Aula 1: métodos probabilísticos aplicados.

↳ base teórica para poder aplicar na análise de dados (principalmente nas matérias estatísticas).

O que é um modelo probabilístico

- Eq. matemática para descrever incertezas sobre os resultados de um experimento, atribuindo um valor de probabilidade para cada resultado possível.

moeda

$P(face) = 1/2$

face = cara ou coroa

dado

$P(face) = 1/6$

face = 1, 2, 3, 4, 5 ou 6

bola colorida em uma caixa

$$P(cor) = \begin{cases} 2/9, & cor = amarela \\ 4/9, & cor = azul \\ 3/9, & cor = vermelha \end{cases}$$

Prob. número x de caras em 5 jogadas de uma moeda

$$P(X=x) = \binom{5}{x} \cdot \left(\frac{1}{2}\right)^x \cdot \left(\frac{1}{2}\right)^{5-x}, \quad x = 0, 1, 2, 3, 4, 5.$$

Prob. do número x de bolas vermelhas em 2 bolas retiradas sem reposição da caixa (9 bolas: 3 vermelhas e 6 outra cor):

$$P(X=x) = \frac{\binom{3}{x} \cdot \binom{6}{2-x}}{\binom{9}{2}}, \quad x = 0, 1, 2$$

* modelo hipergeométrico

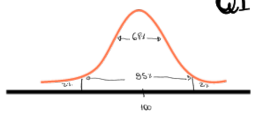
número de medicamentos usados por idosos: modelo Poisson

Parametros, p/ esse caso é a média da amostra

$$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, \quad x = 0, 1, 2, 3, 4 \ldots$$

Modelo Gaussiano (normal)

QI



▶ Problema a ser resolvido até o final da aula

→ Teorema de Bayes

Exame de sangue p/ diagnosticar certa doença

* dá resultado positivo em 90% das pessoas doentes.
* dá resultado negativo em 90% das pessoas não doentes.

O exame foi aplicado a uma pessoa selecionada ao acaso em uma cidade na qual 6% das pessoas têm a doença, e o resultado foi positivo

Experimento

Espaço amostral: conj de todos resultados

Probabilidade

Caso 1: os elementos do espaço amostral são equiprováveis (têm mesma probabilidade)

Razão das probabilidades.

Caso 2: os elementos do espaço amostral não são equiprováveis

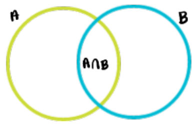
é a soma das probabilidades dos elementos que estão em A.

$$A \cap B = \{3, 9\} = 2/36 + 4/36 = 6/36 = 1/6$$

Regra Aditiva

Se A e B são dois eventos, então

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Ex:

$P(A) = 1/2$

$P(B) = 1/3$

$P(A \cap B) = 1/6$

Regra Aditiva

$P(A \cup B) = 1/2 + 1/3 - 1/6$

$P(A \cup B) = 2/3$

Axiomas da Probabilidade

Para um evento A qualquer do espaço amostral S: $0 \leq P(A) \leq 1$

$P(S) = 1$

Evento vazio → $P(\emptyset) = 0$

Se $A \cap B = \emptyset$ $\quad P(A \cup B) = P(A) + P(B)$

Probabilidade Condicional

↳ A probabilidade de um evento B ocorrer quando sabemos que A ocorreu

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

* Dois eventos são independentes se e somente se: $P(B|A) = P(B)$ ou $P(A|B) = P(A)$

Regra Multiplicativa

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A \cap B) = P(A) \cdot P(B|A)$$

$$P(A \cap B) = P(B) \cdot P(A|B)$$

Aula 2:

Continuação da Regra Multiplicativa

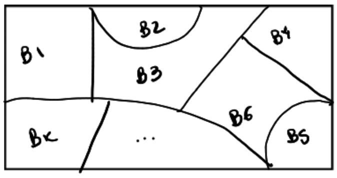
* Dois eventos são **independentes** se e somente se:

