

Análise da influência das condições meteorológicas e desempenho do ser humano em ocorrências aeronáuticas na aviação civil do Brasil e suas regiões.

Adailson J. T. Silva¹, Bruno L. Maia¹,
Caroline G. Cavalcante¹, Jadeilson J. R. Campos¹, João J. A. S. D'Amorim¹

¹Departamento de Informática e Estatística
Universidade Federal Rural de Pernambuco (UFRPE)
Pernambuco – PE – Brazil

$\left\{ \begin{array}{l} \text{adailson.jose.tavares, brunolinsm,} \\ \text{carol.gcavalcante, jadei, joao.damorim} \end{array} \right\}@gmail.com$

Abstract. *The objective of the study was to analyze the data provided by the Center for Research and Prevention of Accidents Aeronautics - (CENIPA) in order to generate results and samples, analyzing them and obtaining specific conclusions about the most influential factors in aeronautical occurrences in Brazil and its regions, raising the following hypotheses: meteorological conditions and human performance influence or not influence in civil aviation occurrences in Brazil and its regions.*

Resumo. *O objetivo do trabalho foi analisar os dados disponibilizados pelo Centro de Investigação e Prevenção de Acidentes Aeronáuticos - (CENIPA) com intuito de gerar resultados e amostras, analisando-os e obter conclusões específicas sobre os fatores mais influentes nas ocorrências aeronáuticas no Brasil e suas regiões, levantando as seguintes hipóteses: condições meteorológicas e desempenho do ser humano influenciam ou não influenciam nas ocorrências aeronáuticas civis no Brasil e suas regiões.*

Palavras-chave: *Data Mining, Data Science, Pandas, Teste de Hipótese, Aviação, Ocorrências Aeronáuticas, Condições Meteorológicas.*

1. Introdução

Na aviação é empregada tecnologias bastantes avançadas, tornando-a um dos meios de transporte mais seguros existentes no mundo. “...em 2017, foi transportado um total de 112,5 milhões de passageiros pagos no país, 90,6 milhões em voos domésticos e 21,8 milhões em voos internacionais.”(Anuário do Transporte Aéreo, ANAC, 2017). Porém, no mesmo ano foram registrados casos de acidentes e incidentes aéreos, ou seja, mesmo com tamanha segurança, a aviação não está imune ao desempenho do ser humano, condições meteorológicas e outros fatores.

Entre os fatores contribuintes de acidentes aeronáuticos analisados pelo CENIPA, o desempenho do ser humano é o mais frequente, sendo explicitado nesse trabalho através de testes de hipóteses. Essa análise foi gerada com o objetivo de identificar, o nível de criticidade do desempenho do ser humano e das condições meteorológicas em ocorrências

aeronáuticas no Brasil e suas regiões, a fim de promover resultados para auxiliar decisões tomadas na aviação civil brasileira, com o objetivo de diminuir a quantidade das ocorrências aeronáuticas na aviação civil brasileira.

Uma hipótese foi levantada inicialmente, antes da análise dos dados, "As condições meteorológicas influenciam nas ocorrências aeronáuticas civis do Brasil e suas regiões.". A partir dessa hipótese as análises dos dados foram realizadas, a fim de aceitar ou rejeitar essa afirmação, através da análise dos resultados obtidos por técnicas, como a Mineração de Dados, *Data Science* e testes de hipóteses estatísticos (t-Student).

Uma das técnicas mais utilizadas é a Mineração de Dados (MD), sendo esta um processo que tem o objetivo em extrair padrões de bases de dados. Esses padrões não são observados facilmente, visto que os dados estão desorganizados e seu volume nem sempre é possível ser analisado diretamente por seres humanos, sendo necessário o uso de técnicas de MD para se obter padrões em relação a um tipo de negócio. Uma das técnicas aplicadas é a normalização dos dados com o propósito de melhorar a qualidade dos dados; a etapa adiante é processar com algoritmos os dados já tratados. A intenção dessa última etapa é descobrir determinados padrões que podem auxiliar na análise negócio.

Com o uso do *Data Science* pode-se estudar as informações e seus processos de captura, transformação, e posteriormente realizar à análise dos dados, utilizando-se de disciplinas como Computação, Estatística e Matemática. Proporcionando *insights* para novas soluções, para os problemas levantados sobre o negócio.

Dado o exposto, aplica-se neste trabalho, pesquisa dos dados abertos do CENIPA e condições meteorológicas, análise dos dados utilizando a linguagem de programação Python com suas bibliotecas e teste T-Student, para aceitar ou rejeitar as duas hipóteses nulas levantadas no trabalho.

1.1. Trabalhos correlatos

1º Como as Condições Meteorológicas Impactam a Aviação e Quais Medidas São Adotadas Para Reduzir Seus Efeitos: "O objetivo deste trabalho é apresentar como as condições meteorológicas impactam a aviação e quais medidas são por ela adotadas para reduzir os efeitos negativos."[1]

2º Tomada de decisão em condições meteorológicas adversas: "o presente artigo procura demonstrar como os avanços tecnológicos contribuíram no julgamento e no processo de decisão dos pilotos em condições meteorológicas adversas e a consequência de todos esses fatores na segurança de voo. A metodologia utilizada para tal fim se dá por meio de pesquisa bibliográfica e documental que demonstra os avanços na área de segurança de voo oriundos dos estudos de fatores humanos, das condições adversas de tempo e dos avanços tecnológicos e os efeitos destes nas tecnologias de radares meteorológicos embarcados, na tomada de decisão em condições meteorológicas adversas."[2]

3º Fatores contribuintes aos acidentes aeronáuticos: "O objetivo do presente estudo foi comparar os resultados de investigações de acidentes aeronáuticos brasileiros da aviação geral pelo Cenipa com o modelo (Human Factors Analysis and Classification System – HFACS)."[3]

4º Caracterização dos Fatores Contribuintes em Acidentes de Pequenas Aeronaves da Aviação Brasileira Utilizando Inteligência Artificial: O objetivo do estudo

foi: "Identificar padrões nos acidentes da aviação geral entre os anos de 2005 e 2013."[4]

Em resumo, os trabalhos correlatos aqui explicitados, evidenciam a importância da prevenção de acidentes aéreos com o foco no fator humano, o qual é destacado nos trabalhos como sendo um dos maiores fatores influenciadores de acidentes aéreos. Em nossa pesquisa, também objetivamos comprovar que os fatores humanos estão entre os principais causadores de acidentes aéreos através de testes de hipóteses, não somente com revisão bibliográfica e análise parcial de dados, como foi feito na grande maioria dos trabalhos citados nesta pesquisa. Além disso, levantou-se outro viés, que os fatores meteorológicos tem significância em relação aos acidentes aéreos. Esse viés foi observado em um dos trabalhos, porém com enfoque apenas em revisão bibliográfica.

