

Calculando a probabilidade

Transcrição

[0:00] Ok, continuando aqui, você lembra que eu deixei probleminha no ar, no vídeo anterior, a gente vai voltar agora, vamos tentar identificar se conseguimos resolver esse cara utilizando a distribuição binomial, e se conseguirmos, vamos resolver ele, vamos mostrar duas alternativas para fazer isso.

[0:17] Então, reler o problema, em um concurso para preencher uma vaga de cientista de dados, temos um total de 80 questões de múltipla escolha, com 3 alternativas cada.

[0:27] Cada questão tem o mesmo valor, e supõe que um candidato maluco resolva se aventurar sem ter estudado absolutamente nada e resolva chutar todas as questões de forma aleatória.

[0:34] Assumindo que a prova vale 10 pontos e 5 seja a nota de corte, ou seja, 5 para cima passa, menos que isso é reprovado, qual a probabilidade do candidato passar para a próxima etapa nesse processo seletivo? Vamos lá, a primeira coisa é tentar identificar se esse cara é um experimento binomial.

[1:02] Aqui, temos as perguntas que fizemos aqui em cima, aqui, experimento binomial, realização de ensaios idênticos.

[1:11] O que temos na nossa questão? 80 questões de múltipla escolha com 3 alternativas cada, ou seja, são ensaios idênticos.

[1:19] O N aqui, a gente conseguiu identificar, 80, então vamos lá, acabei de falar sem querer, qual o número de ensaios? $N = 80$, temos 80 questões, então vamos lá, N vale 80, vamos rodar aqui.

[1:34] Mais uma perguntinha antes de seguir, os ensaios são independentes? Sim? Eu deixei essa resposta aqui, sim.

[1:46] A opção escolhida em cada questão não influencia em nada na outra questão, o cara resolveu chutar aleatoriamente, está de olho fechado marcando X em cada questão então não influencia em nada o que ele marcou numa questão, o que ele marcou na outra, independentes.

[2:03] Segundo ponto, vamos lá, somente dois resultados são possíveis em cada ensaio? Aqui o cara pode ficar em dúvida, por que? Tem 80 questões de múltipla escolha com 3 alternativas possíveis cada, mas aqui a gente tem que perceber que dessas três alternativas, somente uma é correta, e ele está tentando escolher entre acertar a questão ou não, ou seja, cada ensaio desse aqui tem duas opções, ou ele acerta a questão, ou ele erra, existem probabilidades diferentes de acertar ou errar, ok, mas esse não é o problema, não é o X da questão, o que está sendo perguntado é somente dois resultados são possíveis em cada ensaio? Em uma questão, quando você marca um X, você só pode ou acertar ou errar, então, realmente, sim, o candidato tem duas possibilidades, acertar ou errar a questão.

[2:59] Vamos lá, estamos realmente diante de um experimento binomial, só falta definirmos a probabilidades de sucesso e fracasso, e verificar se essas probabilidades se repetem em cada experimento, está resolvido.

[3:14] Vamos calcular P, como calculamos P? Bem simples, vamos começar aqui criando uma variável, número de alternativas por questão.

[3:29] Ok, número de alternativas por questão, número de variável e nome.

[3:35] Qual o número de alternativas ou questão? Acabamos de falar, 3, e esse é o espaço amostral do experimento, só tem 3 alternativas possíveis dentro do experimento, ou marco A, ou B, ou C.

[3:50] Muito bem, P , que é a probabilidade de sucesso, de acerto, que é o que estou procurando, é o que? É obtido da seguinte forma, ele tem uma chance em 3 de obter sucesso, ou seja, de acertar a questão, então, 3 aqui, número de alternativas por questão, e conseguimos definir P , que é 0,333.

[4:14] É um dividido por 3, ok? O que é a probabilidade de fracasso? Justamente o que? A gente definiu isso também, Q é 1 menos P , perfeito? Está lá, 0,6666, no final, 7.

[4:33] Essas duas probabilidade, logicamente, como temos cada questão com 3 alternativas cada, elas não vão se alterar, não tem diferença entre as questões, então estamos diante de um experimento que a gente consegue resolver utilizando a distribuição binomial.

[4:49] O único ponto que falta da nossa formula aqui em cima, olha ela aqui, da nossa fórmula, é o que? Achar o número de eventos desejados que tenham sucesso.

[5:02] Como descobrimos isso no nosso problema? O que estou dizendo aqui, está incutido no problema, não está bem na cara, mas é simples de perceber, temos 80 experimentos, aqui está dizendo que essas 80 questões tem o mesmo valor, não importa o valor da questão, o que me importa é isso aqui, a prova vale 10 pontos, e estou dizendo que se o cara acertar metade da prova, ele passa.

