2. Probabilidade e distribuição

1. **Um dado justo de seis faces é lançado duas vezes. Qual é a probabilidade de obter um número par no primeiro lançamento e um número ímpar no segundo lançamento?**

2. **Um dado honesto é lançado repetidamente até que um número maior que 4 seja observado. Qual é a probabilidade de você precisar de 4 jogadas para obter um número maior que 4?**
3. **Em uma escola, 60% dos alunos estudam Matemática e 40% estudam Ciências. Sabendo que 30% dos alunos estudam tanto Matemática quanto Ciências, qual é a probabilidade de selecionar um aluno aleatoriamente e ele estudar apenas uma das duas disciplinas?**
4. **Em um baralho padrão de 52 cartas, qual é a probabilidade de tirar uma carta de ouros e, em seguida, uma carta de espadas, se as cartas são retiradas com reposição?**
5. **Uma empresa de tecnologia produz dois modelos de smartphones: A e B. Sabe-se que 60% dos clientes escolhem o modelo A e 40% escolhem o modelo B. Além disso, sabe-se que 5% dos clientes que compram o modelo A relatam problemas técnicos, enquanto 10% dos clientes que compram o modelo B relatam problemas técnicos. Se um cliente escolhido aleatoriamente relatar problemas técnicos, qual é a probabilidade de que ele tenha escolhido o modelo A?**
6. **Uma clínica médica realiza dois testes (Teste A e Teste B) para diagnosticar uma doença. Sabe-se que o Teste A detecta corretamente a doença em 95% dos casos em que ela está presente, mas também pode dar um resultado positivo em 10% dos casos em que a pessoa está saudável. Já o Teste B tem uma taxa de acerto de 90% para detecção da doença e uma taxa de falsos positivos de 5%. Se um paciente testar positivo em ambos os testes, qual é a probabilidade de ele realmente ter a doença?**

   Considere que a probabilidade de uma pessoa desenvolver a doença é 0.1
7. **Uma loja de eletrônicos oferece três modelos de smartphones (A, B e C) e os clientes podem escolher um dos três modelos. Sabe-se que 50% dos clientes escolhem o modelo A, 30% escolhem o modelo B e 20% escolhem o modelo C. Além disso, a loja tem uma política de garantia estendida, onde 10% dos clientes que escolhem o modelo A compram a garantia estendida, 15% dos clientes que escolhem o modelo B compram a garantia estendida e 20% dos clientes que escolhem o modelo C compram a garantia estendida. Se um cliente é selecionado aleatoriamente, qual é a probabilidade de ele comprar a garantia estendida?**

8. Plus: **Qual a diferença entre probabilidade e verossimilhança (likelihood)?**

9. **Imagine que Jeremy participou de um exame. O teste está tendo uma pontuação média de 160 e um desvio padrão de 15. Se o escore z de Jeremy for 1,20, qual seria sua pontuação no teste? Escreva o racional do cálculo**

Respostas

1. **Um dado justo de seis faces é lançado duas vezes. Qual é a probabilidade de obter um número par no primeiro lançamento e um número ímpar no segundo lançamento?**

Par = {2, 4, 6} = 3/6 = 1/2 Ímpar = {1, 3, 5} = 3/6 = 1/2

1/2 * 1/2 = 1/4 = 0.25 = 25%

2. **Um dado honesto é lançado repetidamente até que um número maior que 4 seja observado.**

O dado tem 6 faces, e queremos que o número observado seja maior que 4. Portanto, há duas faces que satisfazem essa condição: 5 e 6.

Vamos analisar as possibilidades de sequências de jogadas para alcançar esse resultado em exatamente 4 jogadas. O cenário que nos interessa:

1. Primeiro lançamento: 1, 2, 3 ou 4 (não maior que 4)
2. Segundo lançamento: 1, 2, 3 ou 4 (não maior que 4)
3. Terceiro lançamento: 1, 2, 3 ou 4 (não maior que 4)
4. Quarto lançamento: 5 ou 6 (maior que 4)

Agora, vamos calcular a probabilidade de cada etapa.

A probabilidade de obter um número menor ou igual a 4 em cada lançamento é 4/6, pois há quatro números entre 1 e 4 nas seis faces do dado.

A probabilidade de obter um número maior que 4 no quarto lançamento é 2/6, pois há duas faces (5 e 6) que satisfazem essa condição.

A probabilidade total de precisar de exatamente 4 jogadas para obter um número maior que 4 é o produto das probabilidades em cada etapa:

4/6*4/6*4/6*2/6 = 0.098 = 9.8%

3. Em uma escola, 60% dos alunos estudam Matemática e 40% estudam Ciências. Sabendo que 30% dos alunos estudam tanto Matemática quanto Ciências, qual é a probabilidade de selecionar um aluno aleatoriamente e ele estudar apenas uma das duas disciplinas?

