Métodos Probabilísticos para Engenharia Informática

2018-2019

Aulas
3 e 4

Axiomática das probabilidades
Probabilidade condicional
Regra de Bayes
Independência

Definição axiomática de probabilidade

- Em 1933, Kolmogorov estabeleceu a DEFINIÇÃO DE PROBABILIDADE POR AXIOMATIZAÇÃO
  – na sua obra Foundations of the Theory of Probability.

O que é uma axiomática?

Em determinado ponto da evolução de uma teoria de pensamento matemático, torna-se imperioso ordenar, sistematizar e relacionar todos os conhecimentos entretanto nela reconhecidos, isto é, proceder à sua **AXIOMATIZAÇÃO**

O que é uma axiomática?

- **Axiomatizar** consiste em escolher algumas afirmações que podem ser feitas sobre os objectos matemáticos em estudo, na área considerada.

- Delas, **por processo dedutivo**, obter todas as demais proposições que constituem o corpo de conhecimento da teoria em causa.

O que é uma axiomática?

Essas afirmações, das quais deduzimos todas as outras, são os **AXIOMAS** e o seu conjunto constitui uma **AXIOMÁTICA**.

Axiomas

Os axiomas, além de se basearem numa aceitação por evidência, devem ser :

- **logicamente independentes**
  - isto é, nenhum deles deve ser passível de se obter dos restantes
- **Compatíveis**
  - isto é, os axiomas não podem, por dedução lógica, conduzir a proposições contraditórias

Definição axiomática de probabilidade

Foi com base nas <u>propriedades das frequências relativas</u> e das <u>operações sobre conjuntos</u> que **Kolmogorov** concebeu a primeira construção **AXIOMÁTICA GERAL** para a **TEORIA DAS PROBABILIDADES.**

Axiomática de probabilidades

- Axioma 1- probabilidades são não-negativas
  P(A) > = 0
- Axioma 2 – normalização (S tem probabilidade 1)
  P(S) =1

- Axioma 3a – Se A e B forem mutuamente exclusivos
  P(A U B) = P(A) + P(B)

- Axioma 3b – Se A1, A2, … for uma sequência infinita de acontecimentos mutuamente exclusivos ($\forall_{i \neq j} Ai \cap Aj = \emptyset$)

$$P(\cup_{k=1}^{\infty} Ak) = \sum_{i=1}^{\infty} P(Ak)$$

Teoremas

- Como sabem, às afirmações que se obtêm dedutivamente a partir dos axiomas, ou de outras já deles obtidas por dedução, chamamos TEOREMAS.

Teoremas / Corolários : Prob. do acontecimento complementar

- $P(\bar{A}) = 1 - P(A)$

Demonstração:

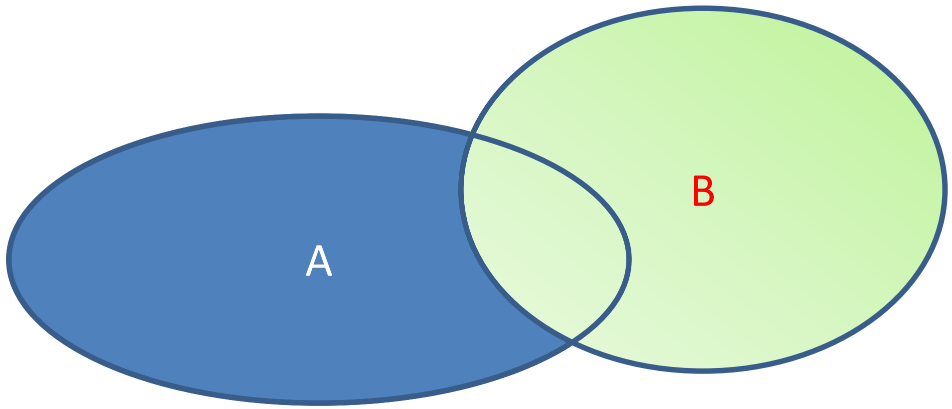
- Como $A \cap \bar{A} = \emptyset$
- E $A \cup \bar{A} = S$

- Pelo axiomas 2 e 3:
- $P\,(A \cup \bar{A}) = P(A) + P(\bar{A}) = P(S)$ ...
- $P(A) + P(\bar{A}) = 1$
- $P(\bar{A}) = 1 - P(A)$

