ANALISANDO O INTERVALO DE CONFIANÇA ASSINTÓTICO DA DISTRIBUIÇÃO EXPONENCIAL A PARTIR DE AMOSTRAS ALEATÓRIAS

Carlos Lisboa Duarte [1]

RESUMO

O presente trabalho tem como objetivo analisar alguns resultados importantes que podemos observar a partir do intervalo de confiança assintótico de uma distribuição exponencial com parâmetro $(1/(\lambda))$, utilizando, para isso, resultados obtidos a partir de amostras aleatórias coletadas por meio de simulações computacionais no software RStudio, vale ressaltar ainda, que as amostras aleatórias serão construídas utilizando réplicas de Monte Carlo com 5000 repetições. Dessa forma, levando-se em consideração a importância do intervalo de confiança em pesquisas estatísticas, bem como a construção do mesmo, o referido estudo procura apresentar de maneira detalhada e simplificada, o passo a passo que é necessário percorrer até a obtenção do intervalo de confiança assintótico para a distribuição exponencial. Assim sendo, o referido trabalho foi construído a partir de uma revisão bibliográfica pautada nas contribuições de autores como: Larsen e Marx, Walpole entre outros. Além disso, neste estudo são apresentadas algumas considerações acerca da avaliação das propriedades do Estimador de Máxima Verossimilhança (EMV) da distribuição exponencial com parâmetro $(1\ /\ (\lambda))$, analisando, computacionalmente, se o estimador obtido é assintoticamente não viesado, assintoticamente consistente e assintoticamente eficiente. Para tal, foram consideradas amostras aleatórias de diferentes tamanhos, com isso foi possível observar que quando aumentamos o tamanho da amostra nas simulações computacionais, obtemos melhores resultados para as propriedades do estimador. E por último, são descritos os resultados acerca das taxas de cobertura e não cobertura do intervalo de confiança construído, considerando, para esse fim, os níveis de confiança $(1-\alpha=95\%)$ e $(1-\alpha=99\%)$, assim, é possível notar que no caso da confiança do intervalo, o aumento do tamanho da amostra aleatória, faz com que a taxa de cobertura do intervalo se aproxime cada vez mais do real valor do nível de confiança considerado.

Palavras-chave: Distribuição exponencial, Estimador de máxima verossimilhança, Intervalo de confiança assintótico, Nível de confiança, Avaliação de estimadores.

INTRODUÇÃO

O referido trabalho abordará o tema do intervalo de confiança assintótico da distribuição de probabilidade exponencial com parâmetro $(1\ /\ (\lambda))$ para amostras aleatórias, bem como, discutirá algumas informações importantes acerca dessa distribuição, tais como: Estimador de Máxima Verossimilhança (EMV), construção do intervalo de confiança assintótico, avaliação do estimador (EMV), utilizando, para isso, simulações computacionais com a linguagem de programação R. Dessa forma, a partir desses dados serão apresentadas conclusões importantes acerca do (EMV) e do intervalo de confiança assintótico dessa distribuição.

[1] Mestrando do Curso Modelagem Matemática e Computacional da Universidade Federal da Paraíba - UFPB, carlos_lisboatf@hotmail.com;

Assim sendo, o trabalho apresenta-se estruturado da seguinte forma, a saber: no primeiro momento será apresentada a metodologia empregada na construção do trabalho; Logo em seguida, no referencial teórico, temos uma breve revisão a respeito das simulações de Monte Carlo, bem como, a fundamentação teórica descrevendo algumas informações acerca da distribuição de probabilidade exponencial e o processo de obtenção do (EMV) dessa distribuição com parâmetro $(1 / (\lambda))$, além da construção do intervalo de confiança assintótico e por fim, as conclusões e considerações observadas a partir das simulações computacionais no *R*, sobre a avaliação do estimador e os resultados da taxa de cobertura e não cobertura do intervalo.

