Session #23

CMSC 409: Artificial Intelligence

http://www.people.vcu.edu/~mmanic/

Virginia Commonwealth University,
Fall 2023,
Dr. Milos Manic
(mmanic@vcu. edu)

1

CMSC 409: Artificial Intelligence

Session # 23

Topics for today

- Announcements
- Previous session review
- Probability Theory
  - *Probabilistic logic, theory, inferencing*
  - *Related fields, applications*
  - *Definitions, interpretations, epostemology, resources*
- Subjective probability - Bayes' Rule
  - *History, definition*
  - *P(R|X) vs. P(R&X), examples*
  - *Who to reward problem*
  - *Principle of Conditionalization*
  - *Monty Hall problem*

2

CMSC 409: Artificial Intelligence

Announcements

Session # 23

- IMPORTANT:
  - *Course materials (slides, assignments) are copyrighted by instructor & VCU. Sharing/posting/chatGPT/similar is copyright infringement and is strictly prohibited. Such must be immediately reported.*
- Canvas
  - *Prev. session slides updated*
- TAs
  - *Victor Cobilean <cobileanv@vcu.edu>, Harindra Sandun Mavikumbure mavikumbureh@vcu.edu*
  - *TA office hours: Thursdays, 3:30 - 4:30pm (Zoom)*
- Project #4
  - *Deadline is Nov. 9*
- Paper (optional)
  - *The 4th draft (final submission) due Nov. 28*
  - *In addition to previous draft, it should contain a technique (or selection thereof), you plan on using to solve the selected problem (check out the class paper instructions for the 4th draft)*
- Subject line and signature
  - *Please use [CMSC 409] Last_Name Question*

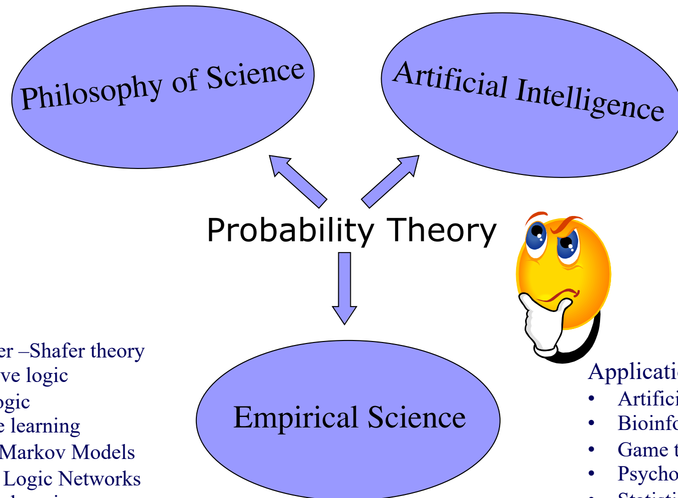
3

- *Probability Theory*
  - *probabilistic logic, theory, inferencing*
  - *related fields, applications*
  - *definitions, interpretations, epostemology, resources*
- *Subjective probability - Bayes' Rule*
  - *history, definition*
  - *P(R|X) vs. P(R&X), examples*
  - *Who to reward problem*
  - *Principle of Conditionalization*
  - *Monty Hall problem*

4

Probability Theory

Interpretations…



Related:
- Dempster –Shafer theory
- Subjective logic
- Fuzzy logic
- Machine learning
- Hidden Markov Models
- Markov Logic Networks
- Machine learning

Applications:
- Artificial Intelligence
- Bioinformatics
- Game theory
- Psychology
- Statistics

Probability Theory

Pierre-Simon Laplace (1749–1827).

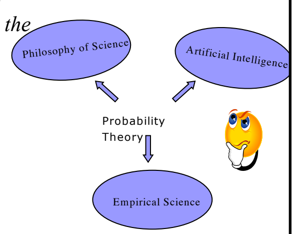
Interpretations…

Probability: *The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible.*

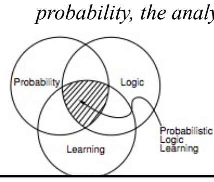
— Pierre-Simon Laplace, A Philosophical Essay on Probabilities, 1814

We cannot precisely predict random events, but..