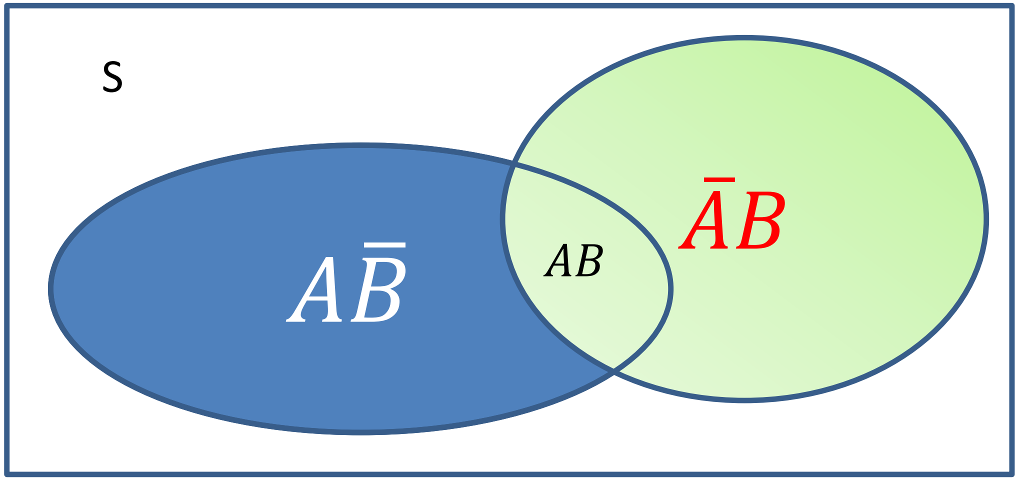
Teorema/Corolário: prob. da união

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

Com AB$\equiv A \cap B$



Demonstração



$A \cup B = A\bar{B} \cup AB \cup \bar{A}B$, disjuntos

Logo (axioma 3):

$$\mathrm{P}(A \cup B) = P(A\bar{B}) + P(AB) + P(\bar{A}B)$$

Adicionando e subtraindo P(AB)

$$\begin{aligned}&\mathrm{P}(A \cup B)\\ &= \big(P(A\bar{B}) + P(AB)\big) + P\big((\bar{A}B) + P(AB)\big) - P(AB)\\ &= P(A) + P(B) - P(AB)\end{aligned}$$

Probabilidade em espaços de amostragem não contáveis

- Exemplo:
  - Escolha de um número real no intervalo [a,b]
- Seja o acontecimento A "número pertencer a [c,d]"

- P(A)= (d-c)/ (b-a)

- A probabilidade de qualquer ponto $x \in [a,b]$ é $igual\ \ a\ 0$
  - Ter, por exemplo,]c,d[dará igual

Probabilidade em espaços de amostragem não contáveis

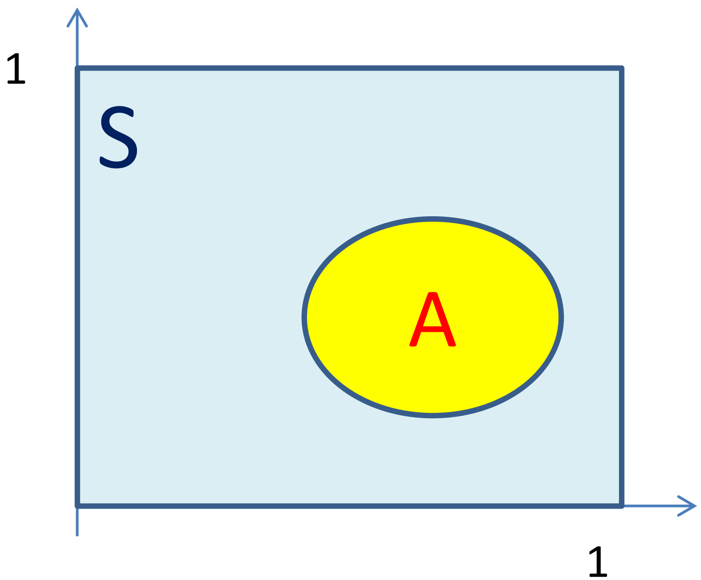
- Exemplo:
  - Escolha de um número real no intervalo [0,90] relativo ao atraso de chegada a uma aula de 1h30m
- Seja o acontecimento A “chegar dentro da tolerância”, i.e. [0,15[

- P(A)= (15-0)/ (90-0) = 0.16(6)
  - Assumindo que podem chegar em qualquer altura da aula ☹
    - O que não é válido

Probabilidade em espaços de amostragem não contáveis

- No caso de um par de números reais x,y entre 0 e 1

$$S = \{x, y: x \in [0,1] \cap y \in [0,1]\}$$



$$P(A) = \frac{Área\ (A)}{Área\ (S)}$$

A axiomática é compatível com as teorias anteriores ?

