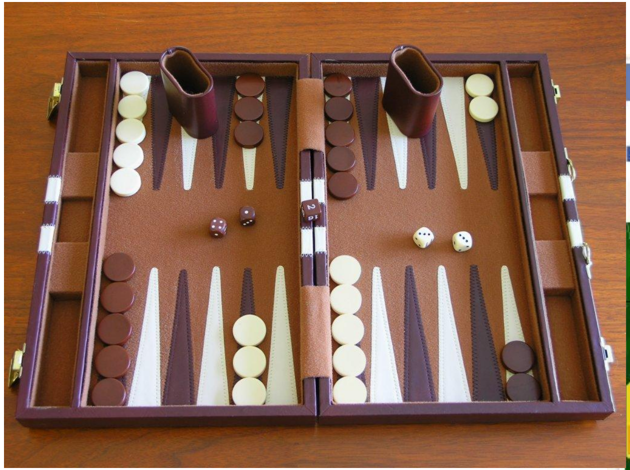
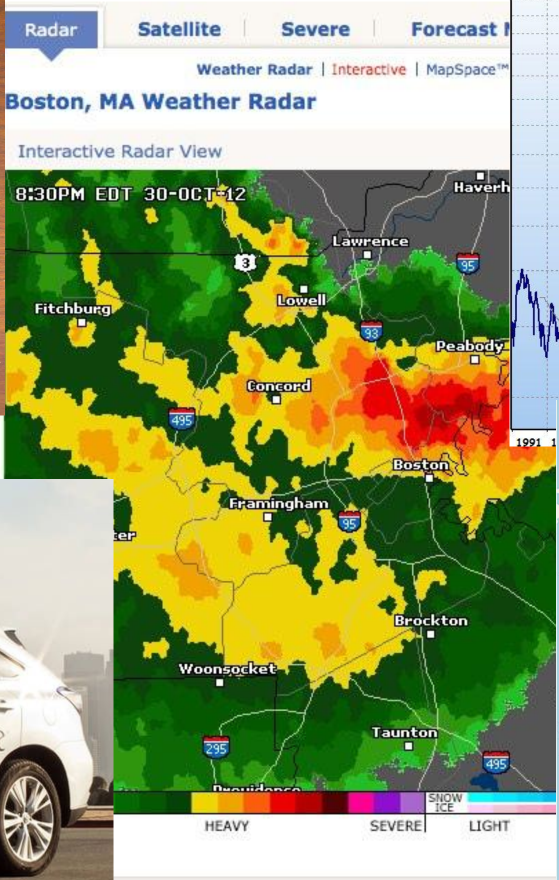
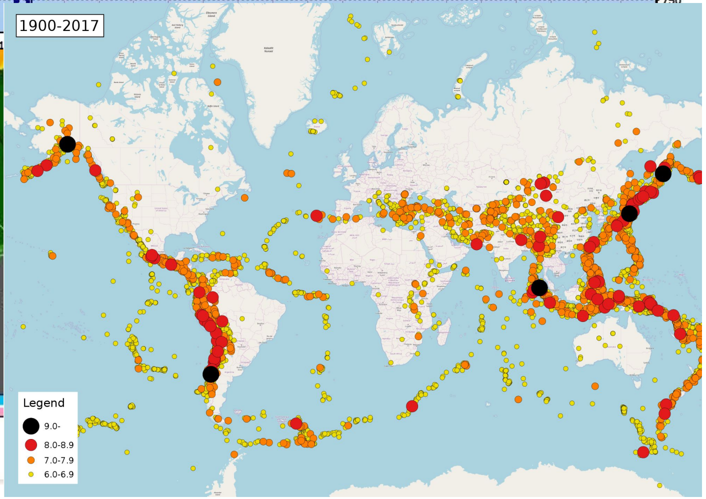
Reasoning with Knowledge and Probability Theory

willie




Radar
Satellite
Severe
Forecast
Weather Radar | Interactive | MapSpace™
Boston, MA Weather Radar
Interactive Radar View
8:30PM EDT 30-OCT-12
Haverh
Lawrence
Lowell
Fitchburg
Peabody
Concord
Boston
Framingham
Brockton
Woonsocket
Taunton
HEAVY
SEVERE
SNOW
ICE
LIGHT