Probabilistic logic *(also probability logic and probabilistic reasoning) combines the capacity of probability theory to handle uncertainty with the capacity of deductive logic to exploit structure.*



Probability theory *is the branch of mathematics concerned with probability, the analysis of random phenomena.*

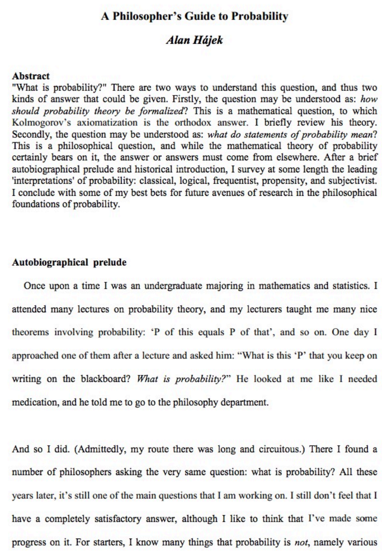


Luc De Raedt, Kristian Kersting,
ACM SIGKDD, July 2003,
http://people.csail.mit.edu/kersting/pll_icml04/pll.pdf

Probability Theory

Resources

- Resources will categorize probability theory into: classical, logical, frequentist, propensity, and subjectivist….
- A Philosopher's Guide to Probability, Alan Hájek: http://philosophy.anu.edu.au/sites/default/files/documents/Philosophers%20Guide%20to%20Pr.final_.pdf
- https://pdfs.semanticscholar.org/d505/db497c988a5022e83d6f807949d78ec6d8ce.pdf?_ga=2.107698050.125558133.1572354838-784886463.1572354838



A Philosopher's Guide to Probability

Alan Hájek

Abstract
"What is probability?" There are two ways to understand this question, and thus two kinds of answer that could be given. Firstly, the question may be understood as: *how should probability theory be formalized*? This is a mathematical question, to which Kolmogorov's axiomatization is the orthodox answer. I briefly review his theory. Secondly, the question may be understood as: *what do statements of probability mean*? This is a philosophical question, and while the mathematical theory of probability certainly bears on it, the answer or answers must come from elsewhere. After a brief autobiographical prelude and historical introduction, I survey at some length the leading 'interpretations' of probability: classical, logical, frequentist, propensity, and subjectivist. I conclude with some of my best bets for future avenues of research in the philosophical foundations of probability.

Autobiographical prelude

Once upon a time I was an undergraduate majoring in mathematics and statistics. I attended many lectures on probability theory, and my lecturers taught me many nice theorems involving probability: 'P of this equals P of that', and so on. One day I approached one of them after a lecture and asked him: "What is this 'P' that you keep on writing on the blackboard? *What is probability?*" He looked at me like I needed medication, and he told me to go to the philosophy department.

And so I did. (Admittedly, my route there was long and circuitous.) There I found a number of philosophers asking the very same question: what is probability? All these years later, it's still one of the main questions that I am working on. I still don't feel that I have a completely satisfactory answer, although I like to think that I've made some progress on it. For starters, I know many things that probability is *not*, namely various

Kolmogorov axioms (1936):

• *For any event A, **P(A) ≥ 0.** In English, that's "For any event A, the probability of A is greater or equal to 0".*

• *When S is the sample space of an experiment; i.e., the set of all possible outcomes, **P(S) = 1**. In English, that's "The probability of any of the outcomes happening is one hundred percent", or—paraphrasing— "anytime this experiment is performed, something happens".*

• *If A and B are mutually exclusive outcomes, **P(A ∪ B) = P(A) + P(B).** Here ∪ stands for 'union'. We can read this by saying "If A and B are mutually exclusive outcomes, the probability of either A or B happening is the probability of A happening plus the probability of B happening*

https://www.statisticshowto.com/axiomatic-probability/

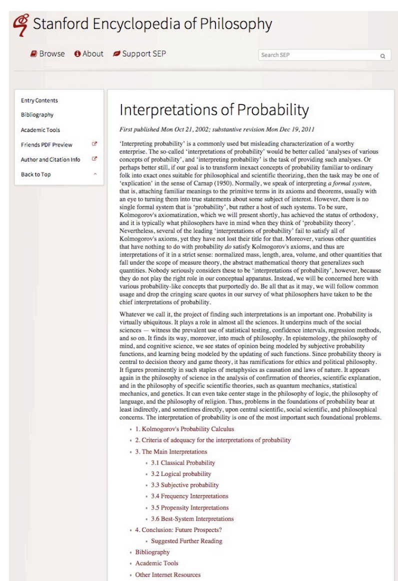
7

Probability Theory

Resources

- Interpretations of Probability, Stanford Encyclopedia of Philosophy: http://plato.stanford.edu/entries/probability-interpret/

*"**Subjective probability**…We may characterize subjectivism (also known as personalism and subjective Bayesianism) with the slogan: 'Probability is degree of belief'. We identify probabilities with degrees of confidence, or credences, or "partial" beliefs of suitable agents."*



Stanford Encyclopedia of Philosophy

Browse About Support SEP

Search SEP

Entry Contents
Bibliography
Academic Tools
Friends PDF Preview
Author and Citation Info
Back to Top

