E o que isso tem a ver com inferência bayesiana...

Fernando de Pol Mayer (LEG/DEST/UFPR) 2022-02-03 (última atualização 2022-02-03)



Material baseado em:

• Hilborn, R; Mangel, M. The ecological detective: confronting models with data. Princeton University Press, 1997.

Objetivos

- Entender as lógicas básicas do pensamento humano e científico
- Entender o procedimento padrão do método científico dominante na estatística (e na ciência em geral)
- Conhecer visões alternativas do método científico e estatístico

Método científico

Método científico

Aprendizado dedutivo-indutivo

- O processo **dedutivo-indutivo** de aprendizado é orientado pelo cérebro humano
- É conhecido desde o tempo de Aristóteles e faz parte de nossa experiência cotidiana

O aprendizado avança conforme ilustrado na figura abaixo (extraída de Box, Hunter, and Hunter, 2005):

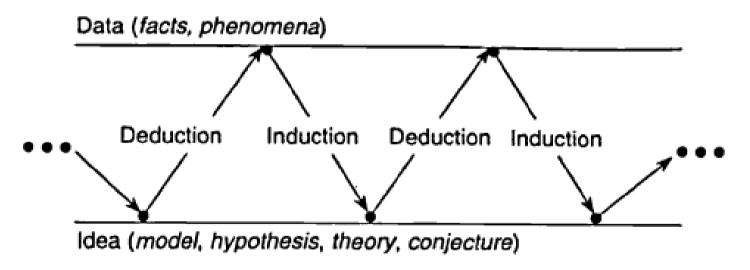


Figure 1.1. Iterative learning process.

Método dedutivo

- Esse método parte do geral e, a seguir, desce para o particular
- O protótipo do raciocínio dedutivo é o silogismo
 - A partir de duas proposições chamadas premissas, retira uma terceira chamada conclusão
- Se as duas premissas são verdadeiras, então a conclusão é verdadeira.
- Parte-se de princípios reconhecidos como verdadeiros e indiscutíveis, possibilitando chegar a conclusões de maneira puramente formal, em virtude de sua lógica (clássica).
- Este método tem larga aplicação na Matemática e na Física, cujos princípios podem ser enunciados por leis.

Exemplo:

Todo mamífero tem um coração.

Ora, todos os cães são mamíferos.

Logo, todos os cães têm um coração.

Método indutivo

- Parte do particular para o geral
- 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 ocorre determinado fenômeno
 - Os casos em que o fenômeno não se verifica
 - Os casos em que o fenômeno apresenta intensidade diferente
- Portanto, por meio da indução chega-se a conclusões que são apenas prováveis.

Exemplo:

Antônio é mortal.

Benedito é mortal.

Carlos é mortal.

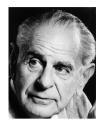
Zózimo é mortal.

Ora, Antônio, Benedito, Carlos, ... e Zózimo são homens.

Logo, (todos) os homens são mortais.

Método hipotético-dedutivo

- Método definido por Karl Popper, a partir de suas críticas ao método indutivo ("problema da indução")
- Para ele, o método indutivo não se justifica, pois o salto de "alguns" para "todos" exigiria que a observação de fatos isolados fosse infinita.
- Pode ser explicado a partir do seguinte esquema:
 - Problema
 - Hipóteses
 - Dedução das consequências observadas
 - Tentativa de falseamento
 - Corroboração
- Para tentar explicar um problema, são formuladas hipóteses; destas deduzem-se consequências que deverão ser testadas ou falseadas.



- Falsear significa tentar tornar falsas as consequências deduzidas das hipóteses.
- Enquanto no método dedutivo se procura confirmar a hipótese, no método hipotético-dedutivo se procuram evidências empíricas para derrubá-la.
- Quando n\u00e3o se consegue derrubar a hip\u00f3tese, temse sua corrobora\u00e7\u00e3o
- Segundo Popper, a hipótese se mostra válida, pois superou todos os testes
- Porém ela não é definitivamente confirmada, pois a qualquer momento poderá surgir um fato que a invalide.

