Inteligência Artificial – ACH2016
Aula 15 – Tratamento de incerteza
Probabilidade

Norton Trevisan Roman
(norton@usp.br)

5 de maio de 2019

Incerteza

- Vimos como raciocinar com fatos absolutos

Incerteza

- Vimos como raciocinar com fatos absolutos
  - Usando lógica proposicional e de primeira ordem, com ontologias e taxonomias

Incerteza

- Vimos como raciocinar com fatos absolutos
  - Usando lógica proposicional e de primeira ordem, com ontologias e taxonomias
- Vimos também como raciocinar à luz de imprecisão

Incerteza

- Vimos como raciocinar com fatos absolutos
  - Usando lógica proposicional e de primeira ordem, com ontologias e taxonomias
- Vimos também como raciocinar à luz de imprecisão
  - Usando lógica nebulosa

Incerteza

- Vimos como raciocinar com fatos absolutos
  - Usando lógica proposicional e de primeira ordem, com ontologias e taxonomias
- Vimos também como raciocinar à luz de imprecisão
  - Usando lógica nebulosa
- Mas e quando não temos certeza de algo?

Incerteza

- Vimos como raciocinar com fatos absolutos
  - Usando lógica proposicional e de primeira ordem, com ontologias e taxonomias
- Vimos também como raciocinar à luz de imprecisão
  - Usando lógica nebulosa
- Mas e quando não temos certeza de algo?
  - Como quando o resultado de uma ação pode sofrer de influências aleatórias, ou quando os antecedentes de uma regra estão sujeitos a variação...

Incerteza

- Vimos como raciocinar com fatos absolutos
  - Usando lógica proposicional e de primeira ordem, com ontologias e taxonomias
- Vimos também como raciocinar à luz de imprecisão
  - Usando lógica nebulosa
- Mas e quando não temos certeza de algo?
  - Como quando o resultado de uma ação pode sofrer de influências aleatórias, ou quando os antecedentes de uma regra estão sujeitos a variação...
  - Nossa base de conhecimento no máximo fornece um grau de crença nas sentenças relevantes

Incerteza

- Como podemos tomar decisões racionais em meio à incerteza?

Fonte: http://www.structural-science.net/___Decision_theory.html

Incerteza

- Como podemos tomar decisões racionais em meio à incerteza?
- Apelando à **Teoria da Decisão**

Fonte: http://www.structural-science.net/___Decision_theory.html

Incerteza

- Como podemos tomar decisões racionais em meio à incerteza?
- Apelando à **Teoria da Decisão**
  - Uma ação é racional se for aquela com a maior utilidade esperada, tomando-se a média de todos seus possíveis resultados (**Princípio da Utilidade Máxima Esperada**)

Fonte: http://www.structural-science.net/___Decision_theory.html

Incerteza

- Como podemos tomar decisões racionais em meio à incerteza?
- Apelando à **Teoria da Decisão**
  - Uma ação é racional se for aquela com a maior utilidade esperada, tomando-se a média de todos seus possíveis resultados (**Princípio da Utilidade Máxima Esperada**)
  - Ou seja, a média das utilidades de cada resultado da ação, ponderada pela probabilidade de ocorrência de cada um desses resultados

Fonte: http://www.structural-science.net/___Decision_theory.html

Incerteza

Teoria da Decisão – Utilidade

- E o que significa “utilidade”?

Incerteza

Teoria da Decisão – Utilidade

- E o que significa "utilidade"?
  - A utilidade de um resultado depende das preferências de quem decide $\rightarrow$ os resultados que preferimos

Incerteza

Teoria da Decisão – Utilidade

- E o que significa "utilidade"?
  - A utilidade de um resultado depende das preferências de quem decide $\rightarrow$ os resultados que preferimos
- Usamos então a **Teoria da Utilidade** para raciocinar com preferências

Incerteza

Teoria da Decisão – Utilidade

- E o que significa "utilidade"?
  - A utilidade de um resultado depende das preferências de quem decide $\rightarrow$ os resultados que preferimos
- Usamos então a **Teoria da Utilidade** para raciocinar com preferências
  - Todo resultado tem um grau de utilidade para nós $\rightarrow$ preferimos o estado com maior utilidade

Incerteza

Teoria da Decisão – Utilidade

- E o que significa "utilidade"?
  - A utilidade de um resultado depende das preferências de quem decide $\rightarrow$ os resultados que preferimos
- Usamos então a **Teoria da Utilidade** para raciocinar com preferências
  - Todo resultado tem um grau de utilidade para nós $\rightarrow$ preferimos o estado com maior utilidade
    - Para isso, precisamos associar valores numéricos aos resultados possíveis

Incerteza

Teoria da Decisão – Utilidade

- E o que significa "utilidade"?
  - A utilidade de um resultado depende das preferências de quem decide → os resultados que preferimos
- Usamos então a **Teoria da Utilidade** para raciocinar com preferências
  - Todo resultado tem um grau de utilidade para nós → preferimos o estado com maior utilidade
    - Para isso, precisamos associar valores numéricos aos resultados possíveis
  - Precisamos também supor que dois resultados quaisquer podem ser comparados → podemos dizer qual preferimos

Incerteza

Teoria da Decisão – Probabilidade

- Então Teoria da Decisão $=$ Teoria da Probabilidade $+$ Teoria da Utilidade

Incerteza

Teoria da Decisão – Probabilidade

- Então Teoria da Decisão = Teoria da Probabilidade + Teoria da Utilidade
  - Dado nosso conhecimento atual, podemos analisar a probabilidade dos resultados de cada ação e então selecionar aquela com maior utilidade esperada

Incerteza

Teoria da Decisão – Probabilidade

- Então Teoria da Decisão = Teoria da Probabilidade + Teoria da Utilidade
  - Dado nosso conhecimento atual, podemos analisar a probabilidade dos resultados de cada ação e então selecionar aquela com maior utilidade esperada
- Não há, contudo, como racionalizar gosto ou preferência

Incerteza

Teoria da Decisão – Probabilidade

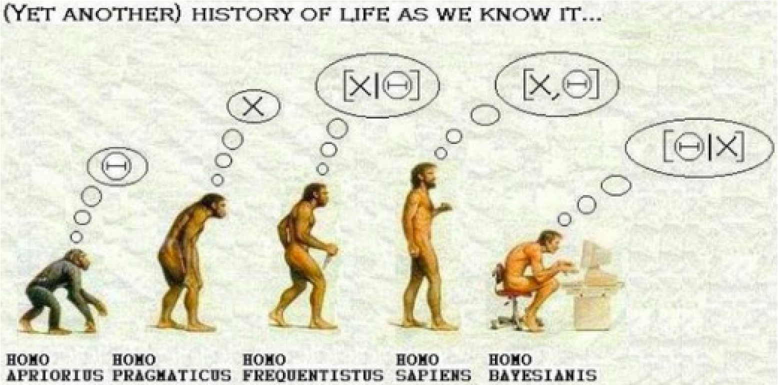
- Então Teoria da Decisão = Teoria da Probabilidade + Teoria da Utilidade
  - Dado nosso conhecimento atual, podemos analisar a probabilidade dos resultados de cada ação e então selecionar aquela com maior utilidade esperada
- Não há, contudo, como racionalizar gosto ou preferência
  - O grau de utilidade de algo

Incerteza

Teoria da Decisão – Probabilidade

- Então Teoria da Decisão $=$ Teoria da Probabilidade $+$ Teoria da Utilidade
  - Dado nosso conhecimento atual, podemos analisar a probabilidade dos resultados de cada ação e então selecionar aquela com maior utilidade esperada
- Não há, contudo, como racionalizar gosto ou preferência
  - O grau de utilidade de algo
  - Nos resta então o cálculo da probabilidade

Teoria da Probabilidade



Fonte: https://conversionxl.com/blog/bayesian-frequentist-ab-testing/

Probabilidade (Revisão)

- Associe a cada sentença da base um grau de certeza numérico (sua probabilidade), entre 0 e 1

Probabilidade (Revisão)

- Associe a cada sentença da base um grau de certeza numérico (sua probabilidade), entre 0 e 1
  - Probabilidade 0 $\rightarrow$ corresponde a uma crença inequívoca de que a sentença é falsa

Probabilidade (Revisão)

- Associe a cada sentença da base um grau de certeza numérico (sua probabilidade), entre 0 e 1
  - Probabilidade 0 $\rightarrow$ corresponde a uma crença inequívoca de que a sentença é falsa
  - Probabilidade 1 $\rightarrow$ corresponde a uma crença inequívoca de que a sentença é verdadeira

Probabilidade (Revisão)

- Associe a cada sentença da base um grau de certeza numérico (sua probabilidade), entre 0 e 1
  - Probabilidade 0 $\rightarrow$ corresponde a uma crença inequívoca de que a sentença é falsa
  - Probabilidade 1 $\rightarrow$ corresponde a uma crença inequívoca de que a sentença é verdadeira
  - Outras probabilidades correspondem a graus intermediários de crença na veracidade da sentença

Probabilidade (Revisão)

- Associe a cada sentença da base um grau de certeza numérico (sua probabilidade), entre 0 e 1
  - Probabilidade 0 $\rightarrow$ corresponde a uma crença inequívoca de que a sentença é falsa
  - Probabilidade 1 $\rightarrow$ corresponde a uma crença inequívoca de que a sentença é verdadeira
  - Outras probabilidades correspondem a graus intermediários de crença na veracidade da sentença
- Note que a sentença por si só ou é verdadeira ou falsa

Probabilidade (Revisão)

- Associe a cada sentença da base um grau de certeza numérico (sua probabilidade), entre 0 e 1
  - Probabilidade 0 $\rightarrow$ corresponde a uma crença inequívoca de que a sentença é falsa
  - Probabilidade 1 $\rightarrow$ corresponde a uma crença inequívoca de que a sentença é verdadeira
  - Outras probabilidades correspondem a graus intermediários de crença na veracidade da sentença
- Note que a sentença por si só ou é verdadeira ou falsa
  - Não cabe imprecisão – o que varia é nossa crença nisso