METODOLOGIA

No que concerne aos procedimentos metodológicos, o presente artigo foi fruto de uma revisão bibliográfica sustentada nas contribuições de autores especialistas na área de estatística, mais precisamente, na parte que diz respeito a intervalo de confiança assintótico e simulações computacionais em linguagem de programação *R*. Além disso, foram feitas consultas a materiais disponíveis na internet. Dessa forma, o trabalho teve nas contribuições de autores como: (BOLFARINE; SANDOVAL, 2020); (WALPOLE, 2011); (ASSUNÇÃO, 2017) entre outros, o embasamento teórico necessário ao estudo.

REFERENCIAL TEÓRICO

Simulações de Monte Carlo

No que tange ao entendimento do conceito de simulações, é possível observar que o significado da palavra *simular*, traz considerações acerca de algo que aparenta ser real, mas que não o é, passando a ideia de fingir. Dessa forma, em estudos do campo da ciência dos dados, simulações são procedimentos que buscam imitar o comportamento ou as características de um sistema probabilístico, fazendo para isso, o uso de um gerador de números aleatórios num computador. Assim sendo, este processo é denominado de *simulação Monte Carlo* (Assunção, 2017).

Nas simulações de Monte Carlo, quando o número de informações cresce, o número de previsões também aumenta, esse fato nos permite projetar resultados com maior precisão. Por exemplo, ao calcular a probabilidade do lançamento de dois dados não-viciados, temos 36 combinações de possíveis resultados. Assim, podemos calcular manualmente a probabilidade de um resultado específico. Porém, com a Simulação de Monte Carlo, é possível simular um número alto de jogadas para obter previsões mais precisas para um experimento aleatório.

De forma resumida, temos ainda que as simulações de Monte Carlo, configuram-se como sendo

> [⋯] um procedimento numérico que utiliza números aleatórios para a obtenção de valores não necessariamente aleatórios, com base na Lei dos Grandes Números e no Teorema do Limite Central. Este método tem como princípio a geração de números aleatórios de qualquer distribuição de probabilidade com o objetivo de avaliar de forma numérica, indireta ou artificialmente um modelo matemático que permite estimar o comportamento de um sistema ou processo que envolve variáveis estocásticas (PAIXÃO, p. 2, 2021).

De modo específico, no trabalho foram utilizadas réplicas de Monte Carlo com 5000 repetições para amostras aleatórias de tamanhos 50, 100 e 200.

Estimador de Máxima Verossimilhança da Distribuição Exponencial

A distribuição exponencial está ligada ao estudo de variáveis aleatórias contínuas, que procuram analisar, por exemplo, o tempo que se gasta para realizar uma dada atividade ou tempo da vida útil de algum aparelho. Dessa forma, podemos ainda mencionar que a função densidade da exponencial, pode ser "aplicada para modelar: o tempo de vida de componentes que falham sem efeito de idade; o tempo de espera entre sucessivas chegadas de fótons; emissões de elétrons de um cátodo; chegadas de consumidores e duração de chamadas telefônicas, entre outros"(CAMPOS; RÊGO; MENDONÇA, 2012, p. 193).

Vale destacar, que para a construção do referencial teórico foram consultados alguns autores e materiais, que proporcionaram ao estudo a fundamentação teórica necessária, tais como: (BONFARINE; SANDOVAL, 2020) e (WALPOLE, 2011) entre outros.

