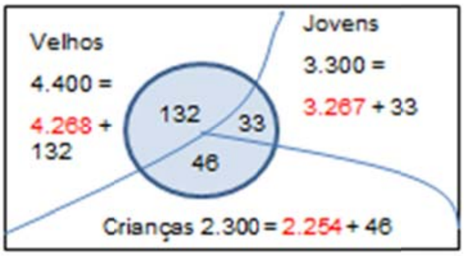
Entendendo o Teorema de Bayes

Vejamos o assunto por meio de um exemplo:

A cidade de Diabetelândia tem 10.000 habitantes, sendo 4.400 velhos, 3.300 jovens, 2.300 crianças. Os dados mostram que nesta cidade está havendo uma desertificação urbana em virtude da falta de empregos.

Foi realizada uma pesquisa pela Secretaria de Saúde onde se constatou uma grande incidência de diabetes na população. A pesquisa revelou que 132 velhos são portadores da doença, 33 jovens são portadores e 46 crianças são portadoras. Assim:

Diabetelândia		
	Habitantes	Portadores de Diabete
Velhos	4.400	132
Jovens	3.300	33
Crianças	2.300	46
TOTAL	10.000	211



Os valores em vermelhos são os habitantes das categorias e que não são portadores de diabetes.

Muitas vezes esta tabela vem dada em porcentagens. Assim:

Diabetelândia		
	Habitantes	Portadores de Diabete na população
Velhos	44,00%	1,32%
Jovens	33,00%	0,33%
Crianças	23,00%	0,46%
TOTAL	100,00%	2,11%

Observe que os velhos **E** portadores de diabetes representam 1,32% da população da cidade Diabetelândia (132 ÷ 10.000 = 0,0132 ou 1,32%). Portanto, nesta cidade a probabilidade (ou chance) de chegar ao hospital um velho **E** portador de diabetes é 1,32%! É pouco provável aos médicos do hospital atender um velho **E** portador de diabetes entre os pacientes em geral que chegam ao hospital.

A tabela mostra as chances para as outras categorias.

Apenas 2,11% dos habitantes tem diabetes.

Porém, quando chega ao hospital um velho (um pedaço da população), a chance dele ser diabético é encontrada dividindo-se 132 por 4.400 (132 ÷ 4.400 = 0,03 ou 3%). A chance aumentou porque restringimos o atendimento apenas aos velhos.

Matematicamente escrevemos esta probabilidade CONDICIONAL (ser portador de diabetes **E** velho) da seguinte maneira: P(X|V). Aqui X significa ser diabético e V significa ser velho. Lemos este símbolo assim: a probabilidade P de ser diabético **dado** que seja velho.

Note que além de ser velho, tem que ser diabético (coitado!). Isto mostra que colocamos uma condição entre os velhos. Lembre-se que nesta cidade existem muitos velhos que *não* são diabéticos. Estes, como mostra a figura acima totalizam 4.400 – 132 = 4.268 velhos sem a doença.

Como vimos esta probabilidade foi calculada tomando todos os velhos E diabéticos (132) que estão no conjunto intersecção [Velho ∩ Diabéticos] e dividido pela quantidade de velhos da cidade (4.400). Então,

$$P(X|V) = \frac{n(X \cap V)}{n(V)} = \frac{132}{4.400} = 0{,}03 \; ou \; 3\%.$$

Lembrem-se que, os eventos ser velho e ser diabético são INDEPENDENTES. Isto significa que ser velho não precisa ser diabético e ser diabético não precisa ser velho ou, em outras palavras, nem todo velho é diabético e nem todo diabético é velho.

A expressão acima poderia ser calculada de outra maneira:

$$P(X|V) = \frac{n(X \cap V)}{n(V)} = \frac{\frac{n(X \cap V)}{N}}{\frac{n(V)}{N}} = \frac{P(X \cap V)}{P(V)} = \frac{1{,}32\%}{44\%} = 0{,}03 \text{ ou } 3\%.$$

Repetindo isto para as outras categorias podemos montar a seguinte tabela:

Diabetelândia		
	Habitantes	Portadores de Diabetes na categoria $P(X\|A_i)$
Velhos	44,00%	3,00%
Jovens	33,00%	1,00%
Crianças	23,00%	2,00%
TOTAL	100,00%	2,11%

Cabe aqui a pergunta: **Qual a probabilidade de um velho ser diabético?**

Resposta: 3%. Esta é a chance de aparecer em qualquer lugar da cidade um velho que é diabético. Não precisa ter medo deles!!!! A chance é pequena e a doença não infecto-contagiosa!

Esta pergunta requer que encontremos P(X|V), certo?

Outra pergunta: **Qual a probabilidade de um diabético ser velho?**

É a mesma coisa que antes? 3%?

Posto de outra forma, chega ao hospital um diabético, qual a chance dele ser velho?

Agora queremos a probabilidade condicional P(V|X), ou seja, o contrário de antes. Imporemos agora uma CONDIÇÃO entre todos os diabéticos – eles devem ser velhos. Observando a figura anterior, temos que dividir nº de velhos diabéticos pelo total de diabéticos

$$P(V|X) = \frac{n(X \cap V)}{n(X)} = \frac{\frac{n(X \cap V)}{N}}{\frac{n(X)}{N}} = \frac{P(X \cap V)}{P(X)} = \frac{1{,}32\%}{2{,}11\%} = 0{,}6256 \text{ ou } 62{,}56\%.$$

Se atentarmos para os diabéticos da cidade, a chance dele ser velho é 62,56%. É, nesta cidade, entre os diabéticos temos muitos velhos (ver figura acima).

Diabetelândia			
	Habitantes	Portadores de Diabetes na categoria $P(X\|A_i)$	A Categoria dos Portadores de Diabetes $P(Ai\|X)$
Velhos	44,00%	3,00%	62,56%
Jovens	33,00%	1,00%	15,64%
Crianças	23,00%	2,00%	21,80%
TOTAL	100,00%	2,11%	100,00%

Médicos, ao chegar um paciente diabético, é muito grande (62,56%) dele ser velho.

Fica bem claro que:

I. $P(X \mid A_i) \neq P(A_i \mid X)$
II. $P(X \cap A_i) = P(A_i \cap X)$
III. Pela definição de probabilidade condicional

$$P(X|A_i) = \frac{P(X \cap A_i)}{P(A_i)}$$

Temos que: $P(X \cap A_i) = P(X|A_i)\, P(A_i)$.
Dessa forma podemos fazer:

$$P(A_i|X) = \frac{P(X \cap A_i)}{P(X)} = \frac{P(X \cap A_i)}{P(X \cap A_1) + P(X \cap A_2) + \cdots + P(X \cap A_n)}$$

Ou, usando a definição de probabilidade condicional:

$$P(A_i|X) = \frac{P(X \cap A_i)}{P(X)} = \frac{\mathrm{P(X|A_i)\,P(A_i)}}{\mathrm{P(X|A_1)\,P(A_1) + P(X|A_2)\,P(A_2) + \cdots + P(X|A_n)\,P(A_n)}} = \frac{P(X \cap A_i)}{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{P(X|A_1)\,P(A_1)}}$$

Este resultado recebe o nome de **Teorema de Bayes**.

