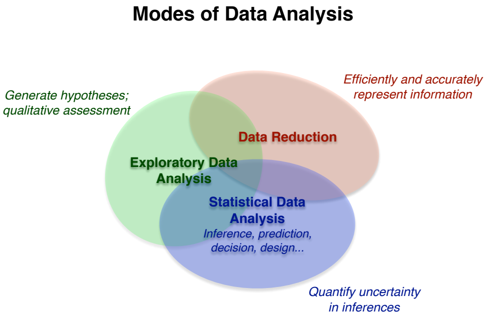
STSCI 4780/5780: Probability theory as logic

Tom Loredo, CCAPS & SDS, Cornell University

2022-01-27

Recap & motivation



Data analysis

Building and appraising data-based arguments about science/engineering models

Statistical inference

Quantifying uncertainty in measurement-based arguments about mathematical models, using probability

Two interpretations of probability

Frequentist probability

Probability describes *variability* and necessarily is *a property of an ensemble of outcomes/events* (*"frequentia"* = "crowd")

⇒ Frequentist statistics quantifies uncertainty via *long-run performance* of a procedure in repeated application across an ensemble

Bayesian probability

Probability describes *uncertainty*, understood as a measure of how strongly evidence supports hypotheses about *the case at hand*

⇒ Bayesian statistics quantifies uncertainty using probability theory as a generalization of logic, providing an abstract measure of *strength of a data-based argument*. Argument strength is related to long-run performance, but not always in a simple way.

Arguments

Argument $\equiv$ Collection of statements comprising an act of reasoning from *premises* to a *conclusion*

- All humans are mortal.
- Socrates is a human.

$\therefore$ Socrates is mortal.

- Today is Monday or Tuesday.
- Today is not Tuesday.

$\therefore$ Today is Monday.

- The NWS predicts sunshine tomorrow.

$\therefore$ Tomorrow will be sunny.

- Globally averaged atmospheric temperature measurements show a recent, rapid warming trend coinciding with industrialization.
- Average ocean surface temperatures show a similar trend.
- Polar ice is melting at unprecedented rates.
- Measurements of human-produced CO_2 show increases at times coinciding with changes in temperature trends.
- Indirect temperature and CO_2 measurements for centuries into the past show no similar periods.
- ...

$\therefore$ There is significant anthropogenic global warming likely to harm ecosystems and human civilization.

$$\Downarrow$$

- There is significant anthropogenic global warming likely to harm ecosystems and human civilization.
- The costs of damages X, Y, Z would be \$$A$, \$$B$, \$$C$.
- The costs of remediations would be...
- ...

$\therefore$ We should undertake the following remediations...

Propositional logic—propositions

"Logic can be defined as *the analysis and appraisal of arguments*"
—Gensler, *Intro to Logic*

Propositions

Proposition $\equiv$ Statement that may be true or false

$\mathcal{C}$: Universe can be modeled with ΛCDM

A : $\Omega_{\rm tot} \in [0.9, 1.1]$ (space-time is nearly flat)

B : Ω_Λ is not 0 (cosmic expansion is accelerating)

Language vs. meaning

These different *statements* express the same *proposition*:

- It is snowing.
- There is snow falling.
- Está nevando.

Building complex/compound propositions

Build complex propositions with simple propositions and logical operators/connectives:

Unary operation

$\mathcal{C}$: Universe can be modeled with ΛCDM

B : Ω_Λ is not 0

$\overline{B}$: "not B," i.e., $\Omega_\Lambda = 0$ (called *denial*)

Binary operations: connectives & compound propositions

$A \wedge B$: A and B are *both* true; *conjunction*

$A \vee B$: A or B is true, or both are; *disjunction*

Logical and/*or* often may differ from *colloquial* "and" / "or"
E.g., inclusive vs. exclusive "or"
Note: Symbols are not universal!

Truth values, truth tables

Truth value: Indicates whether a statement is true (T or 1) or false (F or 0)

Truth-functional statement: *Compound* statement whose truth is entirely determined by the truth values of its components

Conjunction is truth-functional: $A \wedge B$

Truth table: Lists the truth values of a compound statement as a function of the truth values of its components (various conventions!)

Truth-functional statements have truth tables determined entirely by their *form* (the pattern of operations/connectives)

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \wedge \mathcal{B}$
T	T	T
T	F	F
F	T	F
F	F	F

$\mathcal{A}$	$\wedge$	$\mathcal{B}$
T	T	T
T	F	F
F	F	T
F	F	F

$\wedge$	T	F
T	T	F
F	F	F

a	b	a×b
1	1	1
1	0	0
0	1	0
0	0	0

a	×	b
1	1	1
1	0	0
0	0	1
0	0	0

×	1	0
1	1	0
0	0	0

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \vee \mathcal{B}$
T	T	T
T	F	T
F	T	T
F	F	F

$\mathcal{A}$	$\vee$	$\mathcal{B}$
T	T	T
T	T	F
F	T	T
F	F	F

$\vee$	T	F
T	T	T
F	T	F

Not all connectives are truth-functional—propositional logic doesn't explain all rational reasoning!