2. Referencial Teórico

2.1. Metodologia

Para o desenvolvimento do atual trabalho, visando a coleta, tratamento e análise dos dados, foram utilizadas as ferramentas *Python* e *Pandas*, para os métodos foram utilizados mineração de dados, *Data Science* (ciência de dados), os quais serão detalhados a seguir:

- *Pandas* é uma biblioteca de *software* escrita para a linguagem de programação *Python* para manipulação e análise de dados. Em particular, oferece estruturas de dados e operações para manipular tabelas numéricas e séries temporais.
- Ciência de dados é uma área interdisciplinar voltada para o estudo e a análise de dados, estruturados ou não, que visa a extração de conhecimento ou *insights* para possíveis tomadas de decisão, de maneira similar à mineração de dados.
- Mineração de dados é o processo de explorar grandes quantidades de dados à procura de padrões consistentes, como regras de associação ou sequências temporais, para detectar relacionamentos sistemáticos entre variáveis, detectando assim novos subconjuntos de dados.
- Teste de hipóteses, teste estatístico ou teste de significância é um procedimento estatístico que permite tomar uma decisão (aceitar ou rejeitar a hipótese nula H_0) entre duas ou mais hipóteses (hipótese nula H_0 ou hipótese alternativa H_1 , utilizando os dados observados de um determinado experimento).[5]
- As condições meteorológicas definem o estado instantâneo da atmosfera ou a sequência temporal de estados da atmosfera. O comportamento da atmosfera num dado local pode ser descrito por um número de quantidades que caracterizam o estado físico do ar, tais como a temperatura, pressão, conteúdo em umidade, movimento, etc.
- Uma biblioteca em linguagens de programação consiste em um conjunto de arquivos que contém funções que podem ser incorporadas a outros programas (Introdução a programação, UFPB, 2014).

2.2. Experimentos

Os dados utilizados para os experimentos foram obtidos da base de dados Opendata AIG Brazil - Ocorrências Aeronáuticas na Aviação Civil Brasileira - 2008 a 2017 - Gerenciada pelo CENIPA, essa base é de dados abertos, podendo ser facilmente acessada em, <http://dados.gov.br/dataset/ocorrencias-aeronauticas-da-aviacao-civil-brasileira>, os dados estão devidamente documentados e salvos em arquivos do tipo .csv, também há um dicionário explicando a relação entre os dados, que está no formato .pdf.

Antes da análise e experimentação dos dados, algumas hipóteses foram levantadas para continuidade do nosso trabalho, são elas: **“As condições meteorológicas influenciam nas ocorrências aeronáuticas na aviação civil do Brasil e suas regiões?”** e **“O desempenho humano influencia nas ocorrências aeronáuticas na aviação civil do Brasil e suas regiões?”**, partindo dessas hipóteses, processamos e analisamos os dados para que as hipóteses fossem confirmadas ou refutadas.

Os dados foram processados através de uma biblioteca do *Python* chamada **Pandas**. Esta biblioteca fornece ferramentas para análise e estruturas de dados com alta performance, juntamente com as operações para manipular tabelas numéricas e séries temporais. Por ser de fácil utilização, tornou-se referência na área de análise de dados até o presente momento da escrita desse artigo.

Em conjunto com a biblioteca *Pandas*, foram utilizados outras bibliotecas *Python*, para tratamento dos dados e cálculos matemáticos, utiliza-se *Numpy*; geoprocessamento e georreferenciamento de dados, utiliza-se *Folium*; otimização e integração numérica, utiliza-se *Scipy*; geração e plotagem de gráficos, utiliza-se *Matplotlib* e *Seaborn*.

Buscou-se, separar os dados por nome dos fatores, como por exemplo: julgamento de pilotagem, condições meteorológicas, etc. Ao todo são 55 fatores, onde condições meteorológicas representa um desses e desempenho do ser humano representa diversos outros fatores.

2.3. Análise do Experimento

O Experimento foi realizado, a fim de obter-se respostas diante das hipóteses criadas. O trabalho foi realizado em três etapas, na primeira etapa foi a obtenção dos dados do CENIPA sobre ocorrências aeronáuticas civis no Brasil e sua análise sobre os fatores que mais influenciam nas ocorrências e modelos de aeronaves com mais ocorrências, limitando aos dados do sudeste do Brasil. Na segunda etapa, foi realizado o estudo mais aprofundado do fator condições meteorológicas, sendo relacionada uma nova base de dados com as estações meteorológicas do Sudeste, identificando a estação mais próxima de cada ocorrência com fator de condições meteorológicas no Sudeste, levando ao estudo mais detalhado sobre esse determinado fator.