USANDO TABELA

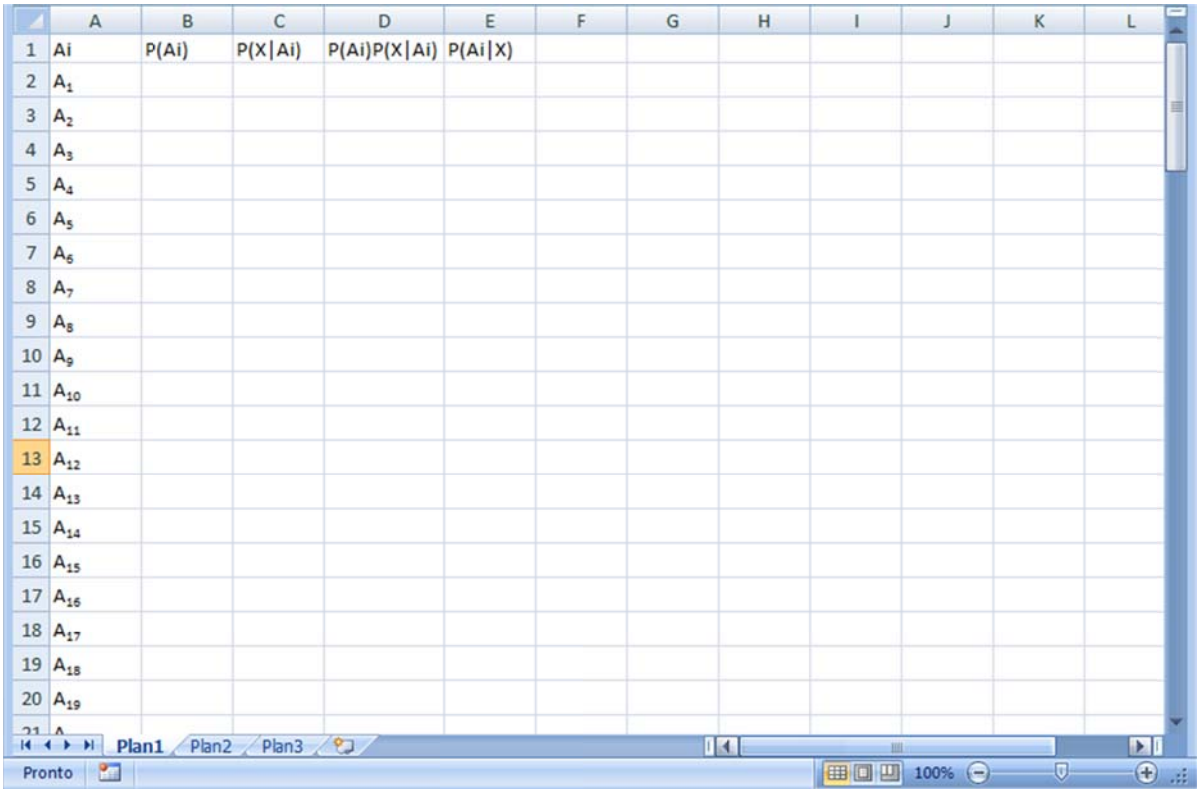
De acordo com o desenvolvimento anterior poderemos realizar os cálculos por meio de uma tabela. Assim

Espaço Amostral				
Partição do Espaço Amostral A_i	Participação da Partição no Espaço Amostral $P(A_i)$	$P(X\|A_i)$ conhecido	$P(X \cap Ai) = P(X\|Ai).P(A_i)$	$P(A_i\|X)$
A_1	$P(A_1)$	$P(X\|A_1)$	$P(A_1) * P(X\|A_1)$	$[P(A_1) * P(X\|A_1)]/SOMA$
A_2	$P(A_2)$	$P(X\|A_2)$	$P(A_2) * P(X\|A2)$	$[P(A_2) * P(X\|A_2)]/SOMA$
A_3	$P(A_3)$	$P(X\|A_3)$	$P(A_3) * P(X\|A_3)$	$[P(A_3) * P(X\|A_3)]/SOMA$
...				
...				
...				
A_n	$P(A_n)$	$P(X\|A_n)$	$P(A_n) * P(X\|A_n)$	$[P(A_n) * P(X\|A_n)]/SOMA$
TOTAL	100,00%		SOMA	

USANDO O EXCEL

Já que podemos fazer uma tabela, podemos também realizar tudo no Excel.

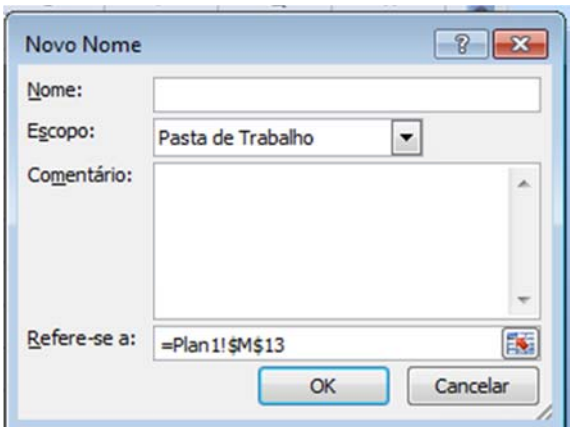
Abra uma pasta de trabalho e introduza os títulos conforme a figura abaixo:



Faça 100 partições do espaço amostral (acho que fica o suficiente, não?). Com isso precisamos selecionar o intervalo A2:A101 e colocar os títulos A_i.

Para criarmos uma planilha com recursos avançados do Excel, vamos definir um nome ao intervalo de células onde existirem valores, deixando o resto das células em branco, evitando assim colunas com ZEROS desnecessários, pois não faremos cálculos naquelas linhas.

Para se nomear um intervalo de células no Excel, procedemos selecionando o intervalo de células G2:G101, depois na guia Formatar, no grupo Nomes Definidos, escolhemos Definir Nome. Aparecerá a janela:



Na caixa Nome: **ProbInterseção**

Na caixa Refere-se a: introduza a seguinte fórmula: =DESLOC(PLAN1!$G$2;0;0;CONT.NÚM(Plan1!$G$2:$G$101;1)). Dê OK.

Selecione o intervalo G2:G101 e escolha a cor branca para a fonte. Isso para não poluir a nossa planilha com dados e precisamos disto para evitarmos as referencias circulares do Excel ao introduzir a função SOMA na célula D101..

Vamos entender o que fizemos. Primeiro, por que usar a função DESLOC?

Esta função embutida do Excel retorna uma referência a um intervalo que possui um número específico de linhas e colunas com base em um referência especificada (no nosso caso se houver número e a existência ou não de números é identificado com a função CONT.NÚM que falaremos abaixo). A sintaxe da função DESLOC é: =DESLOC(**ref**;**Lins**;**cols**;**altura**;largura). Os argumentos em negrito são obrigatórios e os outros são opcionais.

A função CONT.NÚM calcula o nº de células em um intervalo que contém números. Sua sintaxe é: CONT.NÚM(**valor1**;valor2;....). Novamente os argumentos em negrito são obrigatórios. Aqui usamos valor2 = 1, para não retornar zero quando não encontrar número e com isso causando um erro de **altura** na função DESLOC.

Dessa forma a função DESLOC nomeará o intervalo na coluna G que tiver números e com isso não serão introduzidos zeros quando a célula estiver em branco na coluna D que apresenta a fórmula SOMA(DADOS) na célula D101.

Voltemos à célula D2 e introduzimos a fórmula: =SE(B2="";"";B2*C2) e na célula E2, introduzimos: =SE(D2="";"";D2/$D$102).

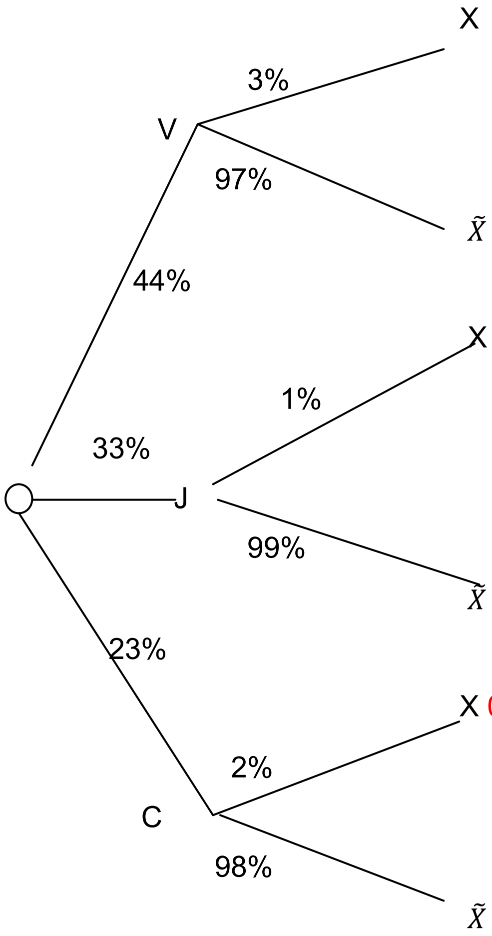
A planilha ficou pronta. Agora é só salvar e guardar com carinho para quando precisar.

DIAGRAMA DE ÁRVORE

O diagrama de árvore ajuda a montar o problema e fazer as contas.

