

## Márcio Ozório de Jesus - Nanodegree de Analista de Dados

### Estatística: A Ciência das Escolhas

#### P1: Testando um Fenômeno de Percepção - Efeito Stroop

Em um teste Stroop, participantes são apresentados a uma lista de palavras, e cada uma das palavras está de uma cor. A tarefa dos participantes é dizer em voz alta a cor em que a palavra foi exibida.

A tarefa tem duas condições: uma **condição de palavras congruentes e uma condição de palavras incongruentes**. Na condição de palavras congruentes, a palavra exibida são nomes de cores cujos nomes são a mesma cor em que a palavra foi exibida: por exemplo, **RED**, **BLUE**.

Na condição de palavras incongruentes, as palavras apresentadas são nomes de cores cujos nomes não são as mesmas cores em que as palavras foram apresentadas: por exemplo, **PURPLE**, **ORANGE**.

Em cada caso, medimos o tempo que demora para o indivíduo falar os nomes em uma lista das cores de tamanhos iguais. Cada participante realizará o teste e tem o tempo medido para a realização do teste nas duas condições.

#### Questões para Investigação

1. Qual é a nossa variável independente? Qual é a nossa variável dependente?
  - Variável Independente: Condição de palavras congruente e incongruentes
  - Variável Dependente: Tempo de demora para a resposta
2. Qual seria um conjunto apropriado de hipóteses para essa tarefa? Que tipo de teste estatístico você espera executar? Justifique suas escolhas.

Hipoteses:

$H_0$  (hipótese nula) : Não há diferença entre a média de resposta da população para as condições congruente e incongruente.

$$H_0: \mu_C = \mu_I$$

$H_a$  (hipótese alternativa): Existe uma diferença significativa entre a média de resposta da população entre as condições congruente e incongruente.

$$H_a: \mu_C \neq \mu_I$$

Iremos realizar um teste de hipótese utilizando os dados das amostras para fazer inferências sobre a população.

O teste de hipótese irá nos indicar se a diferença do tempo de resposta entre a condição congruente e a condição incongruente é ou não estatisticamente significativo, ou seja, se a forma que o teste é realizado (congruente ou incongruente), influência ou não o tempo de resposta dos participantes.

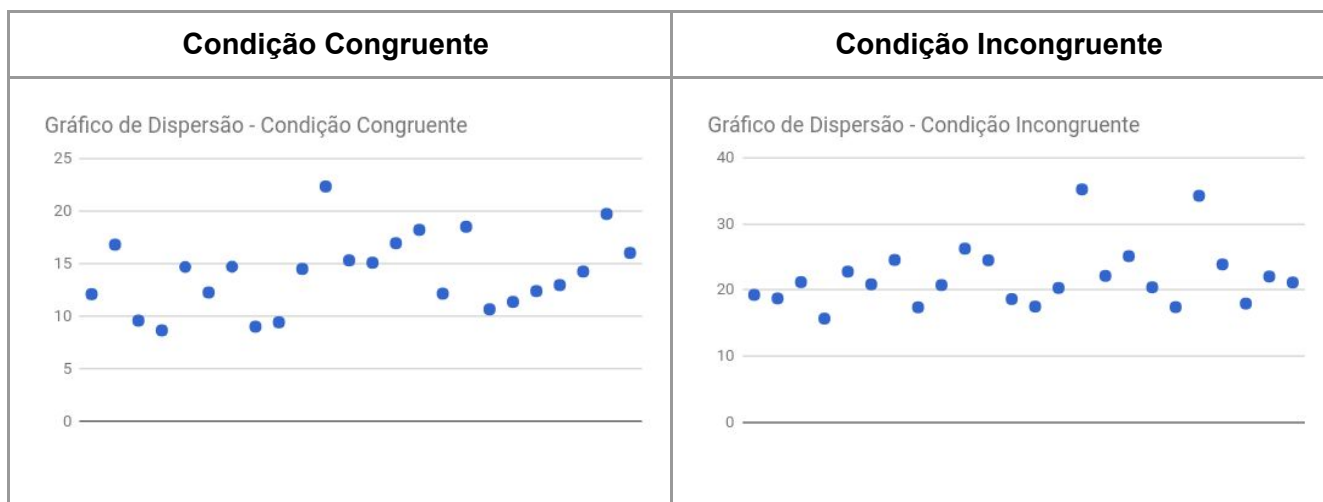
O teste a ser executado é o Test t de amostras dependentes tipo duas condições, pois não temos nenhum parâmetro populacional para comparação, nosso tamanho de amostra é pequena e os testes são realizados pelo mesmo sujeito.