A definição frequencista de probabilidade satisfaz a axiomática

- Satisfaz o axioma 1: «p(A) não negativo»
  - pois se as frequências são números não negativos também convergem para um número não negativo.

A definição frequencista de probabilidade satisfaz a axiomática

- Satisfaz o axioma 2: «p(E)=1»
  - pois as frequências relativas de um acontecimento certo são sempre 1, logo, tendem para 1.

A definição frequencista de probabilidade satisfaz a axiomática

- Satisfaz o axioma 3: «Se (A e B são incompatíveis) então a probabilidade da reunião de A com B é igual à soma das probabilidades de A e de B»
  - pois se os acontecimentos A e B são incompatíveis não têm resultados comuns, a frequência relativa de AUB é a frequência relativa de A mais a frequência relativa de B e o limite da soma das duas sucessões é a soma dos limites.

A Lei de Laplace verifica a axiomática

- Lei de Laplace
  - Num espaço finito de resultados equiprováveis a probabilidade de um acontecimento A é igual ao quociente entre o número de resultados favoráveis a A (#A) e o número de resultados possíveis (#E)

A Lei de Laplace verifica a axiomática

- Satisfaz o axioma 1: «p(A) não negativo»
  - Pois $p(A)=\#A\ /\ \#E$ o que significa que p(A) é o quociente entre um número real não negativo e um número positivo.

A Lei de Laplace verifica a axiomática

- Satisfaz o axioma 2: «p(E)=1»
  - Pois   p(E)= #E / #E é o quociente entre dois números iguais.

A Lei de Laplace verifica a axiomática

- Satisfaz o axioma 3: «Se (A e B são incompatíveis) então a probabilidade da reunião de A com B é igual à soma das probabilidades de A e de B»

  – Se A e B são disjuntos

$$\#(A \cup B) = \#A + \#B$$

E então:

$$p(A \cup B) = \frac{\#(A \cup B)}{\#E} = \frac{\#A + \#B}{\#E} = \frac{\#A}{\#E} + \frac{\#B}{\#E} = p(A) + p(B)$$

Exemplo de Aplicação

k ocorrências em n experiências

- n lançamentos de moeda de 1 Euros

  P(Face) = P(F) = p   ; P(Verso) = P(V) = 1-p

- P(FVVFFF)= P(F)xP(V) … = p (1-p) (1-p) p p p

- P(FVVFFF) = $p^{\#\,Faces}\,(1-p)^{\#\,Versos}$

$$= p^4(1-p)^2$$

…

- P(“k Faces”) = ?
- P(“k Faces”) = $\sum_{sequências\ com\ k\ Faces} P(sequência)$

$$=(\#\ sequências\ de\ k\ Faces).\, p^k(1-p)^{n-k}$$

- Num de sequências = Combinações com k faces em n
- P(“k Faces”)= $\binom{n}{k}.\, p^k(1-p)^{n-k}$

$$= \frac{n!}{(n-k)!\,k!}.\, p^k(1-p)^{n-k}$$

Para aprender mais …

- Links para material online:
  - http://www.stat.berkeley.edu/~stark/SticiGui/Text/probabilityAxioms.htm
  - https://www.youtube.com/course?list=PL10921DED3A8BFF53

- Capítulos iniciais do Livro “Probabilidades e Processos Estocásticos”, F. Vaz, Universidade de Aveiro
- Capítulos iniciais do Livro “O Acaso” , Joaquim Marques de Sá, Gradiva

Probabilidade condicional

- Muitas vezes interessa-nos saber a probabilidade de um evento assumindo outro evento

- Exemplos:
  - Probabilidade de amanhã estar sol **dado que hoje esteve sol**
  - Probabilidade de ter uma doença **sabendo que tenho um sintoma**

Exemplo 1 - Manchas e Sarampo

- P(“doente ter Manchas dado que tem Sarampo”) =?

