O falso dilema: Estatística Bayesiana vs Clássica

Fernando de Pol Mayer
2022-10-19

Roteiro

- Dedução/indução
- Métodos (ou lógica): dedução, indução
- O problema da indução
- Método hipotético-dedutivo
- Estatística como o estudo da incerteza
- Medidas de incerteza
    - Probabilidade, possibilidade, necessidade
- Estatística clássica
    - Tipos de probabilidade (frequentista, Kolmogorov)
    - Inferência (distribuição amostral, p-valores, etc)
- Estatística Bayesiana
    - Tipos de probabilidade (subjetiva, lógica)
    - Teorema de Bayes (atualização de crenças)
- Visões alternativas do método científico
    - Kuhn, Polanyi, Lakatos
    - Confronto de hipóteses
- Estatística clássica como Popperiana, Bayesiana como Lakatosiana
- Conclusão
    - O tipo de probabilidade é importante
    - Saiba em qual visão científica você está
    - Uma vez que isso é aceito, não deve haver mais discussão sobre os métodos
    - Se você quiser criticar algum método, critique a filosofia subjacente a ele, não o método em si
- Outras abordagens: Verossimilhança, Richard Royall

Raciocínio Lógico

Aprendizagem Dedutiva-Indutiva

- Guiado pelo cérebro humano
- Conhecido desde Aristóteles
- Parte de nossa experiência cotidiana

O processo é genericamente ilustrado nesta figura de Box, Hunter, and Hunter (2005)

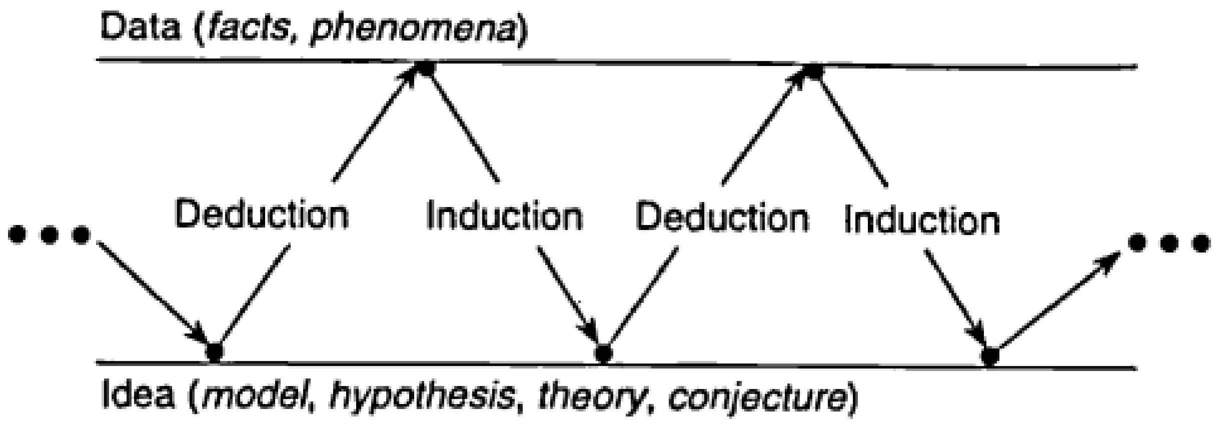


Figure 1.1. Iterative learning process.

Método Dedutivo

- Vai de afirmações **gerais** para afirmações **particulares** (uma definição muito simplificada)
- O protótipo do raciocínio dedutivo é o **silogismo**
  - A partir de duas proposições (**premissas**) chegamos a uma terceira, chamada de **conclusão**
- SE E SOMENTE SE as duas premissas são verdadeiras, então a conclusão também é verdadeira
- As premissas são afirmadas ou assumidas como verdadeiras, de modo que a conclusão será verdadeira pela lógica (dedutiva) pura
- O raciocínio dedutivo é amplamente aplicado em, *e.g.*, Matemática e Física, onde princípios podem ser enunciados como teoremas ou leis

Estrutura:

> Premissa maior.
>
> Premissa menor.
>
> Conclusão.

Exemplo:

> Todo mamífero tem um coração.
>
> Todos os cães são mamíferos.
>
> Todos os cães têm um coração.

Método Indutivo

- Vai de *observações* **particulares** para afirmações **gerais** (novamente uma definição simplificada)
- A partir de observações, é possível formular uma **hipótese explicativa** da causa do fenômeno.
- O método indutivo considera
  - As circunstâncias e a **frequência** com que um evento ocorre
  - Os casos em que o evento **não ocorre**
  - Os casos em que o evento apresenta **intensidades diferentes**
- Portanto, por indução, chegamos a uma **conclusão que é apenas provável**

Exemplo:

> Antônio é mortal.
>
> Benedito é mortal.
>
> Carlos é mortal.
>
> Zózimo é mortal.
>
> Antônio, Benedito, Carlos, ... e Zózimo são homens.
>
> Portanto, todos os homens são mortais.

Método Indutivo

O que acontece neste caso?

O cisne 1 é branco.

O cisne 2 é branco.

O cisne 3 é branco.

...

O cisne n é branco.

Portanto, **todos** os cisnes são brancos.

O cisne 1 é branco.

O cisne 2 é branco.

O cisne 3 é branco.

...

O cisne n é **preto**.

Portanto, **a maioria** dos cisnes é branca.