Método científico

Um loop de feedback (Box, Hunter, and Hunter, 2005)

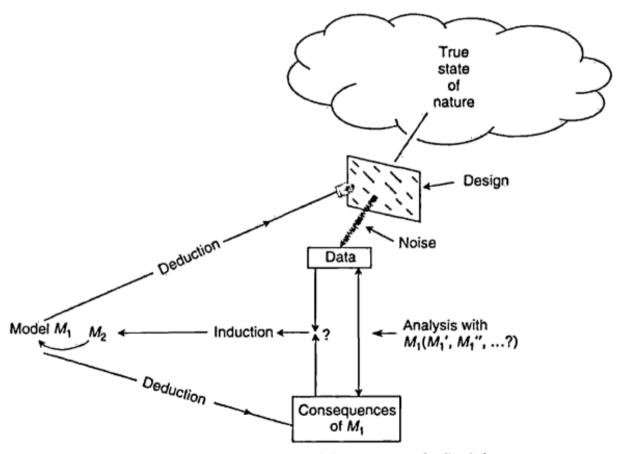


Figure 1.2. Iterative problem solving seen as a feedback loop.

Como a ciência funciona

Como a ciência funciona

A Ciência é um processo de aprendizado da natureza, onde ideias concorrentes sobre como funciona o mundo são medidas contra observações.

Richard Feynman, 1965

- Descrições: incompletas
- Observações: incertas e imprecisas
- Métodos para avaliar a concordância entre as ideias e as observações = Estatística

- Nenhum cientista consegue ser verdadeiramente **neutro**
- Todos os cientistas atuam dentro de uma visão fundamental filosófica
- Isso intefere:
 - Nas ferrmentas estatísticas que serão utilizadas
 - Nos tipos de experimentos que deverão ser realizados

Como a ciência funciona (atualmente)

Árvore de aprendizado (Platt, 1964)

- 1. Conceber **hipóteses alternativas**.
- 2. Conceber um **experimento crucial** (ou vários deles)
- 3. Realizar o experimento de forma a obter resultados mais confiáveis possíveis.
- 4. Reciclar o procedimento
 - Criar sub-hipóteses ou hipóteses sequenciais para refinar as possibilidades que restam

 Esta visão de Platt é naturalmente uma extensão lógica do trabalho de Popper

Uma hipótese **não pode ser provada**, apenas **desprovada**

- A essência do método Popperiano é "desafiar" uma hipótese repetidamente.
 - Se a hipótese permanece válida então ela não é validada, mas adquire um certo "grau de confiança"

Como a ciência funciona (atualmente)

Coincidindo com esta filosofia de Popper está o trabalho estatístico de **Ronald Fisher**, **Karl Pearson**, **Jerzy Neyman** e outros, que desenvolveram grande parte da teoria estatística atual associada a **testes de hipótese**.

Em um teste de hipótese, nós concentramos em uma única hipótese ("hipótese nula" ou H_0) e calculamos a probabilidade dos dados terem sido observados **assumindo que a hipótese nula seja verdadeira** (cuidado com essa condicional)

$$P(\mathrm{Dados}|H_0 \ \mathrm{verdadeira}) = lpha^\star = p\text{-valor}$$

Se essa probabilidade:

- For muito baixa (usualmente abaixo de 0.05 ou 0.01), então **rejeitamos** a hipótese nula.
- For alta (usualmenta acima de 0.05), então **não rejeitamos** a hipótese nula.

Lembre-se que, segundo a filosofia de Popper:

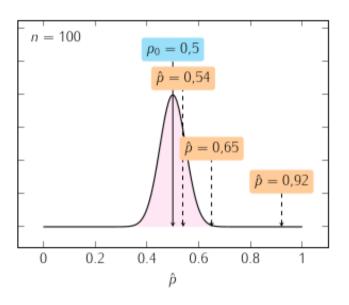
- Uma hipótese não pode ser provada, apenas desprovada.
- Por isso, não podemos aceitar uma hipótese (já que ela não pode ser provada)

aceitar \neq não rejeitar

Um exemplo: proporção sexual em peixes

- Deseja-se estudar a proporção de peixes machos e fêmeas de uma mesma espécie em uma lagoa.
- Sem nenhuma informação prévia, supõe-se que a proporção sexual é de 50% (p=0.5).
- Se, em uma amostra de 100 peixes:
 - 54 forem fêmeas.
 - 65 forem fêmeas.
 - 92 forem fêmeas.
- Qual a evidência necessária para concluir que a proporção de fêmeas é maior que a de machos nessa população?

$$H_0: p = 0.5 \ H_a: p > 0.5$$



Proporções amostrais supondo p=0.5 na população.

