

05

Variância

Transcrição

[0:00] Ok, continuando com medidas de dispersão, vamos falar da variância, que como o desvio médio absoluto que a gente viu no vídeo anterior também é uma média dos desvios, só que no caso da variância, a solução pra aquele problema da soma dos desvios em relação à média ser igual a 0, não é pegar somente os valores absolutos, ou seja, os valores positivos. Aqui, no caso, os desvios são elevados ao quadrado, pra resolver o problema do sinal.

[0:24] Aqui, uma média, a soma dos quadrados dos desvios dividido pelo número total. No caso da variância, temos um diferencial no cálculo dela, quando estamos trabalhando com dados populacionais e amostrais, e no caso da estatística, geralmente estamos trabalhando com dados amostrais, então geralmente vamos usar a fórmula de baixo. No caso do populacional, é representado por σ^2 , que é essa letra grega, sigma, igual ao somatório dos desvios em relação a média populacional, representado pelo μ , dividido por n .

[1:03] No caso do amostral, a gente divide por $n-1$, que é conhecido como fator de correção de Bessel.

[1:09] Eu não vou entrar nos detalhes matemáticos, nas provas matemáticas desse fator de correção, porque como nosso curso tem uma pegada mais de aplicação, pegada mais prática, acho que não faz sentido a gente mexer com isso agora. Pros curiosos, que quiserem saber mais sobre esse assunto, só procurar sobre fator de correção de Bessel nos livros de referência, estatística, na internet, você vai ter uma ideia de por que utilizaram isso aqui. A ideia básica que temos que saber é que utilizando esse cara aqui a gente tem uma estimativa mais precisa do parâmetro populacional, da variância populacional, ok, basicamente é isso que temos que saber. Vamos ao cálculo, utilizando nosso DF que fizemos anteriormente. Vamos pegar aqui em cima, notas fulano desvio, o que é justamente isso que vou criar, você que acompanhou o vídeo anterior viu como fizemos esses cálculos.

[2:11] Quero criar agora o desvio em relação à média mas elevado ao quadrado, vou colocar aqui esse cara, um chapeuzinho, ao quadrado.

[2:22] É igual a notas desvio, no caso do Pandas, ele usa a função `POW2`, elevando o valor dos desvios que calculamos no vídeo anterior, ao quadrado.

[2:37] Esqueci de mostrar aqui, notas fulano, mostrar pra ter uma ideia do que foi calculado, aqui. O que tínhamos feito na aula anterior, esses três primeiros, nota, o desvio, o desvio absoluto, valores positivos, e agora o desvio ao quadrado.

[2:56] Como calculamos variância no caso da amostral?

[3:01] Vou mostrar aqui pra você, pegar esse cara aqui que é o desvio ao quadrado, não vou aplicar direto a média nele, porque se eu aplicar, como fiz no vídeo anterior, ele vai usar como denominador o N , eu quero que use $n-1$, então vou fazer na mão, no braço, vou somar, e vou dividir isso, abro e fecho um parêntese, vou colocar o $n-1$ aqui, notas fulano, colocar aqui. Não escrevi nada errado.

[3:43] Não vou usar o `shape` não, vamos fazer com `Len`, temos também o número de registros do dataframe usando isso aqui, `len-1`. Ou seja, $n-1$, e aqui, tenho o valor da variância das notas de fulano, ok, como calculo isso de forma extremamente mais simples utilizando o pandas?

[4:12] Variância é igual a esse cara só que quero fazer em cima das notas dele direto, `fulano.var`, é só isso. Novamente, esquecendo de mostrar o resultado pra vocês, e tá aqui, o mesmo valorzinho da variância.

[4:37] A variância tem uma característica, como elevei todos os desvios ao quadrado, elevo também as unidades dele ao quadrado, por exemplo, o que quero dizer, vamos supor que eu esteja calculando a variância da variável altura, que a gente viu, tem no nosso dataset, a gente vai utilizar no final do curso, pra fazer a análise descritiva. Está em metros.

[5:00] Se eu calculo a variância, vou obter um resultado em metros ao quadrado, deixa de ser, estamos analisando área. Metro ao quadrado, estamos vendo unidade de área, que é difícil de interpretar, por isso, pra resolver o problema, o desvio padrão apareceu, e a gente vai conhecer ele no próximo vídeo, beleza? Até lá, abraço.