Assim, seja $\{X_1, X_2, \cdots, X_n\}$ uma amostra aleatória de uma população X que segue distribuição exponencial com parâmetro $\frac{1}{\lambda}$, isto é, $X \sim EXP\left(\frac{1}{\lambda}\right)$. Então, vamos obter o Estimador de Máxima Verossimilhança (EMV) para o parâmetro $\frac{1}{\lambda}$ da função densidade da distribuição exponencial. Então, dado que a densidade da exponencial é

$$f(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \tag{1}$$

com $0 \leq x < \infty$ e $\lambda > 0$. Teremos que a função de máxima verossimilhança para (1) será obtida por

$$L\left(\frac{1}{\lambda}\right) = \prod_{i=1}^{n} f(x_i; \lambda)$$

$$L\left(\frac{1}{\lambda}\right) = \prod_{i=1}^{n} \frac{1}{\lambda} e^{-\frac{x_i}{\lambda}}$$

Agora, analisando a distribuição conjunta das variáveis, temos que

$$L\left(\frac{1}{\lambda}\right) = \frac{1}{\lambda} e^{-\frac{x_1}{\lambda}} \times \frac{1}{\lambda} e^{-\frac{x_2}{\lambda}} \times \ldots \times \frac{1}{\lambda} e^{-\frac{x_n}{\lambda}}$$

Daí, teremos que

$$L\left(\frac{1}{\lambda}\right) = \lambda^{-n} e^{-\frac{1}{\lambda}\sum_{i=1}^{n} x_i} \qquad (2)$$

A seguir aplicando (ln) em ambos os lados de (2), teremos

$$\ln\left[L\left(\frac{1}{\lambda}\right)\right] = \ln\left[\lambda^{-n} e^{-\frac{1}{\lambda}\sum_{i=1}^{n} x_i}\right]$$

Aplicando as propriedades de logaritmos,

$$\ln\left[L\left(\frac{1}{\lambda}\right)\right] = -n\ln(\lambda) - \frac{1}{\lambda}\sum_{i=1}^{n} x_i \qquad (3)$$

Por sua vez, iremos calculando a 1ª derivada de (3) em relação a λ. Dessa forma,

$$\frac{\partial \ln\left[L\left(\frac{1}{\lambda}\right)\right]}{\partial \lambda} = -\frac{n}{\lambda} + \frac{1}{\lambda^2}\sum_{i=1}^{n} x_i \qquad (4)$$

Como o interesse é obter um estimador que maximize o parâmetro λ, calculando a 2ª derivada de (3) e analisando o resultado, é possível observar que a mesma é menor que zero, isto é, um valor de máximo.

Por fim, igualando a equação (4) a zero, poderemos isolar o parâmetro λ e, assim, obter o (EMV) $\hat{\lambda}$. Logo, teremos que

$$-\frac{n}{\lambda} + \frac{1}{\lambda^2}\sum_{i=1}^{n} x_i = 0$$

$$\frac{1}{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}$$

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n}$$

resultando em

$$\hat{\lambda} = \overline{X}.$$

Portanto, podemos concluir o estimador de máxima verossimilhança para a distribuição exponencial com parâmetro $\frac{1}{\lambda}$, será o $\overline{X}$.

Construindo o Intervalo de Confiança Assintótico

A construção de um intervalo de confiança permite analisar a possibilidade de erro em uma amostragem. Considerando que os dados da amostra são variáveis aleatórias, o intervalo de confiança nem sempre apresenta o valor real do parâmetro. Entretanto, podemos obter a probabilidade com que um parâmetro é observado em um intervalo. Por exemplo, se consideramos uma confiança de 95% ,podemos afirmar que, em 95% dos experimentos, o intervalo abrange o valor verdadeiro do parâmetro.

Para construir o intervalo de confiança da distribuição exponencial com parâmetro $\frac{1}{\lambda}$ e com estimador de máxima verossimilhança $\hat{\lambda} = \overline{X}$, utilizaremos o conceito de intervalo de confiança assintótico (aproximado). Dessa forma, considerando uma amostra aleatória $\{x_1, x_2, \cdots, x_n\}$, da variável aleatória X com função densidade $f\left(x; \frac{1}{\lambda}\right)$, temos que para o caso de uma amostra grande, podemos aplicar algumas propriedades assintóticas do estimador de máxima verossimilhança, daí teremos que:

$$\frac{\hat{\lambda} - E(\hat{\lambda})}{\sqrt{var(\hat{\lambda})}} \sim \mathcal{N}(0,1), n \to \infty.$$