Analisando os dados meteorológicos, vimos que não existia um padrão nas medidas mais influentes, como dados do vento, umidade e temperatura. Foi levantada uma nova proposta, para descartar essa nova base com as estações meteorológicas e realizar um estudo mais específico em cima das bases do CENIPA. Na terceira e última etapa, foram criadas duas hipóteses baseadas nos dados obtidos, são elas: **“As condições meteorológicas influenciam nas ocorrências aeronáuticas na aviação civil do Brasil e suas regiões?”** e **“O desempenho humano influencia nas ocorrências aeronáuticas na**

aviação civil do Brasil e suas regiões? ”, partindo dessas hipóteses, processamos e analisamos os dados para que as hipóteses fossem confirmadas ou refutadas. A seguir estão as informações e testes obtidos da terceira etapa do trabalho, com os testes de hipóteses realizados para o Brasil e suas regiões. Foi usado o teste T-Student para confirmar as hipóteses propostas, através dos valores do p-value gerados, mediante os dados obtidos das análises.

- Condições Meteorológicas (Brasil)

- Hipótese nula: As condições meteorológicas não influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA no Brasil.
- Hipótese alternativa: As condições meteorológicas influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA no Brasil.

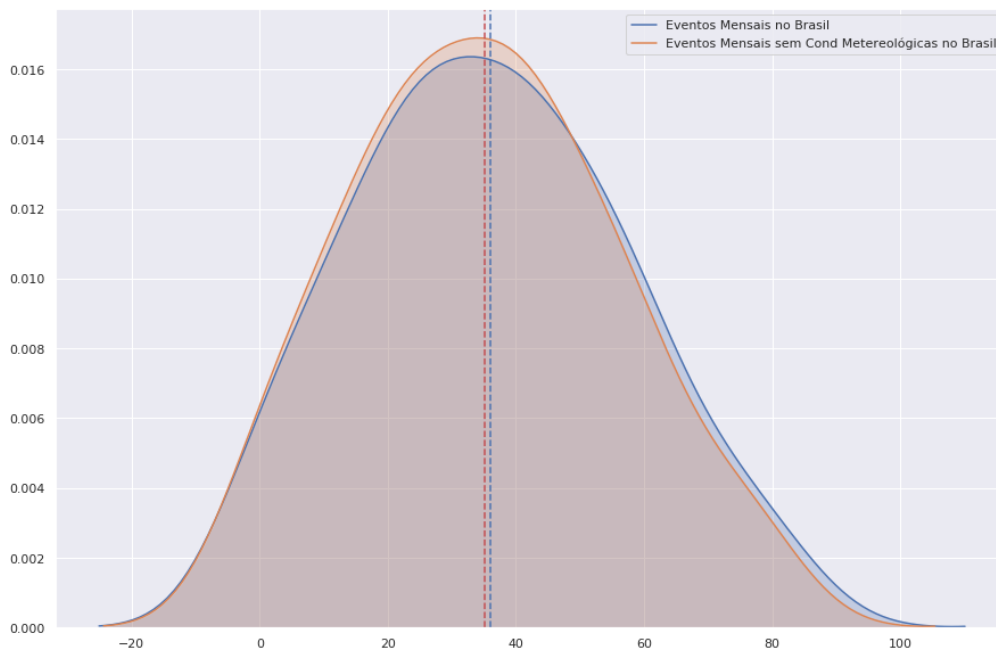


Figura 1. Densidades das distribuições de eventos totais e eventos totais sem dados de condições meteorológicas no Brasil (com valores médios tracejados)

Resultado p-value: 0.6895068690139022

- Desempenho do ser Humano (Brasil)

- Hipótese nula: O desempenho do ser humano não influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA no Brasil.
- Hipótese alternativa: O desempenho do ser humano influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA no Brasil.

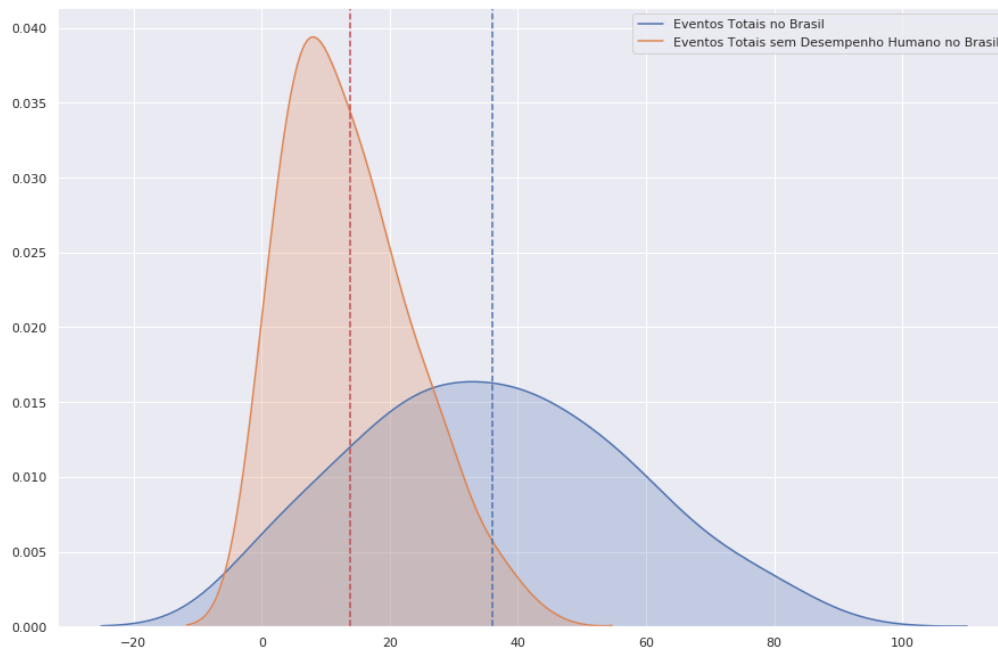


Figura 2. Densidades das distribuições de eventos totais e eventos totais sem dados de desempenho humano no Brasil (com valores médios tracejados)

Resultado p-value: 1.0172776473036971e-23

- Condições Meteorológicas (Norte)

- Hipótese nula: As condições meteorológicas não influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Norte do Brasil.
- Hipótese alternativa: As condições meteorológicas influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Norte do Brasil.