Mas, na lógica Clássica (Aristotélica), uma conclusão (ou afirmação) só pode ser **verdadeira** ou **falsa**, *i.e.*, com probabilidades 1 ou 0, respectivamente.

Método Indutivo

O problema da indução

- Devido a **David Hume**, a partir de seu *Tratado da Natureza Humana* (1739)
- Hume questiona com base em que fundamentos chegamos às nossas crenças sobre o não observado a partir de inferências indutivas
- A partir de seus argumentos, não há possibilidade de qualquer raciocínio das premissas para a conclusão de uma inferência indutiva
- Para Hume, o problema permanece sobre como explicar por que formamos quaisquer conclusões que vão além das instâncias passadas das quais tivemos experiência
- Ele não diz que não é possível tirar conclusões de inferências
- O desafio é entender o **fundamento** da inferência (ou a **lógica** dela)
- O problema da indução de Hume é essencialmente uma questão filosófica **aberta**, alvo de debate ainda nos dias de hoje

Inductive method

O método indutivo é na verdade um **problema de inferência**, com o qual os Estatísticos lidam todos os dias.

Então, em nosso mundo cotidiano, como você responderia a estas perguntas primárias que surgem na prática?

1. Qual é o número necessário de observações para se chegar a uma **conclusão provável**?
2. De fato, o que significa uma "**conclusão provável**"?
    - Uma conclusão com probabilidade $\geq 0,99$?
    - Uma conclusão com probabilidade $\geq 0,95$??
    - Uma conclusão com probabilidade $> 0,5$???

Um Estatístico **frequentista** ("Clássico") responderia a estas duas perguntas facilmente com:

1. Quanto mais observações, melhor (ainda melhor se o tamanho amostral tender ao infinito
2. Use intervalos de confiança e/ou testes de hipóteses (p-valores) para medir evidências nos dados

Ok, mas isso **não é uma solução** para o problema da indução

Estes métodos são baseados em outro modelo, o **Hipotético-dedutivo**, conforme definido por Karl Popper

Método Hipotético-Dedutivo

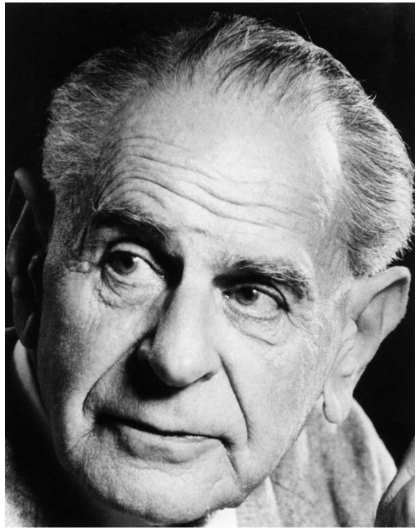
- Seguindo Hume, **Karl Popper** não aceitou o método indutivo como um caminho formal para a inferência

- Para ele, o método indutivo não é justificável, pois o "salto" de *alguns* para *todos* exigiria que as **observações fossem infinitas**

- De forma simplificada, o método segue o esquema geral:

    - Problema
    - **Hipótese** (*singular*)
    - **Dedução** de consequências (dada a hipótese como verdadeira)
    - Elaboração de um **experimento crítico**
    - **Falsificação** da hipótese (evidências contrárias à hipótese)
    - Corroboração (ou não)

- Enquanto o Método Dedutivo tenta **verificar** uma hipótese, o Método Hipotético-Dedutivo tenta **falsificá-la** (provar que está errada)



- Se a hipótese não for falsificada, então ela é **corroborada**

- A essência do método Popperiano é "**desafiar**" uma hipótese repetidamente

    - Se ela permanecer válida, então **não é validada**, mas adquire um **grau de crença**

A feedback loop (Box, Hunter, and Hunter, 2005)