Interpretations of Probability

First published Mon Oct 21, 2002; substantive revision Mon Dec 19, 2011

'Interpreting probability' is a commonly used but misleading characterization of a worthy enterprise. The so-called 'interpretations of probability' would be better called 'analyses of various concepts of probability', and 'interpreting probability' is the task of providing such analyses. Or perhaps better still, if our goal is to transform inexact concepts of probability familiar to ordinary folk into exact ones suitable for philosophical and scientific theorizing, then the task may be one of 'explication' in the sense of Carnap (1950). Normally, we speak of interpreting *a formal system*, that is, attaching familiar meanings to the primitive terms in its axioms and theorems, usually with an eye to turning them into true statements about some subject of interest. However, there is no single formal system that is 'probability', but rather a host of such systems. To be sure, Kolmogorov's axiomatization, which we will present shortly, has achieved the status of orthodoxy, and it is typically what philosophers have in mind when they think of 'probability theory'. Nevertheless, several of the leading 'interpretations of probability' fail to satisfy all of Kolmogorov's axioms, yet they have not lost their title for that. Moreover, various other quantities that have nothing to do with probability *do* satisfy Kolmogorov's axioms, and thus are interpretations of it in a strict sense: normalized mass, length, area, volume, and other quantities that fall under the scope of measure theory, the abstract mathematical theory that generalizes such quantities. Nobody seriously considers these to be 'interpretations of probability', however, because they do not play the right role in our conceptual apparatus. Instead, we will be concerned here with various probability-like concepts that purportedly do. Be all that as it may, we will follow common usage and drop the cringing scare quotes in our survey of what philosophers have taken to be the chief interpretations of probability.

Whatever we call it, the project of finding such interpretations is an important one. Probability is virtually ubiquitous. It plays a role in almost all the sciences. It underpins much of the social sciences — witness the prevalent use of statistical testing, confidence intervals, regression methods, and so on. It finds its way, moreover, into much of philosophy. In epistemology, the philosophy of mind, and cognitive science, we see states of opinion being modeled by subjective probability functions, and learning being modeled by the updating of such functions. Since probability theory is central to decision theory and game theory, it has ramifications for ethics and political philosophy. It figures prominently in such staples of metaphysics as causation and laws of nature. It appears again in the philosophy of science in the analysis of confirmation of theories, scientific explanation, and in the philosophy of specific scientific theories, such as quantum mechanics, statistical mechanics, and genetics. It can even take center stage in the philosophy of logic, the philosophy of language, and the philosophy of religion. Thus, problems in the foundations of probability bear at least indirectly, and sometimes directly, upon central scientific, social scientific, and philosophical concerns. The interpretation of probability is one of the most important such foundational problems.

8

Probabilistic Reasoning

Resources

- A Priori Rules: Wittgenstein on the Normativity of Logic by Peter Railton: http://www.nyu.edu/gsas/dept/philo/courses/rules/papers/RailtonAPriori.pdf
- *In Paul Boghossian & Christopher Peacocke (eds.),* New Essays on the a Priori. Oxford University Press 170--96 (2000).

But where's the problem? After all, we often enough find ourselves in the wrong, sometimes embarrassingly so…
So, there's the problem: adaptiveness. The a priori is not as such just one more area of inquiry…
The distinctive content of a claim of apriority is not a matter of being right, or justified…

A Priori Rules: Wittgenstein on the Normativity of Logic

Peter Railton

The University of Michigan

Introduction

Like many, I have long been uneasy with the category of the a priori. Perhaps I have simply misunderstood it. It has seemed to me, at any rate, that asserting a claim or principle as a priori is tantamount to claiming that we would be justified in ruling out alternatives in advance, no matter what the future course of experience might hold. Yet in my own case, I have felt it would be mere bluffing were I to lodge such a claim. I certainly could not discover in myself any sense of the requisite authority, nor even any clear idea of where to look for guidance in forming it.

Contemplating widely-used examples of "propositions true a priori" did not remove my worry. For a start, there was the shadow of history. A claim like "logical truth is a priori" or "the attribution of rationality is a priori in intentional explanation" kept sounding, to my ears, as if they echoed "the Euclidean geometry of space is a priori" or "the principle of sufficient reason is a priori in physical explanation". And these echoes awakened just sort of the worry that had initially unsettled me: I would pronounce myself satisfied that certain claims, at least, were safe from the threat of contrary experience, just on the eve of developments in our on-going view of the world that would lead a sensible person to want to reopen the question. So I would emerge looking like the (perhaps apocryphal) fellow who claimed, in the wake of the great inventions of the nineteenth century, that the US Patent Office could now be closed, since all the really new

9

Probabilistic Reasoning

Resources

- Brief Introduction to Bayes' Rule by Kevin Murphy: http://www.cs.ubc.ca/~murphyk/Bayes/bayesrule.html

A brief introduction to Bayes' Rule

by Kevin Murphy.

Intuition

Here is a simple introduction to Bayes' rule from an article in the Economist (9/30/00).