Manchas / Sarampo	M	$\bar{M}$	
S	x casos		s casos
$\bar{S}$			
			N casos

- **P(S e M)** = ?
- **P(S e M)** = $\frac{x}{N} = \frac{x}{s} \times \frac{s}{N}$

$$= \boldsymbol{P(M \mid S) \times P(S)}$$

Também pode ser : $P(\,S \mid M) \times P(M)$

- $\boldsymbol{P(M \mid S)}$**= P(S e M) /** $\boldsymbol{P(S)}$

Exemplo 2

Voltando à família com 2 filhos…

Sabendo que um dos filhos é rapaz, qual a probabilidade do outro ser também rapaz ?

Resolução

Como vimos $P(A \mid B)$= P(A e B) / $P(B)$

Considerando os eventos

A="outro filho ser do género Masculino"

B="um filho ser do género Masculino"

$$P(\text{"}outro\ \ filho\ Masc\text{"} \mid \text{"}um\ \ filho\ Masc\text{"}) = \frac{P(um\ M \cap outro\ M)}{P(um\ M)}$$

$$= \frac{1/4}{3/4}$$

$$= \frac{1}{3}$$

=P({MM} se {MF,FM, MM})

outro→ Um dos filhos	M	F
M		
F		

Probabilidade condicional

- Por vezes – como nos exemplos anteriores – dois acontecimentos estão relacionados
  - A ocorrência de um depende ou faz depender a ocorrência do outro

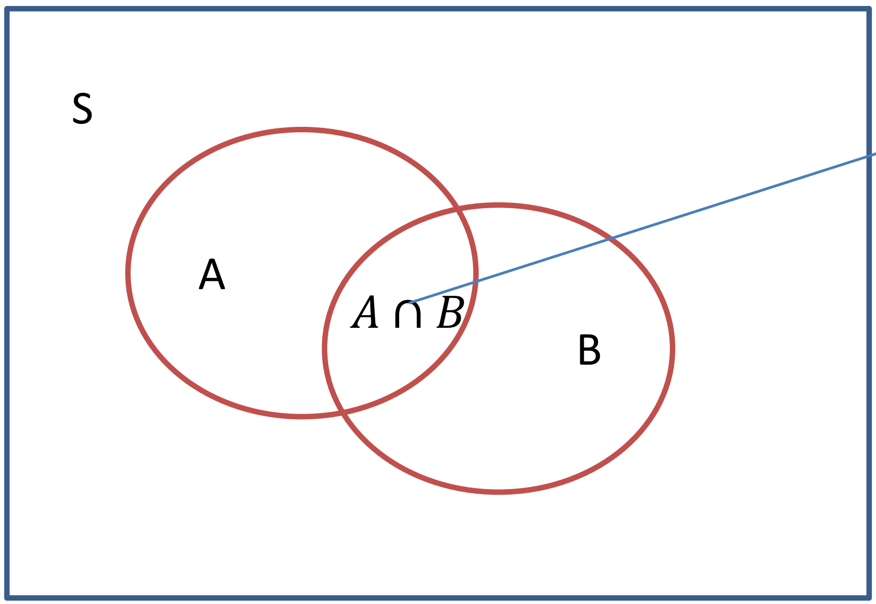
- A Probabilidade de um evento A com a condição que o evento B ocorreu é a designada **PROBABILIDADE CONDICIONAL** de A dado B
  - Definida por:

$$P(A|B) = \frac{\boldsymbol{P(AB)}}{P(B)} \; se \; P(B) \neq 0$$

Indefinida se P(B)=0

Interpretação da probabilidade condicional

- $P(A|B) = \frac{P(AB)}{P(B)}$ $se\ P(B) \neq 0$



Zona onde A se realiza,
Sabendo que B ocorreu

Exemplo de aplicação

- 2 números de 1 a 4, N1 e N2

- Evento B = "min(N1,N2)=2"
- Evento M = "max(N1,N2)"

