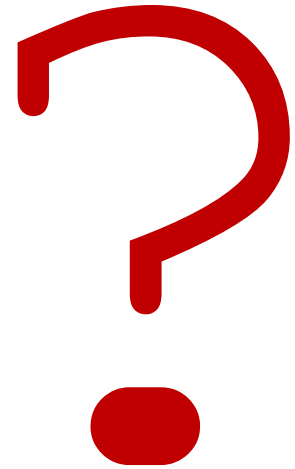
Mathematical Conceptions of the Unknown (an omnivore's guide)

Matthew Roughan

matthew.roughan@adelaide.edu.au

prossibly@gmail.com

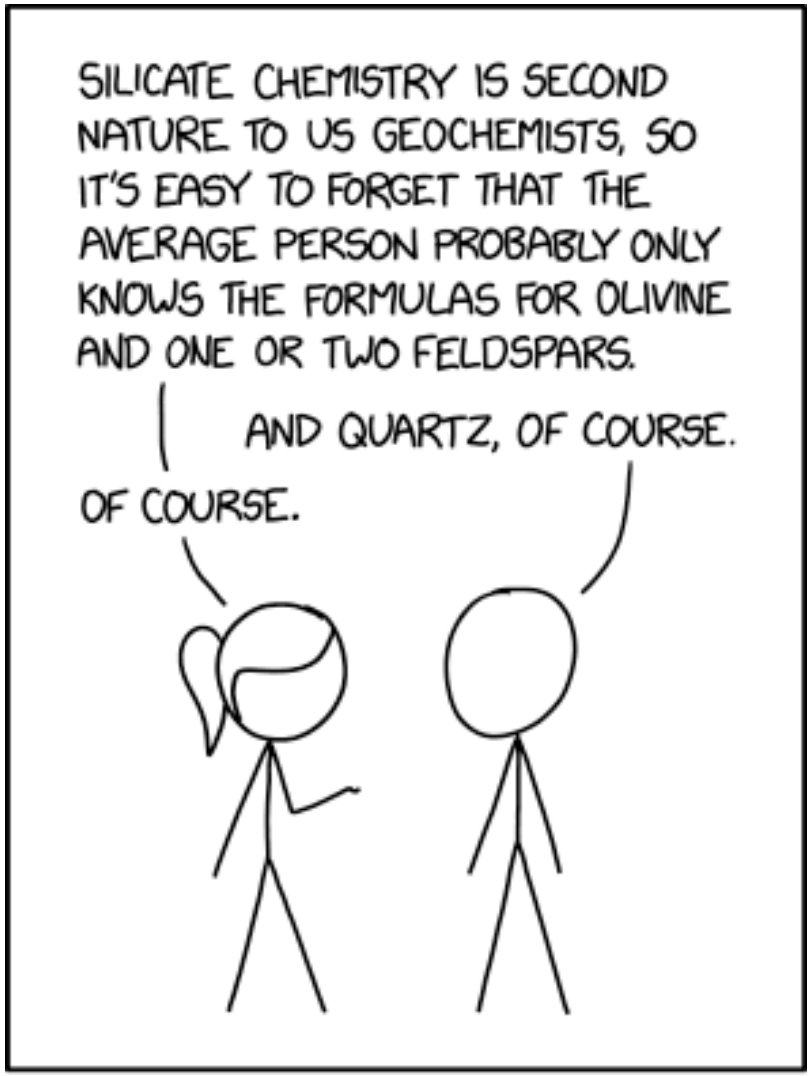


First an apology or two

Terminology

- Risk – quantifiable (probabilistically)
- Uncertainty – not quantifiable

But I haven't used that distinction for most of my life so I will probably mess up at some point.