Probabilidade (Revisão)

Definindo valores

- Valores de probabilidade são obtidos a partir da análise de frequência de dados, ou a partir de regras estabelecidas

Probabilidade (Revisão)

Definindo valores

- Valores de probabilidade são obtidos a partir da análise de frequência de dados, ou a partir de regras estabelecidas
  - Ex: Podemos não saber o que aflige o paciente, mas cremos ter uma chance de 80% (0,8) de ser gripe

Probabilidade (Revisão)

Definindo valores

- Valores de probabilidade são obtidos a partir da análise de frequência de dados, ou a partir de regras estabelecidas
  - Ex: Podemos não saber o que aflige o paciente, mas cremos ter uma chance de 80% (0,8) de ser gripe

    (A cada 100 pacientes com o mesmo tipo de reclamação, teríamos esse problema em 80 deles)

Probabilidade (Revisão)

Definindo valores

- Valores de probabilidade são obtidos a partir da análise de frequência de dados, ou a partir de regras estabelecidas
  - Ex: Podemos não saber o que aflige o paciente, mas cremos ter uma chance de 80% (0,8) de ser gripe

    (A cada 100 pacientes com o mesmo tipo de reclamação, teríamos esse problema em 80 deles)
- Fortemente relacionadas a evidências

Probabilidade (Revisão)

Definindo valores

- Valores de probabilidade são obtidos a partir da análise de frequência de dados, ou a partir de regras estabelecidas
  - Ex: Podemos não saber o que aflige o paciente, mas cremos ter uma chance de 80% (0,8) de ser gripe

    (A cada 100 pacientes com o mesmo tipo de reclamação, teríamos esse problema em 80 deles)
- Fortemente relacionadas a evidências
  - Novas evidências levam a novas probabilidades

Probabilidade (Revisão)

Variáveis aleatórias

Probabilidade (Revisão)

Variáveis aleatórias

- São os elemento básico da probabilidade

Probabilidade (Revisão)

Variáveis aleatórias

- São os elemento básico da probabilidade
- Parte do mundo cujo status é inicialmente desconhecido

Probabilidade (Revisão)

Variáveis aleatórias

- São os elemento básico da probabilidade
- Parte do mundo cujo status é inicialmente desconhecido
  - Ex: gripe pode se referir ao fato de eu ter ou não gripe

Probabilidade (Revisão)

Variáveis aleatórias

- São os elemento básico da probabilidade
- Parte do mundo cujo status é inicialmente desconhecido
  - Ex: gripe pode se referir ao fato de eu ter ou não gripe
- Toda variável tem um domínio

Probabilidade (Revisão)

Variáveis aleatórias

- São os elemento básico da probabilidade
- Parte do mundo cujo status é inicialmente desconhecido
  - Ex: gripe pode se referir ao fato de eu ter ou não gripe
- Toda variável tem um domínio
  - Seu valor é obtido desse domínio

Probabilidade (Revisão)

Eventos Atômicos

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:
  - Variáveis: Gripe e Alergia

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:
  - Variáveis: Gripe e Alergia
  - Eventos atômicos:

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:
  - Variáveis: Gripe e Alergia
  - Eventos atômicos:
    - $Gripe = verdadeiro \wedge Alergia = falso$

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:
  - Variáveis: Gripe e Alergia
  - Eventos atômicos:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:
  - Variáveis: Gripe e Alergia
  - Eventos atômicos:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$
    - $Gripe = falso \wedge Alergia = falso$

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:
  - Variáveis: Gripe e Alergia
  - Eventos atômicos:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$
    - $Gripe = falso \wedge Alergia = falso$
    - $Gripe = falso \wedge Alergia = verdadeiro$

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:
  - Variáveis: Gripe e Alergia
  - Eventos atômicos:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$
    - $Gripe = falso \wedge Alergia = falso$
    - $Gripe = falso \wedge Alergia = verdadeiro$

Nossa convenção: Variáveis com inicial maiúscula e valores em minúscula

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:
  - Variáveis: Gripe e Alergia
  - Eventos atômicos:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$
    - $Gripe = falso \wedge Alergia = falso$
    - $Gripe = falso \wedge Alergia = verdadeiro$

Por conveniência, em vez de $Gripe = verdadeiro$ podemos escrever apenas *gripe*

Probabilidade (Revisão)

Eventos Atômicos

- Especificação completa do estado do mundo sobre o qual estamos incertos
  - Associação de valores a todas as variáveis que compõem esse mundo
- Exemplo:
  - Variáveis: Gripe e Alergia
  - Eventos atômicos:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$
    - $Gripe = falso \wedge Alergia = falso$
    - $Gripe = falso \wedge Alergia = verdadeiro$

E em vez de $Gripe = falso$ podemos escrever apenas $\neg gripe$

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Mutuamente exclusivos

Eventos Atômicos – Propriedades

- Mutuamente exclusivos
  - No máximo um pode valer

Eventos Atômicos – Propriedades

- Mutuamente exclusivos
  - No máximo um pode valer
  - Exemplo:

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Mutuamente exclusivos
  - No máximo um pode valer
  - Exemplo:
    - $Gripe = verdadeiro \wedge Alergia = falso$

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Mutuamente exclusivos
  - No máximo um pode valer
  - Exemplo:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Mutuamente exclusivos
  - No máximo um pode valer
  - Exemplo:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$
    - Ambos não podem acontecer ao mesmo tempo

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Mutuamente exclusivos
  - No máximo um pode valer
  - Exemplo:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$
    - Ambos não podem acontecer ao mesmo tempo
- Exaustivos

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Mutuamente exclusivos
  - No máximo um pode valer
  - Exemplo:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$
    - Ambos não podem acontecer ao mesmo tempo
- Exaustivos
  - Pelo menos um deve valer (cobrem todas as possibilidades)

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Mutuamente exclusivos
  - No máximo um pode valer
  - Exemplo:
    - $Gripe = verdadeiro \wedge Alergia = falso$
    - $Gripe = verdadeiro \wedge Alergia = verdadeiro$
    - Ambos não podem acontecer ao mesmo tempo
- Exaustivos
  - Pelo menos um deve valer (cobrem todas as possibilidades)
  - Um $\vee$ entre eles deve ser sempre verdadeiro

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Qualquer evento atômico particular acarreta a veracidade ou não de toda proposição

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Qualquer evento atômico particular acarreta a veracidade ou não de toda proposição
  - Não há proposição cuja veracidade não possa ser determinada por um evento atômico, pois ele cobre todas as variáveis

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Qualquer evento atômico particular acarreta a veracidade ou não de toda proposição
  - Não há proposição cuja veracidade não possa ser determinada por um evento atômico, pois ele cobre todas as variáveis
- Toda proposição é logicamente equivalente à disjunção de todos os eventos atômicos que acarretam sua veracidade

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Qualquer evento atômico particular acarreta a veracidade ou não de toda proposição
  - Não há proposição cuja veracidade não possa ser determinada por um evento atômico, pois ele cobre todas as variáveis
- Toda proposição é logicamente equivalente à disjunção de todos os eventos atômicos que acarretam sua veracidade
  - $\text{Gripe} \equiv (\text{Gripe} \wedge \text{Alergia}) \vee (\text{Gripe} \wedge \neg\text{Alergia})$

Probabilidade (Revisão)

Eventos Atômicos – Propriedades

- Qualquer evento atômico particular acarreta a veracidade ou não de toda proposição
  - Não há proposição cuja veracidade não possa ser determinada por um evento atômico, pois ele cobre todas as variáveis
- Toda proposição é logicamente equivalente à disjunção de todos os eventos atômicos que acarretam sua veracidade
  - $\text{Gripe} \equiv (\text{Gripe} \land \text{Alergia}) \lor (\text{Gripe} \land \neg\text{Alergia})$

    $(\text{Gripe} \land \text{Alergia} \models \text{Gripe})$ e $(\text{Gripe} \land \neg\text{Alergia} \models \text{Gripe})$

Probabilidade Prévia ou Incondicional

Probabilidade (Revisão)

Probabilidade Prévia ou Incondicional

- Crença antes da evidência ser obtida

Probabilidade (Revisão)

Probabilidade Prévia ou Incondicional

- Crença antes da evidência ser obtida
  - Grau de crença na ausência de qualquer outra informação

Probabilidade (Revisão)

Probabilidade Prévia ou Incondicional

- Crença antes da evidência ser obtida
  - Grau de crença na ausência de qualquer outra informação
  - Probabilidade inicial de que a hipótese h valha, antes de observados os dados

Probabilidade Prévia ou Incondicional

- Crença <u>antes</u> da evidência ser obtida
  - Grau de crença na ausência de qualquer outra informação
  - Probabilidade inicial de que a hipótese h valha, antes de observados os dados
    - Reflete qualquer conhecimento prévio que temos sobre a chance de h ser a hipótese correta

Probabilidade (Revisão)

Probabilidade Prévia ou Incondicional

- Crença antes da evidência ser obtida
  - Grau de crença na ausência de qualquer outra informação
  - Probabilidade inicial de que a hipótese h valha, antes de observados os dados
    - Reflete qualquer conhecimento prévio que temos sobre a chance de h ser a hipótese correta
  - Deve ser usada somente quando não houver outra informação disponível

Probabilidade (Revisão)

Probabilidade Prévia ou Incondicional

- Crença antes da evidência ser obtida
  - Grau de crença na ausência de qualquer outra informação
  - Probabilidade inicial de que a hipótese h valha, antes de observados os dados
    - Reflete qualquer conhecimento prévio que temos sobre a chance de h ser a hipótese correta
  - Deve ser usada somente quando não houver outra informação disponível
- Exemplo:

Probabilidade (Revisão)

Probabilidade Prévia ou Incondicional

- Crença antes da evidência ser obtida
  - Grau de crença na ausência de qualquer outra informação
  - Probabilidade inicial de que a hipótese *h* valha, antes de observados os dados
    - Reflete qualquer conhecimento prévio que temos sobre a chance de *h* ser a hipótese correta
  - Deve ser usada somente quando não houver outra informação disponível
- Exemplo:
  - Se de antemão creio que a probabilidade de ter gripe é 0,1, então P(Gripe=verdadeiro) = P(gripe) = 0,1

Probabilidade Prévia ou Incondicional

De onde vem essa crença?