"The essence of the Bayesian approach is to provide a mathematical rule explaining how you should change your existing beliefs in the light of new evidence. In other words, it allows scientists to combine new data with their existing knowledge or expertise. The canonical example is to imagine that a precocious newborn observes his first sunset, and wonders whether the sun will rise again or not. He assigns equal prior probabilities to both possible outcomes, and represents this by placing one white and one black marble into a bag. The following day, when the sun rises, the child places another white marble in the bag. The probability that a marble plucked randomly from the bag will be white (ie, the child's degree of belief in future sunrises) has thus gone from a half to two-thirds. After sunrise the next day, the child adds another white marble, and the probability (and thus the degree of belief) goes from two-thirds to three-quarters. And so on. Gradually, the initial belief that the sun is just as likely as not to rise each morning is modified to become a near-certainty that the sun will always rise."

In symbols

Mathematically, Bayes' rule states

```
                 likelihood * prior
posterior = ----------------------------
                marginal likelihood
```

or, in symbols,

```
              P(e | R=r) P(R=r)
P(R=r | e) = -----------------
                    P(e)
```

where P(R=r|e) denotes the probability that random variable R has value r given evidence e. The denominator is just a normalizing constant that ensures the posterior adds up to 1; it can be computed by summing up the numerator over all possible values of R, i.e.,

```
P(e) = P(R=0, e) + P(R=1, e) + ... = sum_r P(e | R=r) P(R=r)
```

This is called the marginal likelihood (since we marginalize out over R), and gives the prior probability of the evidence.

Example of Bayes' rule

Here is a simple example, based on Mike Shor's Java applet. Suppose you have been tested positive for a disease; what is the probability that you actually have the disease? It depends on the accuracy and sensitivity of the test, and on the background (prior) probability of the disease.

Let P(Test=+ve | Disease=true) = 0.95, so the false negative rate, P(Test=-ve | Disease=true), is 5%. Let P(Test=+ve | Disease=false) = 0.05,, so the false positive rate is also 5%. Suppose the disease is rare: P(Disease=true) = 0.01 (1%). Let D denote Disease (R in the above equation) and "T=+ve" denote the positive Test (e in the above equation). Then

```
                             P(T=+ve|D=true) * P(D=true)
P(D=true|T=+ve) = ------------------------------------------------------------
                   P(T=+ve|D=true) * P(D=true)+ P(T=+ve|D=false) * P(D=false)

                 0.95  * 0.01          0.0095
               = -------------------  = -------      = 0.161
                 0.95*0.01 + 0.05*0.99   0.0590
```


10

- Probability Theory
  - probabilistic logic, theory, inferencing
  - related fields, applications
  - definitions, interpretations, epostemology, resources
- Subjective probability - Bayes' Rule
  - history, definition
  - P(R|X) vs. P(R&X), examples
  - Who to reward problem
  - Principle of Conditionalization
  - Monty Hall problem

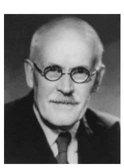
Probabilistic Reasoning

Thomas Bayes

Bayes' Theorem is ascribed to the Reverent Thomas Bayes. After his death, two unpublished essays were discovered and forwarded to the Royal Society. This work made little impact, however, until it was independently discovered a few years later by the great French mathematician Laplace (1749 - 1827).

*English mathematician then quickly **rediscovered** Bayes' work.*

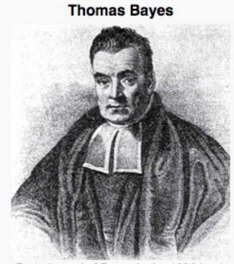
Bayes' theorem and in particular its emphasis on prior probabilities has caused considerable controversy.



Sir Harold Jeffreys was English mathematician, statistician, geophysicist and astronomer (1891-1989). He put Bayes' algorithm and Laplace's formulation on an axiomatic basis. His book Theory of Probability, which first appeared in 1939, played an important role in the revival of the Bayesian view of probability.

Sir Jeffreys wrote that Bayes' theorem "is to the theory of probability what Pythagoras's theorem is to geometry"

Thomas Bayes



Portrait used of Bayes in the 1936 book *History of Life Insurance*; it is dubious whether it actually depicts Bayes.[1] No earlier portrait or claimed portrait survived.