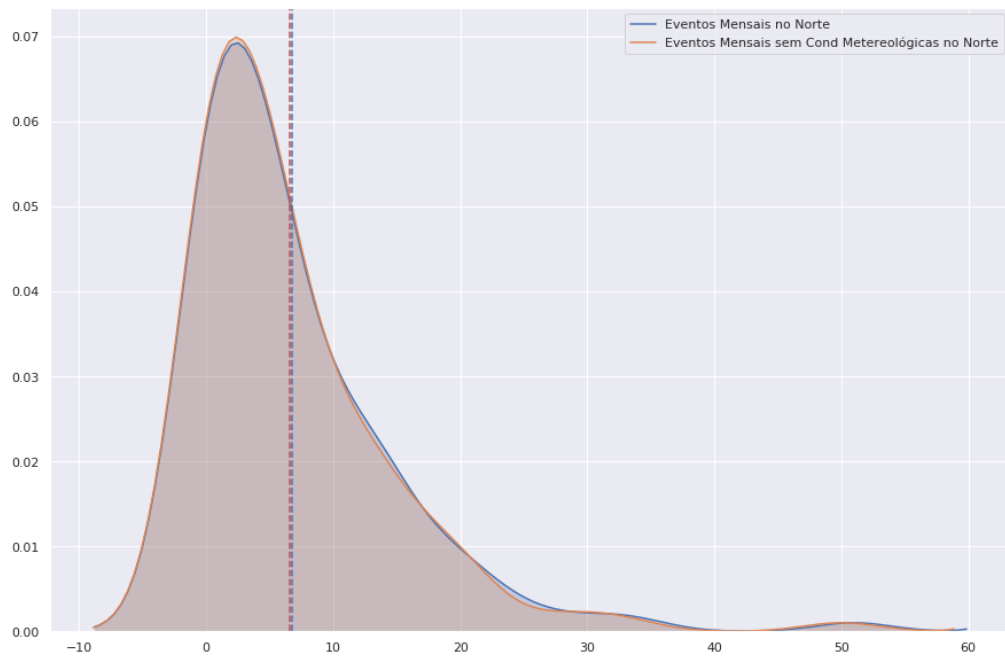


Figura 3. Densidades das distribuições de eventos totais e eventos totais sem dados de condições meteorológicas na região Norte (com valores médios tracejados)

Resultado p-value: 0.8521235265402242

- Desempenho do ser Humano (Norte)

- Hipótese nula: O desempenho do ser humano influencia na ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Norte do Brasil.
- Hipótese alternativa: O desempenho do ser humano influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Norte do Brasil.

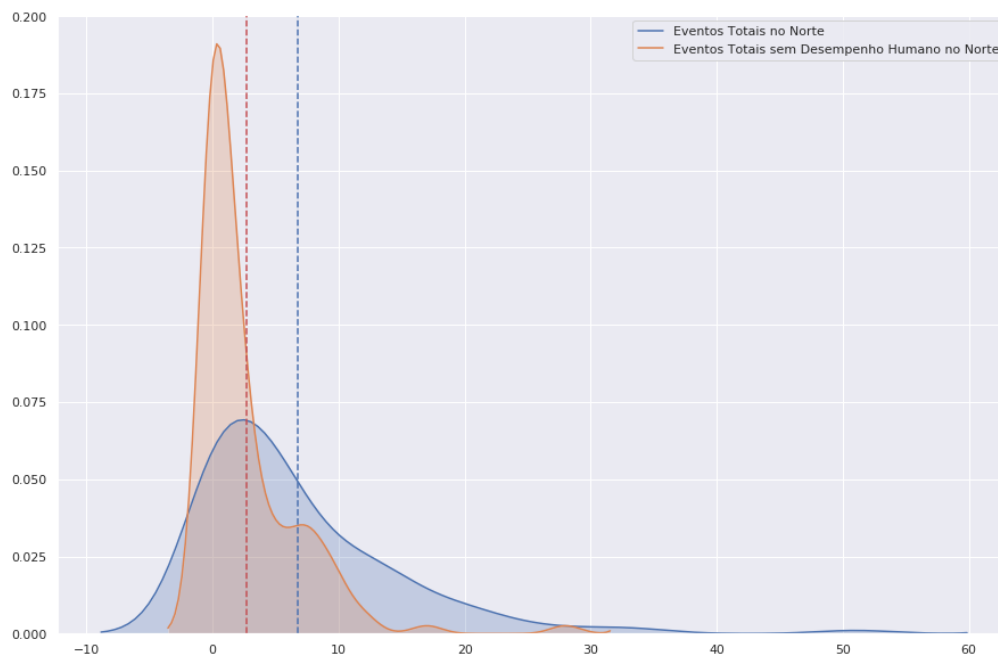


Figura 4. Densidades das distribuições de eventos totais e eventos totais sem dados de desempenho humano na região Norte (com valores médios tracejados)

Resultado p-value: $4.923987593097707e-12$

- Condições Meteorológicas (Nordeste)

- Hipótese nula: As condições meteorológicas não influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Nordeste do Brasil.
- Hipótese alternativa: As condições meteorológicas influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Nordeste do Brasil.