Tipos de hipótese

Hipótese nula H_0

- É uma afirmativa de que o valor de um parâmetro populacional é **igual** a algum valor especificado.
- Exemplos:

$$egin{aligned} \mu &= 10 \ p &= 0.5 \ \sigma^2 &= 4. \end{aligned}$$

• O termo *nula* é usado para indicar nenhuma mudança ou nenhum efeito.

Hipótese alternativa H_a

- É uma afirmativa de que o parâmetro tem um valor que, de alguma forma, **difere** da hipótese nula.
- Exemplos:

$$egin{aligned} \mu
eq 10 \ p > 0.5 \ \sigma^2 < 4. \end{aligned}$$

Tipos de hipótese

- Se você está fazendo um estudo e deseja usar um teste de hipótese para apoiar sua afirmativa, esta deve ser escrita de modo a se tornar a hipótese alternativa.
- Você nunca pode apoiar uma afirmativa de que um parâmetro **seja igual** a algum valor específico.
- Nesse contexto de se tentar apoiar o resultado de pesquisa, a hipótese alternativa é, algumas vezes, chamada de hipótese de pesquisa.

No exemplo anterior, a hipótese de pesquisa é a de que

a proporção de fêmeas é maior do que a proporção de machos

• Supondo inicialmente que a proporção de fêmeas é de 50% (p=0.5), então as **hipóteses estatísticas** são

$$H_0: p = 0.5 \quad {
m vs} \quad H_a: p > 0.5$$

- Com isso, deseja-se que a **hipótese nula** p=0.5 seja rejeitada, de modo que a **hipótese alternativa** p>0.5 seja apoiada.
- Apoiar a hipótese alternativa de que p>0.5 é o mesmo que apoiar a afirmativa de a proporção de fêmeas na população é maior do que a de machos.

Como a ciência funciona (atualmente)

Depois de testar a hipótese de que a proporção de fêmeas é maior do que a proporção de machos

- Nós continuaríamos subindo na "árvore de aprendizado" do Platt
- Para que lado da árvore nós vamos dependenderá se o resultado foi "estatisticamente significativo" ou não (i.e rejeitamos ou não a hipótese nula)

Os pontos chave para essa visão da ciência são:

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

Popper forneceu a filosofia, e Fisher, Pearson e colegas forneceram a estatística.

Como a ciência funciona (atualmente)

Se a hipótese nula for rejeitada

- Temos uma evidência científica
- Existe apoio à hipótese de pesquisa
- Hipóteses alternativas podem ser induzidas (sobe na árvore)
- Se esse experimento for realizado repetidas vezes, e o resultado permanece o mesmo, então essa hipótese de pesquisa vai adquirindo seu grau de confiança (corroboração)

Se a hipótese nula **não** for rejeitada

- Não existe apoio à hipótese de pesquisa
- O que fazer?
 - Reformular a hipótese?
 - Coletar mais dados?
 - o ?

Apesar de ser a mais utilizada na estatística (e na ciência em geral), a visão Popperiana do método científico é apenas uma entre várias alternativas.

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

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



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
 - o 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
 - o 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)



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*



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



De maneira geral

- o falsificacionismo de Popper
- a ciência normal de Kuhn
- a república da ciência de Polanyi
- o cinturão de hipóteses auxiliares de Lakatos

são descrições diferentes da mesma atividade científica.

Será muito raro que as ideias gerais como

- a teoria da evolução natural
- a teoria da relatividade

sejam verdadeiramente testadas.

Na verdade, a maior parte do trabalho de um cientista será em um nível muito mais "mundano", que, dependendo do filósofo, será

- falsificando hipóteses, segundo Popper
- quebrando paradigmas, segundo Kuhn
- fazendo ciência normal, segundo Polanyi
- testando hipóteses auxiliares, segundo Lakatos

Visões alternativas na estatística

A teoria estatística mais utilizada/ensinada atualmente é aquela baseada na visão filosófica da ciência de Popper

- Falseamento de uma única hipótese
- Árvore de aprendizado

Será que existem alternativas?

