**04 - Aplicando p-valor**

[0:00] Já finalizamos o nosso teste de comparação de duas médias. Concluímos rejeitando o H0. Ou seja, a renda do sexo masculino é maior do que a renda do sexo feminino, por algum motivo, e a gente conseguiu concluir isso a um nível de confiança de 99%,beleza?

[0:17] O que eu vou mostrar para vocês agora é aquele critério mais simples. A gente na mão para entender como todo esse processo funciona e agora que já sabemos o que está acontecendo, será possível usar uma ferramenta de linha de código rápido para executar um teste desse tipo e entender o que aquele resultado representa.

[0:35] Então, vamos lá. Do mesmo jeito que a gente fez antes, mas com algumas novidades, a gente está aplicando no Ztest também. Aqui eu já importei.

from statsmodels.stats.weightstats import DescrStatsW, CompareMeans

COPIAR CÓDIGO

[0:44] Vou mostrar duas formas de fazer: statsmodels.stats.weightstats, aquela mesma coisa, import DescrStatsW, o mesmo que a gente já usou. Agora vem algo novo, o CompareMeans.

[0:56] Vamos abrir as ajudas disponíveis [neste link](https://www.statsmodels.org/dev/generated/statsmodels.stats.weightstats.DescrStatsW.html) e [neste outro link](https://www.statsmodels.org/dev/generated/statsmodels.stats.weightstats.CompareMeans.ttest_ind.html), para visualizarmos algumas informações. Então, vamos começar.

[1:05] A primeira coisa é criar aqueles objetos. Eu vou usar primeiro o DescrStatsW. Só que, nesse caso, como eu tenho duas amostras, eu tenho que criar dois objetos e compará-los.

[1:17] O primeiro, eu vou usar nessa ordem, porque definimos que o Mi1 era a média dos homens e o Mi2 a média das mulheres. Então a gente tem que manter essa ordem para aplicar esse teste aqui.

[1:29] Vou chamar de test\_homens\_H e vou passar aquela mesma coisa, homens. Lembrando que homens é uma amostra com os dados do sexo masculino e mulheres é uma amostra com os dados do sexo feminino.

homens = dados.query('Sexo == 0').sample(n = 500, random\_state = 101).Renda

mulheres = dados.query('Sexo == 1').sample(n = 500, random\_state = 101).Renda

COPIAR CÓDIGO

[1:49] Considerando as fórmulas do H0 e o H1 que eu falei, Mi1 é a média dos homens e o Mi2 é a das mulheres. Manter essa ordem.

[1:58] Então, eu já criei a dos homens, chamei de test\_H.

test\_H = DescrStatsW(homens)

COPIAR CÓDIGO

Agora basta apertar shift + Enter. Vou criar um para as mulheres. Vou colar aqui e mudar só de H para M, e a parte que está entre parênteses eu vou chamar de mulheres.

test\_M = DescrStatsW(mulheres)

COPIAR CÓDIGO

[2:12] Feito isso, perfeito. O primeiro método que eu vou utilizar é o DescrStatsW(). Então, o que eu tenho que fazer? Eu vou chamar otest novamente, como a gente vem fazendo nos outros vídeos.

[2:24] Ele vai ser um objeto que vai fazer a comparação entre esses dois objetos e depois aplicamos um teste neles.

[2:32] Então, eu passo, mantendo a ordem, primeiro o homem, portanto, test = test\_H.get\_compare() e agora eu passo o objeto com os dados das mulheres, perfeito?