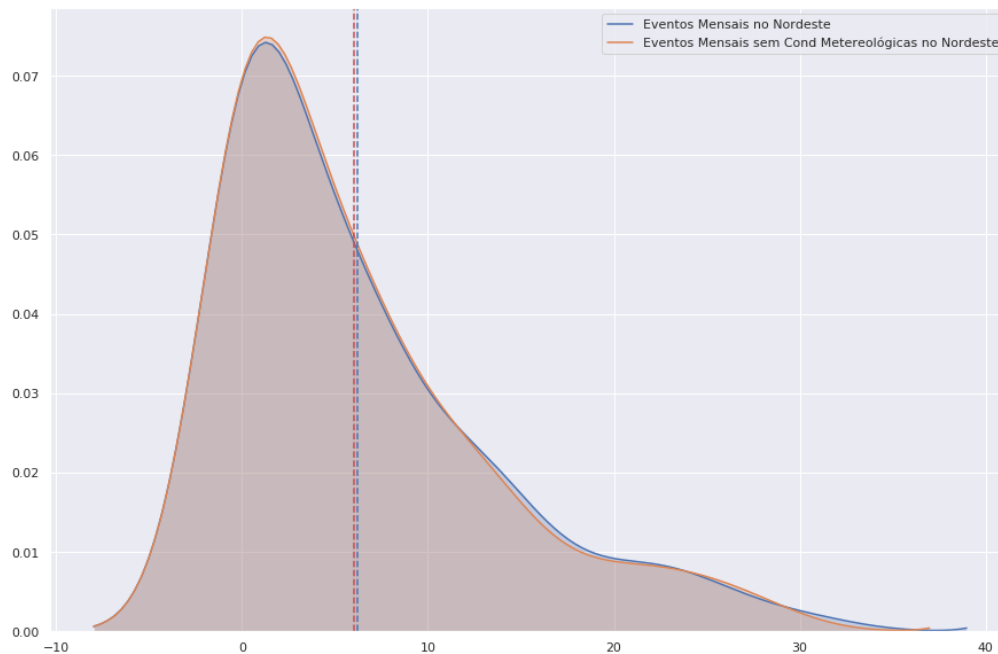


Figura 5. Densidades das distribuições de eventos totais e eventos totais sem dados de condições meteorológicas na região Nordeste (com valores médios tracejados)

Resultado p-value: 0.8764183149087306

- Desempenho do ser Humano (Nordeste)

- Hipótese nula: O desempenho do ser humano não influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Nordeste do Brasil.
- Hipótese alternativa: O desempenho do ser humano influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Nordeste do Brasil.

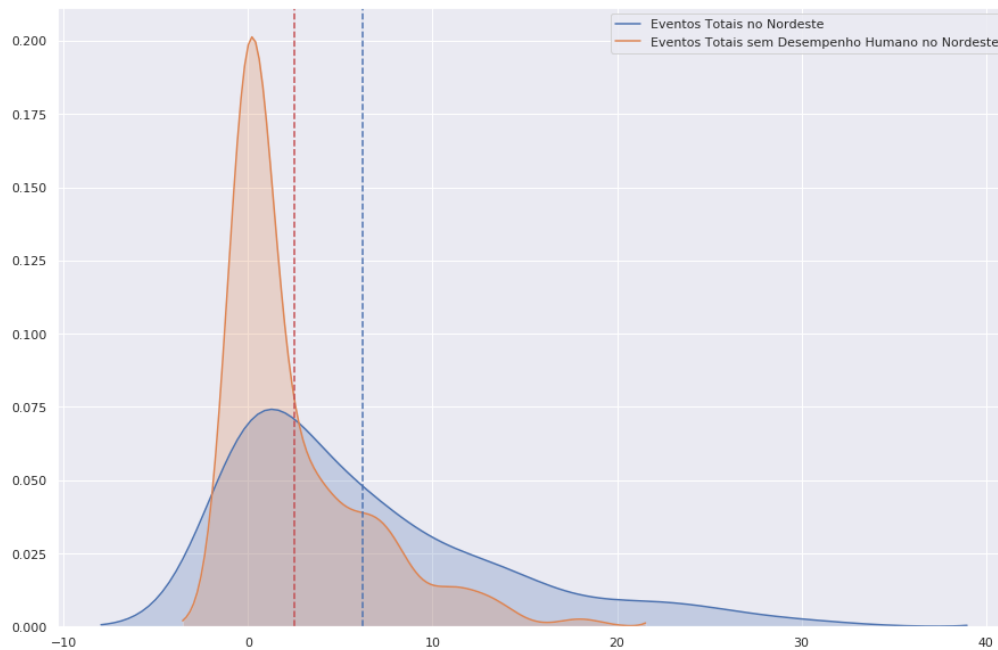


Figura 6. Densidades das distribuições de eventos totais e eventos totais sem dados de desempenho humano na região Nordeste (com valores médios tracejados)

Resultado p-value: $3.4841650184852977e-07$

- Condições Meteorológicas (Centro-Oeste)

- Hipótese nula: As condições meteorológicas não influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Centro-Oeste do Brasil.
- Hipótese alternativa: As condições meteorológicas influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Centro-Oeste do Brasil.

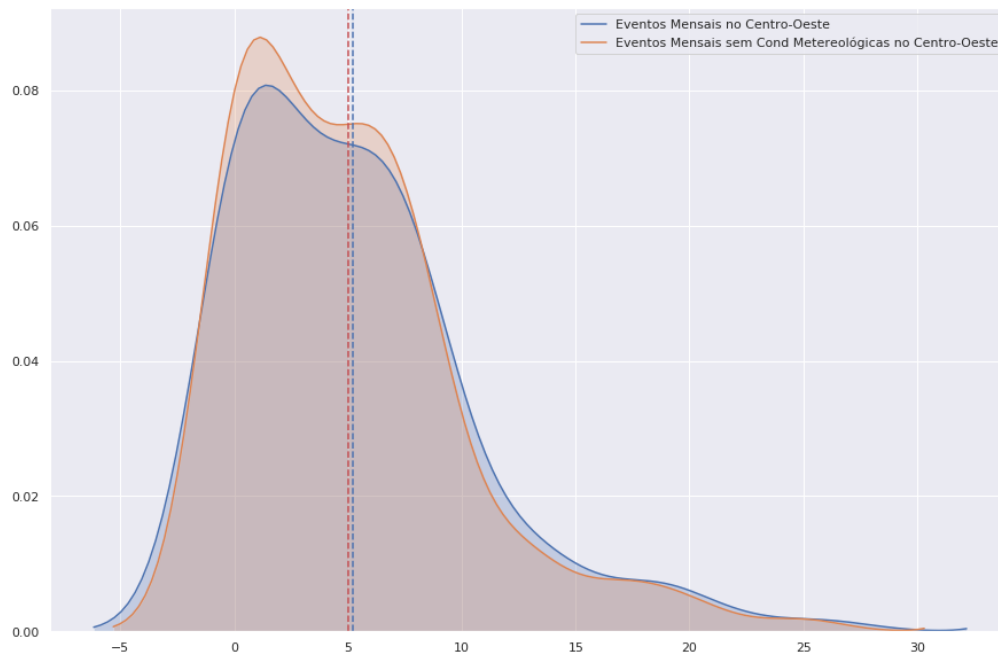


Figura 7. Densidades das distribuições de eventos totais e eventos totais sem dados de condições meteorológicas na região Centro-Oeste (com valores médios tracejados)

