



## Inferência causal

A busca elusiva da causalidade

## Desafio

- Dificuldade em definir quando um evento realmente causa outro

X --> Y ?

- Por milhões de anos apenas deuses, homens ou animais atuavam como causas
- Objetos ou processos físicos eram consequências e não causas



## Tempestade – fúria dos deuses



- Matemática – equação – ciência

$$g = \frac{G \times M}{R^2}$$

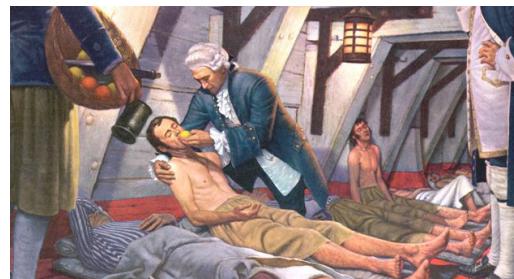
## Epidemiologia

- Primeiras tentativas de explicar o mundo
- Observação x experimentação
- Correlação ≠ Causa

### Escorbuto



### Escorbuto - Lind



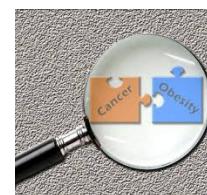
### Vacina - Jenner



Jenner inoculou o pus de uma ordenhadora que tinha contraído vaccinia em James Phipps, 8 anos. Ele teve doença branda. Em seguida foi inoculado com pústula de varíola humana e não ficou doente.

### Evidência - Epidemiologia

A associação observada entre a exposição e o efeito sugere relação de causa e efeito?



### Passos

- Associação estatística – descartar acaso
- Viés? Seleção, aferição, confundimento
- Julgar causalidade



### Causa

- Ocorrência de um evento sempre e invariavelmente segue a ocorrência de um outro evento.
- Física – MODELO DETERMINÍSTICO
- Causa é suficiente e necessária

### Causa suficiente

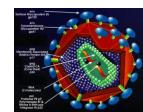
- Agindo sozinha sempre produzirá o evento



### Causa necessária

- Sem ela o evento não acontece

- HIV x aids



### Tuberculose

- Bacilo de Koch é causa necessária mas não suficiente
- Contribuinte - desnutrição



### Saúde

- As causas dos eventos não são necessárias ou suficientes
- Relação temporal não é clara



### Causa contribuinte

- Ou fator de risco
- Fator quando presente aumenta a probabilidade de ocorrência do evento

#### MODELO PROBABILÍSTICO

- Efeito rebote da dieta



### Modelo

- Formas de pensar ou de representar uma ideia
- Representação da realidade

*George Box* - Todos os modelos são errados, mas alguns são úteis



## Modelos de causalidade (explicativos)

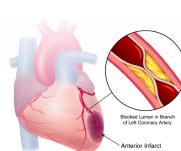
- Determinístico – evento tem uma causa ou causas determinadas - CERTEZA
- Probabilístico (estocástico) – não é possível ou não temos como determinar todas as causas do evento – o sistema pode até ser determinístico - INCERTEZA - acaso

## Modelos preditivos

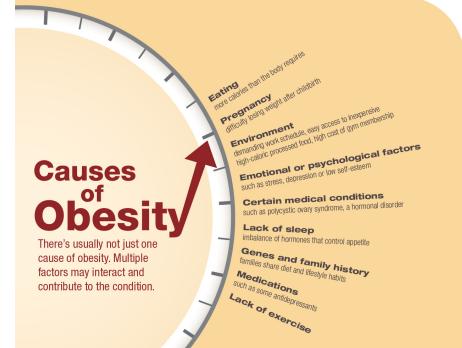
- É possível prever eventos a partir da correlações
- BIG DATA
- Buscas a respeito de sintomas de gripe – predição de epidemias de gripe antes das notificações

## INFARTO

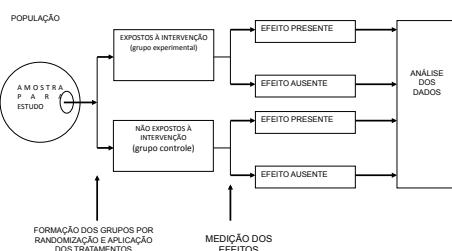
- O que causa o infarto?
- Quais as causas do infarto?
- Fatores de risco??