Então, teremos que a quantidade pivotal será dada por

$$\frac{\hat{\lambda} - E(\hat{\lambda})}{\sqrt{var(\hat{\lambda})}}. \tag{5}$$

Para calcularmos $var(\hat{\lambda})$, iremos utilizar o fato que variância do estimador é igual ao inverso da informação de Fisher. Assim sendo, temos que a medida da informação de Fisher pode ser calculada por

$$IF(\lambda) = -E\left[\frac{\partial^2 \ln\left[L\left(\frac{1}{\lambda}\right)\right]}{\partial \lambda^2}\right].$$

daí, teremos

$$IF(\lambda) = -E\left[\frac{n}{\lambda^2} - \frac{2}{\lambda^3}\sum_{i=1}^{n} x_i\right]. \tag{6}$$

aplicando as propriedades de esperança em (6), temos

$$IF(\lambda) = -E\left[\frac{n}{\lambda^2}\right] + E\left[\frac{2}{\lambda^3}\sum_{i=1}^{n} x_i\right] = -\frac{n}{\lambda^2} + \frac{2}{\lambda^3}\; E\left[\sum_{i=1}^{n} x_i\right].$$

$$IF(\lambda) = -\frac{n}{\lambda^2} + \frac{2}{\lambda^3} nE[X] = -\frac{n}{\lambda^2} + \frac{2n\lambda}{\lambda^3}$$

$$IF(\lambda) = \frac{n}{\lambda^2}. \tag{7}$$

Então, o inverso da informação de Fisher obtida a partir de (7), será igual a

$$[IF(\lambda)]^{-1} = \frac{\lambda^2}{n}.$$

Assim, considerando que $\hat{\lambda} = \overline{X}$, $E(\hat{\lambda}) = \lambda$ e dado que $var(\hat{\lambda}) = [IF(\lambda)]^{-1} = \frac{\lambda^2}{n}$, e que pelo princípio da invariância, $var(\hat{\lambda}) = \frac{\hat{\lambda}^2}{n}$, teremos então, que a quantidade pivotal em (5), será

$$\frac{\overline{X}-\lambda}{\sqrt{\frac{\hat{\lambda}^2}{n}}}. \tag{8}$$

Logo, a partir da quantidade pivotal (8), podemos construir o intervalo de confiança assintótico $(1-\alpha)100\%$ da forma

$$1-\alpha = P\left(-z_{\frac{\alpha}{2}} \leq \frac{\overline{X}-\lambda}{\sqrt{\frac{\hat{\lambda}^2}{n}}} \leq z_{\frac{\alpha}{2}}\right), \tag{9}$$

isolando λ na equação (9), temos o intervalo de confiança assintótico $(1-\alpha)100\%$ para a distribuição exponencial com parâmetro $\frac{1}{\lambda}$ será dado por

$$\left[\overline{X} - \frac{\overline{X}z_{\frac{\alpha}{2}}}{\sqrt{n}}, \overline{X} + \frac{\overline{X}z_{\frac{\alpha}{2}}}{\sqrt{n}}\right].$$

RESULTADOS E DISCUSSÃO

Avaliando os Estimadores

Agora, serão apresentados os resultados observados a partir das simulações de Monte Carlo, para então avaliar as propriedades do Estimador de Máxima Verossimilhança da distribuição exponencial com parâmetro $\frac{1}{\lambda}$.

Vale ressaltar que, as simulações de Monte Carlo foram feitas considerando 5000 repetições para amostras aleatórias de tamanhos $n = 50$, $n = 100$ e $n = 200$.

Tabela 1: Avaliação dos Estimadores com MC = 5000 e n = 50.

n	*lambda.v*	*lambda.hat*	*vies*	*var.est*	*EQM*
50.00	2.00	1.99423	-0.00576	0.07767	0.07767

Fonte: Elaborada pela autor.