- **P(M=1|B) =**

  P("max()=1" & "min()=2") / P("min()=2") =

  …

  =0

- P(M=2|B) = …

  = 1/5

N2→	1	2	3	4
1				
2		B / 2	B /3	B /4
3		B /3		
4		B /4		

P(AB), P(ABC) …

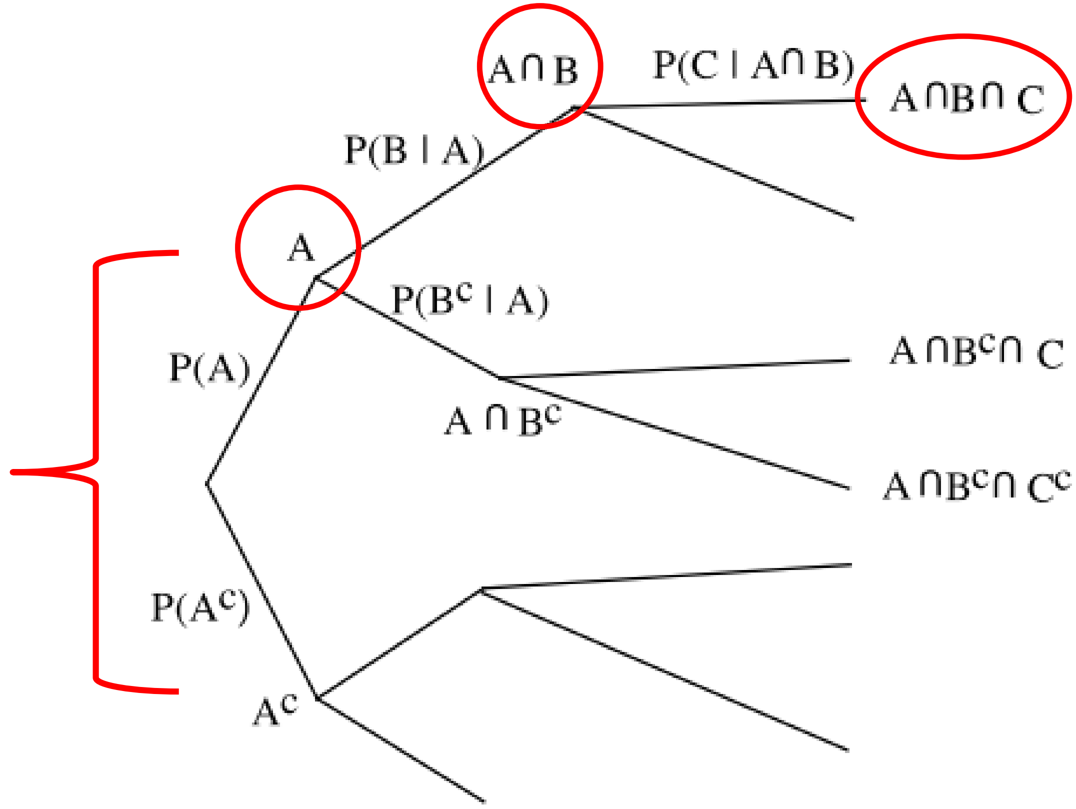
- $P(AB) = P(A|B) \times P(B)$

- Aplicando sucessivamente temos (regra da cadeia)

$$P(A_1 A_2 A_3 \dots A_n) \\ = P(A_1|A_2 \dots A_n) \times P(A_2 A_3 \dots A_n) \\ = P(A_1|A_2 \dots A_n) \\ \times P(A_2|A_3 \dots A_n) \dots P(A_{n-1}|A_n)$$

Regra da cadeia / multiplicação

- $P(ABC) = P(A)P(B|A)\,P(C|AB)$



Problema com 2 urnas...

- Consideremos 2 urnas, designadas por X e Y, contendo bolas brancas e pretas:
  - X contém 4 brancas e 5 pretas e
  - Y contém 3 brancas e 6 pretas.
- Escolhe-se uma urna ao acaso e extrai-se uma bola. Qual a probabilidade de sair bola branca?

- **P("bola branca")**

= P("branca da urna X OU branca da urna Y")

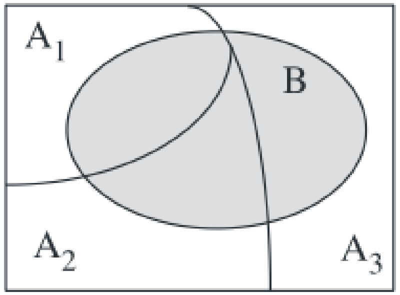
= P("branca E urna X")+P("branca E urna Y")

= P("branca"|"urna X") x P("urna X") + P("branca"|"urna Y") x P("urna Y")

= (4/9)x(1/2)+(3/9)x(1/2) ≈ 0,39

Lei da Probabilidade total

- Dividir para conquistar
- Partição do espaço de amostragem A_1, A_2, A_3
- Ter $P(B|A_i)$, para todos os is



