

Curso: Foundations

Módulo: Noções de Inferência

Lista de Aplicações

- 1) Discuta a validade da amostragem nos casos a seguir:
 - a) Para avaliar a eficácia de uma campanha de vacinação no Estado do Rio de Janeiro, 150 mães de recém-nascidos de uma maternidade de Copacabana foram entrevistadas a respeito da última vez em que vacinaram seus filhos.
 - b) Para verificar a audiência de um programa de TV, a emissora contabilizou a quantidade de pessoas que tuitaram a hashtag #ProgramaDoGugu.
 - c) Uma amostra de sangue foi retirada de um paciente com suspeita de uma doença.
 - d) Para avaliar a intenção de voto para presidente dos brasileiros, 2.000 pessoas foram entrevistadas no Nordeste.
- 2) Utilizando o arquivo “**estimadores.xlsx**”, escolha o melhor estimador para a média populacional. Justifique sua escolha.
- 3) Utilizando o arquivo “**comprovacao_tlc.xlsx**”, execute os seguintes passos:
 - a) Crie uma cópia da guia “Uniforme” e altere seu nome para “Lognormal”
 - b) Altere a fórmula das células J11:DE1010 para a distribuição de probabilidades Lognormal. Utilize a fórmula: $\text{INV.LOGNORMAL}(\text{ALEATÓRIO}();M;DP)$, onde M é a média e DP o desvio padrão. Escolha uma média e um desvio padrão.
 - c) Construa os histogramas para as médias estimadas com tamanho da amostra de $n = 1$ a $n = 100$.
 - d) Analisando os gráficos, o mesmo efeito observado com a distribuição Uniforme ocorre com a distribuição Lognormal?
 - e) Quais são as mudanças na tabela de estatísticas quando o n aumenta?
- 4) Um fabricante de automóveis decidiu exportar o novo modelo produzido para a Suécia. Para verificar se os ajustes do banco do motorista e passageiro dianteiro desenvolvidos no Brasil são adequados à altura da população Sueca, solicitou a uma consultoria um estudo e a coleta de uma amostra da altura da população que lá vive. Com os dados da amostra coletada no arquivo “**altura_suecia.xlsx**”, calcule:
 - a) Estimativa Pontual da altura média dos habitantes da Suécia.
 - b) Intervalos de Confiança considerando o coeficiente de confiança igual a 90% e 99%.
 - c) Por que o Intervalo de Confiança com coeficiente de confiança igual a 99% é maior do que o com coeficiente de confiança igual 90%?

- d) Esse novo modelo foi desenvolvido para a altura média do brasileiro, que segundo o IBGE, é de 1,73m. Nesse caso, será necessária uma adaptação do ajuste dos bancos para exportação?
- 5) Preocupado com a estimativa da altura média dos Suecos, o fabricante de automóveis decidiu solicitar uma nova análise para ter mais acurácia na tomada de decisão de adaptar ou não o ajuste dos bancos para exportação de seu novo modelo. Para isso solicitou os seguintes cálculos (considere o mesmo arquivo “**altura_suecia.xlsx**” do exercício anterior e os valores já calculados):
- a) A Margem de Erro para os dois Intervalos de Confiança calculados no exercício anterior.
 - b) O número de pessoas que devem ter sua altura medida para que a Margem de Erro seja inferior a 1 cm, considerando os coeficientes de confiança de 90% e 99%.
- 6) Uma empresa de produtos fitness gostaria de realizar uma pesquisa de interesse pela sua nova barra de cereal com cobertura de chocolate e zero calorias para estimar qual o percentual do público de academias compraria o produto. Para isso contratou uma consultoria que se comprometeu em calcular o tamanho da amostra necessária para que a margem de erro fosse de no máximo 5%.
- a) Calcule o tamanho da amostra considerando o coeficiente de confiança 95%.
- Dica:** como se trata de um problema relacionado com percentual e não existem dados de uma amostra piloto, considere $S^2 = 0,25$.
- b) Devido a problemas operacionais, não foi possível realizar a pesquisa com o número de pessoas calculado anteriormente. A consultoria conseguiu entrevistar 300 pessoas e a variância dessa amostra foi de $S^2 = 0,22$. Qual é a margem de erro com esse tamanho de amostra, considerando o coeficiente de confiança também 95%?
- 7) Ao analisar a conversão de vendas no site de sua empresa, o analista de dados verificou que houveram 10 vendas em 50 visitas.
- a) Calcule a taxa de conversão e a variância dessa taxa.
 - b) Qual o intervalo de confiança considerando um coeficiente de confiança de 95%?
 - c) Qual a margem de erro?
 - d) Qual seria o tamanho da amostra caso quisermos uma margem de erro de 5p.p?
- 8) Uma empresa quer verificar como está o nível de colesterol de seus colaboradores para um novo programa da área de Saúde. Para isso, coletou amostras de sangue de 220 funcionários selecionados aleatoriamente, esse levantamento está no arquivo **colesterol.xlsx**. Sabendo que é considerado aceitável um nível de colesterol de até 190mg/dL:
- a) Construa o histograma e avalie se a distribuição do colesterol dos colaboradores se aproximam da Normal.
 - b) Realize um Teste de Hipóteses para comparar se a média do colesterol está aceitável ou não.