Voltemos à cidade Diabetelândia e vemos que lembremos que a população foi dividida em 3 categorias de habitantes (velhos, jovens e crianças). Algumas das pessoas de cada categoria eram portadoras de diabetes.

Então, estabelecendo que X é portador e $\tilde{X}$ não é portador, temos



0,44 x 0,03 = 0,0132 ou 1,32% $\quad \frac{0{,}0132}{0{,}0211} = 0{,}6256\ ou\ 62{,}56\%$

0,33 x 0,01 = 0,0033 ou 0,33% $\quad \frac{0{,}0033}{0{,}0211} = 0{,}1564\ ou\ 15{,}64\%$

0,23 x 0,02 = 0,0046 ou 0,46% $\quad \frac{0{,}0046}{0{,}0211} = 0{,}2180\ ou\ 21{,}80\%$

0,0211 ou 2,11%

EXERCÍCIOS RESOLVIDOS

1. Uma urna A contém 3 fichas vermelhas e 2 azuis, e uma urna B contém 2 vermelhas e 8 azuis. Joga-se uma moeda “honesta”. Se a moeda der cara, extrai-se uma ficha da urna A; se der coroa, extrai-se uma ficha da urna B. Uma ficha vermelha é extraída.
   Qual a probabilidade de ter saído cara no lançamento da moeda?

   Solução

A	B
3 V 2 A	2 V 8 A

Queremos encontrar a probabilidade de sair cara dado a bola ser vermelha, isto é P(Ca|V)

P(A) = P(Ca) = (1/2) P(V|A) = (3/5) = 60% = P(V|Ca)

P(B) = P(Co) = (1/2) P(V|B) = (2/10)= 40% = P(V|Co)

Pelo Teorema de Bayes, temos

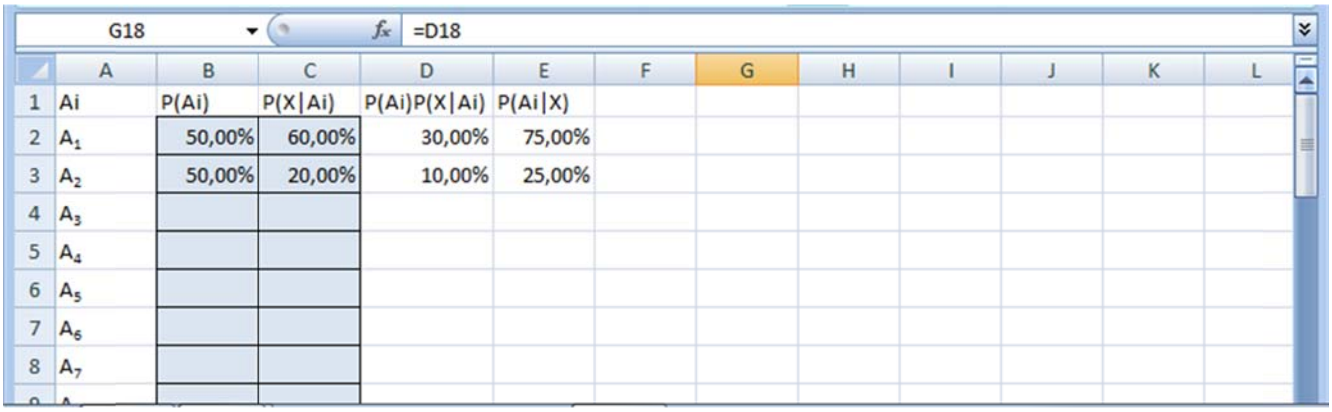
$$P(A_i|X) = \frac{P(X \cap A_i)}{P(X)} = \frac{\mathrm{P(X|A_i)\,P(A_i)}}{\mathrm{P(X|A_1)\,P(A_1) + P(X|A_2)\,P(A_2) + \cdots + P(X|A_n)\,P(A_n)}} = \frac{P(X \cap A_i)}{\sum_{i=1}^{n} \mathrm{P(X|A_1)\,P(A_1)}}$$

$$P(Ca|V) = \frac{P(V \cap Ca)}{P(V)} = \frac{P(V \cap A)}{P(V)} \frac{\mathrm{P(V|A)\ P(A)}}{\mathrm{P(V|A)\ P(A) + P(V|B)\ P(B)}} = \frac{\frac{3}{5} x \frac{1}{2}}{\frac{3}{5} x \frac{1}{2} + \frac{2}{10} x \frac{1}{2}} = \frac{\frac{3}{10}}{\frac{3}{10} + \frac{2}{20}} = \frac{\frac{3}{10}}{\frac{6+2}{20}}$$

$$= \frac{3}{10} x \frac{20}{8} = \frac{60}{80} = \frac{3}{4} = 0{,}75 \ ou \ 75\%$$

Temos, portanto, 75% de probabilidade de que a bola vermelha seja extraída da urna A por ter obtido cara no lançamento da moeda.

Na planilha Excel, preencha apenas a área azul:



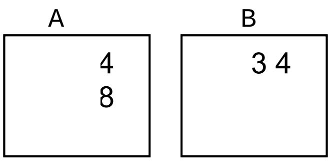
G18 | =D18

	A	B	C	D	E	F	G	H	I	J	K	L
1	Ai	P(Ai)	P(X|Ai)	P(Ai)P(X|Ai)	P(Ai|X)							
2	A_1	50,00%	60,00%	30,00%	75,00%							
3	A_2	50,00%	20,00%	10,00%	25,00%							
4	A_3											
5	A_4											
6	A_5											
7	A_6											
8	A_7											

Faça o exercício usando a árvore.

2. A caixa A tem 9 cartas numeradas de 1 a 9. A caixa B tem 5 cartas numeradas de 1 a 5. Uma caixa é escolhida ao acaso e uma carta é retirada. Se o número é par , qual a probabilidade de que a carta sorteada tenha vindo de A?

Solução



Queremos encontrar a probabilidade da carta vir da urna A dado que o seu número é par, isto é P(A|par).

P(A) = (1/2)   P(par|A) = (4/9)= 44,44%

P(B) = (1/2)   P(par|B) = (2/5)= 40%

Pelo Teorema de Bayes, temos

$$P(A_i|X) = \frac{P(X \cap A_i)}{P(X)} = \frac{\mathrm{P(X|A_i)\ P(A_i)}}{\mathrm{P(X|A_1)\ P(A_1) + P(X|A_2)\ P(A_2) + \cdots + P(X|A_n)\ P(A_n)}} = \frac{P(X \cap A_i)}{\sum_{i=1}^{n} \mathrm{P(X|A_1)\ P(A_1)}}$$

$$P(A|par) = \frac{P(par \cap A)}{P(V)} = \frac{P(A \cap par)}{P(V)} \frac{\mathrm{P(par|A)\ P(A)}}{\mathrm{P(par|A)\ P(A) + P(par|B)\ P(B)}} = \frac{\frac{4}{9} x \frac{1}{2}}{\frac{4}{9} x \frac{1}{2} + \frac{2}{5} x \frac{1}{2}} = \frac{\frac{4}{18}}{\frac{4}{18} + \frac{2}{10}}$$

$$= \frac{\frac{4}{18}}{\frac{40+36}{180}} = \frac{4}{18} x \frac{180}{76} = \frac{40}{76} = \frac{10}{19} = 0{,}5263 \ ou \ 52{,}63\%$$

Temos, portanto, 52,63% de probabilidade de que a carta de número par seja extraída da urna A.

Na planilha Excel, preencha apenas a área azul:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Ai	P(Ai)	P(X|Ai)	P(Ai)P(X|Ai)	P(Ai|X)							
2	A_1	50,00%	44,44%	22,22%	52,63%							
3	A_2	50,00%	40,00%	20,00%	47,37%							
4	A_3											
5	A_4											
6	A_5											
7	A_6											
8	A_7											

3. Num colégio, 4% dos homens e 1% das mulheres têm mais de 1,75 m de altura. 60% dos estudantes são mulheres. Um estudante é escolhido ao acaso e tem mais de 1,75 m. Qual a probabilidade de que seja homem?

Solução

A = evento ter mais de 1,75 m de altura.