Probabilidade Prévia ou Incondicional

De onde vem essa crença?

- Pode refletir nossa intuição

Probabilidade Prévia ou Incondicional

De onde vem essa crença?

- Pode refletir nossa intuição
- Pode advir de experiência passada

Probabilidade Prévia ou Incondicional

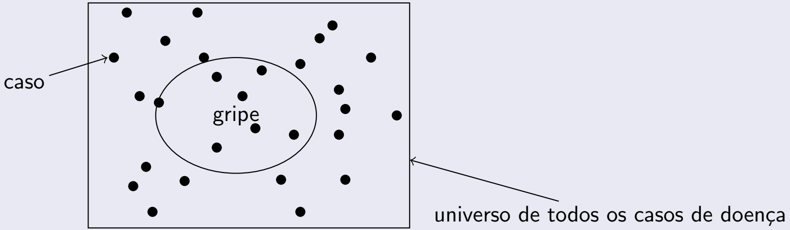
De onde vem essa crença?

- Pode refletir nossa intuição
- Pode advir de experiência passada
  - De todos os casos já vistos, 10% deles eram gripe

Probabilidade Prévia ou Incondicional

De onde vem essa crença?

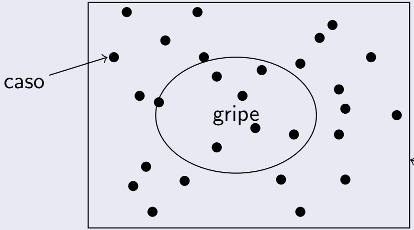
- Pode refletir nossa intuição
- Pode advir de experiência passada
  - De todos os casos já vistos, 10% deles eram gripe
- Diagramas de Venn:



Probabilidade Prévia ou Incondicional

De onde vem essa crença?

- Pode refletir nossa intuição
- Pode advir de experiência passada
  - De todos os casos já vistos, 10% deles eram gripe
- Diagramas de Venn:

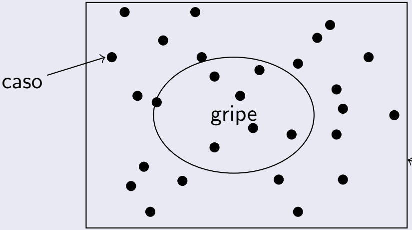


$$P(gripe) = \frac{\textit{Número de casos com gripe}}{\textit{Número total de casos de doença}}$$

Probabilidade Prévia ou Incondicional

De onde vem essa crença?

- Pode refletir nossa intuição
- Pode advir de experiência passada
  - De todos os casos já vistos, 10% deles eram gripe
- Diagramas de Venn:



Probabilidade de, ao retirarmos um caso do universo, ele pertencer a gripe

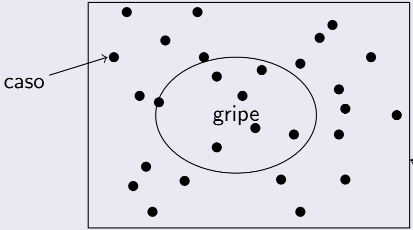
$$P(gripe) = \frac{\textit{Número de casos com gripe}}{\textit{Número total de casos de doença}}$$

universo de todos os casos de doença

Probabilidade Prévia ou Incondicional

De onde vem essa crença?

- Pode refletir nossa intuição
- Pode advir de experiência passada
  - De todos os casos já vistos, 10% deles eram gripe
- Diagramas de Venn:



Probabilidade de, ao retirarmos um caso do universo, ele pertencer a gripe

$$P(gripe) = \frac{\textit{Número de casos com gripe}}{\textit{Número total de casos de doença}}$$

universo de todos os casos de doença

Probabilidade Posterior ou Condicional

Probabilidade Posterior ou Condicional

- Crença <u>após</u> a evidência ser obtida

Probabilidade (Revisão)

Probabilidade Posterior ou Condicional

- Crença <u>após</u> a evidência ser obtida
  - Inclui a evidência explicitamente

Probabilidade (Revisão)

Probabilidade Posterior ou Condicional

- Crença após a evidência ser obtida
  - Inclui a evidência explicitamente
- Reflete a crença dada a nova evidência

Probabilidade (Revisão)

Probabilidade Posterior ou Condicional

- Crença após a evidência ser obtida
  - Inclui a evidência explicitamente
- Reflete a crença dada a nova evidência
  - P(variável = Valor | evidência)

Probabilidade (Revisão)

Probabilidade Posterior ou Condicional

- Crença <u>após</u> a evidência ser obtida
  - Inclui a evidência explicitamente
- Reflete a crença dada a nova evidência
  - P(variável = Valor | evidência)

    (Probabilidade da variável dado que conhecemos a evidência)

Probabilidade (Revisão)

Probabilidade Posterior ou Condicional

- Crença após a evidência ser obtida
  - Inclui a evidência explicitamente
- Reflete a crença dada a nova evidência
  - P(variável = Valor | evidência)

    (Probabilidade da variável dado que conhecemos a evidência)
  - Embora não implique causalidade, às vezes é útil pensar em termos de P(efeito | causa)

Probabilidade (Revisão)

Probabilidade Posterior ou Condicional

- Crença após a evidência ser obtida
  - Inclui a evidência explicitamente
- Reflete a crença dada a nova evidência
  - P(variável = Valor | evidência)

    (Probabilidade da variável dado que conhecemos a evidência)
  - Embora não implique causalidade, às vezes é útil pensar em termos de P(efeito | causa)
  - Ex: $P(\text{gripe}|\text{febre}) = 0{,}8$

Probabilidade (Revisão)

Probabilidade Posterior ou Condicional

- Crença após a evidência ser obtida
  - Inclui a evidência explicitamente
- Reflete a crença dada a nova evidência
  - P(variável = Valor | evidência)

    (Probabilidade da variável dado que conhecemos a evidência)
  - Embora não implique causalidade, às vezes é útil pensar em termos de P(efeito | causa)
  - Ex: $P(\text{gripe}|\text{febre}) = 0{,}8$
    - De todos os casos em que observamos febre (sem levar em conta qualquer outra coisa), 80% eram de pessoas com gripe

Probabilidade (Revisão)

Probabilidade Posterior ou Condicional

- Após coletada a evidência, a probabilidade prévia não se aplica mais

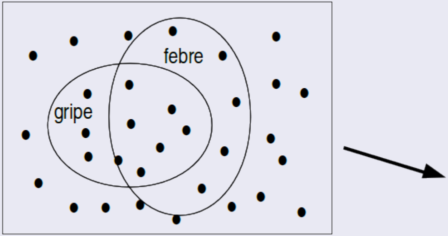
Fonte: `https://www.invespcro.com/blog/bayesian-vs-frequentist-a-b-testing-whats-the-difference/`

Probabilidade Posterior ou Condicional

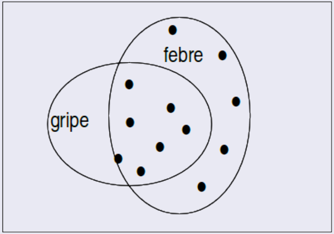
Como calcularmos a probabilidade posterior?

Como calcularmos a probabilidade posterior?

- Suponha o seguinte universo:



Uma vez que identificamos a existência de febre, podemos olhar apenas para esses casos – o universo reduz

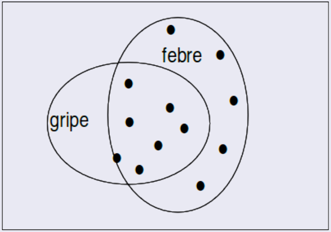


$$P(\text{febre}) = \frac{\text{Número de caso com febre}}{\text{Número total de casos de doença}}$$

$$P(\text{febre} \wedge \text{gripe}) = \frac{\text{Número de caso com febre e gripe}}{\text{Número total de casos de doença}}$$

Probabilidade Posterior ou Condicional

Como calcularmos a probabilidade posterior?



$$P(\text{gripe}|\text{febre}) = \frac{\text{Número de caso com gripe}}{\text{Número total de casos no novo universo}}$$

$$\Downarrow$$

$$P(\text{gripe}|\text{febre}) = \frac{\text{Número de caso com febre e gripe}}{\text{Número total de casos de febre}}$$

Número de casos com febre e gripe = $P(\text{febre} \wedge \text{gripe}) \times$ Número total de casos de doença

Número total de casos com febre = $P(\text{febre}) \times$ Número total de casos de doença

$$P(\text{gripe}|\text{febre}) = \frac{P(\text{febre} \wedge \text{gripe}) \times \text{Número total de casos de doença}}{P(\text{febre}) \times \text{Número total de casos de doença}} = \frac{P(\text{febre} \wedge \text{gripe})}{P(\text{febre})}$$

Probabilidade Posterior ou Condicional

Regra do produto

Probabilidade Posterior ou Condicional

Regra do produto

- Temos que:

$$P(a|b) = \frac{P(a \wedge b)}{P(b)} \text{ (sempre que } P(b) \neq 0)$$

Probabilidade Posterior ou Condicional

Regra do produto

- Temos que:

$$P(a|b) = \frac{P(a \wedge b)}{P(b)} \text{ (sempre que } P(b) \neq 0)$$

- Então (regra do produto)
  - $P(a \wedge b) = P(a|b)P(b)$ e $P(a \wedge b) = P(b|a)P(a)$

