 283–288, 1990.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2020. Disponível em: <http://www.R-project.org/>.
- RIZZO, M. L. *Statistical computing with R*. [S.l.]: CRC Press, 2019.

SANTANA, L.; VILCA, F.; LEIVA, V. Influence analysis in skew-birnbaum–saunders regression models and applications. *Journal of Applied Statistics*, Taylor & Francis, v. 38, n. 8, p. 1633–1649, 2011.

VILCA, F. et al. An extended birnbaum–saunders model and its application in the study of environmental quality in santiago, chile. *Stochastic Environmental Research and Risk Assessment*, Springer, v. 24, n. 5, p. 771–782, 2010.

VILCA, F. et al. Estimation of extreme percentiles in birnbaum–saunders distributions. *Computational Statistics & Data Analysis*, Elsevier, v. 55, n. 4, p. 1665–1678, 2011.