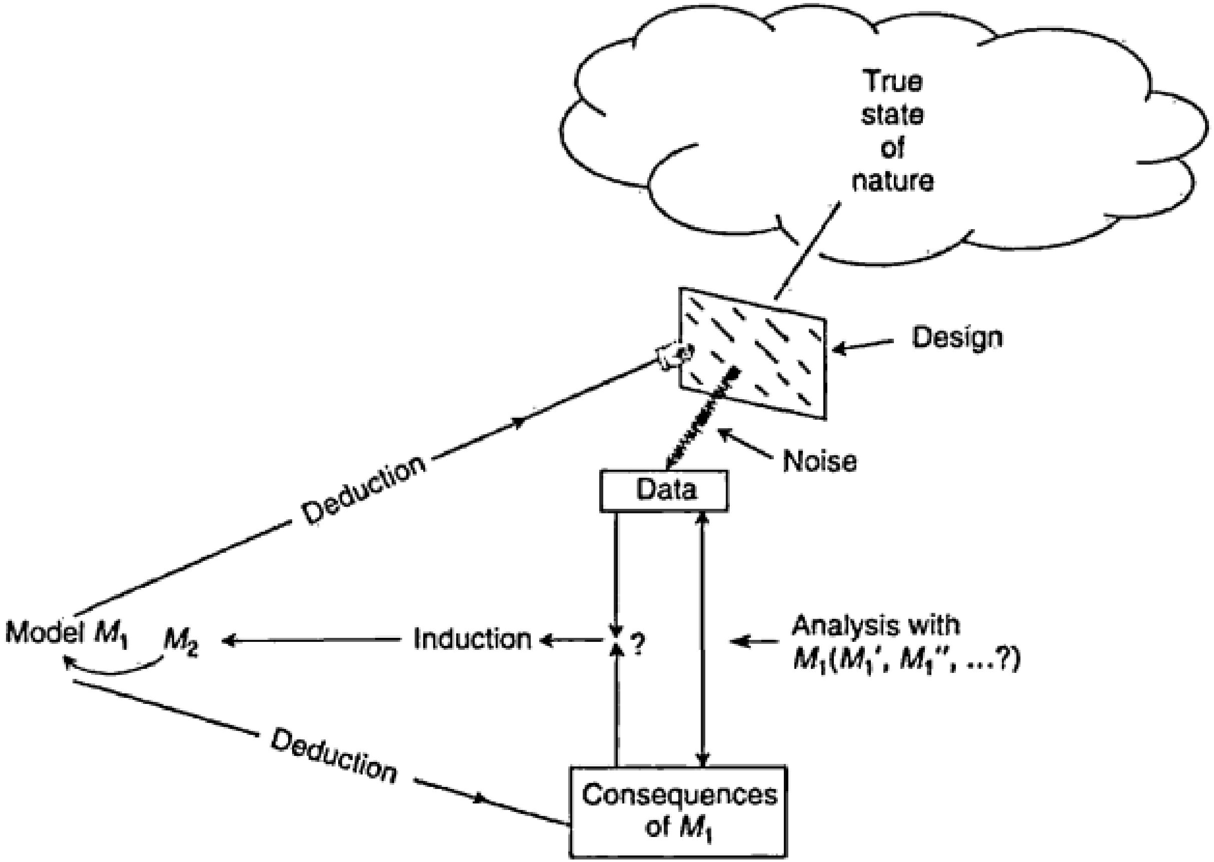


Figure 1.2. Iterative problem solving seen as a feedback loop.

A prática estatística padrão

A prática estatística padrão

A prática estatística padrão para inferência é aquela baseada no **Teste de Significância de Hipótese Nula** (NHST)

- É inteiramente baseada na visão de Karl Popper sobre o método científico
- Popper forneceu a filosofia, e **Ronald Fisher**, **Karl Pearson**, **Jerzy Neyman** (entre outros) forneceram a teoria estatística para sustentá-la
- Em um NHST, focamos em uma **única** hipótese ("**hipótese nula**" ou H_0)
- Em seguida, calculamos a probabilidade de uma estatística ser tão extrema quanto a observada, **assumindo que a nula é verdadeira**

  $$\mathrm{P}(\mathrm{T}(X) \geq \mathrm{T}(x) \mid H_0 \text{ verdadeira}) = \alpha^\star = p\text{-valor}$$

  para qualquer estatística $\mathrm{T}(\cdot)$.

- Isso se baseia em uma **distribuição amostral**, possivelmente derivada de **resultados assintóticos** (quando $n \to \infty$)
- Se essa probabilidade for
  - **Baixa** (geralmente menor que $0,05$ ou $0,01$), então **rejeitamos** a hipótese nula
  - **Alta** (geralmente acima de $0,05$ ou $0,1$), então **não rejeitamos** a hipótese nula
- Seguindo Popper, uma hipótese não pode ser provada, apenas refutada
  - É por isso que não podemos dizer "**aceitar uma hipótese**"

A prática estatística padrão

Se a hipótese nula é **rejeitada**

- Evidência científica
- Suporte à hipótese de pesquisa
- Adquire um grau de crença (se repetida várias vezes e o resultado se mantiver)
- **Artigo é publicado**

Os aspectos-chave aqui são

1. O confronto entre uma **única** hipótese e os dados
2. O **experimento crítico**
3. A **falsificação** como a única "verdade"

Se a hipótese nula **não é rejeitada**

- Sem suporte à hipótese científica
- O que fazer?
  - Reformular a hipótese?
  - Coletar mais dados?
- **Artigo é rejeitado**

Apesar de ser o mais comumente seguido em estatística, o método Popperiano é apenas um entre várias alternativas

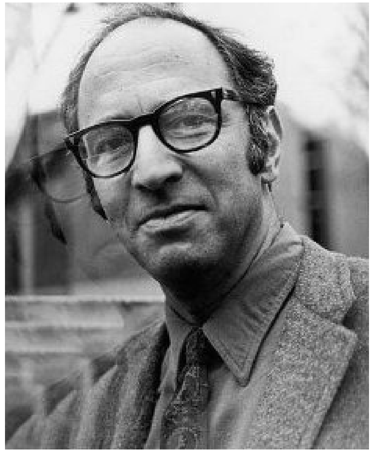
Visões alternativas do método científico

Visões alternativas do método científico

Filósofo	Tema central	Tipo de confronto
Popper	Falsificação da hipótese nula	Uma única hipótese é desprovada em confronto com os dados
Kuhn	Paradigmas, ciência normal, revoluções científicas	Uma única hipótese é utilizada até que existam informações contraditórias suficientes para que ela seja "derrubada" por uma hipótese "melhor"
Polanyi	República da ciência	Múltiplas visões do mundo são permitidas de acordo com a opinião de diferentes cientistas.
Lakatos	Programa de Pesquisa Científica	Confronto entre múltiplas hipóteses, tendo os dados como mediador