Non-truth-functional example: "because" or "therefore" in the cause-effect (CE) sense, vs. the ground-consequent (GC) sense:

I am sad *because*$_{\text{CE}}$ it is raining.
It is raining *therefore*$_{\text{CE}}$ I am sad.
$A = C$ *because*$_{\text{GC}}$ $A = B$ and $B = C$.

S	R	S because R	R because S
T	T	?	?
T	F	F	F
F	T	F	F
F	F	F	F

Note that logic uses familiar words in a more precise way than we do in ordinary language; e.g., "or" and "because$_{\text{GC}}$" (*implication*)

Arguments

Argument: Assertion that an *hypothesized conclusion*, H, follows from *premises*, $\mathcal{P} = \{A, B, C, \ldots\}$ (take "," = "and")

Notation:

$$
\begin{aligned}
&\mathcal{P} \therefore H, \\
&\mathcal{P} \vdash H, \\
&H \mid \mathcal{P} \ : \qquad && \text{Premises } \mathcal{P} \text{ imply } H \\
& && H \text{ may be deduced from } \mathcal{P} \\
& && H \text{ follows from } \mathcal{P} \\
& && H \text{ is true given that } \mathcal{P} \text{ is true}
\end{aligned}
$$

Arguments may be considered to be (compound) propositions

Central role of arguments $\rightarrow$ special terminology for true/false:

- A true argument is *valid*
- A false argument is *invalid* or *fallacious*

Valid vs. sound arguments

- Today is Monday or Tuesday.
- Today is not Tuesday.

$\therefore$ Today is Monday.

Content vs. form

- An argument is *factually correct* iff all of its *premises are true* (it has "good content")
- An argument is *valid* iff its conclusion *follows from* its premises (it has "good form"); otherwise it's *invalid* or *fallacious*
- An argument is *sound* iff it is both *factually correct and valid* (it has good form and content)

Deductive logic (and probability theory) addresses *validity*

We want to make *sound* arguments. There is no formal approach for addressing factual correctness. $\rightarrow$ There is always a subjective element to an argument.

Factual correctness

Passing the buck

> Although logic can teach us something about validity and invalidity, it can teach us very little about factual correctness. The question of the truth or falsity of individual statements is primarily the subject matter of the sciences.
>
> — Hardegree, *Symbolic Logic*

An open issue

> To test the truth or falsehood of premises is the task of science. . . . But as a matter of fact we are interested in, and must often depend upon, the correctness of arguments whose premises are not known to be true.
>
> — Copi, *Introduction to Logic*

Premises

- *Facts* — Things known to be true, e.g. *observed data*
- *"Obvious" assumptions* — Axioms, postulates, e.g., Euclid's first 4 postulates (line segment b/t two points; congruency of right angles . . .)
- *"Reasonable" or "working" assumptions* — E.g., Euclid's fifth postulate (parallel lines)
- *Desperate presumption!*
- *Conclusions from other arguments* → chains of discovery

Every argument has a set of premises defining a *fixed context* in which the argument is assessed

Premises are considered "given"—if only for the sake of the argument!

Deductive and inductive inference

Deduction—Syllogism as a prototype

Premise 1: A implies H
Premise 2: A is true
Deduction: $\therefore H$ is true
$H|\mathcal{P}$ is valid

Induction—Analogy as a prototype

Premise 1: A, B, C, D, E all share properties x, y, z
Premise 2: F has properties x, y
Induction: F has property z
"F has z"$|\mathcal{P}$ is not strictly valid, but may still be rational (likely, plausible, probable); some such arguments are stronger than others

Boolean algebra (and/or/not over $\{0, 1\}$) quantifies deduction

Bayesian probability theory (and/or/not over $[0, 1]$) generalizes this to quantify the strength of inductive arguments

Deductive Logic

Assess arguments by decomposing them into parts via connectives, and assessing the parts:

Validity of $A \wedge B \mid \mathcal{P}$

	$A \mid \mathcal{P}$	$\overline{A} \mid \mathcal{P}$
$B \mid \mathcal{P}$	valid	invalid
$\overline{B} \mid \mathcal{P}$	invalid	invalid

Validity of $A \vee B \mid \mathcal{P}$

	$A \mid \mathcal{P}$	$\overline{A} \mid \mathcal{P}$
$B \mid \mathcal{P}$	valid	valid
$\overline{B} \mid \mathcal{P}$	valid	invalid