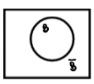
$$P(A\cap B) = P(A).P(B)$$

3 eventos:

$$P(A\cap B\cap C) = P(A).P(B|A).P(C|A\cap B)$$

Teorema da Probabilidade Total





Exemplo:

fábrica com 3 máquinas

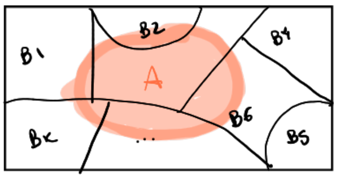
	Produção	Defeito
B_1	0,30	$P(A\mid B_1)$ 0,02
B_2	0,45	$P(A\mid B_2)$ 0,03
B_3	0,25	$P(A\mid B_3)$ 0,02

A: a peça é defeituosa

$$P(A) = P(A\cap B_1) + P(A\cap B_2) + P(A\cap B_3)$$

$$P(A) = P(A|B_1).P(B_1) + P(A|B_2).P(B_2) + P(A|B_3)P(B_3)$$

$$= (0{,}02)(0{,}30) + (0{,}03)(0{,}45) + (0{,}02)(0{,}25) = 0{,}0245$$



$$A = (A\cap B_1)\cup(A\cap B_2)\ldots\cup\ldots(A\cap B_k)$$

Se $P(A) = P(A\cap B_1) + P(A\cap B_2) \ldots + \ldots P(A\cap B_k)$

$$P(A) = P(B_1).P(A|B_1) + P(B_2).P(A|B_2) \ldots + \ldots P(B_k).P(A|B_k)$$

Então $P(A) = \sum_{i=1}^{k} P(B_i).P(A|B_i)$

Teorema de Bayes

o exemplo das máquinas de novo:

fábrica com 3 máquinas

	Produção	Defeito
B_1	0,30	$P(A\mid B_1)$ 0,02
B_2	0,45	$P(A\mid B_2)$ 0,03
B_3	0,25	$P(A\mid B_3)$ 0,02

A: a peça é defeituosa

Se a peça escolhida é defeituosa qual a Prob. de ter sido feita na maq 3?

$$P(B_3|A) = \frac{P(A\cap B_3)}{P(A)} = \frac{P(A|B_3).P(B_3)}{P(A|B_1).P(B_1)+P(A|B_2).P(B_2)+P(A|B_3)P(B_3)} = \frac{0{,}02 . 0{,}25}{0{,}0245} = 0{,}204$$

Problema Inicial

6% de pessoas doentes $P(D) = 0{,}06$

94% não doentes $P(N) = 0{,}94$

Resultado + 80% dos doentes $P(+|D) = 0{,}80$

Resultado − 90% dos não doentes $P(-|N) = 0{,}90$

Prob do resultado + estar errado? $P(N|+)$

$$P(N|+) = \frac{P(N\cap +)}{P(+)} = \frac{P(+|N)\,P(N)}{P(+|D).P(D) + P(+|N)\,P(N)}$$

$$= \frac{0{,}10\,(0{,}94)}{0{,}80\,(0{,}06) + 0{,}10\,(0{,}94)} = \frac{0{,}094}{0{,}142}$$

$= 0{,}662$ prob do diagnóstico + estar errado

Erro do diagnóstico negativo

$$P(D|-) = \frac{P(D\cap -)}{P(-)} = \frac{P(-|D).P(D)}{P(-|D).P(D)+P(-|N).P(N)}$$