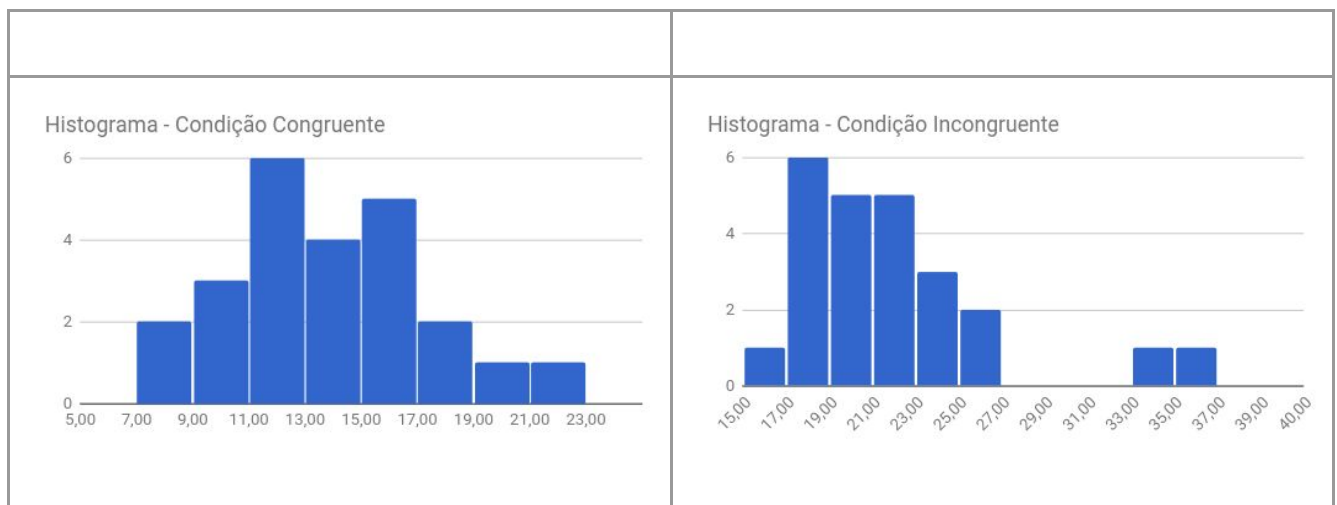
3. Reporte alguma estatística descritiva em relação a esse conjunto de dados. Inclua, pelo menos, uma medida de tendência central de pelo menos uma medida de variabilidade.

Condição Congruente	Condição Incongruente
<b>Mc</b> (Média condição congruente) = <b>14,05</b>	<b>Mi</b> (Média da condição incongruente) = <b>22,02</b>
Mediana = 14,36	Mediana = 21,02
Desvio Padrão = 3,56	Desvio Padrão = 4,80
<b>Md</b> (Diferença das médias = $Mc - Mi$ ) = <b>-7,97</b>	
<b>n</b> = <b>24</b>	
<b>df</b> (grau de liberdade) = <b>23</b>	

4. Forneça uma ou duas visualizações que mostre a distribuição da amostra de dados. Escreva uma ou duas sentenças sobre o que você observou do gráfico ou gráficos.

Visualização da distribuição dos dados:

