

Aplicando a tabela - Caso I

Transcrição

[0:00] Legal galera, voltando na distribuição normal, entendemos mais ou menos como funciona a distribuição normal, falamos sobre as tabelas padronizadas, a variável normal padronizada, então agora vamos voltar no probleminha que propus no começo, e vamos tentar resolver ele utilizando o que aprendemos um pouco.

[0:20] O problema é, em um estudo sobre as alturas dos moradores de uma cidade, verificou-se que o conjunto de dados segue uma distribuição normal, aproximadamente normal, com média 1.7 e desvio padrão 0.1, maravilha, com essas informações obtenha o seguinte conjunto de probabilidade.

[0:41] vamos resolver o primeiro, o A, nos próximos vídeos, vamos resolvendo os outros, o primeiro, probabilidade de uma pessoa selecionada ao acaso, aleatoriamente, ter menos de 1 metro e 80, então aqui fiz um desenho para representar o que estamos procurando, sob a curva normal, que área estamos procurando sob a curva normal, a gente tem a média, 1 metro e 70, o cara que parte a distribuição exatamente ao meio, daqui para cá, eu sei que tem 50%, a probabilidade aqui é 0.5, só que está me faltando saber esse cara, esse eu não sei, de 1 metro e 80, até a média, não sei, é isso que estou querendo saber, daqui eu sei que 50%, porque sou esperto, eu sei que é simétrica em relação a média, então 50% para um lado, 50% pro outro. Agora, repare uma coisa, lembra a figurinha que mostrei para vocês da tabela padronizada? É justamente isso que queremos, de menos infinito até Z, preciso fazer o que? Padronizar essa variável, a gente está aqui, 1 metro e 80, com nossa variável original, para justamente saber isso, essa área, de menos infinito até 1 metro e 80, como que fazemos isso? Primeiro passo, padronizar a variável para poder utilizar nossa tabela, consultar nossa tabela.

[2:04] Então, vamos lá, primeira coisa, a informação que o próprio problema me diz, a média é igual a que? 1 metro e 70, aqui na figura, aqui em cima, vamos colocar 1.7.

[2:20] Média, só para ver de novo, o desvio padrão, desvio padrão, ele me falou também, é 0.1, correto? Vamos olhar lá em cima, desvio padrão 0.1, ok, estamos indo bem, vamos colocar também para mostrar.

[2:38] Como que obtém Z? Voltando aqui, na nossa teoria, está aqui, $Z = x$ menos a média dividido pelo desvio padrão.

[2:49] X é a variável original, 1 metro e 80, é o que estamos procurando, probabilidade da pessoa ter menos de 1 metro e 80, vamos fazer essa conta, Z, abre e fecha parênteses, põe aqui 1.8-média, criamos a variável média aqui em cima, dividido pelo desvio padrão, também calculamos esse aqui, e Z vai ser igual a 1, perfeito? 1, Z é 1.

[3:26] Como que faço para consultar esse cara na tabela, e obter aqui embaixo, está pedindo, utilize a tabela, a solução é utilizar a tabela, a outra vai mostrar como.

[3:38] Probabilidade vai ser igual ao valor que vou consultar na tabela, tenho $Z = 1$, como faço para consultar? Essa é simples, não falei ainda, mas o Z vai ser construído da seguinte forma, tenho que somar o valor da linha com o da coluna, por exemplo, achei um Z de 0.15, nessa linha, para que eu ache o 0.1, mais quanto falta para chegar no 0.15? falta o que? 0.05, então venho aqui, esse mais 0,05, e eu pego esse valorzinho aqui, é o correspondente a $Z = 0,15$, 0,1 mais 0,05, ok? No nosso caso, não vamos nem precisar fazer soma, porque deu 1, ou seja, venho nessa linha, procuro 1, 1 está aqui, 1 com 0, aqui, 1 mais 0, 1, perfeito? Esse aqui, é o valor que estou procurando, 0,8413, venho aqui, simplesmente colo aqui, perfeito? Esse é o valor da probabilidade que estou procurando, justamente esse cara que quero, a chance de eu encontrar um cara que tenha menos de 1 metro e 80 selecionando aleatoriamente é de 84.13%, é isso, uma forma mais simples, utilizando o scipy, vamos importar novamente, já está importado mas vou importar novamente para deixar claro de onde está vindo o nome, não ficar nenhum vazio, eu vou importar esse do scipy stats, import o nome, ele está definido aqui, deixei aqui, esqueci de falar no vídeo anterior.

[5:37] A ajuda, a documentação está aqui, o que preciso é justamente aqui embaixo, encontrar o CDF, cumulative distribution function, justamente quem eu uso para preencher as células do dataframe, tá? Então, vamos lá, onde está? Aqui, norm, o que eu passo pro norm? Bem simples, venho aqui, chamo norm novamente, norm.cdf e passo para ele o Z que encontrei aqui em cima, que é 1, pode ser, vai dar um valor um pouco diferente porque está usando outro tipo, não está usando a tabela que nem estamos usando, ela está arredondada, simplificada, que é uma coisa para simplificar nosso problema, aqui podemos ver que as casas decimais batem, 0,8413, mas o Z não é exatamente 1, é 1,0009, então ele faz um cálculo mais exato, é mais preciso, ok? Mas a gente consegue obter com boa precisão utilizando a tabela manualmente, mas hoje em dia não usamos muito lápis e papel e não ser que estamos fazendo uma prova, então podemos usar o computador, e isso seria a forma de consultar a tabela padronizada utilizando Python, beleza? Próximo vídeo, a gente faz a opção B do nosso problema, que tem outra configuração, temos que pensar melhor como consultar a tabela, vejo você lá, abraço.