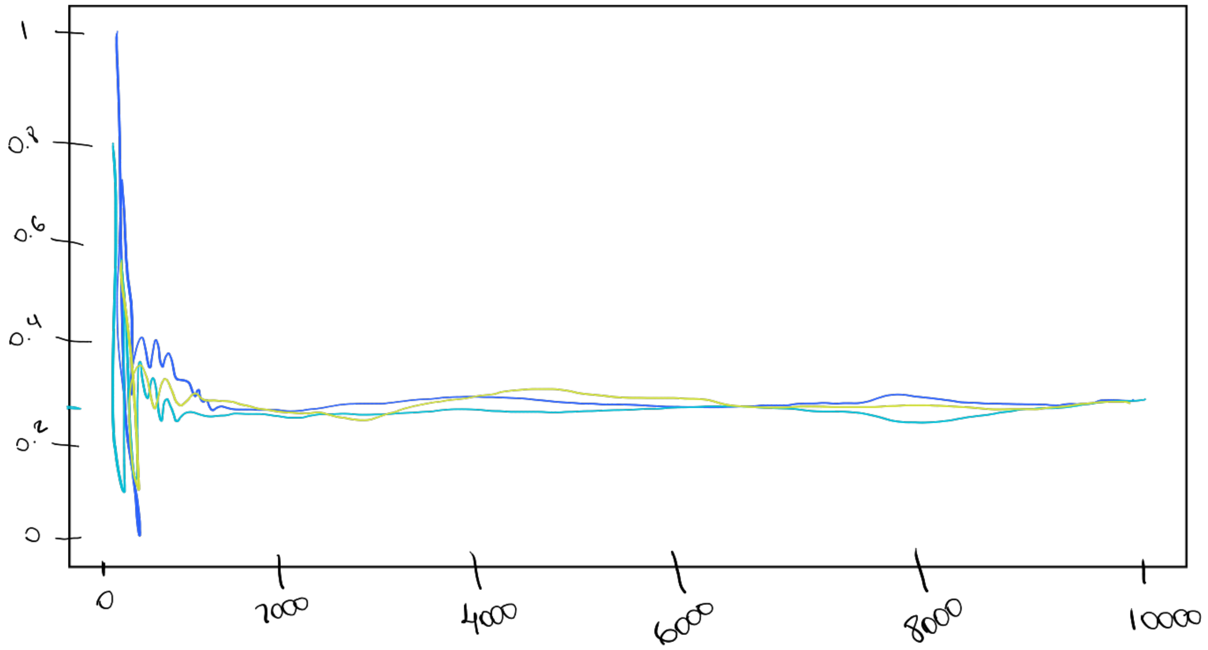
$$= \frac{0,20(0,06)}{0,20.(0,06)+0,90(0,94)} = \frac{0,012}{0,858} = 0,014$$ prob. do diagnóstico negativo estar errado

O teste é ruim? Não, apenas um teste de triagem que é bom em prever quando o paciente é não doente e isso se dá principalmente pela prevalência da doença ser muito baixa: 0,06

Lei dos Grandes Números

A medida que n cresce ($n \to \infty$), a frequência relativa do evento A tende a se estabilizar na vizinhança de um certo valor, ou seja, converge para um valor que é a P(A).

5 moedas viciadas (P(cara) = 0,25) jogadas 10 mil vezes



Exercícios: Lista 1   A e B

1 a)

Dado 1	Dado 2	Produto
1	1	1
1	2	2
1	3	3
1	4	4
1	5	5
1	6	6
2	1	2
2	2	4
2	3	6
2	4	8
2	5	10
2	6	12
3	1	3
3	2	6
3	3	9
3	4	12
3	5	15
3	6	18
4	1	4
4	2	8
4	3	12
4	4	16
4	5	20
4	6	24
5	1	5
5	2	10
5	3	15
5	4	20
5	5	25
5	6	30
6	1	1
6	2	12
6	3	18
6	4	24
6	5	30
6	6	36

Espaço Amostral: "valores possíveis para produto"

$\mathcal{E} = \{1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 16, 18, 20, 24, 25, 30, 36\}$

b)

