

Estatística - Teste de Hipóteses

Considerando a base de dados `experimento_test_ab.csv` responda as questões abaixo:

1. Qual dos cenários tem a maior taxa de conversão?
2. Calcule qual o tamanho da amostra necessária para o desenvolvimento de um teste A/B, seguindo os seguintes critérios:
 - a. O cenário A, da base, como o inicial, que funciona hoje.
 - b. Considere 95% de confiança de que o efeito na conversão não foi aleatório com um nível de significância de 5% ($\alpha = 0,05$).
 - c. Também considere 80% de certeza conseguir capturar o efeito da nova abordagem.
 - d. O aumento para 10% de conversão.

Verifique a base `pacientes.csv`, os dados são fictícios. A sintetiza uma base de dados de um hospital que trata de pacientes com problemas cardíacos. Considere a base como a população, portanto as estatísticas da população são conhecidas. Responda as perguntas abaixo:

3. Considerando uma amostra de 45 números que representam o index do dataframe, índices= ([909, 751, 402, 400, 726, 39, 184, 269, 255, 769, 209, 715, 677, 381, 793, 697, 89, 280, 232, 756, 358, 36, 439, 768, 967, 699, 473, 222, 89, 639, 883, 558, 757, 84, 907, 895, 217, 224, 311, 348, 146, 505, 273, 957, 362]). Considerando essa amostra é possível dizer que a idade média das pessoas com problemas cardíacos é maior que 50 anos? Nível de significância igual a 5%.
4. Queremos entender que tipo de amostra estamos lidando se dividirmos os conjuntos em 2, sendo um com pessoas que têm condições de saúde adicionais e o outro com pessoas saudáveis. Seria dependente ou independente?
5. Agora considere o um conjunto de pessoas aleatória que representam o index do dataframe, índices = ([690, 894, 67, 201, 364, 19, 60, 319, 588, 643, 855, 623, 530, 174, 105, 693, 6, 462, 973, 607, 811, 346, 354,

966, 943, 372]], podemos dizer que a pressão arterial média para pacientes com condições de saúde adicionais é igual à pressão arterial média para pacientes sem condições adicionais de saúde? Considere o nível de significância a 6%.

6. Existe uma diferença significativa na pressão arterial média entre diferentes grupos étnicos nesta população? (Teste ANOVA, α é 5%)
 - a. Hipótese Nula (H_0): A pressão arterial média é a mesma em todos os grupos étnicos.
 - b. Hipótese Alternativa (H_1): Há uma diferença significativa na pressão arterial média entre pelo menos dois grupos étnicos nesta população.
7. Lúcia é uma pesquisadora e tem o objetivo de entender a relação de gênero neste grupo de pacientes. Acredita-se que há uma relação entre o sexo e condições de saúde adicionais. (Teste qui-quadrado)
8. Existe uma associação entre a idade dos pacientes e sua pressão arterial?
 - a. Hipótese Nula (H_0): se a pressão arterial é independente da idade
 - b. Hipótese Alternativa (H_1): a pressão arterial esta associada a idade
9. Qual é o intervalo de confiança para a média da pressão arterial entre os pacientes com condições de saúde adicionais? (nível de confiança 95%)
10. A distribuição da pressão arterial na população segue uma distribuição normal?
 - a. Hipótese Nula (H_0): A distribuição da pressão arterial na população segue uma distribuição normal.
 - b. Hipótese Alternativa (H_1): A distribuição da pressão arterial na população não segue uma distribuição normal.