Probabilidade Posterior ou Condicional

Regra do produto

- Temos que:

$$P(a|b) = \frac{P(a \wedge b)}{P(b)} \text{ (sempre que } P(b) \neq 0)$$

- Então (regra do produto)
  - $P(a \wedge b) = P(a|b)P(b)$ e $P(a \wedge b) = P(b|a)P(a)$
  - Para que a e b sejam verdade, precisamos que b seja verdade, e que a seja verdade, dado que b é verdade

Probabilidade Posterior ou Condicional

Regra do produto

- Temos que:

$$P(a|b) = \frac{P(a \wedge b)}{P(b)} \text{ (sempre que } P(b) \neq 0)$$

- Então (regra do produto)
  - $P(a \wedge b) = P(a|b)P(b)$ e $P(a \wedge b) = P(b|a)P(a)$
  - Para que a e b sejam verdade, precisamos que b seja verdade, e que a seja verdade, dado que b é verdade

- Alternativamente, podemos denotar $P(a \wedge b)$ por $P(a, b)$

Probabilidade Posterior ou Condicional

Exemplo

- Baralho de 52 cartas

Probabilidade Posterior ou Condicional

Exemplo

- Baralho de 52 cartas
  - Retiramos uma carta

Probabilidade Posterior ou Condicional

Exemplo

- Baralho de 52 cartas
  - Retiramos uma carta
- P(ás de espadas) antes de olharmos a carta:

Probabilidade Posterior ou Condicional

Exemplo

- Baralho de 52 cartas
  - Retiramos uma carta
- P(ás de espadas) antes de olharmos a carta:

$$P(\text{ás de espada}) = \frac{\text{Número de ases de espada}}{\text{Número de cartas}} = \frac{1}{52}$$

Probabilidade Posterior ou Condicional

Exemplo

- Baralho de 52 cartas
  - Retiramos uma carta
- P(ás de espadas) antes de olharmos a carta:

$$P(\text{ás de espada}) = \frac{\text{Número de ases de espada}}{\text{Número de cartas}} = \frac{1}{52}$$

- P(ás de espadas) após olharmos a carta:

Probabilidade Posterior ou Condicional

Exemplo

- Baralho de 52 cartas
  - Retiramos uma carta
- P(ás de espadas) antes de olharmos a carta:

$$P(\text{ás de espada}) = \frac{\text{Número de ases de espada}}{\text{Número de cartas}} = \frac{1}{52}$$

- P(ás de espadas) após olharmos a carta:
  - Se era o ás de espadas: 1
  - Se não era: 0

Probabilidade Posterior ou Condicional

- E se houver mais informação disponível?

Probabilidade Posterior ou Condicional

- E se houver mais informação disponível?
  - $P(a \wedge b \wedge c) = P(a|b \wedge c)P(b \wedge c)$

Probabilidade Posterior ou Condicional

- E se houver mais informação disponível?
  - $P(a \wedge b \wedge c) = P(a|b \wedge c)P(b \wedge c)$
    - Podemos mudar a ordem de a, b e c

Probabilidade Posterior ou Condicional

- E se houver mais informação disponível?
  - $P(a \wedge b \wedge c) = P(a|b \wedge c)P(b \wedge c)$
    - Podemos mudar a ordem de a, b e c
  - Note que $P(a|b \wedge a) = 1$

Probabilidade Posterior ou Condicional

- E se houver mais informação disponível?
  - $P(a \wedge b \wedge c) = P(a|b \wedge c)P(b \wedge c)$
    - Podemos mudar a ordem de a, b e c
  - Note que $P(a|b \wedge a) = 1$
    - Pois supõe que a já ocorreu

Probabilidade Posterior ou Condicional

- E se houver mais informação disponível?
  - $P(a \wedge b \wedge c) = P(a|b \wedge c)P(b \wedge c)$
    - Podemos mudar a ordem de a, b e c
  - Note que $P(a|b \wedge a) = 1$
    - Pois supõe que a já ocorreu

Em geral (regra da cadeia)

$$\begin{aligned}
P(x_1, \ldots, x_n) &= P(x_1, \ldots, x_{n-1})P(x_n|x_1, \ldots, x_{n-1}) \\
&= P(x_1, \ldots, x_{n-2})P(x_{n-1}|x_1, \ldots, x_{n-2})P(x_n|x_1, \ldots, x_{n-1}) \\
&\vdots \\
&= \prod_{i=1}^{n} P(x_i|x_1, \ldots, x_{i-1})
\end{aligned}$$

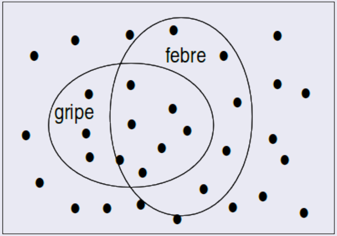
Probabilidade (Revisão)

Disjunção

Disjunção

Primeira aproximação:

$$P(\text{gripe} \vee \text{febre}) = \frac{\text{número de casos de gripe} + \text{número de casos de febre}}{\text{número total de casos}}$$



Problema: contamos 2 vezes a intersecção dos conjuntos (gripe $\wedge$ febre)

$P(\text{gripe} \vee \text{febre}) =$

$$\frac{\text{número de casos de gripe} + \text{número de casos de febre} - \text{número de casos de gripe e febre}}{\text{número total de casos}}$$

$$\boxed{P(\text{gripe} \vee \text{febre}) = P(\text{gripe}) + P(\text{febre}) - P(\text{gripe} \wedge \text{febre})}$$

Axiomas da Probabilidade

Axiomas de Kolmogorov

Axiomas da Probabilidade

Axiomas de Kolmogorov

1. Probabilidades estão entre 0 e 1

Axiomas da Probabilidade

Axiomas de Kolmogorov

1. Probabilidades estão entre 0 e 1
    - $0 \leq \mathsf{P(a)} \leq 1$

Axiomas da Probabilidade

Axiomas de Kolmogorov

1. Probabilidades estão entre 0 e 1
   - $0 \leq P(a) \leq 1$
2. Toda distribuição de probabilidade em uma única variável deve somar 1

Axiomas da Probabilidade

Axiomas de Kolmogorov

1. Probabilidades estão entre 0 e 1
   - $0 \leq \mathrm{P(a)} \leq 1$
2. Toda distribuição de probabilidade em uma única variável deve somar 1
   - $\sum_{i=1}^{n} P(D = d_i) = 1$

Axiomas da Probabilidade

Axiomas de Kolmogorov

1. Probabilidades estão entre 0 e 1
   - $0 \leq \mathrm{P(a)} \leq 1$
2. Toda distribuição de probabilidade em uma única variável deve somar 1
   - $\sum_{i=1}^{n} P(D = d_i) = 1$
   - A probabilidade de algum evento ocorrer é 1 (e de nenhum é 0)

Axiomas da Probabilidade

Axiomas de Kolmogorov

1. Probabilidades estão entre 0 e 1
   - $0 \leq \text{P(a)} \leq 1$
2. Toda distribuição de probabilidade em uma única variável deve somar 1
   - $\sum_{i=1}^{n} P(D = d_i) = 1$
   - A probabilidade de algum evento ocorrer é 1 (e de nenhum é 0)
3. A probabilidade da disjunção é dada por

Axiomas da Probabilidade

Axiomas de Kolmogorov

1. Probabilidades estão entre 0 e 1
   - $0 \leq \mathrm{P(a)} \leq 1$
2. Toda distribuição de probabilidade em uma única variável deve somar 1
   - $\displaystyle\sum_{i=1}^{n} P(D = d_i) = 1$
   - A probabilidade de algum evento ocorrer é 1 (e de nenhum é 0)
3. A probabilidade da disjunção é dada por
   - $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

Axiomas da Probabilidade

Axiomas de Kolmogorov

1. Probabilidades estão entre 0 e 1
   - $0 \leq \text{P(a)} \leq 1$
2. Toda distribuição de probabilidade em uma única variável deve somar 1
   - $\sum_{i=1}^{n} P(D = d_i) = 1$
   - A probabilidade de algum evento ocorrer é 1 (e de nenhum é 0)
3. A probabilidade da disjunção é dada por
   - $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$
   - Se dois eventos são mutuamente exclusivos, então $P(a \vee b) = P(a) + P(b)$

Probabilidade (Revisão)

Independência

Probabilidade (Revisão)

Independência

- Eventos independentes

Independência

- Eventos independentes
  - A ocorrência de um não afeta a ocorrência do outro

Probabilidade (Revisão)

Independência

- Eventos independentes
  - A ocorrência de um não afeta a ocorrência do outro
  - $P(a|b) = P(a)$, se a e b independentes

Probabilidade (Revisão)

Independência

- Eventos independentes
  - A ocorrência de um não afeta a ocorrência do outro
  - $P(a|b) = P(a)$, se a e b independentes
  - $P(b|a) = P(b)$

Probabilidade (Revisão)

Independência

- Eventos independentes
  - A ocorrência de um não afeta a ocorrência do outro
  - $P(a|b) = P(a)$, se a e b independentes
  - $P(b|a) = P(b)$
  - $P(a, b) = P(b|a)P(a) = P(b)P(a)$

Probabilidade (Revisão)

Independência

- Eventos independentes
  - A ocorrência de um não afeta a ocorrência do outro
  - $P(a|b) = P(a)$, se a e b independentes
  - $P(b|a) = P(b)$
  - $P(a, b) = P(b|a)P(a) = P(b)P(a)$
- Independência condicional

Probabilidade (Revisão)

Independência

- Eventos independentes
  - A ocorrência de um não afeta a ocorrência do outro
  - $P(a|b) = P(a)$, se a e b independentes
  - $P(b|a) = P(b)$
  - $P(a, b) = P(b|a)P(a) = P(b)P(a)$
- Independência condicional
  - Ocorre quando duas variáveis possuem a mesma causa, porém não afetam uma a outra diretamente