Born	c. 1701 London, England
Died	7 April 1761 (aged 59) Tunbridge Wells, Kent, England
Residence	Tunbridge Wells, Kent, England
Nationality	English

Signature

T. Bayes.

Courtesy of Wiki

Probabilistic Reasoning

Probability rules (refresher)

"AND" or (intersection)

- Probability of independent events $P(A\,and\,B) = P(A) * P(B)$
- Probability of dependent events $P(A\,and\,B) = P(A) * P(B \mid A)$

(General Multiplication Rule)

Conditional probability $P(B \mid A)$

- P(B|A) – *"probability of B given A"*
  - *or probability of event B occurring given that event A has already occurred.*

- Independent events
  - *...if occurrence of one, does not change the probability of the other event occurring.*
  - *E.g. - rolling a 2 on a die and flipping a head on a coin…*
    - Two events are independent if the occurrence of one does not change the probability of the other occurring. An example would be rolling a 2 on a die and flipping a head on a coin. Rolling the 2 does not affect the probability of flipping the head. If events are independent, then the probability of them both occurring is the product of the probabilities of each occurring.

https://people.richland.edu/james/lecture/m170/ch05-rul.html#gmr

13

Probability (quick refresher)

Conditional Probability

- …is probability of an event, given that another event has occurred.
- *Conditional* probability of *A* given *B*

  Given events A and B, with P(B)>0, the conditional probability of A given B:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$ *(Kolmogorov definition)*

Axiom of probability: $P(A \cap B) = P(A \mid B) \cdot P(B)$

Independent events

- If A and B are independent events, then:
  - *P(A and B) = P(A) * P(B)*
  - *P(A|B) = P(A)*
  - *P(B|A) = P(B)*

14

Probabilistic Reasoning

Bayes’ Rule (theorem)

Prior probability, in Bayesian statistics, is ***the probability of an event before new data is collected****. This is the best rational assessment of the probability of an outcome based on the current knowledge before an experiment is performed. Prior probability can be compared with posterior probability.*

Mathematically, Bayes' rule:

$$posterior = \frac{likelihood \times prior}{marginal\ likelihood}$$

or, in symbols:

$$P(R = r \mid e) = \frac{P(e \mid R = r)\, P(R = r)}{P(e)}$$

- where $P(R = r \mid e)$
  - probability that the random variable *R* has value *r* given evidence *e* (posterior probability);
  - i.e., *e* is the evidence we are using to make inferences about *R*.
- $P(e)$ is a normalizing constant that ensures the posterior adds up to 1.

Probabilistic Reasoning

Bayes’ Rule (cont.)

Prior probability, in Bayesian statistics, is ***the probability of an event before new data is collected****. This is the best rational assessment of the probability of an outcome based on the current knowledge before an experiment is performed. Prior probability can be compared with posterior probability.*

Marginal likelihood is:

$$P(e) = P(R = 0, e) + P(R = 1, e) + \cdots$$

$$= \sum_{i=0}^{r} P(e \mid R = i)\, P(R = i)$$

$P(e)$ is called the marginal likelihood (since we marginalize out over *R*), and gives the *prior probability of the evidence*.

(marginalize out -> “sum out”)

- *P(R|X)* is closely related to *P(R&X)*, but they are not identical:
- *P(R&X)* is the proportion of things that have property *R* and property *X* within all things;
  - *e.g., the proportion of ”cats with disease R and positive X factor” within the group of all cats.*

© M. Manic, CMSC 409: Artificial Intelligence, F23

Session 23, Updated on 11/15/23 12:44:07 PM

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 1 *P(R&X)*:

- If the total number of cats = 10,000 of which 80 have disease *R* and positive *X* factor, then the probability of picking out a random cat from the entire sample, with disease *R* and factor *X* is given by:

  $P(R\&X) = 80/10{,}000 = .8\%$

- On the other hand, $P(R|X)$ is the proportion of things that have property *R* and property *X* within all things that have *X*; e.g., the proportion of "cats with disease *R* and positive *X* factor, within the group of all cats with positive X.

Probabilistic Reasoning

Bayes' Rule (cont.)

- *Again,* $P(R|X)$ *is...*
  - …the proportion of things that have property *R* and property *X* within all things that have *X*; e.g., the proportion of "cats with disease *R* and positive *X* factor, within the group of all cats with positive X.

Example 1 cont.:

- If the total number of cats = 10,000 of which 80 have disease *R* and positive *X* factor, out of 1,000 that have positive *X,* then the probability of picking out a random cat with disease *R* from the entire *sample* with factor *X* is given by:

  $P(R|X) = 80/1{,}000 = 8.0\%$

  (total number of cats = 10,000; 1,000 that have positive *X;* 80 have both disease *R* and positive *X* factor)

- ***In a sense, P(R|X) really means P(R&X | X), but specifying the extra X all the time would be redundant***

Probabilistic Reasoning

Bayes' Rule (cont.)

- *In a sense, P(R|X) really means P(R&X | X), but specifying the extra X all the time would be redundant*

Example 2:

- ***Suppose patient tested positive for a disease; what is the probability that the test is correct?***
- *Depends on the **accuracy** and **sensitivity** of the test, and on the **background (prior) probability** of the disease.*
- Suppose

T is a test (+ or –); D is a disease (t or f)

and

P(T=+ | D=t) = 0.95 (True Pos.)