Visões alternativas do método científico

Thomas Kuhn



- Introduziu as ideias de "ciência normal", "paradigmas científicos", e "revoluções científicas"
- Para Kuhn, os cientistas normalmente atuam dentro de **paradigmas específicos**, que são descrições gerais de como o mundo funciona
- A **ciência normal** envolve a coleta de dados dentro do contexto do paradigma atual
  - Não confronta o paradigma atual, mas sim o executa
  - O paradigma dita o tipo de experimento que deverá ser executado, quais dados serão coletadas, e como serão analisados e interpretados
- Uma verdadeira mudança (quebra de paradigma ou **revolução científica**) ocorrerá quando:
  1. Um grande conjunto de dados contraditórios é acumulado, de maneira que o paradigma atual não consegue mais explicar os dados
  2. Existe um paradigma alternativo, que consegue explicar as discrepâncias entre o antigo paradigma e as observações

Visões alternativas do método científico

Michael Polanyi

- Descreve que a **república de ciência** consiste de uma comunidade de pensadores independentes cooperando livremente entre si

- Para Polanyi, isso representa uma versão simplificada de uma sociedade livre

  - Cientistas são *treinados* por um *mestre* (orientador)
  - Aprendizes observam e participam
  - Os indivíduos constituem a **república** de cidadãos ensinados por essa cadeia de ensino-aprendizado
  - É este sistema que previne que a ciência se torne "rígida" ou "estagnada"

- O aprendiz recebe um elevado padrão de conhecimento científico e desenvolve sua própria capacidade de julgamento para assuntos científicos

- Polanyi argumenta implicitamente que o confronto

  - não é entre **uma** hipótese e os dados
  - mas sim **entre hipóteses** (diferentes descrições de como o mundo funciona) e os **dados** (observações e medidas)

Visões alternativas do método científico

Imre Lakatos

- Lakatos descreve o que chama de **Programa de Pesquisa Científica** (PPC)

    - Um conjunto de regras metodológicas que indicam os caminhos a serem guiados e evitados
    - O "**núcleo rígido**" é o elemento chave de um PPC
        - Ao redor do núcleo, existe um conjunto de hipóteses
        - Esse conjunto de hipóteses é chamado de "**cinturão**", que **protege** o núcleo
    - As hipóteses individuais do cinturão podem ser testadas
    - Raramente o núcleo pode ser desafiado diretamente

- Lakatos aponta que muitas **hipóteses** têm sido consideradas e utilizadas, apesar de suas reconhecidas inconsistências

    - Leis de Newton
    - Teoria da gravidade
    - Modelos gerais de química orgânica

- Apesar disso, elas *são utilizadas por não haver um substituto melhor*

Visões alternativas do método científico

Imre Lakatos

- Um PPC só pode ser substituído por outro PPC
  - *Não podemos rejeitar uma hipótese a menos que exista uma outra melhor para substituí-la*
- Portanto, na visão Lakatosiana, **o confronto deve ser sempre entre hipóteses concorrentes e os dados**
  - Uma hipótese individual **pode** ser *inconsistente* com os dados
  - Mas a menos que exista outra mais consistente, não descartaremos a primeira, pois é necessário que se continue investigando

Visões alternativas em estatística

Visões alternativas em estatística

A **inferência clássica** é usada para sustentar o falsificacionismo de Popper de uma única hipótese

A **inferência Bayesiana** pode ser usada para sustentar a teoria de Lakatos de confronto entre hipóteses concorrentes

Podemos chamar de agora em diante: estatística **Clássica/Popperiana** e **Bayesiana/Lakatosiana**

Inferência Bayesiana

Inteiramente baseada no **Teorema de Bayes**

$$
\begin{aligned}
[H_i|D] &= \frac{[DH_i]}{[D]} \\
&= \frac{[D|H_i][H_i]}{[D]} \\
&\propto [D|H_i][H_i]
\end{aligned}
$$

- $[H_i|D]$ é a probabilidade da hipótese H_i, condicionada aos dados D (**posterior**)
- $[D|H_i]$ é a probabilidade do dado, condicionada à hipótese (**verossimilhança**)
- $[H_i]$ é a distribuição de probabilidade **a priori** para a hipótese
- O resultado será a probabilidade de cada uma das $i = 1, \ldots, k$ hipóteses concorrentes
- É possível incorporar informações prévias sobre uma hipótese, por meio da (distribuição) *a priori*

Suponha que se deseja testar 3 hipóteses concorrentes (H_1, H_2, H_3) para o mesmo conjunto de dados observado D. O resultado será

$$
\begin{aligned}
[H_1|D] &= p_1 \\
[H_2|D] &= p_2 \\
[H_3|D] &= p_3
\end{aligned}
$$

Com isso, é possível determinar qual a **hipótese mais plausível**, de acordo com os dados.