Resultado p-value: 0.716423767430306

- Desempenho do ser Humano (Centro-Oeste)
 - Hipótese nula: O desempenho do ser humano não influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Centro-Oeste do Brasil.
 - Hipótese alternativa: O desempenho do ser humano influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Centro-Oeste do Brasil.

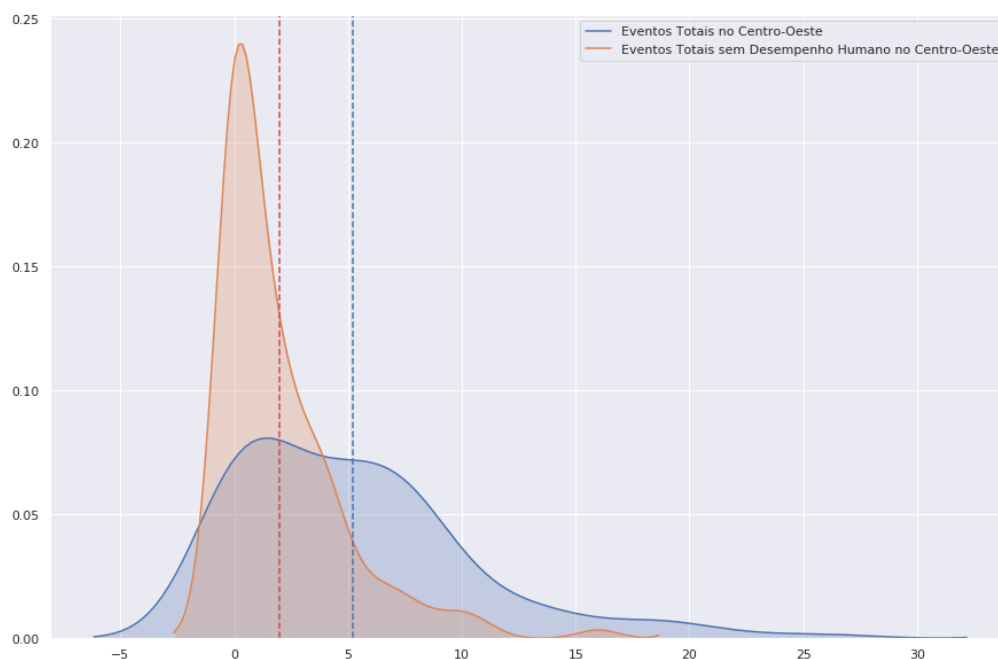


Figura 8. Densidades das distribuições de eventos totais e eventos totais sem dados de desempenho humano na região Centro-Oeste (com valores médios tracejados)

Resultado p-value: $6.88981019654684e-10$

- Condições Meteorológicas (Sudeste)

- Hipótese nula: As condições meteorológicas não influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Sudeste do Brasil.
- Hipótese alternativa: As condições meteorológicas influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Sudeste do Brasil.

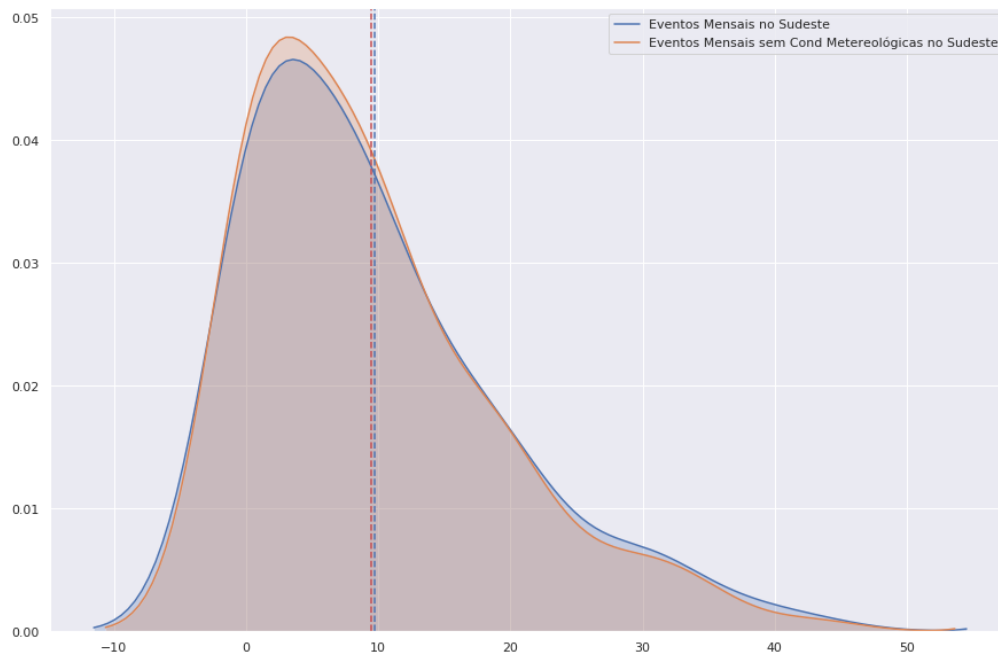


Figura 9. Densidades das distribuições de eventos totais e eventos totais sem dados de condições meteorológicas na região Sudeste (com valores médios tracejados)

Resultado p-value: 0.7910227081746037

- Desempenho do ser Humano (Sudeste)

- Hipótese nula: O desempenho do ser humano não influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Sudeste do Brasil.
- Hipótese alternativa: O desempenho do ser humano influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Sudeste do Brasil.