P(T= –|D=t) = 0.05 (False Neg.)

Remember regression session?

	Predicted 1	Predicted 0
Actual 1	True positives (a)	False negatives (b)
Actual 0	False positives (c)	True negatives (d)

(TP + FN = 1; FP+TN = 1)

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 2 (cont.):

- Let

P(T=+ | D=f) = 0.05 (False Pos.)

Suppose the disease is rare (disease rate):

P(D=t) = 0.01

So, probability of having disease, given patient tested positive is just 16%.

(under stated assumptions)

$$\therefore \quad P(D=t \mid T=+) =$$

$$= \frac{P(T=+ \mid D=t)\, P(D=t)}{P(T=+ \mid D=t)\, P(D=t) + P(T=+ \mid D=f)\, P(D=f)}$$

$$P(D=t \mid T=+) = \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.05 \cdot 0.99} = 0.161$$

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 2 (cont.):

$$P(D = t \mid T = +) = 0.161$$

- *So, probability of having disease given patient tested positive is just 16%.*

- May seem low, but…
  - For disease rate 0.01, 1 in 100 has disease (that patient will test positive)
  - We expect 5% others (5 in 100 people) to test positive by error
  - *So, if 6 people test positive, only 1 will actually have a disease (1/6 ~0.16)*

Starting assumptions:

$P(T{=}+ \mid D{=}f) = 0.05$ (False Pos.)

$P(D{=}t) = 0.01$ (rare disease)

- *Probability Theory*
  - *probabilistic logic, theory, inferencing*
  - *related fields, applications*
  - *definitions, interpretations, epostemology, resources*
- *Subjective probability - Bayes' Rule*
  - *history, definition*
  - *P(R|X) vs. P(R&X), examples*
  - *Who to reward problem*
  - *Principle of Conditionalization*
  - *Monty Hall problem*

Probabilistic Reasoning
Hollywood or reality...☺

The Movie 21 (2008):
https://www.youtube.com/watch?v=6x9DNc_OdNM
https://www.youtube.com/watch?v=Zr_xWfThjJ0

ABOUT 21

Inspired by the true story of MIT students who mastered the art of card counting and took Vegas casinos for millions in winnings. Looking for a way to pay for tuition, Ben Campbell (Jim Sturgess) finds himself quietly recruited by MIT's most gifted students in a daring plot to break Vegas. With the help of a brilliant statistics professor (Kevin Spacey) and armed with fake IDs, intelligence and a complicated system of counting cards, Ben and his friends succeed in breaking the impenetrable casinos. Now, his challenge is keeping the numbers straight and staying one step ahead of the casinos before it all spirals out of control.

SCREENPLAY
Peter Steinfeld, Allan Loeb

BASED ON A WORK BY
Bed Mezrich

EXECUTIVE PRODUCER
Brett Ratner, Ryan Kavanaugh, William S. Beasley

PRODUCER
Kevin Spacey, Dana Brunetti, Michael De Luca

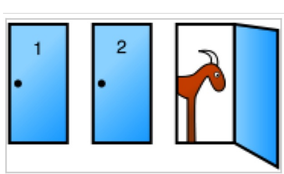
DIRECTOR
Robert Luketic

ACTOR
Laurence Fishburne, Kevin Spacey, Kate Bosworth, Jim Sturgess, Josh Gad, Aaron Yoo, Liza Lapira, Jacob Pitts, Sam Golzari

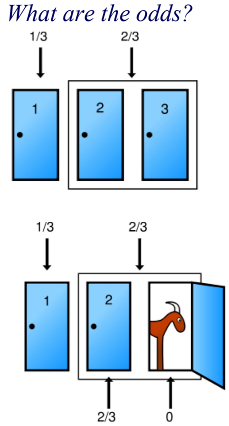
Probabilistic Reasoning

Bayes' Rule (cont.)

Example 6: Monty Hall Problem (Let's make a deal)



What are the odds?



Monty (Halperin) Hall…in 1963, created the popular trading game show Let's Make a Deal, which spawned a mathematical probability puzzle, the "Monty Hall Problem."

Bayes' Rule (cont.)

Example 6: Monty Hall Problem (Let's make a deal)

Application of Bayes' Rule:

- *Consider scenario when door 3 has been chosen, and no other door has been opened:*

  P(C1) that the car is behind door 1

  P(C2) that the car is behind door 2

  P(C3) that the car is behind door 3

  thus

  P(C1) = P(C2) = P(C3) = 1/3

Bayes' Rule (cont.)

Example 6: Monty Hall Problem (cont.)