A inferência bayesiana (ou estatística bayesiana)

- Provê a teoria e as ferramentas para o teste de múltiplas hipóteses concorrentes
- Vai de encontro com a visão filosófica da ciência de Lakatos
 - O confronto deve ser entre hipóteses concorrente e os dados

Para simplificar a comparação vamos chamar essas duas visões de clássica/Popperiana e bayesiana/Lakatosiana

Inferência bayesiana

Baseada no **Teorema de Bayes**

$$egin{aligned} [H_i|D] &= rac{[DH_i]}{[D]} \ &= rac{[D|H_i][H_i]}{[D]} \ &\propto [D|H_i][H_i] \end{aligned}$$

- $[H_i|D]$ 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

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

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

Visões alternativas na estatística

Popperiana/clássica

$$P[D|H_0]$$

- Confronto entre dados e uma única hipótese
- Rejeita/não rejeita a hipótese nula
- Hipótese é **falsificada** (mas nunca provada)
- Experimentos críticos são aqueles que (geralmente) fazem a hipótese nula ser rejeitada
- As probabilidades são atribuídas apenas aos **conjuntos** de uma σ -algebra, e manipuladas por meio dos axiomas de Kolmogorov
 - Não é possível atribuir probabilidades para hipóteses

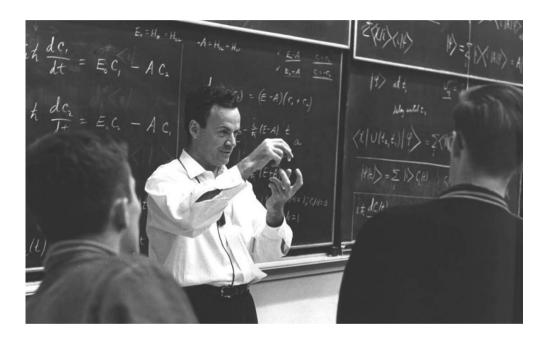
Lakatosiana/bayesiana

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

- Confronto entre dados e **múltiplas** hipóteses
- Grau de plausibilidade (ou confiança) em cada hipótese
- Hipóteses são classificadas em ordem crescente/decrescente de confiança
- Experimentos críticos são aqueles capazes de mudar os graus de confiança entre as hipóteses
- A probabilidade é considerada como extensão da lógica indutiva, e manipuladas por meio dos desideratos de Pòlya e as algebras de RT Cox
 - As hipóteses naturalmente recebem probabilidades (a priori e a posteriori)
- Para Lakatos, uma hipótese não pode ser rejeitada, a menos que exista uma alternativa melhor
- Para Popper, uma hipótese pode, sozinha, ser rejeitada/não rejeitada

As declarações da ciência não são sobre o que é verdade e o que não é verdade, mas declarações do que é conhecido com **diferentes graus de certeza**

@ProfFeynman

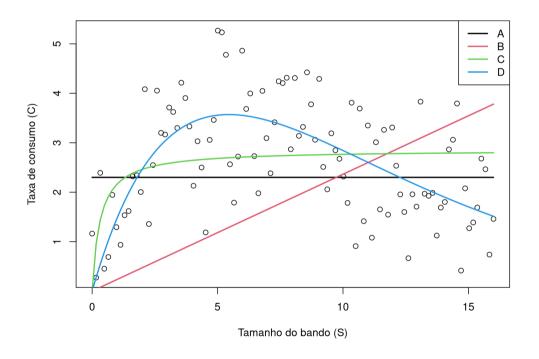


Exemplo

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
Α	C=a	Consumo independente do tamanho do bando (modelo nulo)
В	C=aS	Consumo é proporcional ao tamanho do bando
С	$C = rac{aS}{1+bS}$	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

$$P[D|H_0]$$

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

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

- 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

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

- O confronto seria entre os quatro modelos competidores e os dados
- Pode-se determinar o grau de plausibilidade de cada hipótese/modelo

$$egin{aligned} [H_A|D] &= p_A \ [H_B|D] &= p_B \ [H_C|D] &= p_C \ [H_D|D] &= p_D \end{aligned}$$

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**

Referências

Referências

Box, G. E., S. Hunter, and W. G. Hunter (2005). *Statistics for Experimenters: Design, Innovation, and Discovery.* 2nd ed. Cambridge: Wiley, p. 672.

Platt, J. R. (1964). "Strong inference". In: *Science* 146.3642, pp. 347-353. ISSN: 00368075.