Produto ímpar $A = \{1, 3, 5, 9, 15, 25\}$

Produto é múltiplo de 4 $B = \{4, 8, 12, 16, 20, 24, 36\}$

" " 5 $C = \{5, 10, 15, 20, 25, 30\}$

Produto é primo $D = \{2, 3, 5\}$

" é 13 $E = \{\ \}$

Ímpar múltiplo de 4 $A \cap B = \{\ \}$

" 5 $A \cap C = \{5, 15, 25\}$

Ímpar primo $A \cap D = \{3, 5\}$

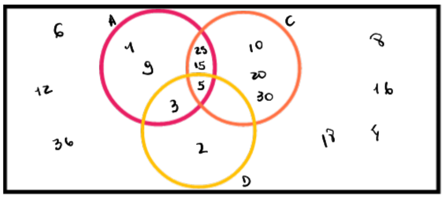
múltiplo de 4 ou 5 $B \cup C = \{4, 5, 8, 10, 12, 15, 16, 20, 24, 25, 30, 36\}$

" " ou primo $B \cup D = \{2, 3, 4, 5, 8, 12, 16, 20, 24, 36\}$

" de 5 ou primo $C \cup D = \{2, 3, 5, 10, 15, 20, 25, 30\}$

Complementar de B $B^c = \{1, 2, 3, 4, 6, 9, 10, 15, 18, 25, 30\}$

Diagrama de Venn de A, C, D



Exercícios:

C)

Elemento	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36
Probabilidade	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

D) $P(A) = (1+2+2+1)/36 = 9/36$

$P(B) = (3+2+4+1+2+2+1)/36 = 15/36$

$P(C) = (2+2+2+2+1+2)/36 = 11/36$

$P(D) = (2+2+2) = 6/36 = 1/6$

$P(E) = 0$

$P(A \cap B) = 0$

$P(A \cap C) = (2+2+1)/36 = 5/36$

$P(A \cap D) = (2+2)/36 = 4/36$

$P(B \cup C) = P(B) + P(C) = 24/36$

$P(B^c) = 1 - P(B) = 1 - 15/36 = 21/36$

E) $P(A \cup B) = 9/36 + 15/36 - 0 = 24/36$

$P(A \cup D) = 9/36 + 6/36 - 4/36 = 11/36$

$P(B \cap C) = 15/36 + 11/36 - 24/36 = 2/36$

F e G

F) $P(A|B) = P(A \cap B) / P(B) = 0$

$P(A|C) = P(A \cap C) / P(C) = 5/36 / 11/36 = 5/11$

$P(D|A) = P(A \cap D) / P(A) = 4/36 / 9/36 = 4/9$

G) A e B são dependentes pq $P(A|B) = 0 \neq P(A) = 9/3$

A e C são dependentes pq $P(A|C) = 5/11 \neq P(A) = 9/3$

A e D são dependentes pq $P(D|A) = 4/9 \neq P(D) = 6/3$

Exercício 2:

10 peças: 7 conformes e 3 defeituosas. 3 peças são tirados

a) $\mathcal{E} = \{C_1, C_2, C_3; D_1, D_2, D_2; C_1, C_2, D_3; C_1, D_2, C_3; D_1, C_2, C_3; D_1, D_2, C_3; D_1, C_2, D_3; C_1, D_2, D_3\}$

b)
1 $C_1C_2C_3 = 7/10 + 6/9 + 5/8 = 210/720$
2 $D_1D_2D_3 = 3/10 + 2/9 + 1/8 = 6/720$
3 $C_1C_2D_3 = 7/10 + 6/9 + 3/8 = 126/720$
4 $C_1D_2C_3 = 7/10 + 3/9 + 6/8 = 126/720$
5 $D_1C_2C_3 = 3/10 + 7/9 + 6/8 = 126/720$
6 $D_1D_2C_3 = 3/10 + 2/9 + 7/8 = 42/720$
7 $D_1C_2C_3 = 3/10 + 7/9 + 6/8 = 42/720$
8 $C_1D_2D_3 = 7/10 + 3/9 + 2/8 = 42/720$

c) Eventos: 1, 3, 4, 5

$P(A) = P(C_1C_2C_3) + P(C_1C_2D_3) + P(C_1D_2C_3) + P(D_1C_2C_3)$

$= \frac{(210+126+126+126)}{720}$

$= \frac{588}{720} = 0{,}82$

Exercício 3

$P(D) = 0,20$ 3 produtos são escolhidos.