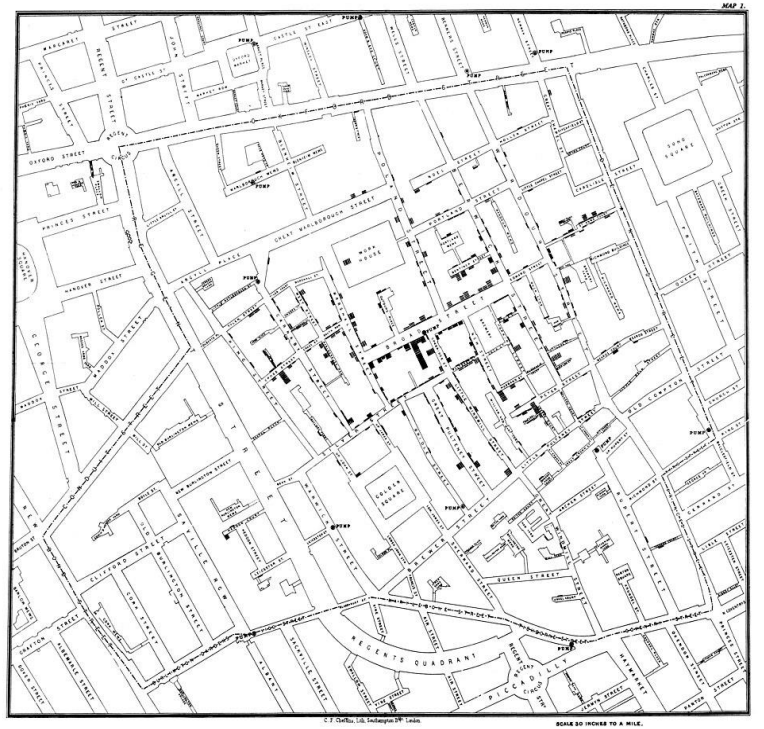
https://xkcd.com/2501/

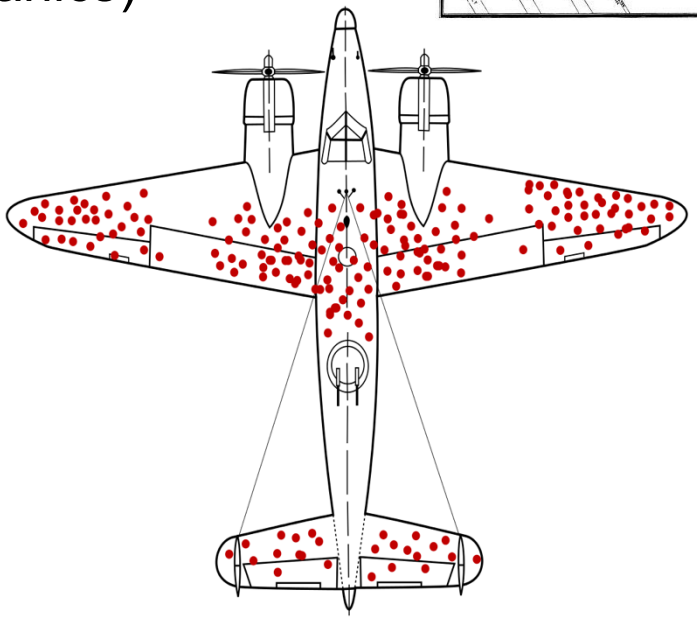
We have a lot of words for snow

accident, belief, bet, capricious, chance, coincidence, conjectural, contingency, credence, degrees, erratic, expectation, fate, fickle, fluke, fortuitous, fortune, frequency, gamble, happenstance, hazard, hedge, hunch, iffy, incalculable, jeopardy, likelihood, long-shot, luck, odds, opportunity, out-of-the-blue, peril, plausibility, possibility, probability, propensity, prospect, questionable, randomness, risk, serendipity, shot, speculate, stochastic, surprising, synchronicity, unanticipated, uncertainty, uncertainty, unconfirmed, unexpected, unfixed, unforeseen, unknown, unlooked-for, unplanned, unpredictable, unreliable, unstable, unsure, wager, windfall, wyrd

Probability&Stats

- **Arguably** the most successful conception of uncertainty (aka risk)
- Definitely one of the most successful mathematical tools
  - Physics (statistical physics, quantum mechanics)
  - Information Theory (and all comms)
  - Operations Research
  - Finance
  - Games
  - Medicine
  - AI
  - *Etc., etc., etc.,*