## SAÚDE - maioria dos eventos têm múltiplas causas



## Grupo controle



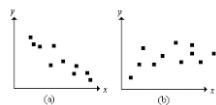
## Risco Relativo (razão de risco)

Exposição	Doença Sim	Não	Total
Sim	a	b	a+b
Não	c	d	c+d
Total	a+c	b+d	a+b+c+d

$$RR = \frac{\text{Coef. Incidência entre expostos}}{\text{Coef. Incidência entre não expostos}} = \frac{a/a+b}{c/c+d}$$

## Como atribuir causalidade?

- X causa Y?
- Correlação não é causa – não há causa sem correlação, mas nem toda correlação é causal
- Confundimento

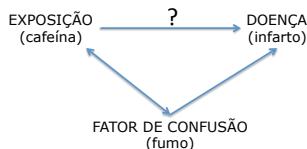


## Características do fator de confusão

- **ser um fator de risco para a doença**
- **associado com a exposição**
- **não ser um mediador entre a exposição e a doença**

Mediador não deve ser controlado na análise - ↴ associação

## Exemplo



Problemas – definição falha em várias ocasiões, o que leva a ajuste inadequado e/ou desnecessário

## Solução

- Controlar confundimento em modelos estatísticos

$$y_i = a + bX_i + e_i$$

- $X_i$  correlacionado com  $e_i$  – sabota interpretação causal – confundimento por variável omitida
- Pressupostos simplificados irrealistas
- Erros de especificação – não linearidade, não realização de testes para interação

## Associação causal

### Critérios de Hill

- Causa precede o efeito
- Força da associação
- Efeito dose-resposta
- Consistência
- Plausibilidade

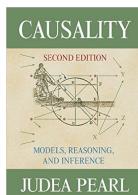


## Consenso

- Detecção de associação sem viés (confundimento, seleção, aferição)
- Aplicação dos critérios de julgamento causal

### Pearl (2009)

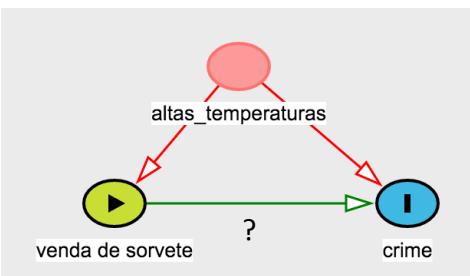
Não vejo nenhum maior impedimento para o progresso científico do que a prática vigente de se concentrar todos os nossos recursos matemáticos em inferências probabilísticas e estatísticas, deixando considerações causais à mercê da intuição e do bom senso.



### Venda de sorvete causa crime?

- Correlação – à medida em que aumenta a venda de sorvete aumenta também a ocorrência de crimes

### Venda de sorvete causa crime?



### O Paradoxo de Simpson

Tabela 6.1.: O paradoxo de Simpson - associação marginal

	D=1	D=0	Total
	n	%	n
T=1	20	50	20
T=0	6	50	6
Total	26	50	52

$RR_{TD} = 20/40 / 6/12 = 1$   
 $OR_{TD} = 20 \times 6 / 20 \times 6 = 1$

RR=Razão de Risco

OR=Odds ratio

Fonte: adaptada de ([Hernan et al., 2011](#))

### O Paradoxo de Simpson

Tabela 6.2.: O paradoxo de Simpson - associação condicional por sexo

C=1 Masculino			C=0 Feminino		
	D=1	D=0		D=1	D=0
	n	%	n	n	%
T=1	5	39	8	12	56
T=0	3	43	4	2	12
Total	8	40	12	14	52

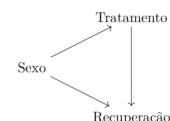
$RR_{TD|C=1} = 5/13 / 3/7 = 0,90$   
 $OR_{TD|C=1} = 5 \times 4 / 8 \times 3 = 0,83$

RR=Razão de Risco

OR=Odds ratio

Fonte: adaptada de ([Hernan et al., 2011](#))

Figura 6.2.: DAG interpretando o paradoxo de Simpson como viés de confundimento



## Conclusão

- Os dados comunicam associações
- Para decifrar causalidade é preciso conhecer a história por trás dos dados, os mecanismos que geraram os dados

- Como decifrar os mecanismos que geram os dados observados?

RESPOSTA

TEORIA