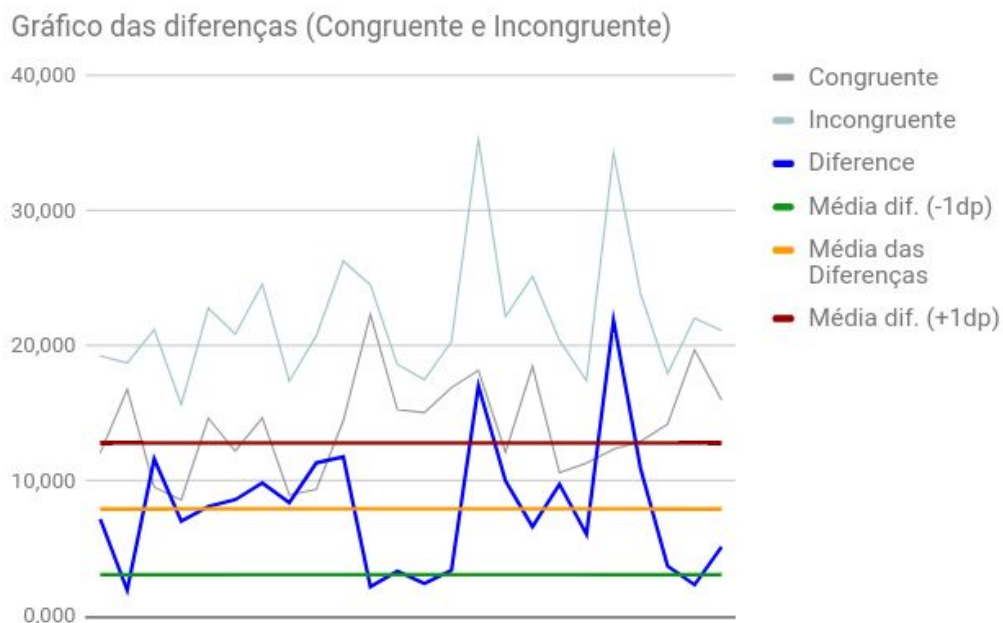




Observações identificadas:

- Pode-se notar nos dois primeiros gráficos que há pouca dispersão nos dois conjuntos de dados.
- A distribuição dos dados da condição congruente, está em um formato que indica uma distribuição normal.
- Já o histograma de condição incongruente, a densidade dos dados está maior do lado esquerdo, formando assim uma distribuição assimétrica positiva.

Segue um gráfico adicional onde é exibida a distribuição dos dados, a diferença entre os valores, a diferença média e um desvio padrão para mais e um desvio padrão para menos da diferença média.



5. Agora desempenhe o teste estatístico e reporte seus resultados. Qual seu nível de confiança e o valor estatístico crítico? Você rejeitou a hipótese nula ou falhou ao tentar rejeitá-la? Encontre uma conclusão em relação ao experimento da tarefa. Os resultados estão de acordo com suas expectativas?

<b>S</b> (Desvio padrão das amostras)	4,86
<b>SEM</b> (Erro Padrão da Média)	0,99
<b>Estatística t</b> (Test $t = M_d / SEM$ )	-8,05
Alfa ( $\alpha$ )	0,05
Tipo do teste	Bicaudal
<b>Intervalo de confiança (CI)</b>	95%
Valor crítico de t (0,025)	2,069
Margem de erro	2,50
Limite inferior	-10,02
Limite superior	-5,92
Valor-p (p-value)	0,0001
d de Cohen (Cohen's d)	-1,64
R <sup>2</sup>	0,76

## CONCLUSÕES

Conforme podemos ver na tabela acima o **Valor-p** (0,0001) é menor que o **Valor crítico de t** (2,069), ou seja, a *diferença dos tempos de resposta* entre o teste de condição congruente e condição incongruente *são estatisticamente significativas*.

Portanto, **rejeitamos a hipótese nula  $H_0$** , confirmando assim que os resultados são influenciados dependendo das condições que são realizadas (congruente ou incongruente) e que esta diferença não ocorre devido ao acaso.

O **intervalo de confiança** de 95% (10,02, -5,92), utilizando um alfa de de 0,05. Isto significa que os participantes quando realizam o teste com condição congruente são entre 10 a 6 segundos mais rápidos que quando realizam o teste na condição incongruente.

Realizei o teste antes através do link indicado e de acordo com a percepção que obtive, os resultados acima estão de acordo com minhas expectativas. Antes da realização inferência, havia notado uma grande diferença entre as médias dos testes, o que causou minha suspeita. Apesar disso, como esta análise não é suficiente para tirarmos qualquer conclusão, realizei os cálculos os estatísticos onde pude confirmar que esta diferença é realmente significativa.

Tamanho do efeito:

- Como complemento de análise da significância estatística, foi calculado o “**d de Cohen** (Cohen's d)” onde foi encontrada uma grande significância entre a diferença das médias: 1,64 (valor em desvios padrão).