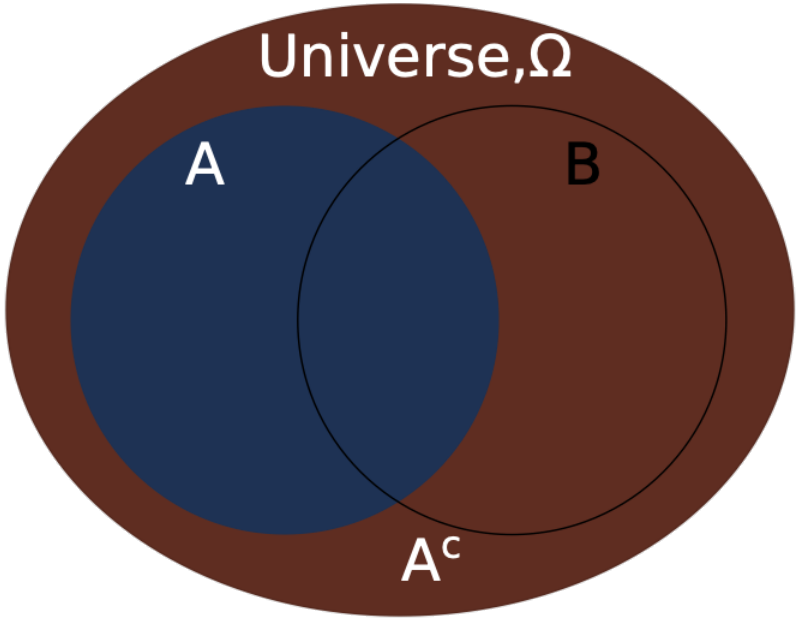


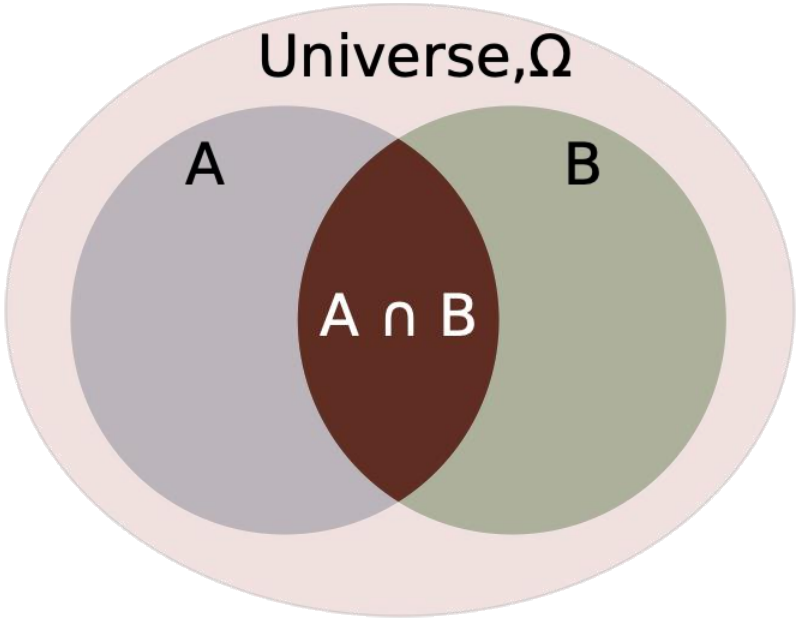


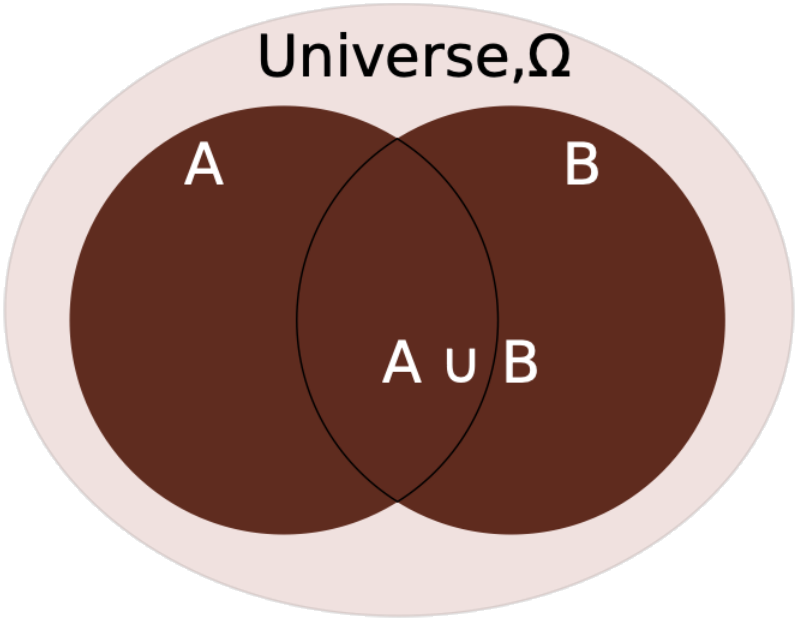
Sometimes Math works, but you have to frame the problem the right way
See Abraham Wald and Survivorship Bias
https://cameronmoll.com/journal/abraham-wald-red-bullet-holes-origin-story

Set Theory

- Many of the ideas of mathematics have set-theory underneath







Philosophy of Probability

Probability is the most important concept in modern science, especially as nobody has the slightest notion of what it means
Bertrand Russell, 1929

- Probability isn't built on rock-solid foundations

- Many philosophical conceptions, but two main branches
  - *frequentist* or *objective* or *aleatory probability*
    - Probabilities are real, objective, measurable and relate to outcome of many events
  - *epistemic* or *Bayesian probability*
    - Probabilities are our constructs, subjective, and describe our state of knowledge

- Even though probability is a theory fundamentally based on randomness, we frequently apply it to deterministic questions

I am ignoring classical or intuitionist probability

Probability does not exist
Bruno de Finetti, Theory of Probability (1970)

But Probability&Stats

- Works!
- It has mathematical rules that we can manipulate according to well-known axioms and theorems and algorithms
- We can and do solve many real problems
  - *e.g.,* use it as a component of opt.
- So, what's my issue?

- I think we over-extend ourselves
- Success leads to hubris, and "When the Gods wish to punish us, they answer our prayers." (Oscar Wilde)

"There are lies, damn lies, and statistics. We're looking for someone who can make all three of these work for us."

The Unreasonable Effectiveness of Mathematics in the Natural Sciences

Richard Courant Lecture in Mathematical Sciences delivered at New York University, May 11, 1959

EUGENE P. WIGNER
Princeton University

"and it is probable that there is some secret here which remains to be discovered." (*C. S. Peirce*)

There is a story about two friends, who were classmates in high school, talking about their jobs. One of them became a statistician and was working on population trends. He showed a reprint to his former classmate. The reprint started, as usual, with the Gaussian distribution and the statistician explained to his former classmate the meaning of the symbols for the actual population, for the average population, and so on. His classmate was a bit incredulous and was not quite sure whether the statistician was pulling his leg. "How can you know that?" was his query. "And what is this symbol here?" "Oh," said the statistician, "this is π." "What is that?" "The ratio of the circumference of the circle to its diameter." "Well, now you are pushing your joke too far," said the classmate, "surely the population has nothing to do with the circumference of the circle."

The Unreasonable Effectiveness of Mathematics, Hamming, 1980

The Unreasonable Effectiveness of Number Theory, Burr, 1993

The Unreasonable Effectiveness of Mathematics in Molecular Biology, Lesk, 2000

The unreasonable **ineffectiveness** of mathematics in economics Velupilla, 2005

The unreasonable effectiveness of deep learning in artificial intelligence Sejnowski, 2020

Probability is too important to be left to the experts
Richard Hamming, The Art of Probability (1991)

Probability problems

- *ignotum per ignotius* (explaining the unknown with the more unknown)
  - Probability is hard – witness the many fallacies and paradoxes

- Neither aleatory or epistemic probability cope with all of the questions that concern us
  - `How likely is Bob to be related to Alice?`
  - `How likely are the Democrats to win the next election?`
  - `Is my dice fair?`
  - `What is the probability my nuclear PP will explode?`

  Particularly problematic for small probabilities (rare events)

- We quickly get to a point where we want to talk about
  "probability distributions of probabilities"

If you already know what recursion is, just remember the answer. Otherwise, find someone who is standing closer to Douglas Hofstadter than you are; then ask him or her what recursion is.
Andrew Plotkin

Probabilities of Probabilities

- In subjective, Bayesian probability this is an intrinsic

- In simple terms, we treat a "probability" like any other property that can be measured, *e.g.*, the speed of light
  - Measurements contain errors
  - We model the errors and arrive at a distribution

- But now we have a loop in our definition
  - What happens if I try to assign probabilities to the probabilities of probabilities of ...