Visões alternativas em estatística

Clássica/Popperiana

$$\mathrm{P}[D|H_0]$$

- Confronto entre uma única hipótese e os dados
- Rejeitar/não rejeitar a hipótese nula
- A hipótese é falsificada
- Experimentos críticos são aqueles que fazem a hipótese nula ser falsificada
- Probabilidades são atribuídas a subconjuntos de uma σ-álgebra, pelos Axiomas de Kolmogorov
  - Não é possível atribuir probabilidade a hipóteses

Bayesiana/Lakatosiana

$$\mathrm{P}[H_i|D], \quad i = 1, 2, \ldots, k$$

- Confronto entre os dados e múltiplas hipóteses
- Graus de plausibilidade para cada hipótese
- Experimentos críticos são aqueles capazes de mudar os graus de crença
- A probabilidade é uma extensão da lógica, atribuída pelos teoremas de Cox
  - Probabilidades devem ser atribuídas a hipóteses

- Para Lakatos, uma hipótese não pode ser rejeitada, a menos que haja uma alternativa melhor
- Para Popper, uma única hipótese pode ser rejeitada (ou não)

Visões alternativas em probabilidade

Visões alternativas em probabilidade

No entanto, lembre-se que a estatística clássica é fundada em probabilidade definida para conjuntos. Se definirmos uma classe de subconjuntos de Ω (o espaço amostral) em que

1. $\Omega \in \mathcal{F}$
2. Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$
3. Se $A_i \in \mathcal{F}, i \geq 1$, então $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$

então $\mathcal{F}$ é uma σ-álgebra.

Uma função $\mathrm{P} \in [0, 1]$, definida na σ-álgebra $\mathcal{F}$ de subconjuntos de Ω, é chamada de probabilidade, se satisfizer os **Axiomas de Kolmogorov**

1. $\mathrm{P}(\Omega) = 1$
2. $\mathrm{P}(A) \geq 0$, $\forall$ subconjunto $A \in \mathcal{F}$
3. $\mathrm{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathrm{P}(A_i), \ \forall A_1, A_2, \ldots \in \mathcal{F}$ onde $A_i \cap A_j = \phi \ (i \neq j)$

Assim, o conjunto $(\Omega, \mathcal{F}, \mathrm{P})$ é chamado de espaço de probabilidade, e **probabilidades são atribuídas apenas aos subconjuntos em $\mathcal{F}$**, que são chamados de eventos.

Por esta definição, afirmações como

$$A \equiv \text{"vai chover hoje"}$$

$$B \equiv \text{"a chuva vai causar enchente"}$$

são meramente afirmações de fatos (incertos), **fora do escopo dos conjuntos**.

Na verdade, a definição (**informal**) de p-valor deve ser tomada com cuidado

$$\mathrm{P}[D|H_0] = \frac{\mathrm{P}[D, H_0]}{\mathrm{P}[H_0]}$$

pois $\mathrm{P}[H_0]$ não está definido neste contexto (e não é uma probabilidade condicional).

Visões alternativas em probabilidade

Portanto, precisamos de outra definição de probabilidade para a Inferência Bayesiana, pois precisamos atribuir probabilidades a hipóteses (ou afirmações)

O que aprendemos é que as probabilidades podem ser definidas como uma de

- Clássica (m de n)
- Frequentista (proporções limítrofes)
- Axiomática (Kolmogorov)