Probabilidade (Revisão)

Independência

- Eventos independentes
  - A ocorrência de um não afeta a ocorrência do outro
  - $P(a|b) = P(a)$, se a e b independentes
  - $P(b|a) = P(b)$
  - $P(a, b) = P(b|a)P(a) = P(b)P(a)$
- Independência condicional
  - Ocorre quando duas variáveis possuem a mesma causa, porém não afetam uma a outra diretamente
  - $P(x, y|z) = P(x|z)P(y|z)$

Probabilidade (Revisão)

Independência

- Eventos independentes
  - A ocorrência de um não afeta a ocorrência do outro
  - $P(a|b) = P(a)$, se a e b independentes
  - $P(b|a) = P(b)$
  - $P(a, b) = P(b|a)P(a) = P(b)P(a)$
- Independência condicional
  - Ocorre quando duas variáveis possuem a mesma causa, porém não afetam uma a outra diretamente
  - $P(x, y|z) = P(x|z)P(y|z)$
    - x e y são condicionalmente independentes, dada z

Probabilidade (Revisão)

Independência Condicional

- x é condicionalmente independente de y dada z se a distribuição de probabilidade que governa x é independente de y dado um valor para z:

Probabilidade (Revisão)

Independência Condicional

- x é condicionalmente independente de y dada z se a distribuição de probabilidade que governa x é independente de y dado um valor para z:
  - $P(X = x_i | Y = y_j, Z = z_k) = P(X = x_i | Z = z_k)$

Probabilidade (Revisão)

Independência Condicional

- x é condicionalmente independente de y dada z se a distribuição de probabilidade que governa x é independente de y dado um valor para z:
  - $P(X = x_i | Y = y_j, Z = z_k) = P(X = x_i | Z = z_k)$
- Também vale para conjuntos de variáveis:

Probabilidade (Revisão)

Independência Condicional

- x é condicionalmente independente de y dada z se a distribuição de probabilidade que governa x é independente de y dado um valor para z:
  - $P(X = x_i | Y = y_j, Z = z_k) = P(X = x_i | Z = z_k)$
- Também vale para conjuntos de variáveis:
  - O conjunto $x_1, \dots, x_l$ é condicionalmente independente do conjunto $y_1, \dots, y_m$, dado o conjunto $z_1, \dots, z_n$ se

    $$P(x_1, \dots, x_l | y_1, \dots, y_m, z_1, \dots, z_n) = P(x_1, \dots, x_l | z_1, \dots, z_n)$$

Probabilidade (Revisão)

Independência Condicional

- x é condicionalmente independente de y dada z se a distribuição de probabilidade que governa x é independente de y dado um valor para z:
  - $P(X = x_i | Y = y_j, Z = z_k) = P(X = x_i | Z = z_k)$
- Também vale para conjuntos de variáveis:
  - O conjunto $x_1, \dots, x_l$ é condicionalmente independente do conjunto $y_1, \dots, y_m$, dado o conjunto $z_1, \dots, z_n$ se

    $$P(x_1, \dots, x_l | y_1, \dots, y_m, z_1, \dots, z_n) = P(x_1, \dots, x_l | z_1, \dots, z_n)$$
  - Ao sabermos os valores de z_i não precisamos mais dos y_i para saber x_i

Independência Condicional

Exemplo

- Suponha que um paciente tenha dor de dente, foi ao dentista e este detectou algo com a broca

Independência Condicional

Exemplo

- Suponha que um paciente tenha dor de dente, foi ao dentista e este detectou algo com a broca
  - A detecção e a dor deveriam ser independentes. Contudo, se a broca prende em algo, deve ser uma cárie, e isso provavelmente causa a dor

Independência Condicional

Exemplo

- Suponha que um paciente tenha dor de dente, foi ao dentista e este detectou algo com a broca
  - A detecção e a dor deveriam ser independentes. Contudo, se a broca prende em algo, deve ser uma cárie, e isso provavelmente causa a dor
  - Causa comum a ambas

Independência Condicional

Exemplo

- Suponha que um paciente tenha dor de dente, foi ao dentista e este detectou algo com a broca
    - A detecção e a dor deveriam ser independentes. Contudo, se a broca prende em algo, deve ser uma cárie, e isso provavelmente causa a dor
    - Causa comum a ambas
- $P(Dor, Broca|Cárie) = (Broca|Dor, Cárie)P(Dor|Cárie)$

Independência Condicional

Exemplo

- Suponha que um paciente tenha dor de dente, foi ao dentista e este detectou algo com a broca
  - A detecção e a dor deveriam ser independentes. Contudo, se a broca prende em algo, deve ser uma cárie, e isso provavelmente causa a dor
  - Causa comum a ambas
- $P(Dor, Broca|Cárie) = (Broca|Dor, Cárie)P(Dor|Cárie)$
  - Se eu sei que tenho uma cárie, a probabilidade da broca detectá-la não depende do fato de eu ter dor

Independência Condicional

Exemplo

- Suponha que um paciente tenha dor de dente, foi ao dentista e este detectou algo com a broca
  - A detecção e a dor deveriam ser independentes. Contudo, se a broca prende em algo, deve ser uma cárie, e isso provavelmente causa a dor
  - Causa comum a ambas
- $P(Dor, Broca|Cárie) = (Broca|Dor, Cárie)P(Dor|Cárie)$
  - Se eu sei que tenho uma cárie, a probabilidade da broca detectá-la não depende do fato de eu ter dor
    - $P(broca|dor, cárie) = P(broca|cárie)$

Independência Condicional

Exemplo

- Suponha que um paciente tenha dor de dente, foi ao dentista e este detectou algo com a broca
  - A detecção e a dor deveriam ser independentes. Contudo, se a broca prende em algo, deve ser uma cárie, e isso provavelmente causa a dor
  - Causa comum a ambas
- $P(Dor,Broca|Cárie) = (Broca|Dor,Cárie)P(Dor|Cárie)$
  - Se eu sei que tenho uma cárie, a probabilidade da broca detectá-la não depende do fato de eu ter dor
    - $P(broca|dor,cárie) = P(broca|cárie)$
    - $P(broca|dor,\neg cárie) = P(broca|\neg cárie)$

Independência Condicional

Exemplo

- Suponha *dor* e *broca* condicionalmente independentes dada *cárie*:

$$\begin{aligned} P(Dor, Broca, Cárie) &= P(Dor, Broca|Cárie)P(Cárie) \\ &= P(Dor|Cárie)P(Broca|Cárie)P(Cárie) \end{aligned}$$

Independência Condicional

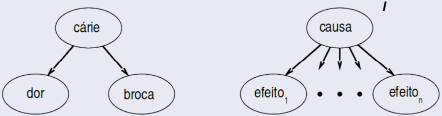
Exemplo

- Suponha *dor* e *broca* condicionalmente independentes dada *cárie*:

$$
\begin{aligned}
P(Dor, Broca, Cárie) &= P(Dor, Broca|Cárie)P(Cárie) \\
&= P(Dor|Cárie)P(Broca|Cárie)P(Cárie)
\end{aligned}
$$

Em geral

$$P(causa, efeito_1, \ldots, efeito_n) = P(causa) \prod_i P(efeito_i|causa)$$



Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Probabilidade

Distribuição de Probabilidade

Distribuição de Variáveis Discretas

Distribuição de Variáveis Discretas

- Considere os 4 possíveis valores para a variável aleatória clima:

Distribuição de Probabilidade

Distribuição de Variáveis Discretas

- Considere os 4 possíveis valores para a variável aleatória clima:
  - $P(\text{Clima} = \text{ensolarado}) = 0.7$

Distribuição de Probabilidade

Distribuição de Variáveis Discretas

- Considere os 4 possíveis valores para a variável aleatória clima:
  - $P(\text{Clima} = \text{ensolarado}) = 0.7$
  - $P(\text{Clima} = \text{chuvoso}) = 0.2$

Distribuição de Probabilidade

Distribuição de Variáveis Discretas

- Considere os 4 possíveis valores para a variável aleatória clima:
  - $P(\text{Clima} = \text{ensolarado}) = 0.7$
  - $P(\text{Clima} = \text{chuvoso}) = 0.2$
  - $P(\text{Clima} = \text{nublado}) = 0.08$

Distribuição de Probabilidade

Distribuição de Variáveis Discretas

- Considere os 4 possíveis valores para a variável aleatória clima:
  - $P(\text{Clima} = \text{ensolarado}) = 0.7$
  - $P(\text{Clima} = \text{chuvoso}) = 0.2$
  - $P(\text{Clima} = \text{nublado}) = 0.08$
  - $P(\text{Clima} = \text{nevando}) = 0.02$

Distribuição de Probabilidade

Distribuição de Variáveis Discretas

- Considere os 4 possíveis valores para a variável aleatória clima:
  - P(Clima = ensolarado) = 0.7
  - P(Clima = chuvoso) = 0.2
  - P(Clima = nublado) = 0.08
  - P(Clima = nevando) = 0.02

Note que a soma de todas as possibilidades é 1

Distribuição de Probabilidade

Distribuição de Variáveis Discretas

- Considere os 4 possíveis valores para a variável aleatória clima:
  - $P(\text{Clima} = \text{ensolarado}) = 0.7$
  - $P(\text{Clima} = \text{chuvoso}) = 0.2$
  - $P(\text{Clima} = \text{nublado}) = 0.08$
  - $P(\text{Clima} = \text{nevando}) = 0.02$

  Note que a soma de todas as possibilidades é 1
- $\mathbf{P}(\text{Clima}) = <0.7, 0.2, 0.08, 0.02>$