It's turtles all the way down ...

Probability problems

- There are many settings where probabilities are little more than badly informed guesses
  - Space shuttle Challenger (1986) (RP Feynman)
  - GFC (Global Financial Crisis) 2007 https://www.rba.gov.au/education/resources/explainers/the-global-financial-crisis.html
  - Intransitive dice
  - Bay of Pigs (1961) https://www.scientificamerican.com/article/why-probability-probably-doesnt-exist-but-its-useful-to-act-like-it-does/
  - Littlewood's Law of Miracles
- People are in the loop
  - People make mistakes
  - Probability is hard to reason about
- Largest source of error in a given settings is often not math it is in the underlying
  - Assumptions
  - Knowledge
  - Framing

And We Take it Too Far

- Secretaries Problem (1949, M Flood)
  - Well known mathematical recreations problem (1960)
  - Aim to hire the best secretary
  - Can only go forward, not back
  - What is the optimal selection criteria (*n/e* law)
- It all goes wrong when people apply it to real life
  - Dating for instance

https://sites.lsa.umich.edu/mje/2024/03/13/mathematically-and-economically-optimal-dating/

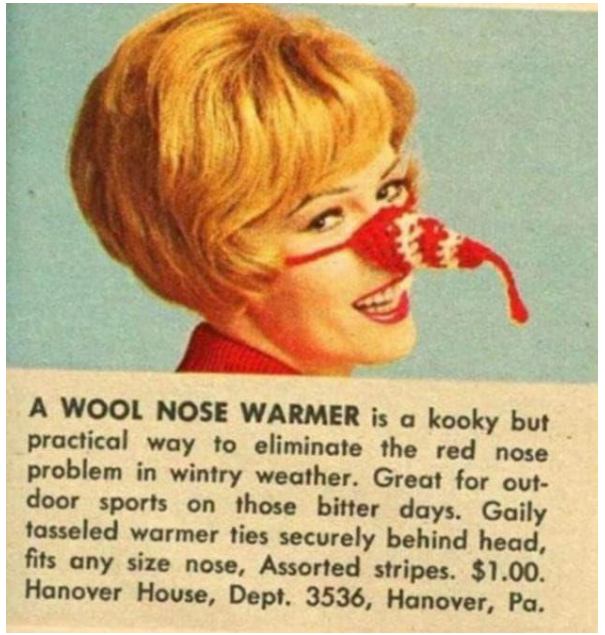
And Probability is Hard

https://xkcd.com/221/

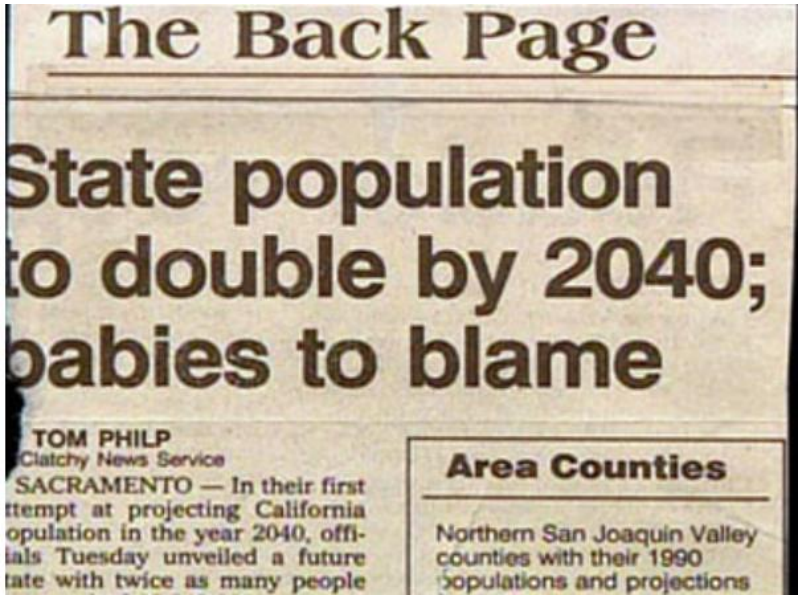
```
int getRandomNumber()
{
    return 4;  // chosen by fair dice roll.
               // guaranteed to be random.
}
```

It's like the tale of the roadside merchant who was asked to explain how he could sell rabbit sandwiches so cheap. "Well" he explained, " I have to put some horse- meat in too. But I mix them 50:50. One horse, one rabbit.
Darrel Huff, How to lie with statistics

Two envelopes problem



A WOOL NOSE WARMER is a kooky but practical way to eliminate the red nose problem in wintry weather. Great for outdoor sports on those bitter days. Gaily tasseled warmer ties securely behind head, fits any size nose, Assorted stripes. $1.00. Hanover House, Dept. 3536, Hanover, Pa.