P(M) = probabilidade de estudar matematica apenas

P(C) = probabilidade de estudar ciencia apenas

P(M e C) = probabilidade de estudar matemática E ciência

P(M ou C) = = probabilidade de estudar matemática OU ciência

Sabemos que:

P(M ou C) = P(M) +P(C) - P(M e C)

Do enunciado, temos uma leve pegadinha:

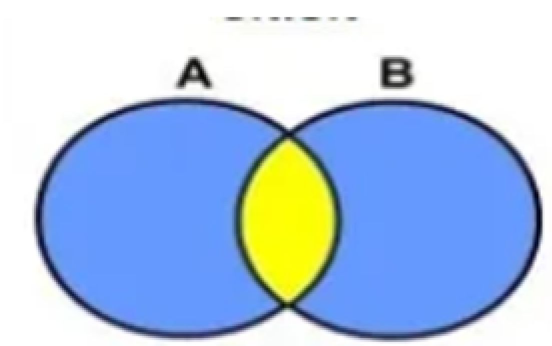
P(M) = 0.6 → cursa apenas matemática OU matemática + ciência

P(C) = 0.4 → cursa apenas ciência OU matemática + ciência

P(M e C) = 0.3 → cursa matemática + ciência

P(M ou C) é a probabilidade de cursas matemática, ciência ou ambas! Por isso, não podemos usar a fórmula P(M ou C) = P(M) +P(C) - P(M e C) diretamente, pois queremos a probabilidade de cursar apenas 1.

Para entender melhor, veja o diagrama abaixo. Cursar apenas 1 disciplina corresponde a área azul (não podemos incluir a área amarela, pois nesse caso ele cursaria ambas)



Vamos supor que A é matemática e B é ciência. Temos a área de ambos, ou seja, P(M) + P(C) = 0.6 + 0.4. Contudo, note que se somarmos dessa forma, contariamos a área amarela 2 vezes (uma quando incluimos o A inteiro e outra quando incluimos o B inteiro). Essa área amarela é P(M e C).

Portanto, para termos unica e exclusivamente a área azul, teríamos que:

P (apenas 1 matéria) = P(M) +P(C) - 2P(M e C) = 0.6 + 0.4 - 2*0.3 = 0.4 = 40%

4. Em um baralho padrão de 52 cartas, qual é a probabilidade de tirar uma carta de ouros e, em seguida, uma carta de espadas, se as cartas são retiradas com reposição?

P(espadas) = 13/52

P(ouros) = 13/52

Sem reposição: um não afeta o outro

P(total) = 13/52*13/52 = 0.0625 = 6.25%

5. Uma empresa de tecnologia produz dois modelos de smartphones: A e B. Sabe-se que 60% dos clientes escolhem o modelo A e 40% escolhem o modelo B. Além disso, sabe-se que 5% dos clientes que compram o modelo A relatam problemas técnicos, enquanto 10% dos clientes que compram o modelo B relatam problemas técnicos. Se um cliente escolhido aleatoriamente relatar problemas técnicos, qual é a probabilidade de que ele tenha escolhido o modelo A?

P(A|Problema) = ?

P(A) = 0.6

P(B) = 0.4

P(Problema|A) = 0.05

P(Problema|B) = 0.10

P(Problema) = P(Problema|A)P(A) + P(Problema|B)P(B) = 0.05*0.6 + 0.1*0.4 = 0.07

P(A|Problema) = P(A)P(Problema|A) /P(Problema) = 0.6*0.05/0.07 = 0.42 = 42%

6. Uma clínica médica realiza dois testes (Teste A e Teste B) para diagnosticar uma doença. Sabe-se que o Teste A detecta corretamente a doença em 95% dos casos em que ela está presente, mas também pode dar um resultado positivo em 10% dos casos em que a pessoa está saudável. Já o Teste B tem uma taxa de acerto de 90% para detecção da doença e uma taxa de falsos positivos de 5%. Se um paciente testar positivo em ambos os testes, qual é a probabilidade de ele realmente ter a doença?

Considere que a probabilidade de uma pessoa desenvolver a doença é 0.1

A = positivo no teste A

B = positivo no teste B

P(doença|A e B) = ?