[5:30] Se não, ele é reprovado, então, é justamente isso que eu quero, quero saber a chance do cara, a probabilidade desse cara, chutando a prova toda, dele acertar 5, só isso, acertou 5, ele passa.

[5:45] Como todas as questões tem o mesmo valor, você concorda que 10 pontos acertar 5 é exatamente metade da prova? Ou seja, 80 dividido por 2, 40 questões, perfeito? Vamos lá, K é o que, vamos botar aqui, a gente deduziu que 80 dividido por 2, metade da prova, então, vamos lá, K é isso, muito bem, temos todos os elementos da nossa fórmula, aqui do lado, tem o K , P , Q , efeito.

[6:19] Vou mostrar para vocês como estou solucionando manualmente, vou mostrar a forma manual, vou aplicar essa fórmula de probabilidade, vamos chamar isso de probabilidade, é igual.

[6:37] Vou separar em três partes, tem a combinação, depois P elevado a Q , depois Q elevado a $N-K$, P elevado a K , desculpa.

[6:45] Então, vamos fazer 3 parenteses aqui, vezes, ok? Só para gente dividir bem nosso raciocínio, não perder nada.

[6:56] O primeiro vamos colocar a combinação de NKK , então, vamos lá, Comb, aquele carinha da Comb lá, N que eu já calculei, vírgula K , também já calculei, está aqui.

[7:09] Qual essa combinação? 80, 40, 40, ok? N , K , K .

[7:18] Perfeito, segunda parte, P elevado aqui, no python, elevado é 2 asteriscos, tá? Elevado a K , voltando na fórmula, P elevado a K , e o próximo é Q elevado a N menos K , vamos escrever aqui, Q asterisco asterisco, vamos fechar o parênteses, N menos K .

[7:46] Perfeito? Essa é a nossa fórmula, esse é o nosso resultado, vou aproveitar esse print que usei aqui, que está bonitinho, vou fazer para ele aqui também, só que aqui vou botar menos de 15, vou botar 8, pode ser? Então vamos lá, calcular a probabilidade desse cara maluco vir, chutar tudo e passar na prova, é bem pequena, 0,00079972, essa é a chance pode transformar esse em percentual, multiplica por 100, fica 0,08% de chance desse doido varrido vim, chutar tudo e ainda conseguir passar, chance bem pequena, ainda bem, para quem estuda.

[8:39] Vou mostrar agora, e você vai ficar até irritado, mas todas as etapas que passei, é importante na fase de aprendizado, porque a gente consegue entender o que está acontecendo, da onde vem N , de onde vem K , como aplica essa fórmula, por que disso e daquilo, mas o Python me oferece uma forma bem simples de calcular isso, vou fazer From, scipy de novo, .stats

agora import binom, ok? Se estiver tudo certinho, vou usar esse cara aqui, posso usar esse probabilidade também, para fazer a segunda solução que quero mostrar para vocês, é igual a binom.pmf, e aqui, eu passo os parâmetros que quero, no caso do pmf, deixei aqui.

[9:38] Vamos lá, rapidamente, a gente está aprendendo, eu deixei os parâmetros do jeito que está na ajuda, tem aqui, fórmula do que conhecemos, aquela fórmula, deixa eu ver se eu vejo os parâmetros do jeito que eu deixei, pmf, está vendo, o pmf, é probability mass function. Knp, é esse que você tem que oferecer pro pmf.

[10:06] Dessa forma, deixei do jeito que precisa passar, o K, definimos o K, que é o número de sucessos que estou procurando, o N que é o número de experimentos, e o P que é a probabilidade de sucesso, Ok? É a mesma coisa, probabilidade, tá, e aqui, deveria dar o número igualzinho ao aqui de cima.

[10:34] Bem simples, lógico, a gente precisa ter o K, o N e o P, precisamos ter, perceber tudo isso, calcular tudo isso, mas depois não precisa calcular um monte de coisa, só tem o K, o N e o P.

[10:53] Passa para função binom.pmf, e ele já te dá a probabilidade que estamos procurando, legal? A maioria dos problemas aqui, vou tentar resolver inicialmente passo a passo, uma coisa meio manual, e depois vou mostrar a forma mais simplificada, vai funcionar, a dinâmica do treinamento, vai ser mais ou menos essa, é legal para você ir entendendo os passos, e depois aprender a forma de fazer bem simples quando você tiver K, e já sabe o que está vindo lá dentro, legal? Vou fazer mais um exemplo, lembra que eu prometi falar da coisa da média aqui no binomial? Esse cara aqui, então, eu vou mostrar um exemplo, no próximo vídeo, de aplicação dessa média, a gente se vê lá, abraço.