Tabela 2: Avaliação dos Estimadores com MC = 5000 e n = 100.

n	*lambda.v*	*lambda.hat*	*vies*	*var.est*	*EQM*
100.00	2.00	1.99993	-0.00007	0.04178	0.04178

Fonte: Elaborada pela autor.

Tabela 3: Avaliação dos Estimadores com MC = 5000 e n = 200.

n	*lambda.v*	*lambda.hat*	*vies*	*var.est*	*EQM*
200.00	2.00	1.99991	-0.00009	0.02027	0.02027

Fonte: Elaborada pela autor.

onde

- *n:* Tamanho da amostra aleatória;
- *lambda.v:* Valor verdadeiro do parâmetro;
- *lambda.hat:* Valor estimado para o parâmetro (lambda chapéu);
- *var.est:* Valor estimado para a variância do parâmetro;
- *vies:* Valor do viés do estimador;
- *EQM:* Valor do erro quadrático médio.

Analisando os dados obtidos a partir das simulações no *R*, é possível observar que os parâmetros utilizados para avaliar os estimadores, tendem a ter melhores resultados à medida que aumentamos o tamanho da amostra. Por exemplo, nos dados apresentados nas tabelas o objeto *lambda.hat* apresenta a média das estimativas para o parâmetro da distribuição a partir das réplicas de Monte Carlo, e foi possível constatar que essa média tende a aproximar-se do valor real *lambda.v = 2*, quando aumentamos o tamanho da amostra.

Contudo, devido aos critérios de arredondamento nos dados das tabelas, não foi possível inserir um número maior de casas decimais nos resultados, porém, os resultados observados para *lambda.hat* com uma amostra n = 200, foi aquele que mais se aproximou do valor real de *lambda.v*. Dessa forma, podemos dizer que o estimador $\hat{\lambda}$ é *assintoticamente não viesado*.

Uma outra informação importante que colabora para a obtenção de melhores resultados dos estimadores, é o *vies*. Essa informação está armazenada no objeto *vies* das tabelas, e têm-se que, quanto menor o seu valor, melhor serão os resultados do estimador. Já que o mesmo, é utilizado para calcular o Erro Quadrático Médio (*EQM*), métrica que avalia a consistência do estimador. Então, dizemos que um estimador é *consistente* quando o limite do *EQM* com $(n \to 0)$ for igual a zero, e pelos resultados apresentados, percebemos que ao aumentar o tamanho da amostra o *EQM* aproxima-se cada vez mais de zero. Podendo afirmar, assim, que o estimador $\hat{\lambda}$ é *assintoticamente consistente*.

No que diz respeito os critério de eficiência, podemos dizer que um (EMV) é *eficiente*, quando a variância do estimador é igual ao inverso da informação Fisher. Além disso, caso não possamos ter acesso a informação de Fisher, podemos analisar a eficiência do estimador a partir dos resultados de sua variância. Assim sendo, dados dois estimadores não viesados $\hat{\lambda}_1$ e $\hat{\lambda}_2$, o estimador $\hat{\lambda}_1$ será mais eficiente que $\hat{\lambda}_2$, se :

$$var(\hat{\lambda}_1) < var(\hat{\lambda}_2)$$

Portanto, quando temos que um estimador cuja a variância esteja próxima de zero, o critério de eficiência também será satisfeito. Dessa maneira, observando os dados das tabelas 1, 2 e 3, notamos que o valor da variância do estimador está se aproximando de zero, à medida que aumentamos o tamanho da amostra, assim, poderemos dizer que o (EMV) para a distribuição exponencial com parâmetro $\frac{1}{\lambda}$ será *assintoticamente eficiente*.