The Back Page

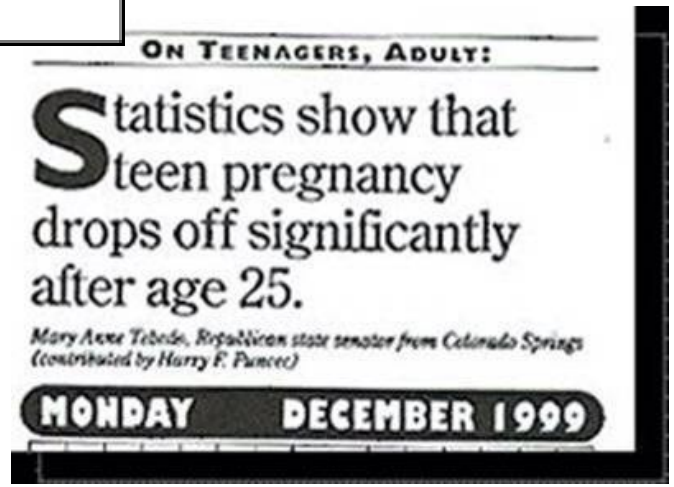
State population to double by 2040; babies to blame

TOM PHILP
Clatchy News Service

SACRAMENTO — In their first ttempt at projecting California opulation in the year 2040, offiials Tuesday unveiled a future tate with twice as many people

Area Counties

Northern San Joaquin Valley counties with their 1990 opulations and projections



ON TEENAGERS, ADULT:

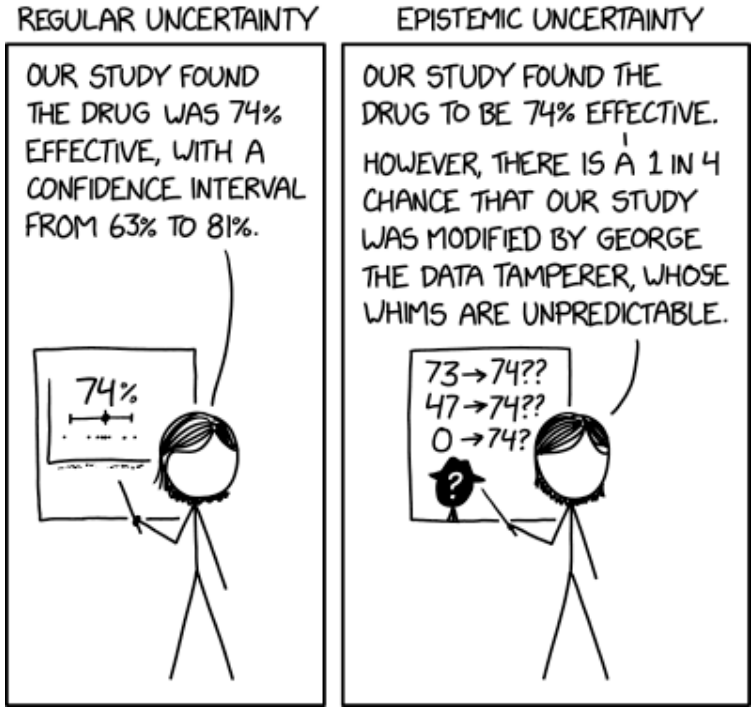
Statistics show that teen pregnancy drops off significantly after age 25.

Mary Anne Tebedo, Republican state senator from Colorado Springs (contributed by Harry F. Pancco)

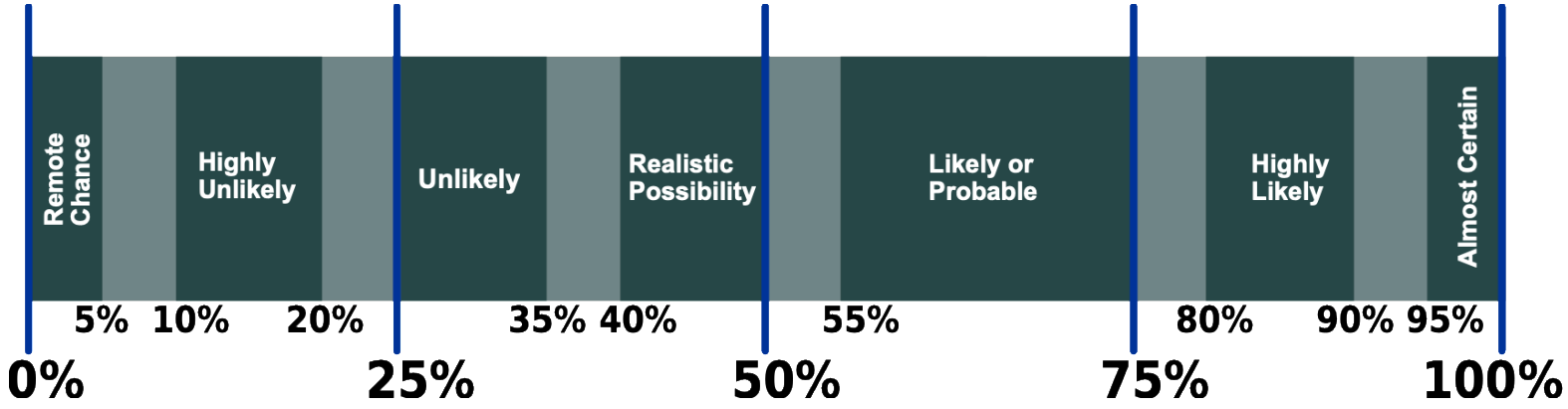
MONDAY DECEMBER 1999

The novice goes astray and says, "The Art failed me."
The master goes astray and says, "I failed my Art."

And then we get confused



https://xkcd.com/2440/



UK Defence Intelligence – Communicating Probabilities

https://www.gov.uk/government/news/defence-intelligence-communicating-probability

And what does probability 0.0 mean?

What else is there?

