

04

Entendendo p-valor

Transcrição

[0:00] Pessoal, continuando o exercício que a gente estava fazendo, vamos falar agora do P valor. Se você já mexeu com estatística provavelmente você já ouviu falar do P valor, e eu vou mostrar para vocês agora como calcular e como interpretar.

[0:10] É uma coisa bem simples. Aqui, lembrando, voltando aqui no nosso quadro do vídeo anterior, onde tem as regras de rejeição de H_0 , aqui tem o P valor, uma delas é o P valor.

[0:20] Esse T a gente vai ver no próximo vídeo. A gente rejeita H_0 se P valor for menor do que Alfa. Alfa é o que? É o nosso nível de significância. Geralmente, a gente ouve falar, é menor que 0,5?

[0:33] Isso quando o nível de significância é de 5%. Se for 1%, ele tem que ser menor do que 0,01. Perfeito? Então preste atenção só nesse detalhe.

[0:43] O que é o P valor? Lembra que a gente calculou a estatística de teste? Ela chegou e caiu aqui, 5,11, ok? O P valor é justamente a probabilidade, a área calculada desse ponto aqui, do 5.

[0:54] Está vendo aqui? Isso aqui é o nível de significância, Alfa sobre dois. Aqui a gente tem uma área, uma probabilidade, para cá, essa areazinha aqui.

[1:07] O P valor é justamente a área, a probabilidade, do valor crítico para cá. O que está querendo dizer aquele teste? É justamente isso daqui. Se essa área aqui for menor do que isso aqui.

[1:18] Isso quer dizer que a minha estatística está dentro da região de rejeição. Ok? Ou seja, quando isso é menor, quando P valor é menor do que Alfa, que é esse cara aqui.

[1:28] Ou seja, ele está em probabilidade. Quando essa probabilidade for menor do que essa aqui, eu rejeito H_0 . Perfeito? É justamente isso, é só uma outra forma mais simples.

[1:38] Porque todo software estatístico vai soltar para você nos testes o P valor, e aí você já tem na sua cabeça que essa regra aqui é a de rejeição da hipótese nula. Perfeito?

[1:49] Então, vamos lá. Vamos calcular o P valor e vamos ver se ele concorda também com essa decisão que a gente tomou no vídeo anterior.

[2:01] Uma das primeiras formas é o que? Eu vou chamar de P valor mesmo. No caso do teste bicaudal, eu tenho que multiplicar por dois. Por motivo que a gente já viu aqui na figura.

[2:12] A gente tem esse cara aqui e esse cara aqui. Ele vai calcular um deles e eu tenho que multiplicar isso por dois, ok? Então duas vezes.

[2:24] Duas vezes. Eu abro e fecho um parêntese aqui. A primeira forma é a seguinte: um menos norm.cdf, Z. Aí eu vou ter um P valor aqui. P valor, esse é o meu P valor.

[2:42] Bem pequeno, não é? Tem um monte de zeros até chegar ao 322. Aqui eu já deixei uma situação formada aqui, que é aquela pergunta: o P valor é menor ou igual a significância, que é 0,05?

[2:57] Aqui eu já consigo responder que é. Se é, eu rejeito H₀, foi a conclusão que a gente chegou aqui em cima também. Então é isso.

[3:07] A outra forma que eu queria mostrar para você, utilizando também o mesmo recurso aqui do Norm, é duas vezes também, só no teste bicaudal, preste atenção nisso.

[3:20] Duas vezes. E aqui eu vou chamar de Norm, também, ponto. Aqui é o SF, e eu passo para ele o Z. O Z, lembrando, que é justamente o cálculo da probabilidade dessa área aqui.

[3:33] Z é esse cara aqui, 5,11. Onde está? Aqui, vamos lá. Não mostrei. P valor.

[3:44] O mesmo carinha aqui. Lógico, tem uma diferencinha aqui. A documentação diz que esse cara aqui, usando o SF, é um pouco mais preciso.

[3:52] Então, você pode escolher porque não vai atrapalhar a sua decisão aqui, lá no final, essa precisão.