- *The probability that the game host will open door 1, P(O1) = 1/2; i.e.*

$$P(O1) = P(C1) \times P(O1 | C1) + P(C2) \times P(O1 | C2) + P(C3) \times P(O1 | C3)$$

$$= \frac{1}{3} \times 0 + \frac{1}{3} \times 1 + \frac{1}{3} \times \frac{1}{2} = \frac{1}{2}$$

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 6: Monty Hall Problem (cont.)

- *When the car is behind door 1, the game host will never open door 1;*
- *When the car is behind door 2, the game host will certainly open door 1;*
- *When the car is behind door 3, the game host will open door 1 or door 2 with probability 1/2*

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 6: Monty Hall Problem (cont.)

- *Thus the probability that the car is behind door 2 when door 1 is opened is*

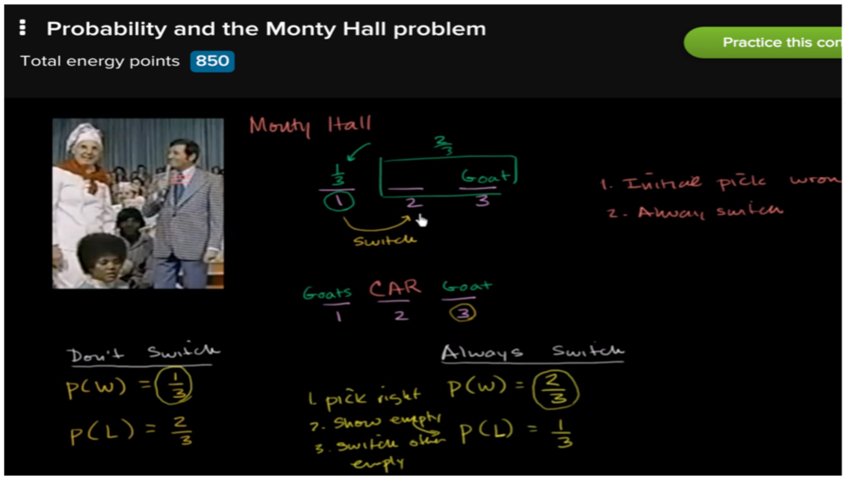
$$P(C2 \mid O1) = \frac{P(O1 \mid C2) \times P(C2)}{P(O1)} =$$

$$= \frac{1 \times \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3} = P(C1 \mid O2)$$

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 6: Monty Hall Problem (cont.)



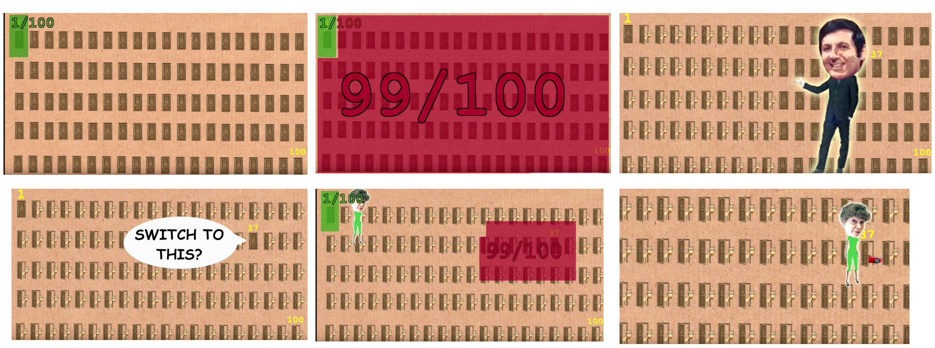
https://www.khanacademy.org/math/precalculus/prob_comb/dependent_events_precalc/v/monty-hall-problem

29

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 6: Monty Hall Problem (cont.)



Numberphile: https://www.youtube.com/watch?v=4Lb-6rxZxx0

30

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 6: Monty Hall Problem (cont.)

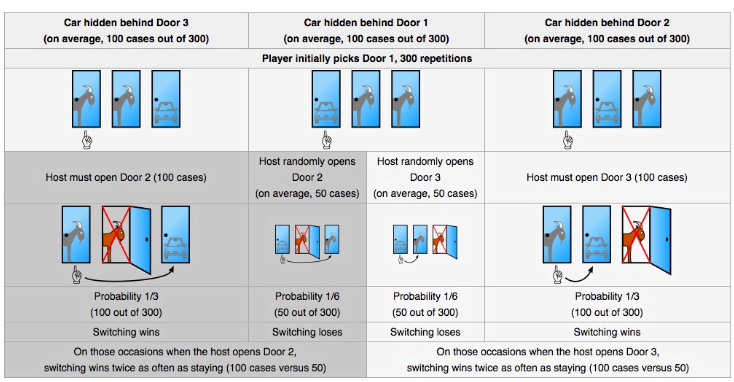
Numberphile: https://www.youtube.com/watch?v=4Lb-6rxZxx0

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 6: Monty Hall Problem (Let's make a deal)

What are the odds?