We could improve our measurements

We could improve probability

Or … we could do something else

Measure what is measurable, and
make measurable what is not so.
Galileo Galilei

Some men went fishing in the sea with a net, and upon examining what they caught they concluded that there was a minimum size to the fish in the sea.
Sir Arthur Eddington

Information Theory and Entropy

- As an aside entropy is used to measure uncertainty (risk)
- But it is about measuring the amount of uncertainty, not the details
- It is defined in terms of probabilities, and pitched towards aleatory probability
- It has its own variants, *e.g.,* Shannon, Renyi and Gibb's entropies

Generalisations of Probability

Goes back to Boole, 1854

- Imprecise or interval probabilities
  - Lower and upper probabilities $\underline{P}(A)$ and $\overline{P}(A)$
    - Maybe based on the odds you would offer, vs take
  - Some axioms of common probability have to be loosened (additivity)
  - Maybe more *accurate* than a spuriously precise representation
  - But how come we can be precise about two numbers when we had trouble with just one!

$$0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1$$

$$\underline{P}(A^c) = 1 - \overline{P}(A)$$

Possibility Theory

- Another type of interval theory
  - Lower bound is called **possibility**
  - Upper bound called **necessity**

    $P(A \cup B) = \max(P(A), P(B))$ (for disjoint A, B)

    - Replaces additivity

    - Necessity $N(A) \leq P(A)$

      $N(A) = 1 - P(A^c)$

Demster-Shafer theory

- A theory for imprecise or interval probabilities
  - Connect up various ideas, *e.g.*, intervals and Bayes
  - Rather a large body of various ideas beyond D-S
- Formalism
  - Mass function assigned to all sets of outcomes
  - From mass derive belief (support) and plausibility s.t.

$$\text{belief} \leq \text{probability} \leq \text{plausibilty}$$

Demster-Shafer theory

- X is the universe, *e.g.*, $X = \{a,b\}$
- Set of all subsets is $2^X = \{ \emptyset, \{a\}, \{b\}, X \}$
- Mass $m: 2^X \rightarrow [0,1]$ such that $m(\emptyset)=0$ and total mass $=1$
- Belief and plausibility
  - $bel(A)$ = sum of mass for all subsets of A
  - $pl(A)$ = sum of mass of subsets that are disjoint from A

    $pl(A) = 1 - bel(A^c)$
- Now we need to build a calculus for working with these

Demster-Shafer theory

Cat in the box

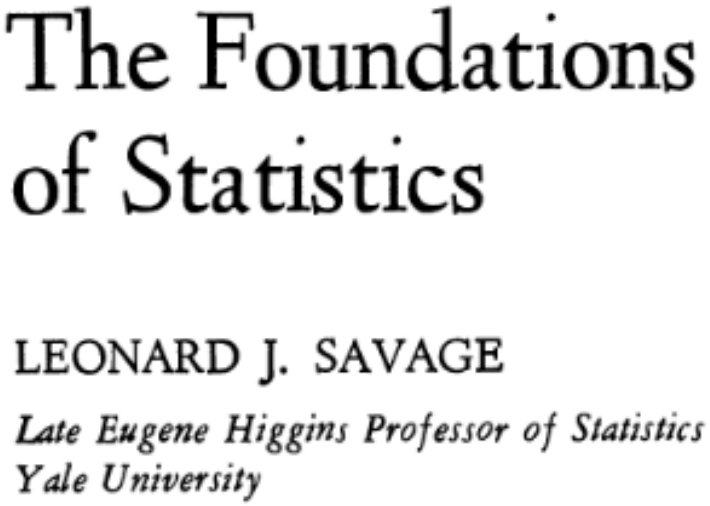
State	Mass	Belief	Plausibility
Neither Alive or Dead	0.0	0.0	0.0
Alive	0.2	0.2	0.5
Dead	0.5	0.5	0.8
Either	0.3	1.0	1.0

Demster-Shafer theory

- Good for fusing disparate “preferences” or “beliefs”
  - Use in data fusion
- Can produce counter-intuitive results

  Suppose that one has two equi-reliable doctors and one doctor believes a patient has either a brain tumor, with a *probability* of 0.99; or meningitis, with a *probability* of only 0.01. A second doctor believes the patient has a concussion, with a *probability* of 0.99, and believes the patient suffers from meningitis, with a *probability* of only 0.01. Applying Dempster's rule to combine these two sets of masses of belief, one gets finally m(meningitis)=1, *i.e.,* **the meningitis is diagnosed with 100 percent of confidence, even though neither doctor really believes it is likely.**

Savage Axioms


The Foundations of Statistics

LEONARD J. SAVAGE

Late Eugene Higgins Professor of Statistics
Yale University

- States ω which are unknown, sets called events
- X denotes a set of consequences
- Acts f,g,… map states to consequences
- Preference f≥g over acts means we prefer act f

Axiom 1: **weak ordering:** that is, ≥ creates a total order

Axiom 2: **sure things:** if you prefer f over g regardless of whether and event E happens, then it doesn't matter what the consequence is when E doesn't happen

AIM: build utility into the interpretation of probability&statistics (we kind-of do this with Bayes in the gambling interpretation)

but folks are not like that

Ambiguity Aversion

RISK, AMBIGUITY, AND THE SAVAGE AXIOMS*

By DANIEL ELLSBERG

I. Are there uncertainties that are not risks? 643. — II. Uncertainties that are not risks, 647. — III. Why are some uncertainties not risks? — 656.

- Preference for known risks over unknown “risks”
- Classic example is Ellsberg Paradox
  - Two urns
    - Urn A has 50 red and 50 black balls
    - Urn B has 100 balls, some black and some red
    - Bet 1 = red, Bet 2 = black
  - People
    - Are indifferent between 1 and 2
    - Prefer A to B

So Probability has Failed Us, What Next?

- If there is a problem, maybe it is a problem in the underlying logic we are applying
- Trad Logic has only two values: TRUE & FALSE
- It has its share of paradoxes and problems

Three-valued logic

See “The Ternary Calculating Machine of Thomas Fowler (1840)”, Glusker et al., 2005.