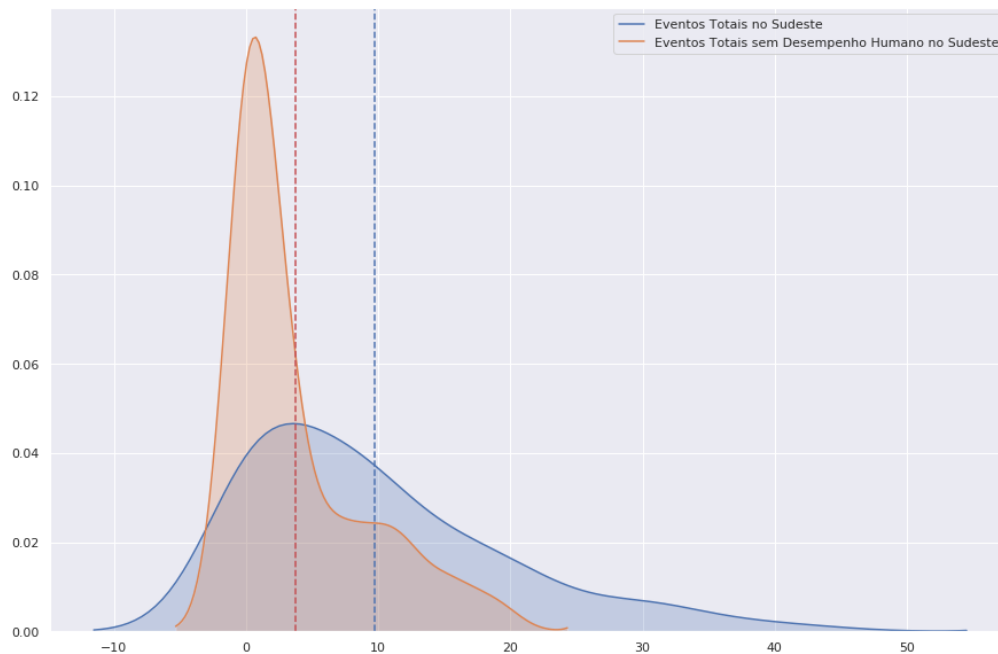


Figura 10. Densidades das distribuições de eventos totais e eventos totais sem dados de desempenho humano na região Sudeste (com valores médios tracejados)

Resultado p-value: 8.309934191863586e-10

- Condições Meteorológicas (Sul)

- Hipótese nula: As condições meteorológicas não influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Sul do Brasil.
- Hipótese alternativa: As condições meteorológicas influenciam nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Sul do Brasil.

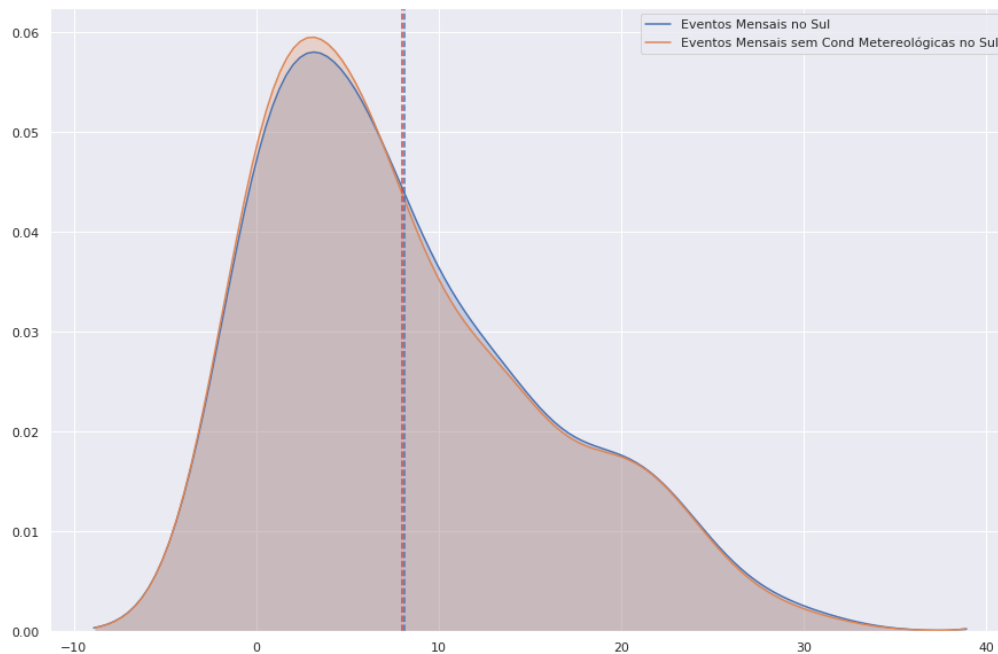


Figura 11. Densidades das distribuições de eventos totais e eventos totais sem dados de condições meteorológicas na região Sul (com valores médios tracejados)

Resultado p-value: 0.8624614612947489

- Desempenho do ser Humano (Sul)
 - Hipótese nula: O desempenho do ser humano não influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Sul do Brasil.
 - Hipótese alternativa: O desempenho do ser humano influencia nas ocorrências aeronáuticas civis analisadas pelo CENIPA na Região Sul do Brasil.