O que temos:

P(A|doença) = 0.95

P(A|nao doença) = 0.10

P(B|doença) = 0.90

P(B|nao doença) = 0.05

P(doença) = 0.1

P(não doença) = 1- P(doença) = 1 - 0.1 = 0.90

Pelo teorema de **Bayes**, temos também que

P(doença|A e B) = P(A e B|doença)*P(doença) /P(A e B)

Também temos que :

P(A e B|doença) = P(A|doença) * P(B|doença)

Da mesma forma

P(A e B|não doença) = P(A| não doença) * P(B|não doença)

Agora, podemos calcular a probabilidade de uma pessoa testar positivo em ambos os testes (Regrinha do “E”. Considerando ambos os cenários, tem ou não a doença):

P(A e B) = P(A e B | doença)*P(doença) + P(A e B | nao doença)*P(não doença)

Substituindo todos esses valores na fórmula inicial do Teorema de Bayes, podemos encontrar a probabilidade desejada. Vamos calcular isso:

P(doença|A e B) = P(A|doença) * P(B|doença)*P(doença) / [P(A e B | doença)*P(doença) + P(A e B | nao doença)*P(não doença)]

P(doença|A e B) = 0.95*0.9*0.01/[P(A|doença) * P(B|doença)*P(doença) +P(A| não doença) * P(B|não doença)*P(não doença)]

P(doença|A e B) = 0.95*0.9*0.1/[0.95*0.9*0.1 + 0.10*0.05*0.90]

P(doença|A e B) =0.95 = 95%

7. Uma loja de eletrônicos oferece três modelos de smartphones (A, B e C) e os clientes podem escolher um dos três modelos. Sabe-se que 50% dos clientes escolhem o modelo A, 30% escolhem o modelo B e 20% escolhem o modelo C. Além disso, a loja tem uma política de garantia estendida, onde 10% dos clientes que escolhem o modelo A compram a garantia estendida, 15% dos clientes que escolhem o modelo B compram a garantia estendida e 20% dos clientes que escolhem o modelo C compram a garantia estendida. Se um cliente é selecionado aleatoriamente, qual é a probabilidade de ele comprar a garantia estendida?

Podemos abordar esse problema usando o conceito de probabilidade condicional e a Lei da Probabilidade Total. Vamos denotar os eventos da seguinte forma:

Queremos calcular a probabilidade de um cliente escolhido aleatoriamente comprar a garantia estendida, ou seja, (P(E).

De acordo com as informações dadas:

- P(A) = 0,50 (probabilidade de escolher o modelo A)
- P(B) = 0,30 (probabilidade de escolher o modelo B)
- P(C) = 0,20 (probabilidade de escolher o modelo C)
- P(E|A) = 0,10 (probabilidade de comprar a garantia estendida dado que escolheu o modelo A)
- P(E|B) = 0,15 (probabilidade de comprar a garantia estendida dado que escolheu o modelo B)
- P(E|C) = 0,20 (probabilidade de comprar a garantia estendida dado que escolheu o modelo C)

Usando a Lei da Probabilidade Total:

P(E) = P(E|A)**P(A) + P(E|B)**P(B) + P(E|C)*P(C)

Substituindo os valores:

P(E) = 0,10 *0,50 + 0,15 * 0,30 + 0,20 *0,20

P(E) = 0,05 + 0,045 + 0,04 = 0,135

Portanto, a probabilidade de selecionar um cliente aleatoriamente e ele comprar a garantia estendida é de 0,135, ou seja, 13,5%.

8. **Qual a diferença entre probabilidade e verossimilhança (likelihood)?**

A probabilidade e a verossimilhança (likelihood) são conceitos relacionados, mas têm significados distintos e são usados em contextos diferentes na estatística.

Probabilidade: A probabilidade é uma medida quantitativa que expressa a chance ou a frequência relativa de ocorrência de um evento. Ela varia de 0 a 1, onde 0 indica impossibilidade e 1 indica certeza. A probabilidade é usada para fazer previsões sobre eventos futuros ou para descrever a incerteza associada a diferentes resultados. Por exemplo, ao lançar um dado justo de seis faces, a probabilidade de obter qualquer número específico é de 1/6, pois há seis resultados igualmente prováveis.

Verossimilhança (Likelihood): A verossimilhança é uma medida usada para avaliar o quão bem um modelo estatístico se ajusta aos dados observados. Em outras palavras, é a probabilidade de observar os dados que temos, dado um conjunto específico de parâmetros do modelo. A verossimilhança é frequentemente usada para estimar os parâmetros desconhecidos do modelo, ajustando-os de forma que os dados observados sejam os mais prováveis possíveis, dadas as suposições do modelo.

A principal diferença entre probabilidade e verossimilhança reside em como são usadas:

- **Probabilidade:** Usada para prever ou calcular a chance de eventos futuros ou a ocorrência de resultados específicos em situações conhecidas.
- **Verossimilhança:** Usada para avaliar a adequação de um modelo estatístico aos dados observados, geralmente para encontrar os parâmetros que melhor explicam esses dados.

Enquanto a probabilidade se concentra na frequência de eventos antes que eles ocorram, a verossimilhança se concentra em como os dados observados podem nos informar sobre a adequação do modelo estatístico após a ocorrência dos eventos.

9. Imagine que Jeremy participou de um exame. O teste está tendo uma pontuação média de 160 e um desvio padrão de 15. Se o escore z de Jeremy for 1,20, qual seria sua pontuação no teste? Escreva o racional do cálculo

$1.20 = (x - 160)/15$ $X = 178$