P(M) = 60%   P(A|M) = 1%

P(H) = 40%   P(A|H) = 4%   Queremos P(H|A)

Pelo Teorema de Bayes, temos

$$P(A_i|X) = \frac{P(X \cap A_i)}{P(X)} = \frac{\mathrm{P(X|A_i)\ P(A_i)}}{\mathrm{P(X|A_1)\ P(A_1) + P(X|A_2)\ P(A_2) + \cdots + P(X|A_n)\ P(A_n)}} = \frac{P(X \cap A_i)}{\sum_{i=1}^{n} \mathrm{P(X|A_1)\ P(A_1)}}$$

$$P(H|A) = \frac{P(H \cap A)}{P(A)} = \frac{P(A \cap H)}{P(A)} = \frac{\mathrm{P(A|H)\ P(H)}}{\mathrm{P(A|H)\ P(H) + P(A|M)\ P(M)}} = \frac{0{,}04x0{,}40}{0{,}04x0{,}40 + 0{,}01x0{,}60} = \frac{0{,}016}{0{,}016 + 0{,}006}$$
$$= \frac{0{,}016}{0{,}022} = \frac{8}{11} = 0{,}7273\ ou\ 72{,}73\%$$

Temos, portanto, 72,73% de probabilidade de que o estudante com mais de 1,75 m de altura escolhido ao acaso seja homem. Embora o colégio tenha mais mulher do que homem, há mais homens com altura superior a 1,75 m do que mulheres.

Na planilha Excel, preencha apenas a área azul:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Ai	P(Ai)	P(X|Ai)	P(Ai)P(X|Ai)	P(Ai|X)							
2	A_1	60,00%	1,00%	0,60%	27,27%							
3	A_2	40,00%	4,00%	1,60%	72,73%							
4	A_3											
5	A_4											
6	A_5											
7	A_6											
8	A_7											
9	A_8											
10	A_9											

4. Uma caixa tem 3 moedas: uma não viciada, outra com 2 caras e uma terceira viciada, de modo que a probabilidade de ocorrer cara nesta moeda é 1/5. Uma moeda é selecionada ao acaso na caixa. Saiu cara. Qual a probabilidade de que a 3ª moeda tenha sido a selecionada?

Solução

A = primeira moeda,

B = segunda moeda,

C = terceira moeda

P(A) = 1/3 P(Ca|A) = 50% ou 1/2

P(B) = 1/3 P(Ca|B) = 100% ou 1 Queremos P(C|Ca)

P(C) = 1/3 P(Ca|C) = 20% ou 1/5

Pelo Teorema de Bayes, temos

$$P(A_i|X) = \frac{P(X \cap A_i)}{P(X)} = \frac{\mathrm{P(X|A_i)\,P(A_i)}}{\mathrm{P(X|A_1)\,P(A_1) + P(X|A_2)\,P(A_2) + \cdots + P(X|A_n)\,P(A_n)}} = \frac{P(X \cap A_i)}{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{P(X|A_1)\,P(A_1)}}$$

$$P(C|Ca) = \frac{P(C \cap Ca)}{P(Ca)} = \frac{P(Ca \cap C)}{P(Ca)} \frac{\mathrm{P(Ca|C)\,P(C)}}{\mathrm{P(Ca|A)P(A) + P(Ca|B)P(B) + P(Ca|C)\,P(C)}}$$

$$= \frac{0{,}20x(1/3)}{0{,}50x\left(\frac{1}{3}\right) + 0{,}20x\left(\frac{1}{3}\right) + 1x(\frac{1}{3})} = \frac{0{,}20}{0{,}50 + 1 + 0{,}20} = \frac{0{,}20}{1{,}70} = \frac{2}{17} = 0{,}1176\ ou\ 11{,}76\%$$

Temos, portanto, 11,76% de probabilidade de que a face cara seja apresentada pela 3ª moeda quando esta for escolhida ao acaso.

Na planilha Excel, preencha apenas a área azul:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Ai	P(Ai)	P(X|Ai)	P(Ai)P(X|Ai)	P(Ai|X)							
2	A_1	33,33%	50,00%	16,67%	29,41%							
3	A_2	33,33%	100,00%	33,33%	58,82%							
4	A_3	33,33%	20,00%	6,67%	11,76%							
5	A_4											
6	A_5											
7	A_6											
8	A_7											
9	A_8											

5. A probabilidade de um indivíduo de classe A comprar um carro de 3/4, da B é de 1/5 e da C é de 1/20. As probabilidades se os indivíduos comprarem um carro da marca x são 1/10, 3/5 e 3/10, dado que sejam de A, B e C, respectivamente. Certa loja vendeu um carro da marca x. Qual a probabilidade de que i indivíduo que comprou seja da classe B?

Solução

A = classe A,

B = classe B,

C = classe C

P(A) = 3/4 P(car|A) = 1/10

P(B) = 1/5 P(car|B) = 3/5 Queremos P(B|car)

P(C) = 1/20 P(car|C) = 3/10

Pelo Teorema de Bayes, temos

$$P(A_i|X) = \frac{P(X \cap A_i)}{P(X)} = \frac{\mathrm{P(X|A_i)\,P(A_i)}}{\mathrm{P(X|A_1)\,P(A_1) + P(X|A_2)\,P(A_2) + \cdots + P(X|A_n)\,P(A_n)}} = \frac{P(X \cap A_i)}{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{P(X|A_1)\,P(A_1)}}$$

$$P(B|car) = \frac{P(B \cap car)}{P(car)} = \frac{P(car \cap B)}{P(Ca)} \frac{\mathrm{P(car|B)\,P(B)}}{\mathrm{P(car|A)P(A) + P(car|B)P(B) + P(car|C)\,P(C)}}$$

$$= \frac{\frac{3}{5}x\frac{1}{5}}{\frac{1}{10}x\frac{3}{4} + \frac{3}{5}x\frac{1}{5} + \frac{3}{10}x\frac{1}{20}} = \frac{\frac{3}{25}}{\frac{3}{40} + \frac{3}{25} + \frac{3}{200}} = \frac{\frac{3}{25}}{\frac{15 + 24 + 3}{200}} = \frac{3}{25}x\frac{200}{42} = \frac{4}{7}$$

$$= 0{,}5714\ ou\ 57{,}14\%$$

Temos, portanto, 57,14% de probabilidade de que o carro foi comprado por um indivíduo da classe B.

Na planilha Excel, preencha apenas a área azul:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Ai	P(Ai)	P(X|Ai)	P(Ai)P(X|Ai)	P(Ai|X)							
2	A_1	75,00%	10,00%	7,50%	35,71%							
3	A_2	20,00%	60,00%	12,00%	57,14%							
4	A_3	5,00%	30,00%	1,50%	7,14%							
5	A_4											
6	A_5											
7	A_6											
8	A_7											
9	A_8											
10	A_9											
11	A_{10}											
12	A_{11}											
13	A_{12}											

6. Um certo programa pode ser usado com uma entre duas sub-rotinas A e B, dependendo do problema. A experiência tem mostrado que a sub-rotina A é usada 40% das vezes e B é usada 60% das vezes. Se A é usada, existe 75% de chance de que o programa chegue a um resultado dentro do limite de tempo. Se B é usada, a chance é de 50%. Se o programa foi realizado dentro do limite de tempo. Se B é usada, a chance é de 50%. Se o programa foi realizado dentro do limite de tempo, qual a probabilidade de que a sub-rotina A tenha sido a escolhida?

Solução

P(A) = 40%   P(resultado|A) = 75%

P(B) = 60%   P(resultado|B) = 50%   Queremos P(A|resultado).

Pelo Teorema de Bayes, temos

$$P(A_i|X) = \frac{P(X\cap A_i)}{P(X)} = \frac{\mathrm{P(X|A_i)\,P(A_i)}}{\mathrm{P(X|A_1)\,P(A_1) + P(X|A_2)\,P(A_2) + \cdots + P(X|A_n)\,P(A_n)}} = \frac{P(X\cap A_i)}{\sum_{i=1}^{n}\mathrm{P(X|A_1)\,P(A_1)}}$$

$$P(A|resultado) = \frac{P(A\cap resultado)}{P(resultado)} = \frac{P(resultado\cap A)}{P(resultado)} = \frac{\mathrm{P(resultado|A)\,P(A)}}{\mathrm{P(resultado|A)P(A) + P(resultado|B)P(B)}}$$
$$= \frac{0{,}75x0{,}40}{0{,}75x0{,}40 + 0{,}50x0{,}60} = \frac{0{,}30}{0{,}30+0{,}30} = \frac{0{,}30}{0{,}60} = \frac{1}{2} = 0{,}5000\ ou\ 50{,}00\%$$

Temos, portanto, 50,00% de probabilidade de que o resultado foi atingido dentro do limite usando a sub-rotina A.