- Podemos dizer que na amostra de 24 pessoas, 76% ( $R^2$ ) da diferença de tempos de resposta é devido a condição incongruente.

**6. Opcional:** O que você acha que é responsável pelo efeito observado? Consegue pensar em uma alternativa ou tarefa similar que resultaria em um efeito parecido? Alguma pesquisa em relação ao problema pode ser útil para pensar nessas duas questões!

Eu acredito que na condição congruente, o cérebro já esteja “treinado” com os padrões de caracteres das palavras apresentadas. Já na condição incongruente, o cérebro encontra dificuldade em desvincular os caracteres da palavra e focar somente na cor da mesma. Isso causa o aumento do tempo de resposta.

Para saber ao certo o que acontece, realizei algumas pesquisas onde descobri que estes estudos foram realizados pelo psicólogo Ridley Stroop, que fez esta investigação década de 1930 nomeando-a como Efeito Stroop.

O efeito Stroop é um aumento no tempo de reação para uma determinada tarefa, quando o cérebro simultaneamente lida com informações conflitantes.

De acordo com suas descobertas, processamento de texto é muito mais rápido e mais difícil para suprimir do que processamento de cores; palavras escritas em uma cor em conflito com o seu significado, como “vermelho” escrito em tinta azul, causam atrasos e erros na resposta.

Para a resposta correta, o cérebro precisa inibir o processo de reconhecimento de palavras mais forte com atenção seletiva. Geralmente é mais fácil de aplicar o Teste de Stroop a crianças jovens, não escolarizadas. Quanto mais alfabetizada a pessoa, maior a tendência em se usar a leitura no teste.

O teste Stroop produz resultados associados com a vitalidade e flexibilidade mental do indivíduo, tornando-se uma ferramenta de diagnóstico amplamente utilizada; erros em reconhecer as cores estão associados com impulsividade e inibição insuficiente no córtex pré-frontal, comum em pacientes com TDAH (Transtorno de déficit de atenção e hiperatividade) e outros distúrbios dependentes de dopamina.

Um efeito semelhante pode ser realizado utilizando imagens de animais. Sobre as imagens colocamos textos referentes ao nome de cada animal. Na primeira condição colocamos as imagens dos animais com os nomes corretos (essa seria a condição congruente), na segunda condição os nomes serão embaralhados, ou seja, o nome escrito não estará relacionado com a imagem do animal (condição incongruente). Em ambas as condições, a pessoa deverá dizer o nome do animal, independente do nome escrito. Segue um exemplo de como ficaria na condição incongruente:



Você deve dizer "Vaca" porque o animal na foto é uma vaca.



Você deve dizer "Gato" porque o animal na imagem é um gato.

## Fontes:

- **Efeito Stroop**  
[https://pt.wikipedia.org/wiki/Efeito\\_Stroop](https://pt.wikipedia.org/wiki/Efeito_Stroop)
- **Efeito Stroop: Teste e explicação da Psicologia**  
<http://psicoativo.com/2016/08/efeito-stroop-teste-e-explicacao-da-psicologia.html>
- **Efeito Stroop: Psicologia Cognitiva**  
[http://www.notapositiva.com/old/pt/trbestsup/psicologia/psicmente/efeito\\_stroop.htm](http://www.notapositiva.com/old/pt/trbestsup/psicologia/psicmente/efeito_stroop.htm)
- **Interactive "Animal Stroop" Effect Experiment**  
<https://faculty.washington.edu/chudler/words.html>  
<https://faculty.washington.edu/chudler/java/readya.html>
- **Discussion Forum - Intro and Statistics**  
<https://discussions.udacity.com/c/nd002-statistics>
- **Encontrando o valor exato de p (p-value):**  
GraphPad Software (<http://www.graphpad.com/quickcalcs/>)  
Opções:
  - Statistical distributions and interpreting P values
  - Calculate P value from z, t, F, r or chi-squarePreenchimento dos parâmetros t e DF identificados no estudo.
- **Cohen's D: Definition, Examples, Formulas**  
<http://www.statisticshowto.com/cohens-d/>
- **Effect Size (Cohen's d)**  
[https://www.vcalc.com/wiki/Caroline4/Effect+Size+\(Cohen's+d\)](https://www.vcalc.com/wiki/Caroline4/Effect+Size+(Cohen's+d))