Distribuição de Probabilidade

Distribuição de Variáveis Discretas

- Considere os 4 possíveis valores para a variável aleatória clima:
  - P(Clima = ensolarado) = 0.7
  - P(Clima = chuvoso) = 0.2
  - P(Clima = nublado) = 0.08
  - P(Clima = nevando) = 0.02

  Note que a soma de todas as possibilidades é 1

- **P**(Clima) = <0.7, 0.2, 0.08, 0.02>
  - Arranjo de valores para as probabilidades de cada estado individual da variável → **Distribuição de probabilidade**

Distribuição de Probabilidade

Caso Discreto: Definição Formal

- A distribuição de probabilidade **P** de uma variável discreta X é a função que dá a probabilidade $P(X = x_i)$ de X assumir o valor x_i, para todo x_i ($\mathbf{P}(X)$ dá os valores de $P(X = x_i)$, $\forall x_i$), satisfazendo as seguintes condições

Distribuição de Probabilidade

Caso Discreto: Definição Formal

- A distribuição de probabilidade **P** de uma variável discreta X é a função que dá a probabilidade $P(X = x_i)$ de X assumir o valor x_i, para todo x_i ($\mathbf{P}(X)$ dá os valores de $P(X = x_i)$, $\forall x_i$), satisfazendo as seguintes condições
  - $0 \leq P(X = x_i) \leq 1$

Distribuição de Probabilidade

Caso Discreto: Definição Formal

- A distribuição de probabilidade **P** de uma variável discreta X é a função que dá a probabilidade $P(X = x_i)$ de X assumir o valor x_i, para todo x_i ($\mathbf{P}(X)$ dá os valores de $P(X = x_i)$, $\forall x_i$), satisfazendo as seguintes condições
  - $0 \leq P(X = x_i) \leq 1$
  - $\displaystyle\sum_{x_i} P(X = x_i) = 1$

Distribuição de Probabilidade

Distribuição de Variáveis Contínuas

Distribuição de Probabilidade

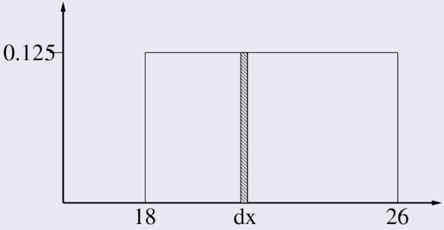
Distribuição de Variáveis Contínuas

- Chamada **função de densidade de probabilidade**

Distribuição de Probabilidade

Distribuição de Variáveis Contínuas

- Chamada **função de densidade de probabilidade**
- Ex: $\mathbf{P}(x) = U[18, 26](x)$ (Uniforme entre 18 e 26)



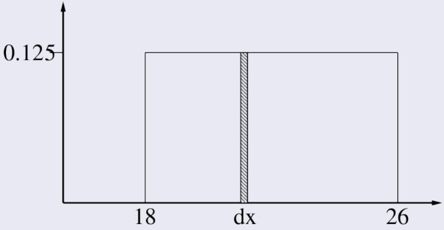
Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Probabilidade

Distribuição de Variáveis Contínuas

- Chamada **função de densidade de probabilidade**
- Ex: $\mathbf{P}(x) = U[18, 26](x)$ (Uniforme entre 18 e 26)

$$\int_{18}^{26} P(x)dx = 1$$



Fonte: Slides de AIMA. Russell & Norvig.

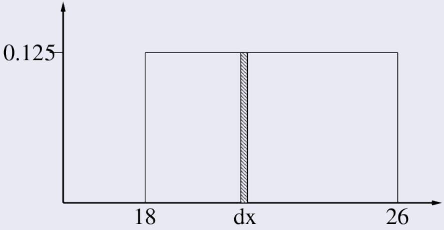
Distribuição de Probabilidade

Distribuição de Variáveis Contínuas

- Chamada **função de densidade de probabilidade**
- Ex: $\mathbf{P}(x) = U[18, 26](x)$ (Uniforme entre 18 e 26)

$$\int_{18}^{26} P(x)dx = 1$$

$$C \int_{18}^{26} dx = 1$$



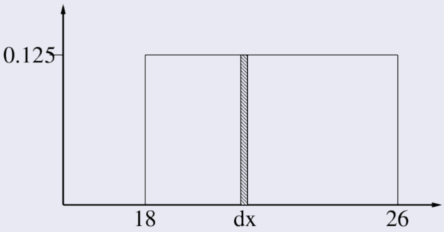
Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Probabilidade

Distribuição de Variáveis Contínuas

- Chamada **função de densidade de probabilidade**
- Ex: $\mathbf{P}(x) = U[18, 26](x)$ (Uniforme entre 18 e 26)

$$\begin{aligned}\int_{18}^{26} P(x)dx &= 1\\ C\int_{18}^{26} dx &= 1\\ C(26-18) &= 1\end{aligned}$$



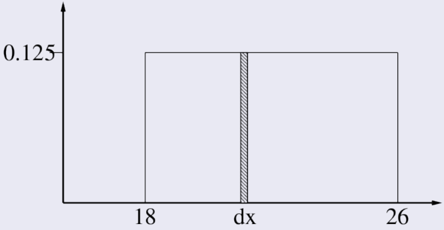
Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Probabilidade

Distribuição de Variáveis Contínuas

- Chamada **função de densidade de probabilidade**
- Ex: $\mathbf{P}(x) = U[18, 26](x)$ (Uniforme entre 18 e 26)

$$\begin{aligned}
\int_{18}^{26} P(x)dx &= 1 \\
C\int_{18}^{26} dx &= 1 \\
C(26-18) &= 1 \\
8C &= 1
\end{aligned}$$



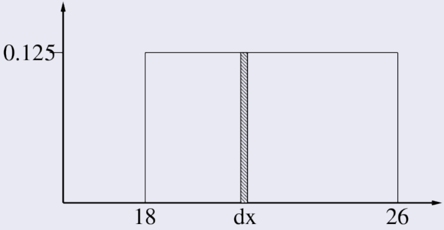
Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Probabilidade

Distribuição de Variáveis Contínuas

- Chamada **função de densidade de probabilidade**
- Ex: $\mathbf{P}(x) = U[18, 26](x)$ (Uniforme entre 18 e 26)

$$\begin{aligned}
\int_{18}^{26} P(x)dx &= 1 \\
C\int_{18}^{26} dx &= 1 \\
C(26-18) &= 1 \\
8C &= 1 \\
C &= 0,125
\end{aligned}$$



Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Probabilidade

Caso Contínuo: Definição Formal

- A distribuição de probabilidade de uma variável contínua X é uma função que satisfaz as seguintes propriedades

Distribuição de Probabilidade

Caso Contínuo: Definição Formal

- A distribuição de probabilidade de uma variável contínua X é uma função que satisfaz as seguintes propriedades
  - A probabilidade do valor de X estar entre dois pontos a e b é
    $$P(a \leq x \leq b) = \int_a^b f(x)dx$$

Distribuição de Probabilidade

Caso Contínuo: Definição Formal

- A distribuição de probabilidade de uma variável contínua X é uma função que satisfaz as seguintes propriedades
  - A probabilidade do valor de X estar entre dois pontos a e b é
    $$P(a \leq x \leq b) = \int_a^b f(x)dx$$
  - $P(x) \geq 0, \forall x \in \mathbb{R}$

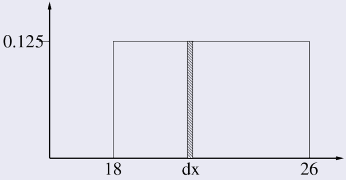
Distribuição de Probabilidade

Caso Contínuo: Definição Formal

- A distribuição de probabilidade de uma variável contínua X é uma função que satisfaz as seguintes propriedades
  - A probabilidade do valor de X estar entre dois pontos a e b é $P(a \leq x \leq b) = \int_a^b f(x)dx$
  - $P(x) \geq 0, \forall x \in \mathbb{R}$
  - A a área total da curva sobre todos os valores possíveis de x é 1 $\Rightarrow \int_{-\infty}^{\infty} f(x)dx = 1$

Distribuição de Variáveis Contínuas

Possuem interpretação diferente...

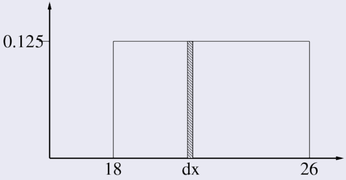


Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Variáveis Contínuas

Possuem interpretação diferente...

- $P(X = 20,5) = \int_{20,5}^{20,5} P(x)dx = 0$

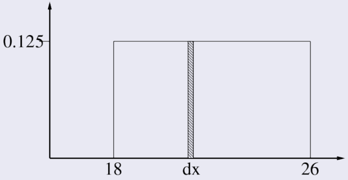


Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Variáveis Contínuas

Possuem interpretação diferente...

- $P(X = 20,5) = \int_{20,5}^{20,5} P(x)dx = 0$
  - Cada variável toma um intervalo infinito de valores

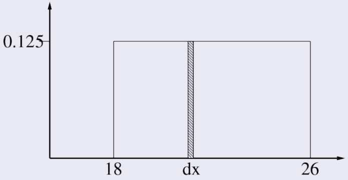


Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Variáveis Contínuas

Possuem interpretação diferente...

- $P(X = 20,5) = \int_{20,5}^{20,5} P(x)dx = 0$
  - Cada variável toma um intervalo infinito de valores
  - $X = 20,5 \Rightarrow$ precisão infinita

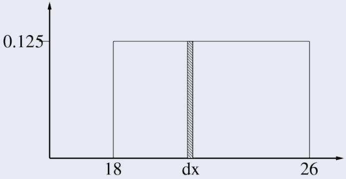


Fonte: Slides de AIMA. Russell & Norvig.

Distribuição de Variáveis Contínuas

Possuem interpretação diferente...

