**04 - Aplicando p-valor**

[0:00] Ok. Já finalizamos o nosso teste de Wilcoxon e concluímos que devemos rejeitar a hipótese nula de que não existe diferença entre os grupos. Ou seja, existe uma diferença e como a calculamos a média dos grupos antes e depois do tratamento, verificamos que antes as pessoas fumava, em média, mais cigarros do que depois do tratamento.

[0:20] E como o nosso teste indicou que existe uma diferença, a gente conclui que essa diferença é boa, ou seja, o tratamento é satisfatório, ajuda a realmente parar de fumar, a diminuir pelo menos.

[0:32] Vou mostrar agora uma forma mais simples, como estamos fazendo ao longo desse treinamento, de calcular, de obter o resultado desse teste sem ter que fazer todas essas contas que a gente acompanhou no vídeo anterior.

[0:45] E por falar nas contas, algo que eu esqueci de falar quando estávamos calculando os postos. Lembra que existe a seguinte fórmula?

fumo['Posto (-)'] = fumo.apply(lambda x: x.Posto if x.Dif < 0 else 0, axis = 1)

fumo

COPIAR CÓDIGO

E eu estou pegando quando o Dif for menor do que zero, não incluo o zero aqui, e também não incluo o zero no trecho a seguir.

fumo['Posto (+)'] = fumo.apply(lambda x: x.Posto if x.Dif > 0 else 0, axis = 1)

fumo

COPIAR CÓDIGO

[1:02] O que eu quero dizer é justamente que eu não quero incluir o zero. Quando não tiver diferença, eu não conto ele em posto nenhum. Então, está certo, é isso mesmo.

[1:11] É só para esclarecer e você não ficar sem saber o que eu fiz com o zero.

[1:18] Vamos fazer isso de maneira mais simples. Em duas linhas é possível resolver esse problema todo.

[1:25] Vamos fazer from scipy.stats import o próprio nome wilcoxon. Está [neste link](https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.wilcoxon.html) a documentação dele.

[1:39] Esse teste realiza algumas outras coisas que a gente não aprendeu, então vamos focar só no que estudamos até agora.

[1:47] Importamos. É bem simples, wilcoxon() e eu passo para ele o meu Dataframe, nesta ordem: fumo. Antes e fumo.Depois.

from scipy.stats import wilcoxon

wilcoxon(fumo.Antes, fumo.Depois)

wilcoxonResult(statistic=22.0, pvalue=1.584310018505865e-06)

p\_valor <= significancia

COPIAR CÓDIGO

[2:13] Ele me retorna a estatística e o p valor. Podemos separar, mas antes eu quero mostrar que a estatística é 22.

[2:22] O que ele está reportando não é aquele Z que a gente calculou. O Z = - 4,80. O que ele está reportando é o nosso T, o nosso T maiúsculo.

T = min(fumo['Posto (+)'].sum(), fumo['Posto (-)'].sum())

T

22.0

COPIAR CÓDIGO

[2:32] Ele é realmente a estatística, que depois, conforme o n aumenta, como eu disse, converge para uma normal.

[2:40] Então, é esse T que ele reporta no teste.

[2:48] Então, vamos retornar ao trecho de código a seguir.

from scipy.stats import wilcoxon

wilcoxon(fumo.Antes, fumo.Depois)

wilcoxonResult(statistic=22.0, pvalue=1.584310018505865e-06)

p\_valor <= significancia

COPIAR CÓDIGO

O que eu posso fazer é colocar o T. Como ele reporta dois valores. Exato, dois valores. O T e o P valor.

[2:58] Vamos imprimir isso fazendo print(t) e print(p\_valor).

T, p\_valor = wilcoxon(fumo.Antes, fumo.Depois)

print(T)

print(p\_valor)

22.0

1.584310018505865e-06

p\_valor <= significancia

True

COPIAR CÓDIGO

[3:08] Está aqui separado e a gente já consegue fazer o teste: se p valor for menor ou igual a significância, rejeitamos H0. True é sinal de rejeição.

[3:21] A conclusão é: Rejeitamos a hipótese de que não existe diferença entre os grupos, isto é, existe uma diferença entre as médias de cigarros fumados pelos pacientes antes e depois do tratamento. E como é possível verificar através das médias de cigarros fumados por dia antes (31.86) e depois (11.2) do tratamento, podemos concluir que o tratamento apresentou resultado satisfatório.

No próximo vídeo vamos estudar outro teste de comparação de amostras, só que agora as amostras independentes.

[3:32] Lembra aquela comparação que a gente fez de renda de homens e mulheres? Vai ser uma coisa bem parecida.

[3:38] Na verdade, a mesma coisa, só que para uma amostra um pouco menor.