$P(C) = 0,80$

a) $\mathcal{E} = C_1C_2C_3$
$D_1D_2D_3$
$C_1C_2D_3$
$C_1D_2C_3$
$D_1C_2C_3$
$D_1D_2C_3$
$D_1C_2D_3$
$C_1D_2D_3$

(b)
$P(C_1C_2C_3) = 0,8 \cdot 0,8 \cdot 0,8 = 0,512$
$P(D_1D_2D_3) = 0,2 \cdot 0,2 \cdot 0,2 = 0,008$
$P(C_1C_2D_3) = 0,8 \cdot 0,8 \cdot 0,2 = 0,128$
$P(C_1D_2D_3) = 0,8 \cdot 0,2 \cdot 0,8 = 0,128$
$P(D_1C_2C_3) = 0,2 \cdot 0,8 \cdot 0,8 = 0,128$
$P(D_1D_2C_3) = 0,2 \cdot 0,2 \cdot 0,8 = 0,032$
$P(D_1C_2D_3) = 0,2 \cdot 0,8 \cdot 0,2 = 0,032$
$P(C_1D_2D_3) = 0,8 \cdot 0,2 \cdot 0,2 = 0,032$

(C) $P(A) = P(C_1C_2C_3) + P(C_1C_2D_3) + P(C_1D_2C_3) + P(D_1C_2C_3)$
$= 0,512 + 0,128 + 0,128 + 0,128$
$= 0,896$

Exercício 4

4 refis Azuis (A) 3 são selecionados
3 refis Vermelhos (V)
2 refis Rosas (R)

(a) $P(A_1 \cap V_2 \cap R_3) = P(A) \times P(V|A) \times P(R|A \cap V) = 4/9 \times 3/8 \times 2/7 = 24/504$

(b) $P(V_1 \cap V_2 \cap V_3) = P(V_1) \times P(V_2|V_1) \times P(V_3|V_1 \cap V_2) = 3/9 \times 2/8 \times 1/7 = 6/504$

(c) Impossível $P = 0$

Exercício 5:

	Prod.	Def.
m_1	10	1
m_2	35	2
m_3	55	3

B_i
$P(B_1) = 0,10$
$P(B_2) = 0,35$
$P(B_3) = 0,55$

A:
$P(A|B_1) = 0,01$
$P(A|B_2) = 0,02$
$P(A|B_3) = 0,03$

$P(A) = P(B_1 \cap A) + P(B_2 \cap A) + P(B_3 \cap A)$
$= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3) \cdot P(B_3)$
$= (0,01)(0,10) + (0,02) \cdot (0,35) + (0,03)(0,55)$
$= 0,0245$

Exercício 6:

(a) $P(B_2|A) = \dfrac{P(B_2 \cap A)}{P(A)} = \dfrac{P(A|B_2)\,P(B_2)}{P(A)} = \dfrac{0,02 \cdot 0,35}{0,0245} = \dfrac{0,007}{0,0245} = 0,29$

(b) $P(B_3|A^c) = \dfrac{P(A^c|B_3)\,P(B_3)}{P(A^c)} = \dfrac{(1 - P(A|B_3)) \cdot P(B_3)}{1 - P(A)} = \dfrac{(1-0,03) \cdot (0,55)}{1 - 0,0245} = 0,55$

Exercício 7

$P(+|D) = 0,90$
$P(-|N) = 0,95$
$P(D) = 0,12$
$P(N) = 0,88$

Para $P(D) = 0,12$

$P(-|D) = \dfrac{P(-|D) \cdot P(D)}{P(-|D) \cdot P(D) + P(-|N)\,P(N)} = \dfrac{(1 - P(+|D)) \cdot P(D)}{(1 - P(+|D)) \cdot P(D) + P(-|N) \cdot P(N)} = \dfrac{1 - 0,90 \cdot 0,12}{1 - 0,90 \cdot 0,12 + 0,95 \cdot 0,88} = 0,014$