- $P(X = 20,5) = \int_{20,5}^{20,5} P(x)dx = 0$
  - Cada variável toma um intervalo infinito de valores
  - $X = 20,5 \Rightarrow$ precisão infinita



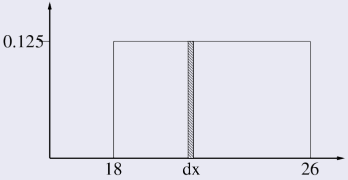
Fonte: Slides de AIMA. Russell & Norvig.

- Pensamos então em P(X) em torno de x

Distribuição de Variáveis Contínuas

Possuem interpretação diferente...

- $P(X = 20,5) = \int_{20,5}^{20,5} P(x)dx = 0$
  - Cada variável toma um intervalo infinito de valores
  - $X = 20,5 \Rightarrow$ precisão infinita



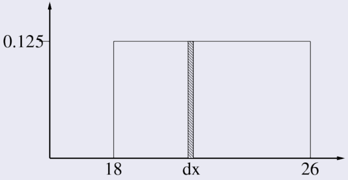
Fonte: Slides de AIMA. Russell & Norvig.

- Pensamos então em P(X) em torno de x
  - $\lim_{a \to b} P(a \leq x \leq b) = \lim_{a \to b} \int_a^b P(x)dx$

Distribuição de Variáveis Contínuas

Possuem interpretação diferente...

- $P(X = 20,5) = \int_{20,5}^{20,5} P(x)dx = 0$
  - Cada variável toma um intervalo infinito de valores
  - $X = 20,5 \Rightarrow$ precisão infinita

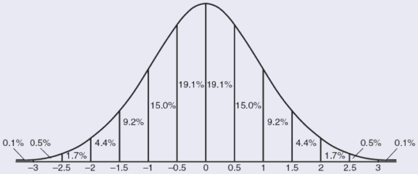


Fonte: Slides de AIMA. Russell & Norvig.

- Pensamos então em P(X) em torno de x
  - $\lim_{a \to b} P(a \leq x \leq b) = \lim_{a \to b} \int_a^b P(x)dx$
  - E $P(X = 20,5)$ passa a significar essa densidade de probabilidade em torno de 20,5

Distribuição de Variáveis Contínuas

Outras distribuições: Gaussiana ou Normal

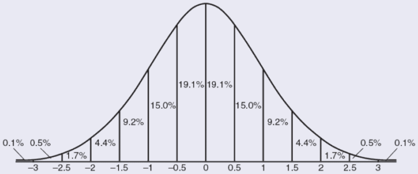


Fonte: https://coloneltedcampbell.blog/2016/02/13/the-800-pound-gorillas/normal67/

Distribuição de Variáveis Contínuas

Outras distribuições: Gaussiana ou Normal

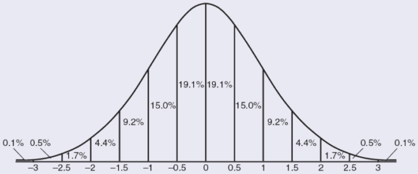
- $\mathbf{P}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



Fonte: https://coloneltedcampbell.blog/2016/02/13/the-800-pound-gorillas/normal67/

Outras distribuições: Gaussiana ou Normal

- $\mathbf{P}(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- Onde
  - μ = valor médio
  - σ = desvio padrão (e portanto σ^2 a variância)

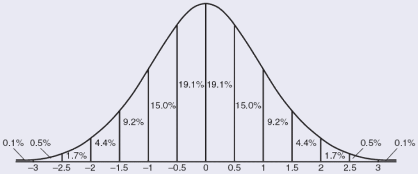


Fonte: `https://coloneltedcampbell.blog/2016/02/13/the-800-pound-gorillas/normal67/`

Distribuição de Variáveis Contínuas

Outras distribuições: Gaussiana ou Normal

- $\mathbf{P}(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- Onde
  - μ = valor médio
  - σ = desvio padrão (e portanto σ^2 a variância)



Fonte: https://coloneltedcampbell.blog/2016/02/13/the-800-pound-gorillas/normal67/

Para uma lista bastante elaborada, consulte
http://en.wikipedia.org/wiki/List_of_probability_distributions

Distribuição de Probabilidade

Distribuição Conjunta de Probabilidade

- E se tivermos mais de uma variável?

Distribuição Conjunta de Probabilidade

- E se tivermos mais de uma variável?
  - Como saber a probabilidade de termos gripe num dia chuvoso, por exemplo?

Distribuição de Probabilidade

Distribuição Conjunta de Probabilidade

- E se tivermos mais de uma variável?
  - Como saber a probabilidade de termos gripe num dia chuvoso, por exemplo?
  - Variáveis: Clima e (ter) Gripe

Distribuição de Probabilidade

Distribuição Conjunta de Probabilidade

- E se tivermos mais de uma variável?
  - Como saber a probabilidade de termos gripe num dia chuvoso, por exemplo?
  - Variáveis: Clima e (ter) Gripe
- Temos uma **distribuição conjunta de probabilidade**

Distribuição de Probabilidade

Distribuição Conjunta de Probabilidade

- E se tivermos mais de uma variável?
  - Como saber a probabilidade de termos gripe num dia chuvoso, por exemplo?
  - Variáveis: Clima e (ter) Gripe
- Temos uma **distribuição conjunta de probabilidade**
  - Com todas as combinações possíveis de valores

Distribuição de Probabilidade

Distribuição Conjunta de Probabilidade

- E se tivermos mais de uma variável?
  - Como saber a probabilidade de termos gripe num dia chuvoso, por exemplo?
  - Variáveis: Clima e (ter) Gripe
- Temos uma **distribuição conjunta de probabilidade**
  - Com todas as combinações possíveis de valores
- Vimos a distribuição de uma única variável:
  - $\mathbf{P}(\text{Clima}) = <0{,}7,\ 0{,}2,\ 0{,}08,\ 0{,}02>$

Distribuição de Probabilidade

Distribuição Conjunta de Probabilidade

- **P**(Clima, Gripe) – distribuição conjunta

Distribuição Conjunta de Probabilidade

- **P**(Clima, Gripe) – distribuição conjunta
  - Denota todas as combinações de valores das variáveis

	Clima			
Gripe	Sol	chuva	nublado	neve
sim	0,144	0,02	0,016	0,02
não	0,576	0,08	0,064	0,08

Distribuição Conjunta de Probabilidade

- **P**(Clima, Gripe) – distribuição conjunta
  - Denota todas as combinações de valores das variáveis

	Clima			
Gripe	Sol	chuva	nublado	neve
sim	0,144	0,02	0,016	0,02
não	0,576	0,08	0,064	0,08

- **Distribuição conjunta completa**:

Distribuição de Probabilidade

Distribuição Conjunta de Probabilidade

- **P**(Clima, Gripe) – distribuição conjunta
  - Denota todas as combinações de valores das variáveis

	Clima			
Gripe	Sol	chuva	nublado	neve
sim	0,144	0,02	0,016	0,02
não	0,576	0,08	0,064	0,08

- **Distribuição conjunta completa**:
  - Envolve o conjunto completo de variáveis

Distribuição de Probabilidade

Distribuição Conjunta de Probabilidade

- **P**(Clima, Gripe) – distribuição conjunta
  - Denota todas as combinações de valores das variáveis

	Clima			
Gripe	Sol	chuva	nublado	neve
sim	0,144	0,02	0,016	0,02
não	0,576	0,08	0,064	0,08

- **Distribuição conjunta completa**:
  - Envolve o conjunto completo de variáveis
    - Todas as variáveis do domínio

Distribuição de Probabilidade

Distribuição Conjunta de Probabilidade

- $\mathbf{P}$(Clima, Gripe) – distribuição conjunta
  - Denota todas as combinações de valores das variáveis

	Clima			
Gripe	Sol	chuva	nublado	neve
sim	0,144	0,02	0,016	0,02
não	0,576	0,08	0,064	0,08

- **Distribuição conjunta completa**:
  - Envolve o conjunto completo de variáveis
    - Todas as variáveis do domínio
    - Especifica a probabilidade de cada evento atômico

Distribuição Conjunta de Probabilidade

- **P**(Clima, Gripe) – distribuição conjunta
  - Denota todas as combinações de valores das variáveis

	Clima			
Gripe	Sol	chuva	nublado	neve
sim	0,144	0,02	0,016	0,02
não	0,576	0,08	0,064	0,08

- **Distribuição conjunta completa**:
  - Envolve o conjunto completo de variáveis
    - Todas as variáveis do domínio
    - Especifica a probabilidade de cada evento atômico
  - **P**(Gripe, Alergia, Clima) $\rightarrow$ tabela $2 \times 2 \times 4$

Distribuição Conjunta de Probabilidade

Também denotam probabilidades condicionais

Distribuição Conjunta de Probabilidade

Também denotam probabilidades condicionais

- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$

Distribuição Conjunta de Probabilidade

Também denotam probabilidades condicionais

- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$
- $\mathbf{P}(X|Y)$ dá os valores de $P(X{=}x_i|Y{=}y_j)$ para todo i,j possível

Distribuição Conjunta de Probabilidade

Também denotam probabilidades condicionais

- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$
- $\mathbf{P}(X|Y)$ dá os valores de $P(X{=}x_i|Y{=}y_j)$ para todo i,j possível
- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$ é um conjunto de equações

Distribuição Conjunta de Probabilidade

Também denotam probabilidades condicionais

- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$
- $\mathbf{P}(X|Y)$ dá os valores de $P(X{=}x_i|Y{=}y_j)$ para todo i,j possível
- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$ é um conjunto de equações
  - $P(X{=}x_1 \wedge Y{=}y_1) = P(X{=}x_1|Y{=}y_1)P(Y{=}y_1)$

Distribuição Conjunta de Probabilidade

Também denotam probabilidades condicionais

- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$
- $\mathbf{P}(X|Y)$ dá os valores de $P(X{=}x_i|Y{=}y_j)$ para todo i,j possível
- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$ é um conjunto de equações
  - $P(X{=}x_1 \wedge Y{=}y_1) = P(X{=}x_1|Y{=}y_1)P(Y{=}y_1)$
  - $P(X{=}x_1 \wedge Y{=}y_2) = P(X{=}x_1|Y{=}y_2)P(Y{=}y_2)$

Distribuição Conjunta de Probabilidade

Também denotam probabilidades condicionais

- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$
- $\mathbf{P}(X|Y)$ dá os valores de $P(X=x_i|Y=y_j)$ para todo i,j possível
- $\mathbf{P}(X,Y) = \mathbf{P}(X|Y)\mathbf{P}(Y)$ é um conjunto de equações
  - $P(X=x_1 \wedge Y=y_1) = P(X=x_1|Y=y_1)P(Y=y_1)$
  - $P(X=x_1 \wedge Y=y_2) = P(X=x_1|Y=y_2)P(Y=y_2)$
  - ...