- $P(B) = P(B|A_1)P(A_1)$
  $+P(B|A_2)P(A_2)$
  $+ P(B|A_3)P(A_3)$

Em geral: $\mathrm{P(B)} = \sum_j P(B|A_j)P\left(A_j\right)$

Condicionamento inverso

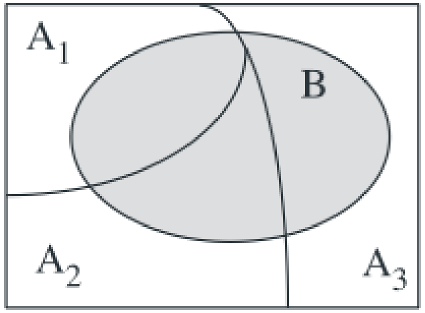
- Continuando com as urnas ...
- **Problema Inverso** (condicionamento inverso)

P("urna X" | "bola branca")

Resolvido pela primeira vez pelo Reverendo Thomas Bayes (1702-1761)

Regra de Bayes

- Probabilidades *a priori* $P(A_i)$
- Sabemos $P(B|A_i)\ \ \forall i$



- Pretendemos calcular $\boldsymbol{P(A_i|B)}$
  - i.e. $P(A_i)$ dado que B ocorreu

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

$= \frac{P(B|A_i)P(A_i)}{P(B)}$

$= \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)}$

Aplicando ao problema das urnas

- **P("urna X" | "bola branca")**

$$= \frac{P(\text{"bola branca"}|\text{"}urna\ X\text{"}) \times P(\text{"}urna\ X\text{"})}{P(\text{"bola branca"})}$$

$$= \frac{\left(\frac{4}{9}\right)\times\left(\frac{1}{2}\right)}{7/18} = \frac{4}{7}$$

- P("urna Y" | "bola branca")=

$$\ldots = \frac{\left(\frac{3}{9}\right)\times\left(\frac{1}{2}\right)}{7/18} = \frac{3}{7}$$

$$= 1 - P(\text{"urna X"}|\text{"bola branca"})$$

Causa e efeito

- No evento "urna X se bola branca" podemos considerar que a saída de bola branca é o EFEITO da causa "urna X"
- A Regra de Bayes, em consequência, pode ser escrita da seguinte forma:

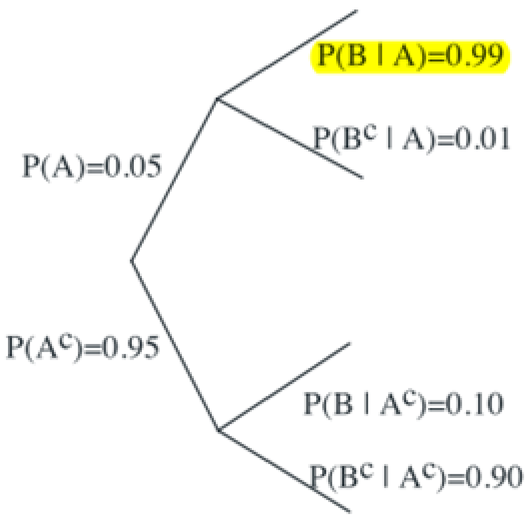
$$\boldsymbol{P}(\text{"causa"}|\text{"efeito"}) = \frac{\boldsymbol{P}(\text{"efeito"}|\text{"causa"}) \times \boldsymbol{P}(\text{"causa"})}{\boldsymbol{P}(\text{"efeito"})}$$

Exemplo de aplicação Regra de Bayes - radar

Evento A: avião voando na zona do radar, P(A)= 0.05

Evento B: Aparece algo no ecrã do radar, P(B|A)= 0.99

P(A | B)= ?



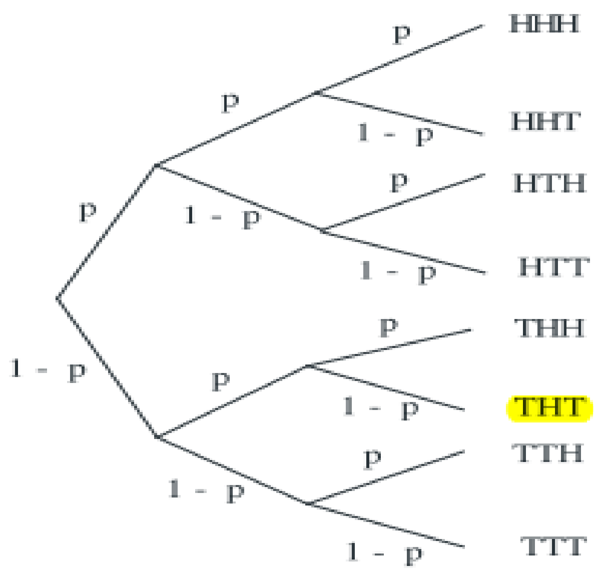
- $P(A \cap B) =$
  $= P(B|A)\,P(A)$
  $= 0{,}99 \times 0{,}05$