[4:03] Então, vamos lá. Fizemos o teste, já vimos o P valor, já entendemos como calcular. Como é que a gente faz isso utilizando o Python, de maneira mais fácil, para não ter que realizar todos aqueles passos? Já que a gente já aprendeu a fazer aquilo.

[4:14] E eu sou um pesquisador experiente, não estou precisando fazer uma prova escrita, como é que eu faço isso usando o Python? Vamos lá.

[4:20] Primeira coisa, eu vou fazer usando esse método aqui. Vou copiar aqui, já dei aqui, a ajuda está ali. Eu vou usar o Ztest, que é justamente isso que a gente está fazendo, um teste Z.

[4:31] Então, From, esse cara aqui, statsmodels.stats.weightstats import Ztest. É isso. Shift + Enter, rodou. Aqui está a ajuda, isso daqui abre essa janela aqui.

[4:51] Eu já vou mostrar as outras. Eu vou mostrar uma segunda opção também de realizar o Ztest, também usando o Statsmodels.

[4:58] E o que eu vou fazer aqui? Chamo o Ztest e passo para ele o que? No caso aqui, X1 vai ser igual a amostra, que é uma amostra que a gente selecionou, está lá em cima aquele Dataframe.

[5:12] E eu passo para ele Value, que é o que? A média que eu estou testando, aqueles 500 ml. Perfeito? Ele vai jogar para a gente o quê? Duas saídas.

[5:23] Uma vai ser a estatística de teste, que é o 5,11 que a gente calculou, e aqui o meu P valor, muito semelhante a esse aqui. 3,2291, E menos 07. São sete zeros, 321.

[5:40] Isso daqui que é o Output do teste. Um outro teste que eu quero mostrar para vocês, também vindo daqui, desse mesmo lugarzinho, só que é esse cara aqui.

[5:53] Eu vou importar esse cara aqui de cima: DescrStatsW. Vou copiar ele aqui, para facilitar a nossa vida, e vou colar ele aqui. Clicando aqui, você vai abrir aquela ajuda lá, do Statsmodels.

[6:16] Para usar essa ferramenta, eu tenho que fazer primeiro o seguinte: eu vou chamar aqui de Test, vou criar um objeto do tipo DescrStatsW e vou passar para ele a minha amostra, a amostra aqui para dentro.

[6:33] Ou seja, eu já estou com esse cara aqui. Com esse objeto teste aqui, aí sim eu vou realizar esse teste aqui, que é o ztest_mean. Posso copiar ele aqui.

[6:46] Pego esse cara aqui, o Test, ponto ztest_mean, passo para ele o quê? O Value, que é igual a média. Por que eu estou passando só o Value?

[6:58] Lembra que eu criei um objeto aqui com a amostra? Ele já tem os meus dados. Eu vou pegar e realizar um teste nesses dados aqui. E ele vai, similar ao de cima, colocar a estatística de teste aqui e o meu P valor aqui.

[7:14] Se eu quiser separar estes dois, eu posso chamar aqui de Z vírgula P valor, vai ser igual a isto aqui. Como ele está dentro de um Array, eu vou fazer um print aqui para o Z, só que eu vou passar o zero aqui, correto?

[7:28] Aí, ele vai me jogar a estatística de teste. Eu posso fazer a mesma coisa para o P valor. O P valor aqui, você lembra que ele estava dentro de um Array, eu quero pegar o elemento que ele estava dentro.

[7:44] Esse aqui, o de cima, era a minha estatística e esse daqui é o meu P valor. Perfeito?

[7:49] Pessoal, era basicamente isso que eu queria mostrar. Aqui já tem um ferramental forte para você utilizar para realizar teste bicaudal.

[8:00] Só mais uma coisinha, antes de a gente finalizar. Eu não configurei nada sobre como ele é bicaudal ou se ele é unicaudal. Por quê? Aqui, esses tipos de testes aqui, essa hipótese aqui Alternative, aqui ele pode dizer two-sided.

[8:14] Ou seja, ele já vem com padrão o bicaudal. No vídeo que vem a gente vai ver o unicaudal e a gente vai alterar essa opção aqui, perfeito?

[8:22] Então, no próximo vídeo teste unicaudal paramétrico, beleza? Até lá.