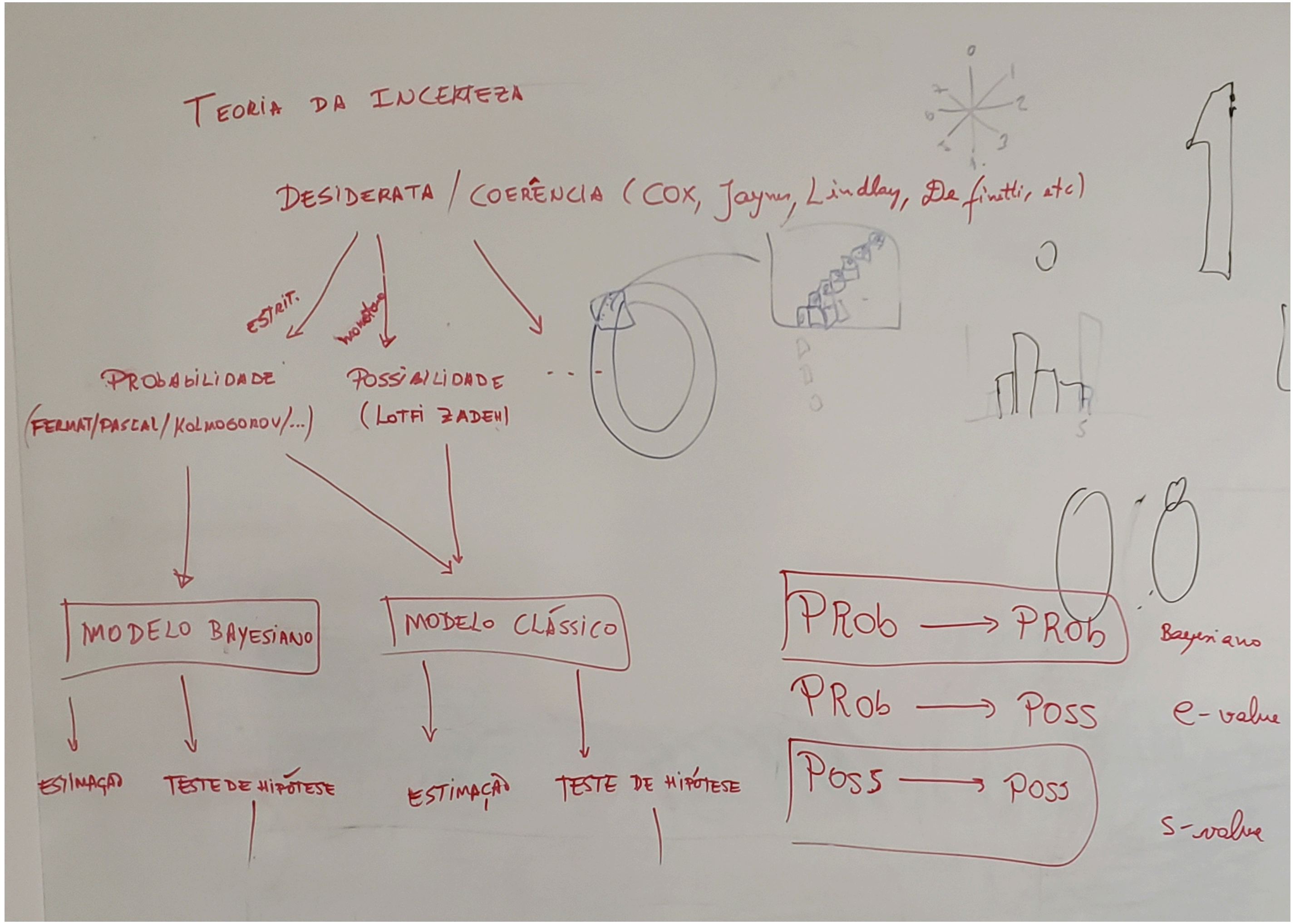
Mas existem outras, como

- Probabilidade subjetiva (de Finetti)
- Probabilidade como extensão da lógica (Jeffreys, Cox, Jaynes)

Outros tipos de medidas de incerteza

- Medidas de possibilidade (Lotfi Zadeh, Didier Dubois)
  - Baseadas em lógica Fuzzy

Teoria da incerteza



Quadro do LEG após discussão com Prof. Alexandre Patriota (IME/USP).

Probabilidade como extensão da lógica

Lembrando a inferência sobre os cisnes:

- **Todos** os cisnes são brancos
- **A maioria** dos cisnes é branca

Na lógica clássica (dedutiva), essas afirmações devem ser classificadas como verdadeiras ou falsas (probabilidade 1 ou 0).

Mas a probabilidade pode ser usada para estender essa lógica, e fornecer um **contínuo** de possibilidades entre esses dois extremos

Portanto, esse tipo de probabilidade pode ser usado diretamente na lógica indutiva (que é o que desejamos em primeiro lugar)

A formalização de probabilidades para quantificar a lógica indutiva é caracterizada por três **regras de racionalidade** básicas.

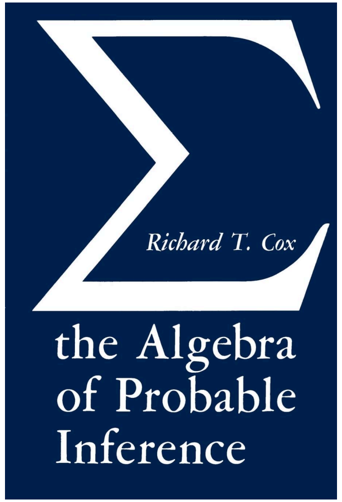
Esses *desiderata* foram propostos pelo matemático **George Pólya** em 1954

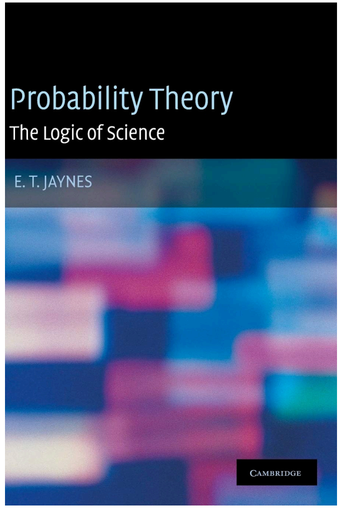
1. Os graus de plausibilidade são representados por números reais.
    - *Com a quantificação, é possível classificar e ordenar proposições*
2. Há uma correspondência qualitativa com o senso comum.
    - *Quanto mais plausível uma proposição, maior deve ser o número a ela associado.*
3. Há consistência nas medições de plausibilidade.
    - *Se a solução pode ser obtida de várias maneiras, todas devem levar ao mesmo resultado.*
    - *Todas as evidências relevantes disponíveis serão levadas em consideração. Nenhuma informação será arbitrariamente ignorada.*
    - *Níveis equivalentes de conhecimento (informação) correspondem a níveis equivalentes de plausibilidade.*

Probabilidade como extensão da lógica

Agora precisamos encontrar um conjunto de operações matemáticas para manipular plausibilidades

- Conforme relatado em Jaynes (2003), Richard T. Cox já havia mostrado em 1946 que esse conjunto de operações matemáticas existe e que o mesmo coincide com o conjunto de operações matemáticas que suportam o cálculo convencional de probabilidades.
    - Os Axiomas de Kolmogorov são um caso especial destes
- Cox também mostrou que qualquer outro conjunto de operações que satisfaça esses desiderata terá necessariamente equivalência completa com o cálculo de probabilidades.
    - Isso exclui medidas de possibilidade, por exemplo
    - Isso também é afirmado nos trabalhos de Dennis Lindley e muitos outros
- Portanto, para incorporar corretamente as incertezas na avaliação de proposições científicas que requerem lógica indutiva, é necessário conhecer os princípios matemáticos do cálculo de probabilidades.