- $P(B) =$
  $= P(B|A) \times P(A) + P(B|\bar{A}) \times P(\bar{A})$
- $P(A \mid B) =$
  $= \frac{P(B|A) \times P(A)}{P(B)}$
  $= \frac{0{,}99 \times 0{,}05}{0{,}1445} = 0.3426$
  (valor baixo)

Outro exemplo

3 lançamentos de uma moeda não honesta

P(H) = p , p(T)=1-p



Calcular:

- P(THT)

=(1-p) p (1-p)

- P("1 H")

= P(HTT)+P(THT)+P(TTH)
= p(1-p)(1-p) + (1-p)p(1-p)+(1-p)(1-p)p
= 3x p(1-p)(1-p)

- P("primeiro lançamento dá H" | "1 H")

= P (primeiro H & 1 H) /p(1H)
… = p(1-p)(1-p) / … = 1/3

Confirmando caso favorável {HTT} e casos possíveis {HTT, THT, TTH}
Nota: os casos possíveis são equiprovaveis

Terceiro exemplo – sistema de comunicação/transmissão

- 0 ou 1 na entrada, transmissão sujeita a erros, entradas equiprováveis.
- No receptor uma decisão é tomada (0 ou 1):
  - Se ε for a probabilidade de erro, qual a entrada mais provável se na saída obtemos 1 ?

- Seja A_k o acontecimento "entrada é k", k=0,1

  $A_0\ eA_1$constituem uma partição de S
- Seja B_1 o acontecimento "saída = 1"

…

- $P(B_1) = P(B_1|A_0) \times P(A_0) + P(B_1|A_1) \times P(A_1)$

$$= \varepsilon\frac{1}{2} + (1-\varepsilon)\frac{1}{2} = \frac{1}{2}$$

- Probabilidade de entrada ter sido zero dado que saída igual a 1 ?
  $P(A_0|B_1) = P(B_1|A_0) \times P(A_0)/P(B_1)$
  $= \frac{\varepsilon}{2} \,/\, \frac{1}{2} = \varepsilon$

- De forma similar $P(A_1|B_1) = \ldots = 1-\varepsilon$

- Se ε <1/2 a entrada mais provável é 1 (se saída for 1)
  – Que é o que se pretende em geral.

Novo problema com filhos

- Novamente com a família de 2 filhos..
- Sabendo que o primeiro filho é rapaz, qual a probabilidade do outro filho ser também rapaz ?

Calculemos

outro→ Primeiro filho	M	$\overline{M}$=F	
M	x casos		m casos
$\overline{M}$=F			N-m casos
			N casos

$$A = P(outro\ M \mid primeiro\ M) = \frac{P(outro\ M \cap primeiro\ M)}{P(primeiro\ M)}$$

$$= \frac{1/4}{2/4} = \frac{1}{2}$$

Voltando um pouco atrás

- As duas variantes do problema dos dois filhos deram resultados (surpreendentemente?) diferentes

- $P(\text{“outro M”} \mid \text{“primeiro M”}) = \frac{1}{2} = P(\text{“outro M”})$

- $P(\text{“outro M”} \mid \text{“um M”}) \neq P(\text{“outro M”})$

- O que significam estes resultados ?

Independência

- 2 acontecimentos são independentes sse P(AB) = P(A)P(B)
  - Simétrico relativamente a A e B
  - Aplica-se mesmo que P(A)=0
  - Implica P(A|B)=P(A)   [mas não é a definição]
    - Ocorrência de B não fornece informação sobre ocorrência de A

- Generalização...

  os acontecimentos $A_1, A_2, A_3 \dots A_n$ são independentes sse

  $$P(A_1 \cap A_2 \cap A_3 \dots \cap A_n) = \mathrm{P}(A_1)P(A_2)P(A_3)\dots P(A_n)$$

Independência vs independência 2 a 2

- 2 lançamentos de moeda
- A: primeira é caras
- B: segunda é caras
- C: mesmo resultado em ambas

HH	HT
TH	TT

- $P(C)$?   P(A) ? P(B) ?

