

❖ **Exercício 1: Nova Feature para um Site**

- Um site criou uma nova funcionalidade (play automático de vídeos) em suas páginas e deseja saber se essa nova funcionalidade deve ser mantida ou não, isto é, se ela gera lucro, prejuízo ou se é indiferente. Para definir se a funcionalidade deve ser mantida, foram fornecidos os dados históricos de utilização do site (antes da implementação da funcionalidade, isto é, a população) e os dados depois da implementação (isto é, a amostra). Esses dados estão relacionados à tempo de navegação no site. Sabendo que a nova funcionalidade custa um preço para ser executada (gasto do site) e, que cada minuto extra de navegação também gera um preço (lucro ao site), o objetivo é descobrir se ela deve ser mantida ou não.
- Formulação da Hipótese: Supor que a média da amostra é igual a média da população, de modo a existir um problema bicaudal, sendo possível calcular se a hipótese H_0 será rejeitada ou não e, caso seja rejeitada, o motivo (abaixo do limite inferior ou acima do limite superior). Ou seja, descrevendo em relação ao problema, é possível verificar se a média da amostra é igual a média da população (H_0 não é rejeitada, mas a funcionalidade gerou prejuízo ao site, dado que ela gasta para ser executada e não aumentou o tempo médio de navegação, isto é, não aumentou o lucro); se a média da amostra é maior do que a média da população (H_0 é rejeitada e a nova funcionalidade gera lucro); se a média da amostra é menor do que a média da população, também trazendo prejuízo implementar a nova funcionalidade (H_0 é rejeitada e a nova funcionalidade gera prejuízo).
- Procedimento
 - Para realizar esse trabalho, foi utilizada a linguagem de programação Python 2.6 e as bibliotecas matplotlib, numpy e scipy.
 - Passos realizados:
 - Ler os dois arquivos csv e carregar os valores de tempo para duas listas (uma para cada arquivo, ou seja, uma para a amostra e uma para população).
 - Calcular a média e o desvio padrão da população, para poder aplicar esses valores no cálculo da estatística z, mais a frente
 - Calcular a média e o tamanho da amostra.
 - Definir o nível de confiabilidade desejado.

- Calcular a estatística z.
 - ◆ $z = (\text{media_amostra} - \text{media_populacao}) / (\text{desvio_padrao_populacao} / \sqrt{\text{tamanho_amostra}})$
- Calcular o pvalor, lembrando que o calculo varia dependendo se a estatística z (pvalor = $1 - P[Z > z]$) é maior ou menor do que zero (pvalor = $P[Z > z]$)
- Comparar o valor encontrado para pvalor com alfa (1 - nível de confiabilidade) / 2 (problema bicaudal) e descobrir em que região essa amostra se encontra, rejeitando ou não H0.
 - O código utilizado para o desenvolvimento dessa questão pode ser encontrado em anexo com este relatório.

➤ **Respostas**

❖ **Exercício 2: Clicks do Site**

- Este exercício visa descobrir se os dados de duas amostras diferentes seguem ou não uma mesma distribuição. Isto é, existem duas amostras: número de clicks em um site sem pop-up e número de clicks em um site com pop-up e o objetivo é descobrir se a existência do pop-up influencia no número de clicks.
- Para isso, foi realizado o teste do qui-quadrado para duas amostras, que tem como hipótese nula que as duas amostras vieram da mesma distribuição. Portanto, se a hipótese H0 for rejeitada, significa que faz diferença utilizar o pop-up.
- Procedimento
 - Classe A = Tela sem pop-up
 - Classe B = Tela com pop-up
 - Grupo I = Usuário clicou
 - Grupo II = Usuário não clicou
 - Criar tabela com a contingência dos dados (tabela observada)

	Classe A	Classe B	Total
Grupo I	X	Y	X + Y
Grupo II	W	Z	W + Z
Total	X + W	Y + Z	T = X + Y + W + Z

- Construir uma tabela com os valores esperados

	Classe A	Classe B
Grupo I	$[(X + Y) * (W + X)] / T$	$[(X + Y) * (Y + Z)] / T$

Grupo II	$[(W + Z) * (W + X)] / T$	$[(Z + Y) * (W + Z)] / T$
----------	---------------------------	---------------------------

- Número de graus de liberdade: 4
- Calcular a estatística X_n^2
- Calcular o pvalor, a partir da estatística X_n^2
- Comparar o pvalor com o alfa definido (1 - nível de confiabilidade)
- Rejeitar ou não a hipótese nula, que diz que as amostras vieram da mesma distribuição

➤ Respostas

❖ **Exercício 3: Produtor de Cinema**