Na planilha Excel, preencha apenas a área azul:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Ai	P(Ai)	P(X|Ai)	P(Ai)P(X|Ai)	P(Ai|X)							
2	A_1	40,00%	75,00%	30,00%	50,00%							
3	A_2	60,00%	50,00%	30,00%	50,00%							
4	A_3											
5	A_4											
6	A_5											
7	A_6											

7. A urna X contém 2 bolas azuis, 2 brancas e 1 cinza, e a urna Y contém 2 bolas azuis, 1 branca e 1 cinza. Retira-se uma bola de cada urna. Calcule a probabilidade de saírem 2 bolas brancas sabendo-se que são bolas de mesma cor?

Solução

X	Y
2 A	2 A
2 B	1 B
1 C	1 C

Após retirar uma bola de cada urna queremos saber $P(B \cap B|\text{mesma cor})$.

A cor da bola que sai da segunda urna não é influenciada pela cor da bola que saiu da primeira urna, isto é os eventos são independentes.

$P(\text{mesma cor}) = P(B \cap B) + P(A \cap A) + P(C \cap C) = P(B).P(B) + P(A).P(A) + P(C).P(C) = (2/5)(1/4) + (2/5)(2/4) + (1/5)(1/4) = (2/20)+(4/20)+(1/20) = (7/20) = 0,35$ ou 35%

Agora

$$P(B \cap B|\text{mesma cor}) = \frac{P(B \cap B)}{P(\text{mesma cor})} = \frac{\frac{2}{20}}{\frac{7}{20}} = \frac{2}{7}$$

8. Uma companhia de transporte urbano tem três linhas numa cidade, de forma que 45% dos ônibus cobrem o serviço da linha 1, 25% cobre a linha 2 e 30% cobre o serviço da linha 3. Sabe-se que a probabilidade de que, diariamente, um ônibus quebre é de 2%,3% e 1%, respectivamente, para cada linha.

a. Calcular a probabilidade de que, em um dia, um ônibus quebre.

b. Calcular a probabilidade de que, em um dia, um ônibus não quebre.

c. De que linha de transporte é mais provável que um ônibus quebre?

Solução

A_1 = Cobre o serviço da linha 1

A_2 = Cobre o serviço da linha 2

A_3 = Cobre o serviço da linha 3

B_1 = Quebra

B_2 = Não quebra

$P(A_1) = 45\% = 0,45$ $\quad P(A_2) = 25\% = 0,25$ $\quad P(A_3) = 30\% = 0,3$

$P(B_1|A_1) = 2\% = 0,02$ $\quad P(B_1|A_2) = 3\% = 0,03$ $\quad P(B_1|A_3) = 1\% = 0,01$

$P(B_2|A_1) = 1 - P(B_1|A_1) = 1 - 0,02 = 0,98$

$P(B_2|A_2) = 1 - P(B_1|A_2) = 1 - 0,03 = 0,97$

$P(B_2|A_3) = 1 - P(B_1|A_3) = 1 - 0,01 = 0,99$

a. A probabilidade de que, em um dia, um ônibus quebre, representada por $P(B_1)$, é encontrada pela probabilidade total:

$P(B_1) = P(A_1 \cap B_1) + P(A_2 \cap B_1) + P(A_3 \cap B_1)$
$P(B_1) = P(A_1).P(B_1|A_1) + P(A_2).P(B_1|A_2) + P(A_3).P(B_1|A_3)$
$P(B_1) = 0{,}45.0{,}02 + 0{,}25.0{,}03 + 0{,}30.0{,}01 = 0{,}0195$ ou 1,95%

b. A probabilidade de que, em um dia, um ônibus não quebre, representada por $P(B_2)$, é encontrada pela probabilidade total:
$P(B_2) = P(A_1 \cap B_2) + P(A_2 \cap B_2) + P(A_3 \cap B_2)$
$P(B_2) = P(A_1).P(B_2|A_1) + P(A_2).P(B_2|A_2) + P(A_3).P(B_2|A_3)$
$P(B_2) = 0{,}45.0{,}98 + 0{,}25.0{,}97 + 0{,}30.0{,}99 = 0{,}9805$ ou 98,05%

Ou também, sabendo-se que $P(B_2) = 1 - P(B_1) = 1 - 0{,}0195 = 0{,}9805$ ou 98,05%.

c. Para encontrarmos a linha de transporte que seja a mais provável de um ônibus quebrar, aplicamos o teorema de Bayes a cada uma das linhas. Assim

$$P(A_i|X) = \frac{P(X \cap A_i)}{P(X)} = \frac{\mathrm{P(X|A_i)\,P(A_i)}}{\mathrm{P(X|A_1)\,P(A_1) + P(X|A_2)\,P(A_2) + \cdots + P(X|A_n)\,P(A_n)}} = \frac{P(X \cap A_i)}{\sum_{i=1}^{n} \mathrm{P(X|A_1)\,P(A_1)}}$$

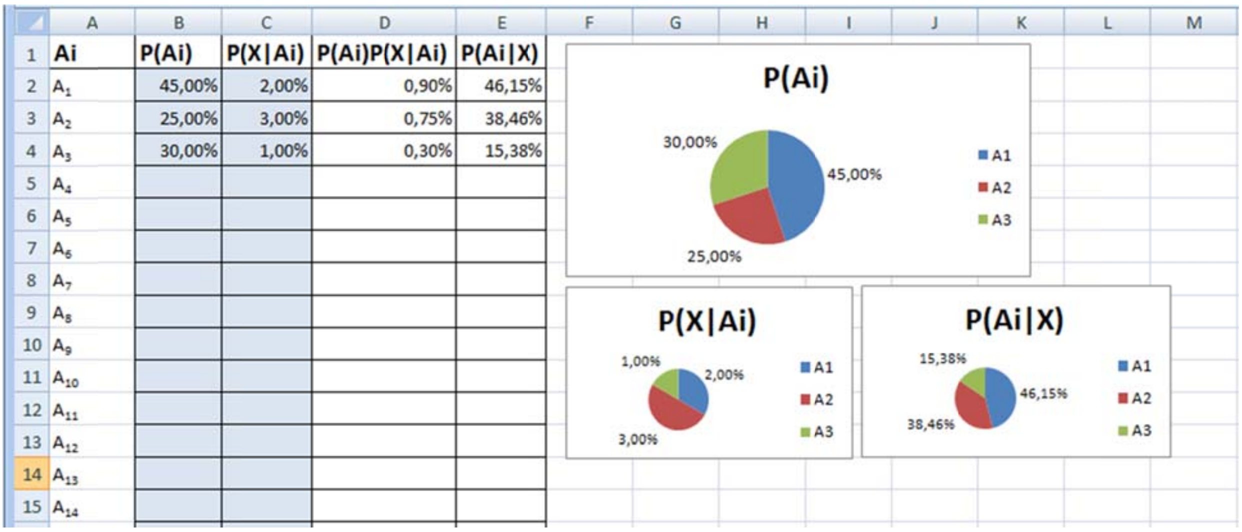
$$P(A_1\,|B_1) = \frac{P(A_1 \cap B_1)}{P(B_1)} = \frac{P(B_1 \cap A_1)}{P(B_1)} = \frac{P(B_1|A_1)\,P(A_1)}{P(B_1|A_1)\,P(A_1) + P(B_1|A_2)\,P(A_2) + P(B_1|A_3)\,P(A_3)} = \frac{0{,}02x0{,}45}{0{,}0195} =$$
$0{,}4615$ *ou* $46{,}15\%$

$$P(A_2\,|B_1) = \frac{P(A_2 \cap B_1)}{P(B_1)} = \frac{P(B_1 \cap A_2)}{P(B_1)} = \frac{P(B_1|A_2)\,P(A_2)}{P(B_1)} = \frac{0{,}03x0{,}25}{0{,}0195} = 0{,}3846 \text{ } ou \text{ } 38{,}46\%$$

$$P(A_3\,|B_1) = \frac{P(A_3 \cap B_1)}{P(B_1)} = \frac{P(B_1 \cap A_3)}{P(B_1)} = \frac{P(B_1|A_3)\,P(A_3)}{P(B_1)} = \frac{0{,}01x0{,}30}{0{,}0195} = 0{,}1538 \text{ } ou \text{ } 15{,}38\%$$

Temos, portanto, a linha A_1 apresentando a maior chance (maior probabilidade) de um ônibus quebrar.