Probabilidade como extensão da lógica

- A probabilidade será um número real obtido como função de dois argumentos: o **evento** incerto E e a **premissa** H, ou $P(E|H)$

- A **premissa** (informação) simbolizada por H é fundamental na definição Bayesiana de probabilidade, para destacar que a probabilidade expressa a incerteza atual sobre E e está **ligada a um observador em posse da informação** H

- Aqui usamos **operadores lógicos** para **manipular as relações entre proposições**, frequentemente também denominadas eventos

- O formalismo completo está definido em Cox (1946) e Jaynes (2003).

- **A Lei da Convexidade**: A probabilidade de qualquer evento E, condicionada a H, é um número real no intervalo $[0,1]$.

$$0 \leqslant P(E|H) \leqslant 1$$

  É conveniente associar números maiores a probabilidades maiores. Portanto, temos que $P(H|H) = 1$ e $P(H'|H) = 0$.

- **A Lei da Adição**: Se E_1 e E_2 são eventos *exclusivos* sob H, então a probabilidade da união lógica $E_1 + E_2$ é igual à soma aritmética de suas probabilidades individuais sob H.

$$P(E_1 + E_2|H) = P(E_1|H) + P(E_2|H)$$

- **A Lei do Produto**: Se E_1 e E_2 são quaisquer eventos, então a probabilidade do produto lógico E_1E_2 sob H é

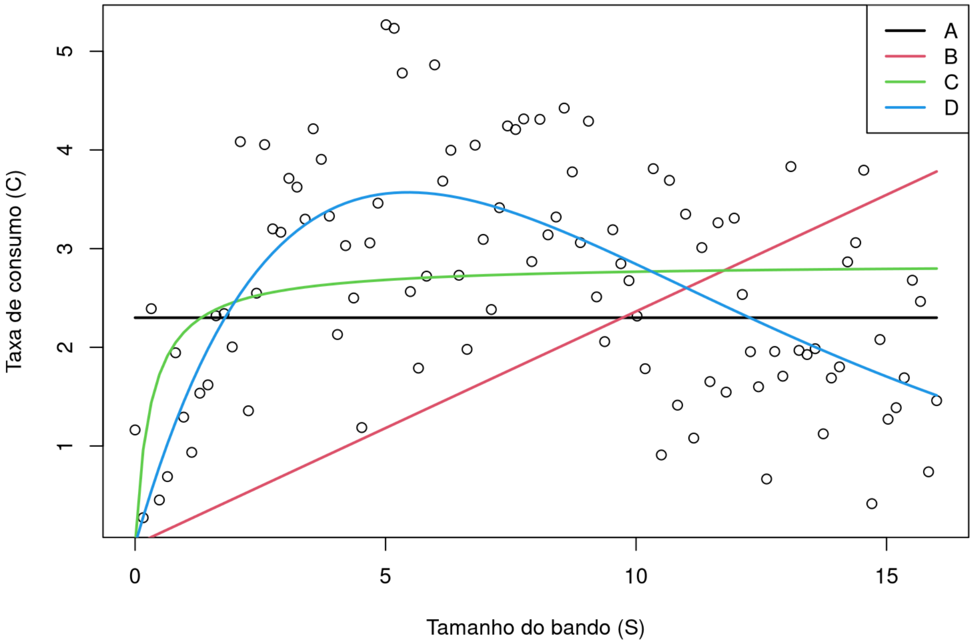
$$P(E_1E_2|H) = P(E_1|H) \cdot P(E_2|E_1H)$$

Exemplo

Hipótese de pesquisa: bandos maiores de pássaros possuem uma taxa maior de consumo (forrageamento)

- Quatro modelos de forrageamento podem ser propostos
- Cada "modelo" aqui se refere à uma hipótese

Hipótese	Modelo	Descrição
A	$C = a$	Consumo independente do tamanho do bando (modelo **nulo**)
B	$C = aS$	Consumo é proporcional ao tamanho do bando
C	$C = \frac{aS}{1+bS}$	Consumo satura à medida que o tamanho da bando aumenta
D	$C = aS\mathrm{e}^{-bS}$	Consumo aumenta e depois diminui com o aumento do tamanho do bando



Exemplo

Popperiana/clássica

$$\mathrm{P}[D|H_0]$$

- Modelos B, C, e D seriam confrontados **individualmente** com o modelo nulo

$$H_0 = H_A \text{ vs } H_B$$
$$H_0 = H_A \text{ vs } H_C$$
$$H_0 = H_A \text{ vs } H_D$$

- Em cada teste há apenas duas possíveis decisões: **rejeitar** ou **não rejeitar** H_0

- As hipóteses não podem (a princípio) serem testadas "par-a-par" pois isso diminui o nível de significância (global) do teste

Lakatosiana/bayesiana

$$\mathrm{P}[H_i|D], \quad i = A, B, C, D$$

- O confronto seria entre os quatro modelos competidores e os dados

- Pode-se determinar o **grau de plausibilidade** de cada hipótese/modelo

$$[H_A|D] = p_A$$
$$[H_B|D] = p_B$$
$$[H_C|D] = p_C$$
$$[H_D|D] = p_D$$

- Se, por exemplo,

$$p_D > p_C > p_B > p_A$$

então assumimos que o modelo D é o mais plausível **entre os 4 testados**

Atualização Bayesiana (*Bayesian updating*)