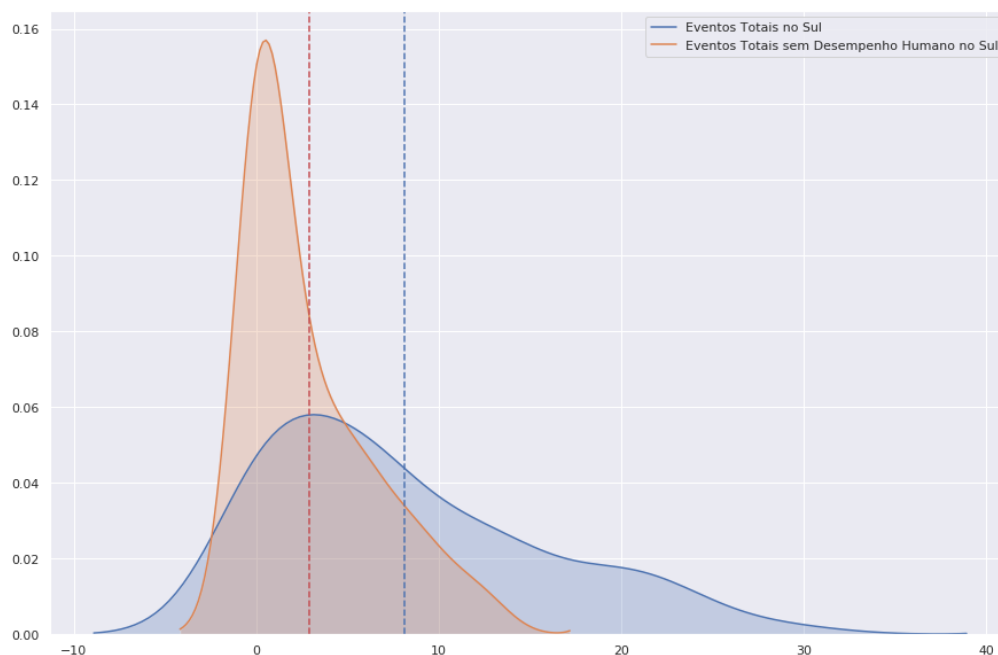


Figura 12. Densidades das distribuições de eventos totais e eventos totais sem dados de desempenho humano na região Sul (com valores médios tracejados)

Resultado p-value: 2.779390005899494e-12

3. Conclusão

Pela observação dos aspectos analisados percebe-se que, no Brasil e em todas as suas regiões, "As condições meteorológicas não influenciam nas ocorrências aeronáuticas civis ocorridas entre 2008 e 2018 analisadas pelo Centro de Investigação e Prevenção de Acidentes Aeronáuticos (CENIPA).", respectivamente para o Brasil e para cada uma de suas regiões, segundo o teste T-Student realizado para verificar as amostras, no Brasil o p-value ficou: 0.6895068690139022, o que confirma a hipótese nula levantada no trabalho, seguindo para a região Norte p-value: 0.8521235265402242, região Nordeste p-value: 0.8764183149087306, região Centro-Oeste p-value: 0.716423767430306, região Sudeste p-value: 0.7910227081746037 e região Sul p-value: 0.8624614612947489, todas confirmam suas hipóteses nulas, visto que todos os p-values deram acima de 0.05 e 0.01, quando isso ocorre fica-se com a hipótese nula e descarta-se a hipótese alternativa.

Na segunda análise realizada para o Brasil e suas regiões, usando T-Student novamente, observa-se que as amostras do Brasil resultaram no seguinte p-value: 1.0172776473036971e-23, seguindo para a região Norte p-value: 4.923987593097707e-12, região Nordeste p-value: 3.4841650184852977e-07, região Centro-Oeste p-value: 6.88981019654684e-10, região Sudeste p-value: 8.309934191863586e-10 e região Sul p-value: 2.779390005899494e-12, todas descartam suas hipóteses nulas, visto que todos os p-values deram abaixo de 0.01, o que representa 99% de confiança no resultado, fica-se com todas as hipóteses alternativas: "O desempenho do ser humano influencia nas ocorrências aeronáuticas civis ocorridas

entre 2008 e 2018 analisadas pelo Centro de Investigação e Prevenção de Acidentes Aeronáuticos (CENIPA).”, respectivamente para o Brasil e para cada uma de suas regiões.

Tendo em vista os aspectos observados percebe-se que o desempenho do ser humano ainda é o principal influenciador nas ocorrências aeronáuticas na aviação civil brasileira. Mostrou-se que as condições meteorológicas têm uma participação insignificativa nas ocorrências, isso pode ser fruto de investimentos em tecnologias aeronáuticas, buscando diminuir as interferências provocadas pelas condições meteorológicas. Dado o exposto, fica evidente que faz-se necessário uma melhor fiscalização em relação aos aspectos humanos envolvidos na aviação civil brasileira, em busca de evitar-se negligência, imperícia e imprudência pelos humanos envolvidos em todas as áreas que compõem a aviação civil como manutenção de aeronaves, voos de instrução, escolas de formação de pilotos, entre outras.

4. Bibliografia

- 1- ABRÃO JUNIOR, F. Como as Condições Meteorológicas Impactam a Aviação e Quais Medidas São Adotadas Para Reduzir Seus Efeitos. In: XII Rio de Transportes, 2014, Rio de Janeiro. XII Rio de Transportes, 2014.
- 2- Revista Brasileira de Geografia Física V. 09 N. 05 (2016) 1335-1345; Tomada de decisão em condições meteorológicas adversas; Gustavo Rodrigues de Oliveira Silva; Humberto César Machado.
- 3- FAJER, Marcia; ALMEIDA, Ildeberto Muniz de and FISCHER, Frida Marina. Fatores contribuintes aos acidentes aeronáuticos. Rev. Saúde Pública [online]. 2011, vol.45, n.2, pp.432-435. Epub Feb 11, 2011. ISSN 0034-8910.
<http://dx.doi.org/10.1590/S0034-89102011005000003>.
- 4- HESPANHOL, Rafael Medeiros. Caracterização dos fatores contribuintes em acidentes de pequenas aeronaves da aviação geral brasileira utilizando inteligência artificial. 2016. xiv, 82 f., il. Dissertação (Mestrado em Transportes)—Universidade de Brasília, Brasília, 2016.
- 5- CONTEÚDO aberto. In: Wikipédia: a enciclopédia livre. Disponível em: https://pt.wikipedia.org/wiki/Testes_de_hipóteses Acesso em : 8 jul 2019.