Na planilha Excel, preencha apenas a área azul:



	A	B	C	D	E
1	Ai	P(Ai)	P(X|Ai)	P(Ai)P(X|Ai)	P(Ai|X)
2	A_1	45,00%	2,00%	0,90%	46,15%
3	A_2	25,00%	3,00%	0,75%	38,46%
4	A_3	30,00%	1,00%	0,30%	15,38%
5	A_4				
6	A_5				
7	A_6				
8	A_7				
9	A_8				
10	A_9				
11	A_{10}				
12	A_{11}				
13	A_{12}				
14	A_{13}				
15	A_{14}				

9. Uma empresa dedicada à comercialização de televisores está considerando comercializar um novo televisor. No passado 90% dos televisores que comercializou tiveram êxito e 10% não. Sabe-se que a probabilidade que .recebera um relatório favorável de pesquisa foi de 85% e 35%, respectivamente.

a. Calcule a probabilidade de que os televisores exitosos recebem um relatório desfavorável de pesquisa

b. Calcule a probabilidade de que os televisores não exitosos recebam um relatório desfavorável de pesquisa.

c. Calcule a probabilidade de que um televisor receba um relatório favorável de pesquisa. $P(B_1) = 0{,}8$

d. Calcule a probabilidade de que um televisor receba um relatório desfavorável de pesquisa. $P(B_2) = 0{,}2$

e. Qual é a probabilidade de que o equipamento do televisor tenha êxito no mercado? $P(A_1|B_1) = 0{,}9563$

Solução

A_1 = Televisores exitosos $P(A_1) = 90\%$

A_2 = Televisores não exitosos $P(A_2) = 10\%$

B_1 = Relatório favorável de pesquisa

B_2 = Relatório desfavorável de pesquisa

$P(B_1|A_1) = 85\%$

$P(B_1|A_2) = 35\%$

a. $P(B_2|A_1) = 15\%$
b. $P(B_2|A_2) = 65\%$
c. $P(B_1) = P(A_1).P(B_1|A_1) + P(A_2).P(B_1|A_2)= 0{,}90x0{,}85 + 0{,}10x0{,}35 = 0{,}8$
d. $P(B_2) = P(A_1).P(B_2|A_1) + P(A_2).P(B_2|A_2)= 0{,}90x0{,}15 + 0{,}10x0{,}65 = 0{,}2$
e. Aqui queremos $P(A_1|B_1)$ e será encontrado pelo teorema de Bayes

$$P(A_1\,|B_1) = \frac{P(A_1 \cap B_1)}{P(B_1)} = \frac{P(B_1 \cap A_1)}{P(B_1)} = \frac{\mathrm{P}(B_1|A_1)\,\mathrm{P}(A_1)}{\mathrm{P}(B_1|A_1)\,\mathrm{P}(A_1) + \mathrm{P}(B_1|A_2)\,\mathrm{P}(A_2)}$$
$$= \frac{0{,}85x0{,}90}{0{,}85x0{,}90 + 0{,}35x0{,}10} = \frac{0{,}7650}{0{,}7650 + 0{,}0350} = \frac{0{,}7650}{0{,}8} = 0{,}9563 \text{ ou } 95{,}63\%$$

No Excel:

	A	B	C	D	E
1	Ai	P(Ai)	P(X|Ai)	P(Ai)P(X|Ai)	P(Ai|X)
2	A_1	90,00%	85,00%	76,50%	95,63%
3	A_2	10,00%	35,00%	3,50%	4,38%
4	A_3				
5	A_4				
6	A_5				

10. A probabilidade de que uma pessoa tenha uma determinada doença é 0,02.

Existem provas de diagnóstico médico disponíveis para determinar se uma pessoa tem realmente a doença. Se a doença estiver realmente presente, a probabilidade de que a prova de diagnóstico indique a presença da doença é de 0,95.

a. Qual é a probabilidade de se ter a doença, se a prova de diagnóstico indica a presença da mesma? Calcular à mão e no Excel.
b. Qual é a probabilidade de não ter a doença, se a prova de diagnóstico não indica a presença da mesma? Calcular à mão e no Excel.

Solução

$P(A_1)$ = Probabilidade de que uma pessoa tenha determinada doença = 0,02 ou 2%.

$P(A_2)$ = Probabilidade de que uma pessoa não tenha determinada doença = 0,98 = 98%.

$P(B_1|A_1)$ = Probabilidade de que o diagnóstico indique a presença da doença dado que a pessoa esteja doente = 0,95 ou 95%.

$P(B_1|A_2)$ = Probabilidade de que o diagnóstico indique a presença da doença dado que a pessoa não esteja doente = 0,05 ou 5%.

$P(B_2|A_1) = 1 - P(B_1|A_1) = 1 - 0{,}95 = 0{,}05$ ou 5%

$P(B_2|A_2) = 1 - P(B_1|A_2) = 1 - 0{,}05 = 0{,}95$ ou 95%

a. O que queremos é a probabilidade de uma pessoa ter a doença dado que a prova de diagnóstico indica a presença da mesma = $P(A_1|B_1)$.
Para isso usamos o teorema de Bayes:

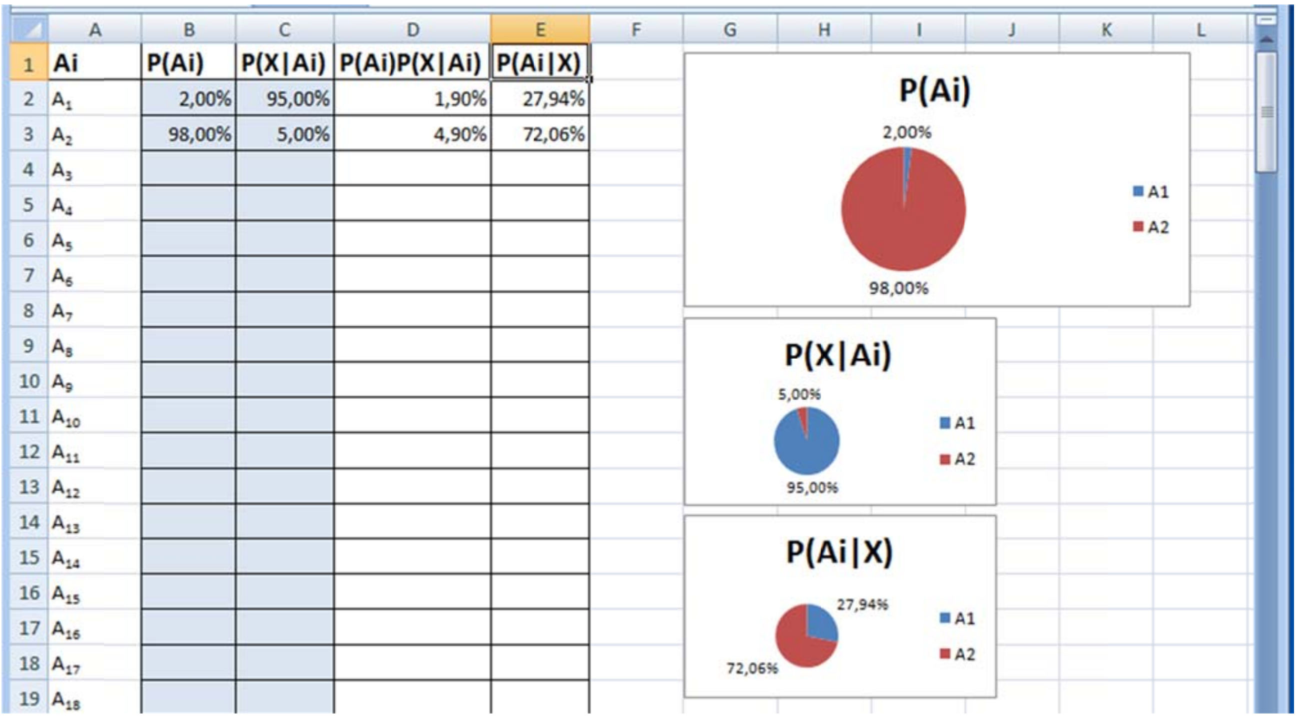
$$P(A_i|X) = \frac{P(X \cap A_i)}{P(X)} = \frac{\mathrm{P(X|A_i)\,P(A_i)}}{\mathrm{P(X|A_1)\,P(A_1) + P(X|A_2)\,P(A_2) + \cdots + P(X|A_n)\,P(A_n)}} = \frac{P(X \cap A_i)}{\sum_{i=1}^{n} \mathrm{P(X|A_1)\,P(A_1)}}$$

