

 04

Usando a tabela padronizada

Transcrição

[0:00] Legal, ainda falando de distribuição normal, vamos falar de tabelas padronizadas.

[0:06] O que eu quero mostrar aqui é a forma de obter a área sob a curva, ou seja, a probabilidade que estamos interessados, sem ter que calcular integrais definidas, como essa aqui de baixo.

[0:17] Isso porque, lembra que a gente falou que a distribuição normal era definida pela sua média e pelo desvio padrão? Então, cada variável aleatória que se distribui normalmente tem uma média e um desvio padrão específicos, ou seja, vão formar curvas de formatos diferentes, formatos específicos, uma mais achatada, outra mais comprida, e por aí vai.

[0:38] Se a gente tiver que calcular essas probabilidades, para cada curva dessa, a gente vai ter que calcular integrais definidas para cada curva dessa, o que torna o trabalho meio complicado.

[0:48] Pra resolver esse problema, a gente pode usar as tabelas padronizadas, onde, para consultar essas tabelas que já têm os valores das probabilidades definidos, tudo certinho, você só precisa transformar sua variável em uma variável normal padronizada, da seguinte forma, como mostra essa formulazinha aqui, a gente vai partir de X , que é nossa variável tradicional, e transformar em Z , como? Pegando X , diminuindo da média da variável e dividindo pelo seu desvio padrão, pronto, isso padroniza a variável e a gente pode consultar a tabela, eu vou mostrar para vocês como consultar essas tabelas a partir do próximo vídeo, onde vamos solucionar aquele probleminha que falamos no começo, que eu propus no começo, antes de começar a falar de distribuição normal, legal? Então, atrás de cada livro de estatística, se você tiver um aí, você pode ir olhar, você tem as tabelas de distribuições de probabilidade, vai ter a da normal, da distribuição T , q^2 , e por aí vai.

[1:49] Eu vou passar para vocês aqui uma forma de obter sua própria tabela, mas se você quiser consultar do livro, vai ser basicamente o mesmo valor, o que estou criando aqui é um dataframe onde eu vou colocar esses valores de probabilidade, vou construir a tabela para você ver, essa tabela que vamos consultar aqui, tem o valor de Z , no próximo vídeo vou mostrar como consultar, o Z está aqui, vamos ter que somar o valor da linha com o da coluna para obter O Z e buscar na tabela.

[2:26] Ele vai de 0 até 4, ou de 3.99, cobre praticamente 100% da distribuição, e esse valor que está aqui dentro, essa probabilidade, essa área sob a curva, ela representa, essa área hachurada, ela vai de menos infinito até o valor da variável padronizada que encontramos, até Z , ok? Algumas tabelas vão estar um pouco diferentes dessa, a não ser o livro, algumas tabelas vão mostrar para você da média que é o valor central, que parte a distribuição, lembra que é simétrica em relação à média, parte ao meio, 50% para cá e para lá, e o valor da tabela vai da média até Z , isso não é um problema, porque para obter essa tabela, você simplesmente tem da média até Z , somaria 0.5, tem o mesmo valor da tabela, tá? Com o tempo você vai pegando os macetes e percebendo o que estou falando.

[3:26] Calma, a gente vai usar a tabela e você vai começar a entender esses pontos, ok? Pessoal, o que eu queria mostrar aqui é justamente a construção dessa tabela, vou dar ela de presente para vocês, vamos usar nos próximos vídeos, aqui não tem nada demais, estou criando um dataframe, a única coisa nova para você vai ser essa do `scipy start import norm`, vamos também ver nos próximos vídeos, vou explicar o que significa.

[4:00] No próximo vídeo, começamos a resolver os exercícios, beleza? Vejo você lá.