Para $P(D) = 0,01$

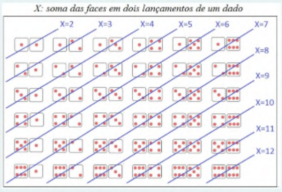
$\dfrac{1 - 0,90 \cdot 0,01}{1 - 0,90 \cdot 0,01 + 0,95 \cdot 0,99} = 0,0011$

Resolução Teste 1

Questão 1

Considere o experimento aleatório de jogar um **dado honesto duas vezes** e observar o número da face de cima em cada jogada.

Veja os resultados possíveis deste experimento na figura abaixo (1a face e 2a face):



Sejam os eventos A e B definidos como:

A = "o número da face na primeira jogada é 1";

B = "a soma dos números das faces nas duas jogadas é maior ou igual a 7".

Notação improvisada: **& simboliza interseção** e **# simboliza união.**

Responda o valor das seguintes probabilidades:

a) P(A) = 1/6 ✔ .

b) P(B) = 7/12 ✔ .

c) P(A&B) = 1/6 ✘ .

d) P(A#B) = 13/18 ✔ .

$P(A) = 6/36 = 1/6$

$P(B) = \frac{6+5+4+3+2+1}{36} = \frac{21^{\div 3}}{36_{\div 3}} = \frac{7}{12}$

$P(A \cap B) = 1/36$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$1/6 + 7/12 - 1/36 =$

$\frac{6 + 21 - 1}{36} = \frac{26^{\div 2}}{36_{\div 2}} = \frac{13}{18}$

13/18 Resultado certo

Questão 2

Uma pesquisa de opinião foi realizada com 1000 clientes (amostra) de uma operadora de celular, selecionados aleatoriamente do cadastro (população). Cada cliente foi identificado pela sua faixa etária (< 60 anos ou ≥ 60 anos) e sua resposta à seguinte pergunta: você está satisfeito com seu plano (sim ou não)?

Os resultados são mostrados na tabela a seguir.

Faixa etária	Satisfação c/ plano		**Total** (na linha)
	Sim	Não	
< 60 anos	60	540	**600**
≥ 60 anos	240	160	**400**
Total (na coluna)	**300**	**700**	**1000**

Suponha que um cliente seja escolhido aleatoriamente do cadastro (população de clientes).

Tomando as frequências relativas nesta amostra de clientes como estimativas das probabilidades na população, expressando os valores de probabilidades entre 0 (zero) e 1 (um), com duas casas decimais, e usando a vírgula como separador de decimais, **calcule as seguintes probabilidades:**

a) Do cliente selecionado ter "idade ≥ 60 anos": ✘ .

b) Do cliente selecionado estar "satisfeito": ✘ .

c) Do cliente selecionado ter "idade ≥ 60 anos" e "estar satisfeito": ✘ .

d) Suponha que o cliente selecionado tenha "idade ≥ 60 anos". A probabilidade deste cliente "estar satisfeito" é: ✘ .

e) Suponha que o cliente selecionado esteja "satisfeito". A probabilidade deste cliente ter ""idade ≥ 60 anos" é: ✘ .

f) Os eventos "ter idade ≥ 60 anos" e "estar satisfeito" são independentes? ✘ .

a) $\frac{400}{1000} = \frac{4}{10}$

c) $P(S \cap +60) = \frac{240}{1000} = \frac{24}{100} = 0{,}24$

b) $\frac{300}{1000} = \frac{3}{10}$

* d) $P(+60 \cap S) = \frac{240}{400} = \frac{24}{40} = 0{,}6$

e) $P(S \cap +60) = \frac{240}{300} = 0{,}8$

f) $P(+60|S) = \frac{P(+60 \cap S)}{P(S)} = \frac{0{,}24}{0{,}3} = 0{,}8 \neq P(+60) = 0{,}4$

→ são dependentes

Questão 3

Uma pessoa é levada para a emergência de um hospital com dor abdominal.