$$P(A_1\,|B_1) = \frac{P(A_1 \cap B_1)}{P(B_1)} = \frac{P(B_1 \cap A_1)}{P(B_1)} = \frac{\mathrm{P}(B_1|A_1)\,\mathrm{P}(A_1)}{\mathrm{P}(B_1|A_1)\,\mathrm{P}(A_1) + \mathrm{P}(B_1|A_2)\,\mathrm{P}(A_2)} = \frac{0{,}95x0{,}02}{0{,}95x0{,}02 + 0{,}05x0{,}98} = \frac{0{,}0190}{0{,}0680} =$$
$0{,}2794 \; ou \; 27{,}94\%$

b. Agora o que queremos é a probabilidade de uma pessoa NÃO ter a doença dado que o diagnóstico NÃO indica a presença da mesma = $P(A_2|B_2)$. Ainda $P(B_2) = 1 - P(B_1) = 1 - 0{,}0680 = 0{,}9320$ ou 93,20%

$$P(A_2\,|B_2) = \frac{P(A_2 \cap B_2)}{P(B_2)} = \frac{P(B_2 \cap A_2)}{P(B_2)} = \frac{\mathrm{P}(B_2|A_2)\,\mathrm{P}(A_2)}{0{,}9320} = \frac{0{,}95x0{,}98}{0{,}9320} = 0{,}9989 \; ou \; 99{,}98\%$$

No Excel:



Ai	P(Ai)	P(X|Ai)	P(Ai)P(X|Ai)	P(Ai|X)
A_1	2,00%	95,00%	1,90%	27,94%
A_2	98,00%	5,00%	4,90%	72,06%
A_3				
A_4				
A_5				
A_6				
A_7				
A_8				
A_9				
A_{10}				
A_{11}				
A_{12}				
A_{13}				
A_{14}				
A_{15}				
A_{16}				
A_{17}				
A_{18}				

11. Uma fábrica de sacolas tem 3 máquinas independentes que produzem o mesmo tipo de sacolas. A máquina 1 produz 15% das sacolas, com 1% delas com defeito. A máquina 2 produz 45% das sacolas, com 3% delas com defeito. A máquina 3 produz 40% das sacolas, com 2% delas com defeito.

a. Se selecionarmos uma sacola aleatoriamente, qual é a probabilidade de que ela seja defeituosa?

b. De qual máquina é mais provável que saia uma sacola, se a sacola não tem defeito?

Solução

$P(A_1)$ = Probabilidade da sacola ser produzida pela máquina 1 = 15%

$P(A_2)$ = Probabilidade da sacola ser produzida pela máquina 2 = 45%

$P(A_3)$ = Probabilidade da sacola ser produzida pela máquina 3 = 40%

$P(B_1|A_1)$ = Probabilidade da sacola com defeito vir da máquina 1 = 1%

$P(B_1|A_2)$ = Probabilidade da sacola com defeito vir da máquina 2 = 3%

$P(B_1|A_3)$ = Probabilidade da sacola com defeito vir da máquina 3 = 2%

$P(B_2|A_1)$ = Probabilidade da sacola boa, vir da máquina 1 = 1 - $P(B_1|A_1)$ = 99%

$P(B_2|A_2)$ = Probabilidade da sacola boa, vir da máquina 2 = 1 - $P(B_1|A_2)$ = 97%

$P(B_2|A_3)$ = Probabilidade da sacola boa, vir da máquina 3 = 1 - $P(B_1|A_3)$ = 98%

a. Queremos a probabilidade de que a sacola seja defeituosa.
   $P(B_1)$?
   $P(B_1) = P(B_1 \cap A_1) + P(B_1 \cap A_2) + P(B_1 \cap A_3)$
   $P(B_1) = P(A_1).\ P(B_1|A_1) + P(A_2).\ P(B_1|A_2) + P(A_3).\ P(B_1|A_3)$
   $P(B_1)$ = 0,15x0,01 + 0,45x0,03 + 0,40x0,02 = 0,0015 + 0,0135 + 0,0080 = 0,0230 ou 2,30%
b. Para sabermos qual a máquina mais provável de produzir uma sacola sem defeito (boa), basta aplicarmos o Teorema de Bayes a todas as máquinas e elegermos a maior probabilidade:
   $P(B_2)$ = 1 – $P(B_1)$ = 1 – 0,0230 = 0,9770

$$P(A_1\ |B_2) = \frac{P(A_1 \cap B_2)}{P(B_2)} = \frac{P(B_2 \cap A_1)}{P(B_2)} = \frac{\mathrm{P}(B_2|A_1)\ \mathrm{P}(A_1)}{0{,}9770} = \frac{0{,}99x0{,}15}{0{,}9770} = 0{,}1520\ ou\ 15{,}20\%$$

$$P(A_2\ |B_2) = \frac{P(A_2 \cap B_2)}{P(B_2)} = \frac{P(B_2 \cap A_2)}{P(B_2)} = \frac{\mathrm{P}(B_2|A_2)\mathrm{P}(A_2)}{0{,}9770} = \frac{0{,}97x0{,}45}{0{,}9770} = 0{,}4468\ ou\ 44{,}68\%$$

$$P(A_3\ |B_2) = \frac{P(A_3 \cap B_2)}{P(B_2)} = \frac{P(B_2 \cap A_3)}{P(B_2)} = \frac{\mathrm{P}(B_2|A_3)\ \mathrm{P}(A_3)}{0{,}9770} = \frac{0{,}98x0{,}40}{0{,}9770} = 0{,}4012\ ou\ 40{,}12\%$$

   A máquina mais provável de produzir uma sacola boa é a máquina 2 com uma probabilidade 44,68%.

12. O primeiro ano de bacharelado no curso de Administração de uma Universidade é integrado por 35 estudantes que se especializarão em Recursos Humanos (RH), 47 em Agronegócios e 40 em Finanças e 38 em Administração Geral. Sabe-se que a probabilidade que um estudante seja reprovado é de 5%, 4%, 3% e 4%, respectivamente. Qual a especialidade é mais provável que seja o estudante, se sabemos que ele ficou reprovado?

Solução

$P(A_1)$=Probabilidade do estudante cursar a especialidade RH=(35/160)=0,2188 ou 21,88%

$P(A_2)$=Probabilidade do estudante cursar a especialidade Agronegócios=(47/160)=0,2938 ou 29,38%

$P(A_3)$=Probabilidade do estudante cursar a especialidade Finanças=(40/160)=0,2500 ou 25,00%

$P(A_4)$=Probabilidade do estudante cursar a especialidade Administração Geral=(38/160)=0,2375 ou 23,75%

$P(B_1|A_1)$ = Probabilidade dos alunos de RH serem reprovados = 5% = 0,05.

$P(B_1|A_2)$ = Probabilidade dos alunos de Agronegócios serem reprovados = 4% = 0,04

$P(B_1|A_3)$ = Probabilidade dos alunos de Finanças serem reprovados = 3% = 0,03.

$P(B_1|A_4)$ = Probabilidade dos alunos de Adm. Geral serem reprovados = 4% = 0,04

A especialidade mais provável de vir um aluno reprovado é encontrada verificando-se o maior dos $P(A_i|B_1)$ (probabilidade da especialidade dado que o aluno é reprovado).