Analisando as Taxas de Cobertura e não Cobertura do Intervalo de Confiança (IC)

A seguir iremos analisar os valores simulados para a taxa de cobertura e não cobertura do intervalo de confiança da distribuição exponencial com parâmetro $\frac{1}{\lambda}$, utilizando o estimador de máxima verossimilhança e considerando os níveis de confiança iguais a $(1-\alpha = 95\%)$ e $(1-\alpha = 99\%)$. Dessa forma, com o objetivo de embasar teoricamente as conclusões que serão apresentadas, foram feitas consultas a autores da área, como: (LARSEN; MARX, 2012) e (COHEN; COHEN, 2008).

CASO I: Níveis de confiança igual a 95% e 99% para uma amostra aleatória de tamanho 50:

Tabela 4: Dados das taxas de cobertura e não cobertura com n = 50 e $\alpha = 5\%$.

Nível confiança	*Taxa cobertura*	*Taxa não de cobertura*	*TB-Inferior*	*TB-Superior*
95.00	94.50	5.50	0.64	4.86

Fonte: Elaborada pela autor.

Tabela 5: Dados das taxas de cobertura e não cobertura com n = 50 e $\alpha = 1\%$.

Nível confiança	*Taxa cobertura*	*Taxa não de cobertura*	*TB-Inferior*	*TB-Superior*
99.00	97.94	2.06	0.02	2.04

Fonte: Elaborada pela autor.

onde

- *TB-Inferior*: Valores verdadeiros do parâmetro que ficaram fora da taxa de cobertura e abaixo do limite inferior do intervalo;
- *TB-Superior*: Valores verdadeiros do parâmetro que ficaram fora da taxa de cobertura e acima do limite superior do intervalo.

Uma informação relevante apresentada nas tabelas, dizer respeito ao percentual de valores verdadeiros do parâmetro que ficaram fora do intervalo de confiança. Assim, podemos ver, por exemplo, que quando aplicamos um nível de confiança igual a 99%, temos que 0,02% dos valores verdadeiros do parâmetro ficaram abaixo do limite inferior do intervalo, e 2,04% ficaram acima do limite superior. Então, podemos concluir que os percentuais dos valores que se encontra fora do intervalo de confiança, podem variar em relação aos que ficam acima ou abaixo do intervalo. Vale evidenciar, que esse comportamento dos valores que ficam fora do intervalo, também acontece para os demais casos de nível de confiança adotado.

CASO II: Níveis de confiança igual a 95% e 99% para uma amostra aleatória de tamanho 100:

Tabela 6: Dados das taxas de cobertura e não cobertura com n = 100 e $\alpha = 5\%$.

Nível confiança	*Taxa cobertura*	*Taxa não de cobertura*	*TB-Inferior*	*TB-Superior*
95.00	94.34	5.66	1.14	4.52

Fonte: Elaborada pela autor.

Tabela 7: Dados das taxas de cobertura e não cobertura com n = 100 e $\alpha = 1\%$.

Nível confiança	*Taxa cobertura*	*Taxa não de cobertura*	*TB-Inferior*	*TB-Superior*
99.00	98.50	1.50	0.02	1.48

Fonte: Elaborada pela autor.

Note que, se compararmos os valores dos níveis de confiança da amostra n = 50 com os valores da amostra n = 100, iremos observar uma leve oscilação da taxa de cobertura do intervalo para o nível de confiança 95%. Contudo, os resultados para o nível de confiança 99% da amostra de tamanho n = 100, foram melhores do que os valores da amostra n = 50 com o mesmo intervalo de confiança.

Assim sendo, o aumento no tamanho da amostra para um mesmo nível de confiança, pode fazer com que os valores da taxas de cobertura sofram oscilações, mas é possível perceber que essas diferenças mantem ainda a taxa de cobertura aproxima do nível de confiança adotado.

CASO III: Níveis de confiança igual a 95% e 99% para uma amostra aleatória de tamanho 200:

Tabela 8: Dados das taxas de cobertura e não cobertura com n = 200 e $\alpha = 5\%$.