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Variáveis:

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Variáveis:
  - Cárie: Cárie $= \{$sim, não$\}$

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Variáveis:
  - Cárie: $Cárie = \{sim, não\}$
  - Dor de dente: $Dor = \{sim, não\}$

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Variáveis:
  - Cárie: Cárie = {sim, não}
  - Dor de dente: Dor = {sim, não}
  - Broca prendendo no dente: Broca = {sim, não}

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Variáveis:
  - Cárie: Cárie = {sim, não}
  - Dor de dente: Dor = {sim, não}
  - Broca prendendo no dente: Broca = {sim, não}
- Distribuição conjunta:

	dor		¬dor	
	prende	¬prende	prende	¬prende
cárie	0,108	0,012	0,072	0,008
¬cárie	0,016	0,064	0,144	0,576

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Observação
  - A soma total das probabilidades é 1

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Observação
  - A soma total das probabilidades é 1
- Cálculo da probabilidade de uma proposição

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Observação
  - A soma total das probabilidades é 1
- Cálculo da probabilidade de uma proposição
  - Identifique os eventos atômicos nos quais a proposição é verdadeira

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Observação
  - A soma total das probabilidades é 1
- Cálculo da probabilidade de uma proposição
  - Identifique os eventos atômicos nos quais a proposição é verdadeira
  - Adicione suas probabilidades

Distribuição Conjunta de Probabilidade

Exemplo – Domínio: odontologia

- Observação
  - A soma total das probabilidades é 1
- Cálculo da probabilidade de uma proposição
  - Identifique os eventos atômicos nos quais a proposição é verdadeira
  - Adicione suas probabilidades
  - Ex:

$$P(\text{cárie} \lor \text{dor}) = 0{,}108 + 0{,}012 + 0{,}072 + 0{,}008 + 0{,}016 + 0{,}064 = 0{,}28$$

	dor		¬dor	
	prende	¬prende	prende	¬prende
cárie	0,108	0,012	0,072	0,008
¬cárie	0,016	0,064	0,144	0,576

Distribuição Conjunta de Probabilidade

Cálculo da probabilidade de uma proposição

- $P(\text{cárie}) = 0{,}108 + 0{,}012 + 0{,}072 + 0{,}008 = 0{,}2$

	dor		¬dor	
	prende	¬prende	prende	¬prende
cárie	0,108	0,012	0,072	0,008
¬cárie	0,016	0,064	0,144	0,576

Distribuição Conjunta de Probabilidade

Cálculo da probabilidade de uma proposição

- $P(\text{cárie}) = 0{,}108 + 0{,}012 + 0{,}072 + 0{,}008 = 0{,}2$

	dor		¬dor	
	prende	¬prende	prende	¬prende
cárie	0,108	0,012	0,072	0,008
¬cárie	0,016	0,064	0,144	0,576

  - Chamada **Probabilidade marginal**

Distribuição Conjunta de Probabilidade

Cálculo da probabilidade de uma proposição

- $\text{P(cárie)} = 0{,}108 + 0{,}012 + 0{,}072 + 0{,}008 = 0{,}2$

	dor		¬dor	
	prende	¬prende	prende	¬prende
cárie	0,108	0,012	0,072	0,008
¬cárie	0,016	0,064	0,144	0,576

  - Chamada **Probabilidade marginal**

- $P(\neg cárie|dor) = \dfrac{P(\neg cárie \wedge dor)}{P(dor)}$

Distribuição Conjunta de Probabilidade

Cálculo da probabilidade de uma proposição

- P(cárie) $= 0{,}108 + 0{,}012 + 0{,}072 + 0{,}008 = 0{,}2$

	dor		¬dor	
	prende	¬prende	prende	¬prende
cárie	0,108	0,012	0,072	0,008
¬cárie	0,016	0,064	0,144	0,576

  - Chamada **Probabilidade marginal**

- $P(\neg cárie|dor) = \dfrac{P(\neg cárie \wedge dor)}{P(dor)}$

	dor		¬dor	
	prende	¬prende	prende	¬prende
cárie	0,108	0,012	0,072	0,008
¬cárie	0,016	0,064	0,144	0,576

$$\frac{0,016 + 0,064}{0,108 + 0,012 + 0,016 + 0,064} = 0,4$$

Distribuição Conjunta de Probabilidade

Regra Geral

- $P(Y) = \sum_{Z} P(Y, Z)$

Distribuição Conjunta de Probabilidade

Regra Geral

- $P(Y) = \sum_{Z} P(Y, Z)$
  - A probabilidade de Y pode ser obtida somando-se todas as outras variáveis de qualquer distribuição conjunta que contenha Y

Distribuição Conjunta de Probabilidade

Regra Geral

- $$P(Y) = \sum_{Z} P(Y, Z)$$
  - A probabilidade de Y pode ser obtida somando-se todas as outras variáveis de qualquer distribuição conjunta que contenha Y
  - A probabilidade de um efeito (Y) pode ser obtida somando-se sua probabilidade de ocorrência com cada uma de suas causas possíveis

Distribuição Conjunta de Probabilidade

Regra Geral

- $$P(Y) = \sum_Z P(Y, Z)$$
  - A probabilidade de Y pode ser obtida somando-se todas as outras variáveis de qualquer distribuição conjunta que contenha Y
  - A probabilidade de um efeito (Y) pode ser obtida somando-se sua probabilidade de ocorrência com cada uma de suas causas possíveis

Alternativamente

- $$P(Y) = \sum_Z P(Y|Z)P(Z)$$

Distribuição Conjunta de Probabilidade

Caso Contínuo

Distribuição Conjunta de Probabilidade

Caso Contínuo

- Pode ser descrita pela função de densidade conjunta $f(X, Y)$, satisfazendo as seguintes condições:

Distribuição Conjunta de Probabilidade

Caso Contínuo

- Pode ser descrita pela função de densidade conjunta $f(X, Y)$, satisfazendo as seguintes condições:
  - $f(x, y) \geq 0$

Distribuição Conjunta de Probabilidade

Caso Contínuo

- Pode ser descrita pela função de densidade conjunta $f(X, Y)$, satisfazendo as seguintes condições:
  - $f(x, y) \geq 0$
  - $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dxdy = 1$

Distribuição Conjunta de Probabilidade

Caso Contínuo

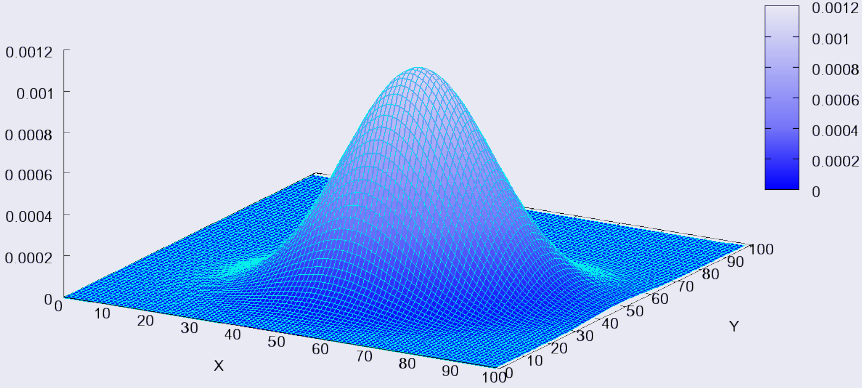
- Pode ser descrita pela função de densidade conjunta $f(X, Y)$, satisfazendo as seguintes condições:
  - $f(x, y) \geq 0$
  - $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dxdy = 1$
- E a probabilidade que $(x, y) \in A$ é dada por

$$P[(x, y) \in A] = \int\int_A f(x, y)dxdy$$

Distribuição Conjunta de Probabilidade

Caso Contínuo

- Ex: Normal bivariada



Fonte: `https://pt.wikipedia.org/wiki/Ficheiro:Multivariate_Gaussian.png`

Referências

- Russell, S.; Norvig P. (2010): Artificial Intelligence: A Modern Approach. Prentice Hall. 3a ed.
  - Slides do livro: `aima.eecs.berkeley.edu/slides-pdf/`
- Mitchell, T.M.: Machine Learning. McGraw-Hill. 1997.
- Murphy, K. P.: Machine Learning: A Probabilistic Perspective. MIT Press. 2012.
- Cover, T.M.; Thomas, J.A.: Elements of Information Theory. 2 ed. Wiley. 2006.
- `http://en.wikipedia.org/wiki/Probability_distribution`

Referências

- https://www.khanacademy.org/math/probability/random-variables-topic
- https://stat.duke.edu/~scs/Courses/STAT102/DecisionTheoryTutorial.pdf
- http://www.umass.edu/preferen/Game%20Theory%20for%20the%20Behavioral%20Sciences/BOR%20Public/BOR%20Decision%20Theory%20and%20Human%20Behavior.pdf
- http://people.kth.se/~soh/decisiontheory.pdf