$P(B_1) = P(B_1 \cap A_1) + P(B_1 \cap A_2) + P(B_1 \cap A_3) + (B_1 \cap A_4)$
$P(B_1) = P(A_1).\ P(B_1|A_1) + P(A_2).\ P(B_1|A_2) + P(A_3).\ P(B_1|A_3) + P(A_4).\ P(B_1|A_4)$

$P(B_1)$ = 0,2188x0,05 + 0,2938x0,04 + 0,2500x0,03 + 0,2375x0,04 = 0,0109 + 0,0118 + 0,0075 + 0,0095 = 0,0397 ou 3,97%

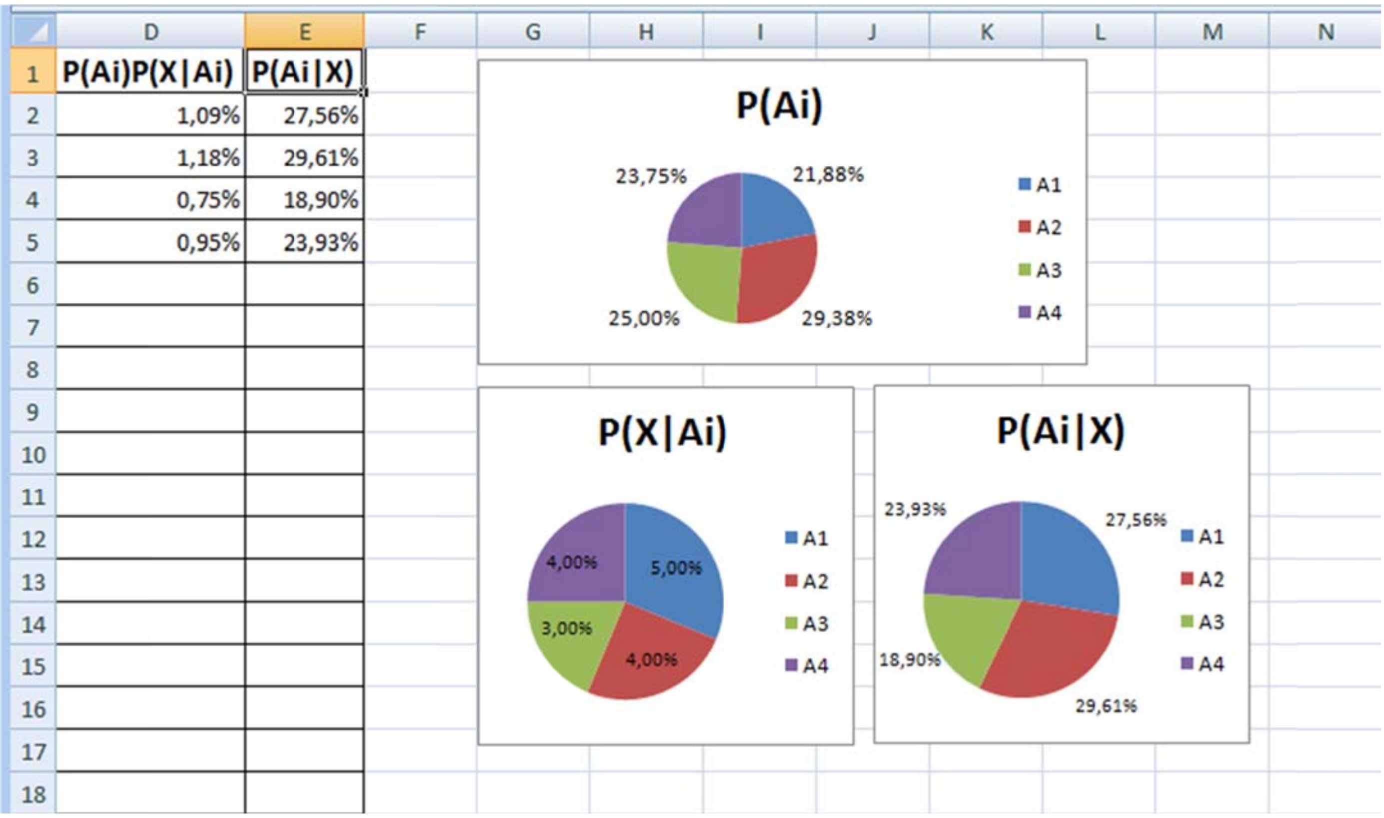
$$P(A_1\,|B_1) = \frac{P(A_1 \cap B_1)}{P(B_1)} = \frac{P(B_1 \cap A_1)}{P(B_1)} = \frac{\mathrm{P}(B_1|A_1)\;\mathrm{P}(A_1)}{0{,}0397} = \frac{0{,}05x0{,}2188}{0{,}0397} = 0{,}2756\;ou\;27{,}56\%$$

$$P(A_2\,|1) = \frac{P(A_2 \cap B_1)}{P(B_1)} = \frac{P(B_1 \cap A_2)}{P(B_1)} = \frac{\mathrm{P}(B_1|A_2)\mathrm{P}(A_2)}{0{,}0397} = \frac{0{,}2938\mathrm{x}0{,}04}{0{,}0397} = 0{,}2960\;ou\;29{,}60\%$$

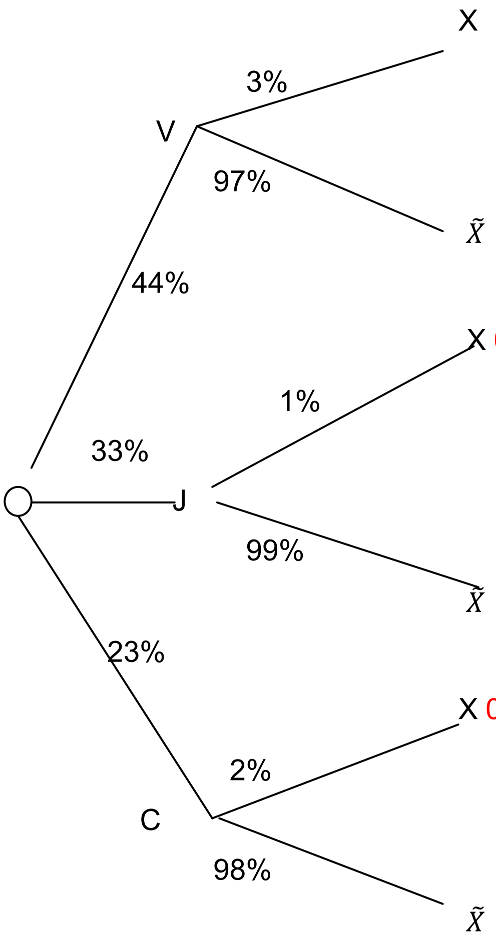
$$P(A_3\,|B_1) = \frac{P(A_3 \cap B_1)}{P(B_1)} = \frac{P(B_1 \cap A_3)}{P(B_1)} = \frac{\mathrm{P}(B_1|A_3)\;\mathrm{P}(A_3)}{0{,}0397} = \frac{0{,}2500\mathrm{x}0{,}03}{0{,}0397} = 0{,}1889\;ou\;18{,}89\%$$

$$P(A_4\,|B_1) = \frac{P(A_4 \cap B_1)}{P(B_1)} = \frac{P(B_1 \cap A_4)}{P(B_1)} = \frac{\mathrm{P}(B_1|A_4)\;\mathrm{P}(A_4)}{0{,}0397} = \frac{0{,}2375\mathrm{x}0{,}04}{0{,}0397} = 0{,}2393\;ou\;23{,}93\%$$

A especialidade mais provável de vir um aluno reprovado é a do Agronegócios.



	D	E
1	P(Ai)P(X|Ai)	P(Ai|X)
2	1,09%	27,56%
3	1,18%	29,61%
4	0,75%	18,90%
5	0,95%	23,93%



0,44 x 0,03 = 0,0132 ou 1,32% $\frac{0,0132}{0,0211} = 0,6256 \text{ ou } 62,56\%$

0,33 x 0,01 = 0,0033 ou 0,33% $\frac{0,0033}{0,0211} = 0,1564 \text{ ou } 15,64\%$

0,23 x 0,02 = 0,0046 ou 0,46% $\frac{0,0046}{0,0211} = 0,2180 \text{ ou } 21,80\%$

0,0211 ou 2,11%