- 9) A Secretaria da Educação solicitou a análise do IDHM Educação de cada um dos municípios para avaliar se a diferença no índice, observada entre as regiões, é **estatisticamente significativa**. Para isso, consolidou a base de dados desse índice para alguns municípios de cada uma das 3 principais regiões a serem avaliadas: Norte, Nordeste e Sul, no arquivo **idhm_educacao.xlsx**.
- a) Construa os histogramas para cada região e avalie se as distribuições se aproximam da Normal.
 - b) Calcule as medidas resumo por região: número de observações, média e desvio padrão.
 - c) Realize um Teste de Hipóteses para comparar se a média do IDHM Educação é diferente entre as regiões:
 - i. Nordeste e Sul
 - ii. Norte e Nordeste

Para cada Teste de Hipóteses realize os seguintes passos:

- Formule as hipóteses a serem testadas.
- Avalie se as variâncias do IDHM entre as regiões podem ser consideradas iguais para selecionar o teste mais adequado.
- Calcule o p-valor.
- Interprete o resultado do teste considerando um coeficiente de confiança de 5%.

- 10) Foi realizado um estudo para avaliar o efeito do álcool no tempo de reação de motoristas. Para isso, foram selecionados 20 motoristas, e mediu-se o tempo de reação de cada um deles em um percurso controlado. Na sequência cada motorista bebeu 2 cervejas e seu tempo de reação após o consumo do álcool foi novamente medido. Para avaliar se as diferenças nos tempos de reação antes e depois do consumo do álcool são estatisticamente significantes, utilize o arquivo **tempo_reacao.xlsx** e realize as seguintes análises:
- a) Construa os histogramas para cada região e avalie se as distribuições se aproximam da Normal.
 - b) Calcule as medidas resumo dos tempos de reação antes e depois do consumo de álcool: número de observações, média e desvio padrão.
 - c) Realize um Teste de Hipóteses T Pareado para comparar se o tempo de reação dos motoristas é maior após o consumo de álcool.

Para o Teste de Hipóteses realize os seguintes passos:

- Formule as hipóteses a serem testadas.
- Calcule o p-valor.
- Interprete o resultado do teste considerando um coeficiente de confiança de 5%.

- 11) Uma empresa presente nos maiores marketplaces do Brasil resolveu fazer um estudo para avaliar o valor dos produtos vendidos e o valor do frete. Para isso, coletou informações de 100 mil vendas com seus respectivos fretes. Utilizando o arquivo **“olist_order_items_dataset.xlsx”**, realize as seguintes análises:

Estimativa Pontual e por Intervalo

- e) Calcule a estimativa pontual do preço médio dos produtos e do valor médio do frete.

- f) Calcule os intervalos de confiança para preço e frete considerando o coeficiente de confiança igual a 95%.
- g) Qual é a probabilidade de que esses intervalos de confiança contenham o verdadeiro valor da média populacional? Argumente.

Margem de Erro e Tamanho da Amostra

- h) Calcule a margem de erro utilizando as estimativas pontuais e por intervalo. Por que a margem de erro do frete é inferior a margem de erro do preço?
- i) Calcule o mínimo tamanho de amostra para que a margem de erro na estimativa do preço seja inferior a R\$ 0,10 e do frete inferior a R\$ 0,05. Considere o mesmo coeficiente de confiança de 95%.

Percentual de pessoas com preço acima da média

- j) Crie uma variável em que o valor seja igual a 1 caso o preço seja maior do que a média, calculada no item a), e 0 caso contrário. Vamos chamá-la de preço > média.
- k) Calcule as estimativas pontuais e por intervalo do percentual de pessoas com compras acima da média. Considerando o valor da estimativa pontual para o percentual de pessoas com compras acima da média, é possível afirmar que o preço médio é maior ou menor do que o preço mediano? Argumente.
- l) Com esse tamanho de amostra é possível obter uma margem de erro inferior a 0,2%? Caso não seja, qual o tamanho da amostra para atingir essa margem de erro desejada?