Ternary (trinary or triadic or trivalent or trilean) logic also allows “sometimes” or “maybe” or “I don’t care” or “I don’t know” or “undecided” or “missing”

NOT(A)

A	¬A
F	T
U	U
T	F

AND(A, B)

A ∧ B	B = F	B = U	B = T
A = F	F	F	F
A = U	F	U	U
A = T	F	U	T

OR(A, B)

A ∨ B	B = F	B = U	B = T
A = F	F	U	T
A = U	U	U	T
A = T	T	T	T

It is used (a lot)

- SQL (TRUE, FALSE, NULL)
- TCAMs in Internet Routers (Hardware Ternary Content Addressable Memory)
- Time Enough for Love, R Heinlein, 1982

And More

- Four-valued logic
  - True
  - False
  - Both
  - Neither
  - IEEE 1364 standard
  - Catuskoti in Buddhist logic

- 9-valued
  - IEEE 1164 standard

Fuzzy logic

Side Note: Is mathematics a subset of logic (Frege, Russel) or is logic a subset of mathematics (Pierce, Wittgenstein)?

Trad logic doesn't work with statements like

`That man is tall.`

Tall is not assessable as True or False because it is subjective

Fuzzy logic allows truth values to lie between 0 and 1

INFORMATION AND CONTROL 8, 338–353 (1965)

Cited >167,000

Fuzzy Sets*

L. A. ZADEH

Department of Electrical Engineering and Electronics Research Laboratory, University of California, Berkeley, California

A fuzzy set is a class of objects with a continuum of grades of membership. Such a set is characterized by a membership (characteristic) function which assigns to each object a grade of membership ranging between zero and one. The notions of inclusion, union, intersection, complement, relation, convexity, etc., are extended to such sets, and various properties of these notions in the context of fuzzy sets are established. In particular, a separation theorem for convex fuzzy sets is proved without requiring that the fuzzy sets be disjoint.

Fuzzy Sets

Why did the programmer get fired? Because he kept saying things like, "The alarm is kind of loud" and "The oven is somewhat hot."

Trad set membership is 0 or 1

Fuzzy sets come with a grade-of-membership mapping, e.g., OLD

m(10) = 0.00 (not included)
m(20) = 0.01
m(30) = 0.05
m(40) = 0.10 (fuzzy member)
m(50) = 0.15
m(57) = 0.00
m(60) = 0.20
...
m(100) = 1.0 (fully included)

New set rules

Complement

$m_A^c(x) = 1 - m_A(x)$

containment

$m_A(x) <= m_B(x)$ for all x

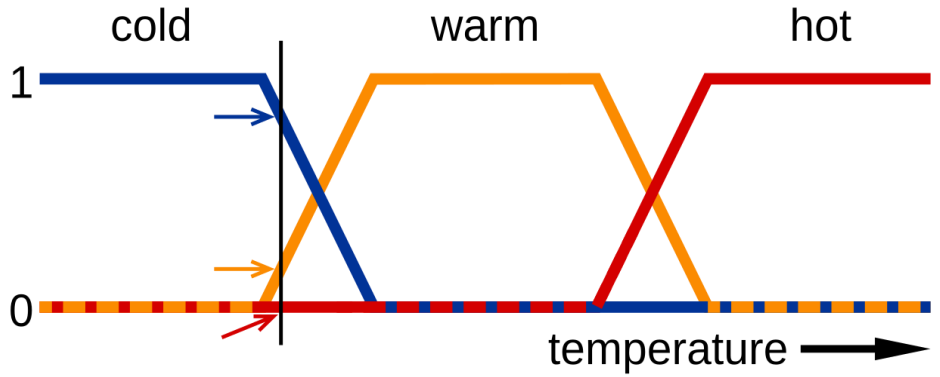
union C = A U B

$m_C(x) = \max(m_A(x), m_B(x))$

Intersection C = A ∩ B

$m_C(x) = \min(m_A(x), m_B(x))$

Standard forms for functions (simple curves)



Fuzzy Control (its killer app)

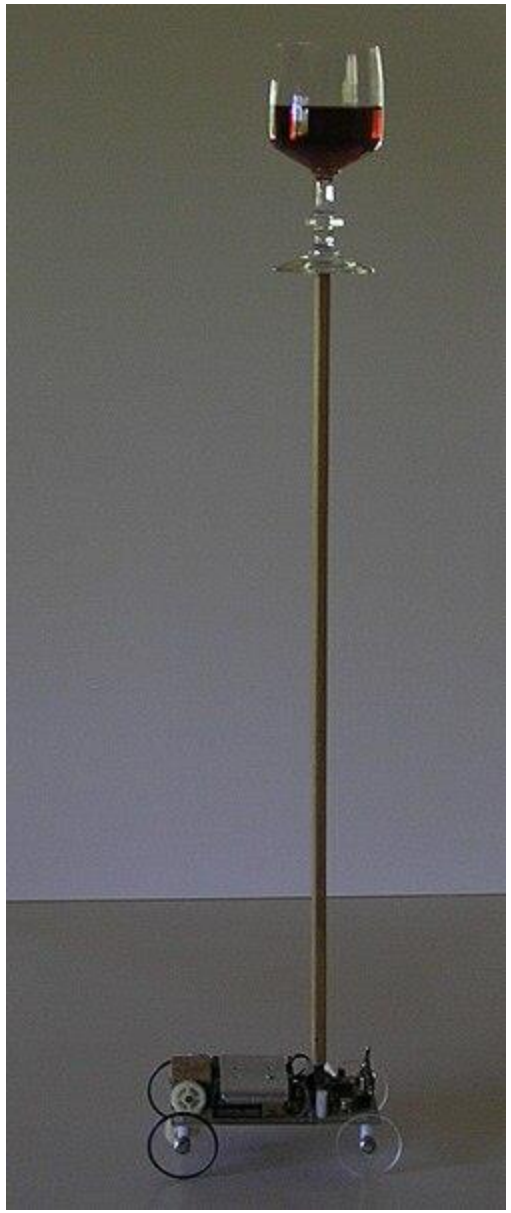
Fuzzy logic is used in many controllers

- Early 80s, Japan, “smart” devices, *e.g.*
  - Washing machines that adjust to load weight
  - Autofocus on cameras
  - A/C

```
IF (temperature is "cold")
THEN turn (heater is "high")
```