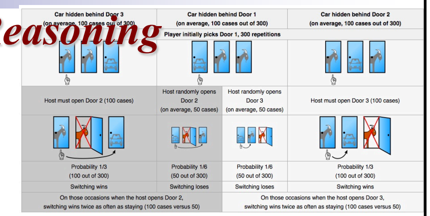


 https://en.wikipedia.org/wiki/Monty_Hall_problem

Probabilistic Reasoning

Bayes' Rule (cont.)

Example 6: Monty Hall Problem

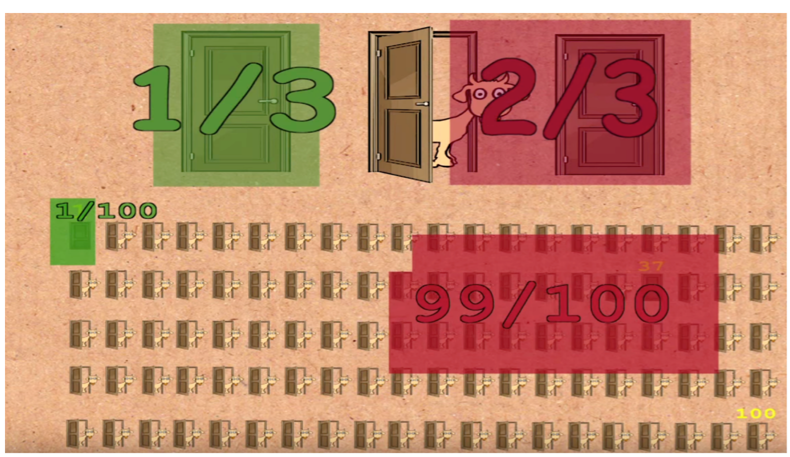


- Initially, **the car is equally likely behind any of the three doors** (odds on doors are 1:1:1 and this remains the case after the player has chosen door 1)
- According to Bayes' rule, **the posterior** odds on the location of the car, given the host opens door 3, **are equal to the prior odds multiplied by the Bayes factor** (likelihood), which is by def. the probability of the new piece of information (host opens door 3) under each of the hypotheses considered (location of the car).
- Now, since the player initially chose door 1, the chance the host opens door 3 is 50% (the car behind door 1), 100% (the car behind door 2), 0% (the car is behind door 3).
- Thus the Bayes factor consists of the ratios 1/2 : 1 : 0 or equivalently 1 : 2 : 0, while the prior odds were 1 : 1 : 1. Thus the posterior odds become equal to the Bayes factor 1 : 2 : 0. Given the host opened door 3, the probability the car is behind door 3 is zero, and it is twice as likely to be behind door 2 than door 1.

Probabilistic Reasoning

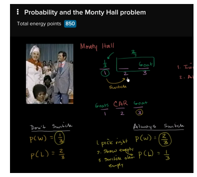
Bayes' Rule (cont.)

Example 6: Monty Hall Problem (cont.)



Numberphile: https://www.youtube.com/watch?v=4Lb-6rxZxx0

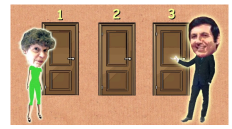
Probabilistic Reasoning

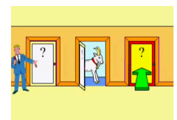


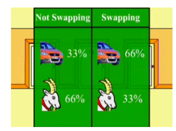
Bayes' Rule (cont.)

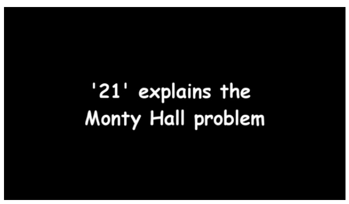
Resources - Monty Hall Problem (cont.)

- *Khan Academy: https://www.khanacademy.org/math/cc-seventh-grade-math/cc-7th-probability-statistics/cc-7th-dependent-probability/v/monty-hall-problem*
- *Numberphile: https://www.youtube.com/watch?v=4Lb-6rxZxx0*
- *Ron Clarke: https://www.youtube.com/watch?v=mhlc7peGlGg*
- *NY Times Interactive: http://www.nytimes.com/2008/04/08/science/08monty.html?_r=0*



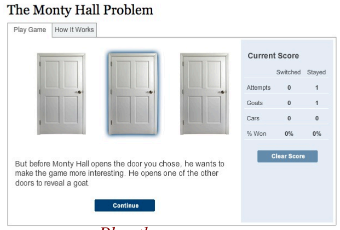






21 (2008): https://www.youtube.com/watch?v=6z9DNe-OdNM
https://www.youtube.com/watch?v=Zr_xWfThjJ0

The Monty Hall Problem



Play the game..