Sejam os eventos:

A: a pessoa tem câncer de fígado;

B: a pessoa precisará de um transplante de fígado;

C: o hospital encontrará um fígado compatível para o transplante.

Pelos registros do hospital: 30% das pessoas com este tipo de dor são diagnosticadas com câncer de fígado; dentre as pessoas diagnosticadas com câncer de fígado, 80% precisaram de um transplante; o hospital encontrou um órgão compatível para 10% das pessoas nesta situação (ter um câncer de fígado que precise de transplante).

Qual a probabilidade desta pessoa fazer o transplante de fígado? ✘ .

(Use três casas decimais e a vírgula como separador decimal na resposta final.)

$P(A^c) = 0{,}70$ $P(B^c) = 0{,}20$ $P(C^c) = 0{,}90$

$P(A) = 0{,}30$ $P(B) = 0{,}80$ $P(C) = 0{,}10$

$P(A \cap B \cap C) = P(A) \times P(B|A) \times P(C|A \cap B)$

$0{,}30 \times 0{,}80 \times 0{,}10 = 0{,}024$

Questão 4

Um consultório odontológico tem três dentistas: Ana, Bruno e Cleo.

Os três dentistas recebem pacientes que querem marcar uma consulta pelo "plano de saúde" ou "particular".

Pelos registros da clínica, a porcentagem de pacientes que marcaram a consulta pelo plano de saúde foi de 20%, 30% e 15%, respectivamente, para os dentistas Ana, Bruno e Cleo.

Ainda pelos registros da clínica, 25% dos pacientes que ligam pedem uma consulta com a dentista Ana, 35% com Bruno e 40% com a Cleo.

Um paciente liga para clínica querendo marcar uma consulta.

Qual a probabilidade deste paciente querer marcar um consulta pelo plano de saúde? ✘ .

(Use três casas decimais nas contas intermediárias e no resultado final. Use a vírgula como separador decimal.)

	Plano	P
Ana	20	25
Bruno	30	35
Cleo	15	40

	Ana	Bruno	Cleo
$\mathcal{E}_0 =$	(0,25 × 0,20)	(0,35 × 0,30)	(0,40 × 0,15)
	0,05	0,105	0,06

A: prob. do paciente marcar pelo plano

$P(A) = 0{,}050 + 0{,}105 + 0{,}060$

$P(A) = 0{,}215$

Questão 5

[Pode-se aproveitar a resposta da Questão 4 para encurtar a solução.]

Um consultório odontológico tem três dentistas: Ana, Bruno e Cleo.

Os três dentistas recebem pacientes que querem marcar uma consulta pelo “plano de saúde” ou “particular”.

Pelos registros da clínica, a porcentagem de pacientes que marcaram a consulta pelo plano de saúde foi de 20%, 30% e 15%, respectivamente, para os dentistas Ana, Bruno e Cleo.

Ainda pelos registros da clínica, 25% dos pacientes que ligam pedem uma consulta com a dentista Ana, 35% com Bruno e 40% com a Cleo.

Um paciente liga para clínica querendo marcar uma consulta pelo plano de saúde.

Qual a probabilidade deste paciente querer marcar uma consulta com o dentista Bruno?

(Use três casas decimais nas contas intermediárias e no resultado final. Use a vírgula como separador decimal.)

$$P(B|A) = \frac{P(B \cap A) \cdot P(B)}{P(B \cap A) \cdot P(B) + P(O \cap A) \cdot P(O)} = \frac{0{,}30 \times 0{,}35}{0{,}30 \times 0{,}35 + 0{,}35 \cdot 0{,}65}$$

$$= \frac{0{,}105}{0{,}105 + 0{,}11} = \frac{0{,}105}{0{,}215} = 0{,}488$$