A estrutura do Teorema de Bayes permite naturalmente a **atualização sequencial** do conhecimento

$$\underbrace{[H|D_1]}_{\text{posteriori}_1} \propto \underbrace{[D_1|H]}_{\text{verossimilhança}_1} \cdot \underbrace{[H]}_{\text{priori}}$$

Quando novos dados D_2 ficam disponíveis, a **posteriori anterior torna-se a nova priori**

$$\underbrace{[H|D_1, D_2]}_{\text{posteriori}_2} \propto \underbrace{[D_2|H]}_{\text{verossimilhança}_2} \cdot \underbrace{[H|D_1]}_{\text{posteriori}_1 = \text{priori}_2}$$

E assim sucessivamente: a cada novo dado, o ciclo se repete

$$\text{priori} \xrightarrow{D_1} \text{posteriori}_1 \xrightarrow{D_2} \text{posteriori}_2 \xrightarrow{D_3} \cdots$$

- A posteriori final **não depende da ordem** em que os dados foram observados
- O resultado é **idêntico** ao obtido com todos os dados analisados de uma só vez
- Isso decorre diretamente da **regra do produto** de probabilidades e da estrutura de **famílias conjugadas**
- A atualização sequencial é a **tradução formal** do processo cotidiano de aprendizado: conforme acumulamos experiências, atualizamos nossas crenças

> *"Yesterday's posterior is today's prior"*
>
> Dennis V. Lindley

E a abordagem clássica?

A atualização sequencial **não é possível** no arcabouço clássico

- p-valores e intervalos de confiança são propriedades da **distribuição amostral** de uma estatística, não do parâmetro
- A distribuição amostral depende do **delineamento inteiro** (tamanho amostral, regra de parada, etc.)
- **Não existe um mecanismo formal** para incorporar resultados de um estudo anterior como "priori" em um estudo posterior
- Ao combinar dois estudos, é necessário recorrer a **meta-análise**, que opera sobre estimativas pontuais, não sobre distribuições
- O conceito de **informação acumulada** não tem tradução direta em termos de p-valores
  - Um p-valor de 0,03 em um estudo **não pode ser combinado** com um p-valor de 0,08 de outro estudo para "atualizar" a conclusão de forma coerente

Por quê?

A razão fundamental é **filosófica**:

- Na visão clássica/Popperiana, probabilidades não são atribuídas a hipóteses
- Portanto, **não existe uma distribuição sobre H** que possa ser atualizada
- Sem uma distribuição que represente o estado de conhecimento, não há o que "atualizar"
- A inferência clássica trata cada experimento como uma **entidade isolada**, com sua própria distribuição amostral e suas próprias conclusões

> A atualização sequencial não é uma "vantagem técnica" da Estatística Bayesiana --- ela é uma **consequência direta** de se adotar probabilidade como medida de incerteza sobre hipóteses

Considerações finais

Considerações finais (de Lindley, 2000)

- A conclusão de que a probabilidade é a única medida de incerteza pode parecer óbvia
- No entanto, a maioria dos Estatísticos (até os Bayesianos) pensa que existe apenas um tipo de probabilidade
- No entanto, isso não é verdade: os estatísticos usam medidas de incerteza que não se combinam de acordo com as regras do cálculo de probabilidades

Considere
$H \equiv$ "um tratamento médico é ineficaz"

- Um Estatístico usa um NHST para testar essa hipótese
- Assumindo que H é verdadeira, calcula-se a probabilidade dos dados observados, ou de dados mais extremos
- Isso é uma medida da credibilidade que pode ser atribuída a H; quanto menor a probabilidade, menor é a credibilidade
- Esse uso contraria os argumentos acima, que afirmam que a incerteza sobre H **precisa ser medida por uma probabilidade** para H.
- **Um nível de significância não é tal probabilidade**
    - **Nível de significância**: a probabilidade de algum aspecto dos dados, dado que H é verdadeira
    - **Probabilidade**: sua probabilidade de H, dados os dados
- Como Jeffreys (1961) colocou: o nível inclui dados que poderiam ter ocorrido, mas não ocorreram

Considerações finais

- Frequentemente, a discussão entre Estatística Clássica vs Bayesiana é focada em "falhas metodológicas"
- Por exemplo
  - Frequentistas não admitem atribuir probabilidade a uma hipótese (*a priori*)
  - Bayesianos não veem sentido no que é um p-valor, ou na interpretação de intervalos de confiança
- Dado que um frequentista está (quase certamente) baseado em uma visão Popperiana e em probabilidade Axiomática, realmente não faz sentido atribuir probabilidades a hipóteses
- Dado que um Bayesiano está (esperançosamente) baseado na visão de Lakatos e em probabilidade como extensão da lógica, um p-valor realmente não tem sentido
- No entanto, na maioria das vezes, ambos pensam que há apenas uma visão da Ciência (a que praticam) e uma maneira de atribuir probabilidades
- Portanto, esse debate está sujeito a uma "guerra" interminável
- Uma vez que cada um conheça sua posição na Ciência, argumento que o debate não deve ser mais metodológico, mas sim filosófico