- Inverted pendulum example
  - There is a mathsy way to solve
  - But what happens to maths if inputs are wrong?

- Fuzzy controllers tend to be flexible, robust and computationally efficient (if not optimal), and intuitive



But we are going into deeper waters

- Are we just falling into the same trap?
- We can instead work with the understanding of the limits of our knowledge
  - Instead of trying to model a distribution, think about the worst case
    - This is common in, for instance, complexity analyses
  - Fudge factors
  - Artificial distinctions, *e.g.,* Defence Intelligence – communicating probability

Knightian Uncertainty

- “Risk, Uncertainty, and Profit,” Frank Knight (economist), 1921
- Not really a mathematical, quantifiable idea
- Makes the distinction
  - Risk applies to situations where we do not know the outcome of a given situation, but can accurately measure the odds
  - Uncertainty, on the other hand, applies to situations where “there is no scientific basis on which to form any calculable probability whatever.”

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4662711

What do we need?

- Definitions
  - That are consistent (maybe axiomatic)
  - That are quantitative/measurable
- Algorithms/methods
  - Ways to estimate from data
  - Ways to compute and build up more complex models
- Semantics
  - What do results mean?
  - Philosophical underpinnings
- Decision support
  - Improves decisions
  - Verification is possible

Too much of this fields seems to think

- All we need is to use different words, e.g., plausibility instead of probability
- Defining ever more complicated math is doing anything helpful

https://procedural-generation.tumblr.com/post/124657726768/substrate-2003-shortly-after-the-turn-of-the

Conclusion

- Long live probability
- Down with probability

No plan survives contact with the enemy
Attrib. Field Marshal Helmuth von Moltke

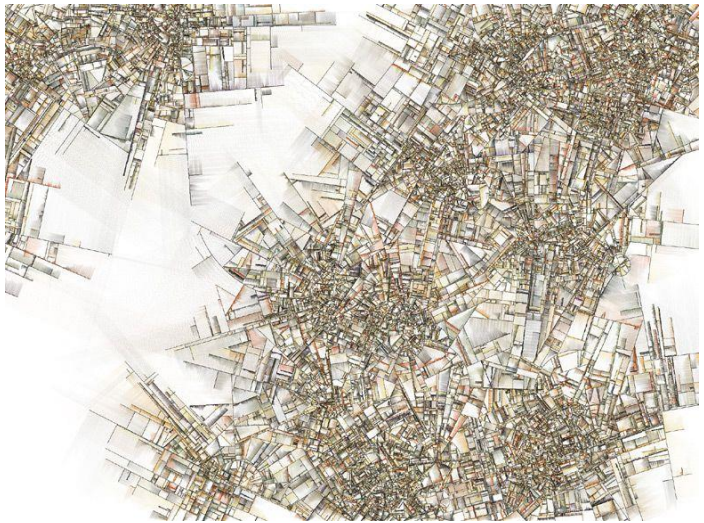
https://quoteinvestigator.com/2021/05/04/no-plan/

Plans are worthless, but planning is everything
Dwight D. Eisenhower.

https://quoteinvestigator.com/2017/11/18/planning/

640K ought to be enough for anybody
Bill Gates, 1981

"It would appear that we have reached the limits of what it is possible to achieve with computer technology, although one should be careful with such statements, as they tend to sound pretty silly in 5 years."
John Von Neumann, c1949



And BTW, long-trad of using randomness in art

- Visual – Dada, surrealism, …
- Music – John Cage, …

Jean (Hans) Arp
Untitled (Collage with Squares Arranged according to the Law of Chance) 1916–17
https://www.moma.org/collection/works/37013

Word stuff

- **Aleatory** from Latin *aleatorius* "pertaining to a gamester" (from *alea*, maybe knuckle bone used as dice)
- **Stochastic** from Greek *stokhastikos* "able to guess, conjecturing"
- **Random** from Old French *randon* "rush, disorder, force, impetuosity"
- **Hazard** (maybe) Arabic *yasara* “he played at dice”, to Spanish, to Old French
- **Chance** from Latin *cadentia* “that which falls out” through Old French
- **Probability** from Latin *probabilitas* “credibility” through Old French
- **Statistics**, Gottfried Achenwall (1948) from Italian *statista* "one skilled in statecraft”
- **Fortune** from Latin *fortuna* “chance, fate, good luck”
- **Epistemological** from the Greek words *episteme* "knowledge"

https://www.etymonline.com/search?q=hazard

Some more refs

- Oberguggenberger, M. (2005). The mathematics of uncertainty: models, methods and interpretations. In: Fellin, W., Lessmann, H., Oberguggenberger, M., Vieider, R. (eds) Analyzing Uncertainty in Civil Engineering. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-26847-2_4
- Wierman, MJ, An introduction to the mathematics of uncertainty, 2010, Honours thesis, Creighton University, https://www.creighton.edu/sites/default/files/2022-03/MOU.pdf