

Distribuição Qui-quadrado

Transcrição

[0:00] Beleza, pessoal? Começando uma seção nova do nosso curso de estatística, vamos falar agora de testes não paramétricos. E o que é isso, testes não paramétricos?

[0:08] Quando a distribuição dos meus dados amostrais não é uma normal ou então eu não tenho elementos suficientes para afirmar que seja.

[0:14] E geralmente isso acontece com amostras pequenas, eu não consigo validar o teorema do limite central. Aquele que a gente viu no curso anterior, a gente fez até uma prova dele.

[0:24] E nesse caso eu não consigo utilizar métodos paramétricos para testar as minhas hipóteses, então eu tenho que utilizar testes não paramétricos. Por quê?

[0:33] Porque esses caras não fazem suposições sobre a distribuição dos meus dados amostrais, então eu consigo utilizar testes não paramétricos para testar as minhas hipóteses.

[0:43] Começando aqui, como sempre, com um probleminha já, para a gente tentar entender o porquê de utilizar esse tipo de teste que a gente está para conhecer, que é o Qui-quadrado. Vamos lá.

[0:52] "Antes de cada partida do campeonato nacional de futebol, as moedas utilizadas pelos árbitros devem ser verificadas para se ter certeza que não são viciadas, ou seja, que não tendam para determinado resultado".

[1:06] Aquela moedinha no começo do jogo que ele joga para o cara escolher o lado do jogo, a bola, testando ela aqui.

[1:12] "Para isso um teste simples deve ser realizado antes de cada partida. Este teste consiste de lançar a moeda do jogo 50 vezes e contar as frequências de caras e coroas obtidas."

[1:26] "A tabela abaixo mostra o resultado obtido no experimento." Aqui embaixo a gente tem uma tabelinha onde tem a cara e a coroa, que é onde a moeda caiu.

[1:32] E aqui, o observado é o quê? Foram os 50 lançamentos que o cara fez para testar, isso é o que foi observado. E aqui embaixo, o esperado, é o que a gente foi espera.

[1:41] Se aquela moeda for legal, se ela foi muito bem organizada, ela vai ser não-viciada, ela vai cair praticamente 25 vezes para um lado e 25 vezes para o outro. Se ela for honesta.

[1:53] Estou olhando aqui, já parece que essa moeda é meio torta, então a gente vai testar justamente isso aqui.

[1:58] Continuando o problema: "A um nível de significância de 5%, é possível afirmar que a moeda não é honesta, isto é, que a moeda apresenta uma probabilidade maior de cair com a face cara voltada para cima?"

[2:12] Testes desse tipo, a gente consegue solucionar utilizando o teste do Qui-quadrado. Que também é conhecido aqui, já deixei uns textinhos para você ler, como teste de adequação ao ajustamento.

[2:23] E ele utiliza uma variável estatística padronizada elevada ao quadrado, que é justamente essa representada por essa letra grega, o Qui, que é esse X mais sofisticado, elevado ao quadrado.

[2:35] Esse teste, está escrito aqui embaixo também para você não perder, ele testa a hipótese nula de que não existe diferença entre as frequências observadas e as esperadas.

[2:48] Justamente, isso que está na nossa tabela. O que a gente vai testar, a nossa hipótese nula, é o quê? Que não existe diferença entre essas duas frequências; isso aqui de cima é igual a essa aqui, a gente está testando isso. Perfeito?

[2:58] Para isso, a gente vai ter que utilizar os valores tabulados dessa variável, Qui-quadrado, dessa estatística padronizada do mesmo jeito que a gente fez com a normal, do mesmo jeito que a gente fez com o T de Student.

[3:14] Eu vou mostrar para vocês a forma de calcular essa tabelinha, que ela também está atrás de todos os livros de estatística. Vou ensinar como usar essa aqui.

[3:22] Eu estou mostrando aqui para você agora, talvez no seu livro você tenha que entender como ela funciona para utilizar ela, então preste atenção como eu estou utilizando.

[3:33] Se for igual no seu livro, então beleza; se não for, você tem que entender o que a tabela está mostrando.

[3:39] Aqui tem a formula da função, é uma coisa meio assimétrica mesmo. Aqui a área de aceitação de H_0 , já vamos começar com isso.

[3:48] Aqui a estatística, o Qui-quadrado, o valor que a gente vai obter do mesmo jeito que a gente fez lá com o T de Student, é a mesma ideia.

[3:55] E aqui é a área de rejeição de H_0 . A forma de construir a tabela vou deixar aqui de presente também.

[4:01] Está tudo importado aqui, o Pandas, o Scipy.stats import chi, que é justamente essa função para obter essa probabilidade.

[4:12] Aqui tem todo um esqueminha, não vou ficar falando disso daqui, você pode ir analisando.

[4:17] Aqui o chi.ppf, aquele PPF que a gente já conhece, igual ao da normal, igual ao do T de Student.

[4:23] Dá uma olhada, dá uma analisada para você entender como que está funcionando. Rodando isso daqui, Shift + Enter, ele vai mostrar essa tabelinha.

[4:31] No próximo vídeo eu vou mostrar como obter isso aqui usando o PPF, seria como se eu tivesse obtendo um ponto na tabela, então não se preocupe de entender isso aqui, isso é só para construir a tabela para você.

[4:44] Aqui é, basicamente, a mesma coisa da T de Student. Aqui tem os graus de liberdade, N menos um, é o grau de liberdade, até esqueci de colocar aqui.

[4:51] Isso daqui é o P. O P é o quê? A imagem aqui embaixo mostra exatamente isso.

[4:55] P é essa areazinha colorida aqui a partir do Qui-quadrado que a gente vai encontrar, que é uma desses pontos que estão aqui dentro da tabela.

[5:07] Por exemplo: se eu estou assumindo um P de 95%, ou seja, um nível de confiança de 95%, com um grau de liberdade um, que é o que a gente vai utilizar aqui no nosso problema, eu vou ter o quê?

[5:23] Um Qui-quadrado de 3,8415. Ou seja, esse valorzinho aqui de baixo é 3,84 e alguns quebrados.

[5:34] Ele divide isso daqui, 95% fica para um lado e os 5% ficam para o outro lado, que é o um menos P. Perfeito?

[5:42] Pessoal, a tabela é isso, a forma de consultar é essa. A gente vai ver isso no nosso próximo vídeo, quando a gente estiver resolvendo esse nosso probleminha simples aqui.

[5:51] Utilizando o teste de Qui-quadrado. Eu vou mostrar para você como funciona na prática e depois como calcular, de forma simples, utilizando o Scipy, beleza? No vídeo que vem a gente resolve o problema. Abraço.