Representing Deduction With $\{0, 1\}$ Algebra

$V(H \mid \mathcal{P}) \equiv$ Validity of argument $H \mid \mathcal{P}$:

$$\begin{aligned} V &= 0 \rightarrow \text{Argument is } \textit{invalid} \\ &= 1 \rightarrow \text{Argument is } \textit{valid} \end{aligned}$$

Then deduction can be reduced to integer multiplication and addition over $\{0, 1\}$ (as in a computer):

$$\begin{aligned} V(A \wedge B \mid \mathcal{P}) &= V(A \mid \mathcal{P})\, V(B \mid \mathcal{P}) \\ V(A \vee B \mid \mathcal{P}) &= V(A \mid \mathcal{P}) + V(B \mid \mathcal{P}) - V(A \wedge B \mid \mathcal{P}) \\ V(\overline{A} \mid \mathcal{P}) &= 1 - V(A \mid \mathcal{P}) \end{aligned}$$

This is a *Boolean algebra* (with a specific interpretation)

Boolean algebra provides a “calculus for deductive inference”

Representing Induction With $[0, 1]$ Algebra

$P(H|\mathcal{P}) \equiv$ strength of argument $H|\mathcal{P}$

$$\begin{aligned} P &= 0 \rightarrow \text{Argument is } \textit{invalid}\text{; premises imply } \overline{H} \\ &= 1 \rightarrow \text{Argument is } \textit{valid} \\ &\in (0,1) \rightarrow \text{Degree of implication/deducibility} \end{aligned}$$

Mathematical model for induction

$$\text{`AND' (product rule):} \quad \begin{aligned} P(A \wedge B|\mathcal{P}) &= P(A|\mathcal{P})\, P(B|A \wedge \mathcal{P}) \\ &= P(B|\mathcal{P})\, P(A|B \wedge \mathcal{P}) \end{aligned}$$

$$\text{`OR' (sum rule):} \quad \begin{aligned} P(A \vee B|\mathcal{P}) &= P(A|\mathcal{P}) + P(B|\mathcal{P}) \\ &\quad - P(A \wedge B|\mathcal{P}) \end{aligned}$$

$$\text{`NOT':} \quad P(\overline{A}|\mathcal{P}) = 1 - P(A|\mathcal{P})$$

The Product Rule

We simply promoted the $V(\cdot)$ algebra to real numbers; the only thing we changed is part of the product rule:

$$
\begin{aligned}
V(A \wedge B|\mathcal{P}) &= V(A|\mathcal{P})\, V(B|\mathcal{P}) \\
P(A \wedge B|\mathcal{P}) &= P(A|\mathcal{P})\, P(B|A \wedge \mathcal{P})
\end{aligned}
$$

Suppose A implies B (i.e., $B|A \wedge \mathcal{P}$ is valid). Then we don't expect $P(A \wedge B|\mathcal{P})$ to differ from $P(A|\mathcal{P})$.

In particular, $P(A \wedge A|\mathcal{P})$ must equal $P(A|\mathcal{P})$!

Such qualitative reasoning satisfied early probabilists that the sum and product rules were worth considering as axioms for a theory of quantified induction. But Boole and Venn criticized the rules in the late 1800s; this partly motivated early 20th century focus on frequency-based interpretations of probability, since these rules trivially are true for frequencies (at least for finite sequences).

Firm foundations

Starting ca. 1950, many different formal lines of argument have been devised that *derive* induction-as-probability from various simple and appealing requirements:

- Consistency with logic + internal consistency (Cox; Jaynes)
- "Coherence"/safe betting (Ramsey; de Finetti; Savage; Lindley)
- Complete class theorems in decision theory (Wald)
- Algorithmic information theory (Rissanen; Wallace & Freeman)
- Optimal information processing (Zellner)
- Avoiding problems with frequentist methods:
  - Avoiding recognizable subsets (Cornfield)
  - Avoiding stopping rule problems → likelihood principle (Birnbaum; Berger & Wolpert — does not quite single out Bayes)

Pierre Simon Laplace (1819)

Probability theory is nothing but *common sense reduced to calculation*.

James Clerk Maxwell (1850)

They say that Understanding ought to work by the rules of right reason. These rules are, or ought to be, contained in Logic, but the actual science of *Logic is conversant at present only with things either certain, impossible, or entirely doubtful*, none of which (fortunately) we have to reason on. Therefore *the true logic of this world is the calculus of Probabilities*, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable man's mind.

Harold Jeffreys (1931)

If we like there is no harm in saying that a probability expresses a degree of reasonable belief. . . . 'Degree of confirmation' has been used by Carnap, and possibly avoids some confusion. But whatever verbal expression we use to try to convey the primitive idea, this expression cannot amount to a definition. *Essentially the notion can only be described by reference to instances where it is used.* It is intended to express *a kind of relation between data and consequence* that habitually arises in science and in everyday life, and the reader should be able to recognize the relation from examples of the circumstances when it arises.