Nível confiança	*Taxa cobertura*	*Taxa não de cobertura*	*TB-Inferior*	*TB-Superior*
95.00	94.78	5.22	1.42	3.80

Fonte: Elaborada pela autor.

Tabela 9: Dados das taxas de cobertura e não cobertura com n = 200 e $\alpha = 1\%$.

Nível confiança	*Taxa cobertura*	*Taxa não de cobertura*	*TB-Inferior*	*TB-Superior*
99.00	98.74	1.26	0.18	1.08

Fonte: Elaborada pela autor.

Por fim, analisando os dados obtidos para os intervalos de confiança de uma amostra aleatória de tamanho n = 200, é possível notar que em comparação com os valores do CASO II, tivemos uma melhora para todos os níveis de confiança, ou seja, as taxas de cobertura e não cobertura do CASO III, foram melhores do que as taxas do CASO II. Então, podemos notar um princípio importante que ocorre em simulações como essas, que o fato do aumento no tamanho da amostra aleatória proporcionar melhores resultados para o estudo, e esse comportamento de melhora dos parâmetros continua para amostras maiores que 200.

CONSIDERAÇÕES FINAIS

Portanto, do ponto de vista estatístico, o intervalo de confiança assintótico de uma distribuição de probabilidade para amostras de variáveis aleatórias, é uma ferramenta extramamente útil, quando temos como objetivo analisar a confiabilidade dos dados que estão sendo investigados para uma população a partir dessas amostras. Neste sentido, temos que a aplicação de simulações computacionais a partir da linguagem de programação *R*, por exemplo, apresenta-se como uma saída viável para realizar o processamento de dados como estes, principalmente, tratando-se de um conjunto grande informações.

REFERÊNCIAS

ASSUNÇÃO, Renato. **Fundamentos Estatísticos de Ciência dos Dados voltado para aplicações**. Published by publisher in BOOK-WEBSITE.COM, 2017. Disponível em: < https://homepages.dcc.ufmg.br/~assuncao/EstatCC/FECD.pdf >. Acesso em: 15 mai. 2022.

BOLFARINE, Heleno. SANDOVAL, Mônica Carneiro. **INTRODUÇÃO À INFERÊNCIA ESTATÍSTICA** - Departamento de Estatística do Instituto de Matemática e Estatística da Universidade de São Paulo - USP, 2020. Disponível em: < https://docs.ufpr.br/~lucambio/CE085/1S2017/LIVRO_Bolfarine_Sandoval.pdf >. Acesso em: 15 mai. 2022.

CAMPOS, Marcilia Andrade. RÊGO, Leandro Chaves. MENDONÇA, Andre Feitoza de. **Métodos Probabilísticos e Estatísticos com Aplicações em Engenharias e Ciências Exatas** - UFPE, 2012.Disponível em:

<https://www.cin.ufpe.br/~aacf/Probabilidade%20e%20Estatistica%20(IF971)/LivroProbabilidade2012.pdf>. Acesso em: 15 mai. 2022.

COHEN, Yosef. COHEN, Jeremiah Y. **Statistics and Data with R: An applied approach through examples**. 1ª ed. John Wiley & Sons Ltd, 2008.

LARSEN, Richard J; MARX, Morris L. **An introduction to mathematical statistics and its applications**. 5ª ed. Pearson Education, 2012.

PAIXÃO, Joelson Lopes da; et al. **Métodos Matemáticos de Modelagem e Otimização: Teoria e Aplicações do Método de Monte Carlo**. In: Jornada de Pesquisa, 26, 2021, Evento Online. Anais. Salão do Conhecimento Unijuí, 2021. Disponível em: <https://publicacoeseventos.unijui.edu.br/index.php/salaoconhecimento/article/view/20767/19478>. Acesso em: 15 mai. 2022.

WALPOLE, Ronald E; et al. **Probability & statistics for engineers & scientists**. Pearson Prentice Hall, 2011.