  2/4=1/2
- $P(C \cap A) =$

  1/4 = 1/2 x 1/2

  C e A indep.
- $P(C \cap B) =$

  1/4 = 1/2 x 1/2

  C e B indep.
- $P(A \cap B) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2}$ .... $A\ e\ B\ ind.$
- $P(C \cap B\ \cap A) =$

$$\frac{1}{4} \neq \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}$$

Independência 2 a 2 não implica independência

Independência e acontecimentos mutuamente exclusivos

- Em geral dois acontecimentos que sejam mutuamente exclusivos e tenham probabilidade não nula não podem ser independentes

$$0 = P(A \cap B) = P(A)\,P(B)$$

implicaria que um deles tenha probabilidade nula

- Em geral considera-se que acontecimentos em experiências distintas são independentes (experiências independentes)

Sequências de experiências independentes

- Se uma experiência aleatória for composta por *n* experiências independentes e se A_k for um acontecimento que diga respeito à experiência *k*, é razoável admitir que os *n* acontecimentos são independentes

- Então:

$$P(A_1 \cap A_2 \cap A_3 \ldots \cap A_n) = \mathrm{P}(A_1)P(A_2)P(A_3) \ldots P(A_n)$$

Experiências de Bernoulli

- Uma **experiência de Bernoulli** consiste em realizar uma experiência e registar se um dado acontecimento se verifica (sucesso) ou não (falha)

Qual a **probabilidade de k sucessos em n ensaios independentes** ?

- Seja *p* a probabilidade de sucesso
  - E (1-*p*) a de falha
- A probabilidade de *k* sucessos e (*n*-*k*) falhas é:

$$p^k \, (1-p)^{n-k}$$

- *k* sucessos em *n* experiências podem ocorrer de C_k^n maneiras

- Então a probabilidade pedida é:

$$\boldsymbol{P_n(k) = C_k^n \;\; p^k \;\; (1-p)^{n-k}}$$

Lei Binomial

Visão frequencista e probabilidade condicional

- Do ponto de vista frequencista, considerando os eventos A e B e N experiências, podemos escrever:

- $P(A|B) \approx \frac{k_{A\ e\ B}/_{N}}{k_{B}/_{N}} = \frac{k_{A\ e\ B}}{k_B}$

- Onde $k_{A\ e\ B}$ é o número de ocorrência de "A e B"
  - Que também se pode denominar por frequência absoluta e representar por f_{AB}

Simulação

- Como fazer para ter P(A|B) ?

- Realizar N experiências

- Contar o número de ocorrências de AB

  Será fAB (frequência absoluta)
- Contar número de ocorrências de B

  fB

- $P \cong \dfrac{fAB/_N}{fB/_N} = \dfrac{fAB}{fB}$

Exemplo de simulação (Independência vs independência 2 a 2)

- Relembremos:

  2 lançamentos de moeda

  A: primeira é caras

  B: segunda é caras

  C: mesmo resultado em ambas

- $P(C \mid A \cap B)$

simulação

```
% P(C | A e B) =  P(C e A e B) / P(A e B)
N= 1e5;
m1= rand(1,N) < 0.5; % registo lançamentos da 1ª moeda
m2= rand(1,N) < 0.5; % registo lançamentos da 2ª moeda


ABC= (m1==m2) & (m1==1) & (m2==1);   %  C: iguais   A: primeira caras                    % B:
segunda caras
fABC=sum(ABC,1);

AB =   (m1==1) & (m2==1);    %  A: primeira caras  B: segunda caras

fAB=sum(AB,1);

p=fABC/fAB
P= 1 ….
```

Tópicos da aula

- Probabilidade Condicional

- 3 Ferramentas muito importantes
  - Regra da multiplicação
  - Teorema da Probabilidade total
  - Regra Bayes

- Aplicação da teoria frequencista a probabilidades condicionais

Não esquecer

- Independência de 2 eventos
- Independência de uma coleção de eventos
- Probabilidade condicional

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

- Regra da multiplicação:

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

- Probabilidade total:

$$P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$$

- Regra de Bayes

$$P(A_i|B) = \frac{P(A_i) \;\; P(B|A_i)}{P(B)}$$

Para aprender mais …

- Capítulos iniciais do Livro "Probabilidades e Processos Estocásticos", F. Vaz, Universidade de Aveiro
  - Disponíveis no Elearning da UC