[2:48] Ou seja Test é igual a Test\_H.get\_compare(test\_), que é o das mulheres. Rodou.

test\_A = test\_H.get\_compare(test\_M)

COPIAR CÓDIGO

[2:56] O que eu vou fazer agora é usar um ztest, só que para esse tipo de teste aqui. Como é? test.ztest\_ind. Estávamos usando o Min, aqui a gente vai usar o Ind.

test.ztest\_ind(alternative='')

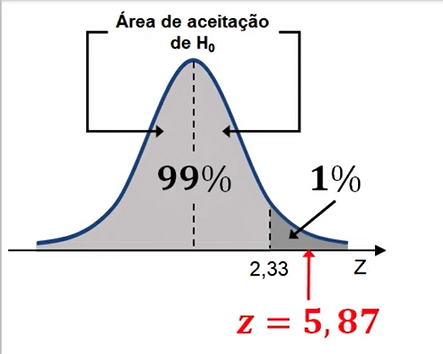
COPIAR CÓDIGO

[3:11] Os parâmetros são os mesmos, alternative igual. A gente fez um teste unicaudal superior, então larger, e o value vai ser igual a zero, porque é o que eu estou testando se os dois são iguais, se a diferença entre os dois é igual a zero versus o contrário disso. Então, vamos lá.

test.ztest\_ind(alternative='larger', value=0)

COPIAR CÓDIGO

[3:37] Rodou e o resultado é a estatística de teste que a gente já obteve: aproximadamente 5,80. Está ela aqui, 5,87.



[4:00] E também o meu p valor, de 2,372 . Eu vou isolar esse p valor, vou chamar esse trecho de código de Z para vir a estatística, seguido do p valor, como a gente tem feito.

z, p\_valor = test\_A.ztest\_ind(alternative='larger', value=0)

COPIAR CÓDIGO

[4:12] Rodou. Ele armazenou o meu p valor. Vou rodar de novo, p\_valor. O p valor está armazenado na variável. Mais abaico, eu já tenho uma comparação com ele, p\_valor <= significancia.

z, p\_valor = test\_B.ztest\_ind(alternative='larger', value=0)

p\_valor

2.2372867859458094e-09

p\_valor <= significancia

True

COPIAR CÓDIGO

[4:22] O p valor é menor ou igual a significância? E, preste atenção, a significância aqui já não é aquele 0,05, é 0,01.

[4:30] Lembra do nosso teste? É 1%. True, ou seja, chegamos à mesma conclusão de que temos que rejeitar H0 e aceitar a hipótese alternativa.

[4:43] Outra forma de fazer é usando o CompareMeans. Como é que eu faço isso? Do mesmo jeito.

[4:50] Vou sobrescrever test igual a CompareMeans, e na ordem, como eu disse, eu vou passar teste homem, test\_H. Vem primeiro, porque eu assumi anteriormente que era primeiro, eu montei o meu teste desse jeito.

[5:03] E teste mulher, test\_M, logo depois. Ele vai rodar, vai criar esse objeto dentro do teste.

test = CompareMeans(test\_H, test\_M)

COPIAR CÓDIGO

[5:11] E a gente pode fazer a mesma coisa embaixo. Vou fazer o seguinte: vou fazer o teste A e aqui eu vou botar teste A, o primeiro que a gente fez.

test\_A = CompareMeans(test\_H, test\_M)

COPIAR CÓDIGO

[5:23] Então vou rodar de novo aqui. Rodou nosso p valor.

[5:27] Aqui vai ser o teste B, só para diferenciar. Não tem sentido nenhum isso daqui, é só para a gente diferenciar os dois, não sobrescrever.

test\_B = CompareMeans(test\_H, test\_M)

COPIAR CÓDIGO

[5:36] Eu rodei o teste B, com o meu p valor igual ao p valor de cima, eu chego à mesma conclusão.

test\_A = test\_H.get\_compare(test\_M)

z, p\_valor = test\_A.ztest\_ind(alternative='larger', value=0)

p\_valor

2.2372867859458094e-09

test\_B = CompareMeans(test\_H, test\_M)

z, p\_valor = test\_B.ztest\_ind(alternative='larger', value=0)

p\_valor

2.2372867859458094e-09

p\_valor <= significancia

True

COPIAR CÓDIGO

[5:45] Duas formas diferentes de chegar ao mesmo objetivo.

[5:49] Era isso que eu queria mostrar, pessoal. Terminamos aqui a nossa seção falando de testes paramétricos.

[5:56] A gente mostrou três formas de fazer isso e, na próxima seção, a gente já vai começar com testes não paramétricos, beleza? Vejo você lá.