5000
4750
4500
4250
4000
3750
3500
3250
3000
2750
2500
2250
2000
1750
1500
1250
1000
750
1991


1900-2017
Legend
9.0-
8.0-8.9
7.0-7.9
6.0-6.9


Google
self-driving car

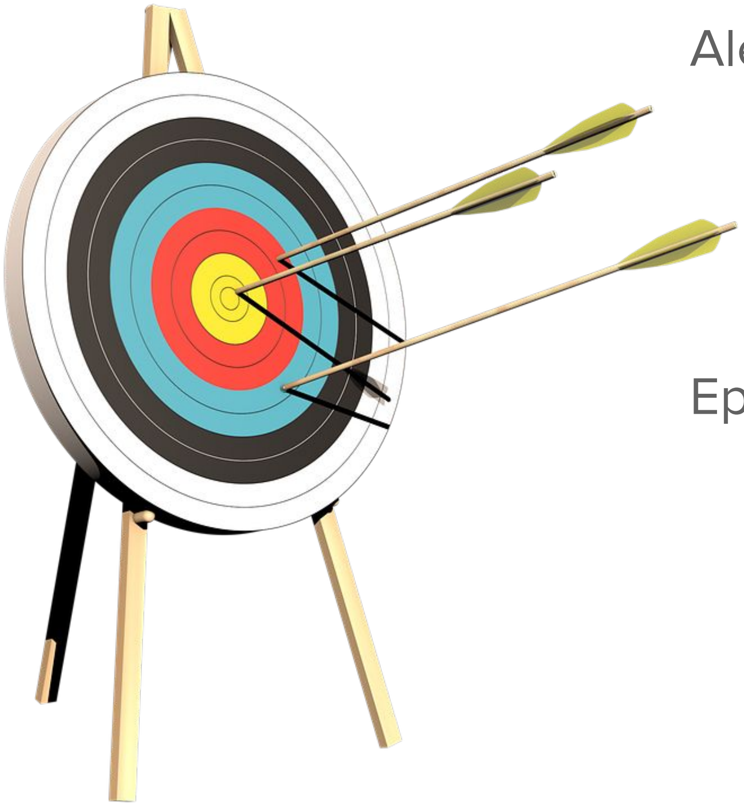
Reasoning with Uncertainty

Pure logic sometimes is insufficient:

- Because the wumpus may not eat you
- Because resetting the modem is not always reliable
- Because the alarm might be caused by an earthquake or a burglar
- Because John said it rained yesterday and Beth said it didn't

Because the real world does not necessarily allow us to deduce everything logically

Categories of uncertainty



Aleatoric uncertainty

- statistical uncertainty
- unknowns that differ in each experiment

Epistemic uncertainty

- systematic uncertainty
- unknowns that could be known

Basic Probability

Probability theory enables us to make rational decisions.

Which mode of transportation is safer:

- Car or Plane?
- What is the probability of an accident?

Basic Probability Theory

- An **experiment** has a set of potential outcomes, e.g., throw a dice
- The **sample space** of an experiment is the set of all possible outcomes, e.g., {1, 2, 3, 4, 5, 6}
  - A **random variable** can take on any value in the sample space
- An **event** is a subset of the sample space.
  - {2}
  - {3, 6}
  - even = {2, 4, 6}
  - odd = {1, 3, 5}

Probability as Relative Frequency

Total Flips: 10
Number Heads: 5
Number Tails: 5

Probability of Heads:
Number Heads / Total Flips = 0.5

Probability of Tails:
Number Tails / Total Flips = 0.5 = 1.0 – Probability of Heads

The experiments, the sample space and the events must be defined clearly for probability to be meaningful

Theoretical Probability

Principle of Indifference — Alternatives are always to be judged equi-probable if we have no reason to expect or prefer one over the other.

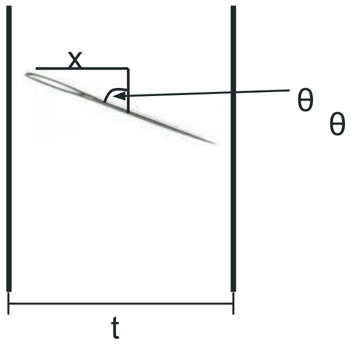
Each outcome in the sample space is assigned equal probability.

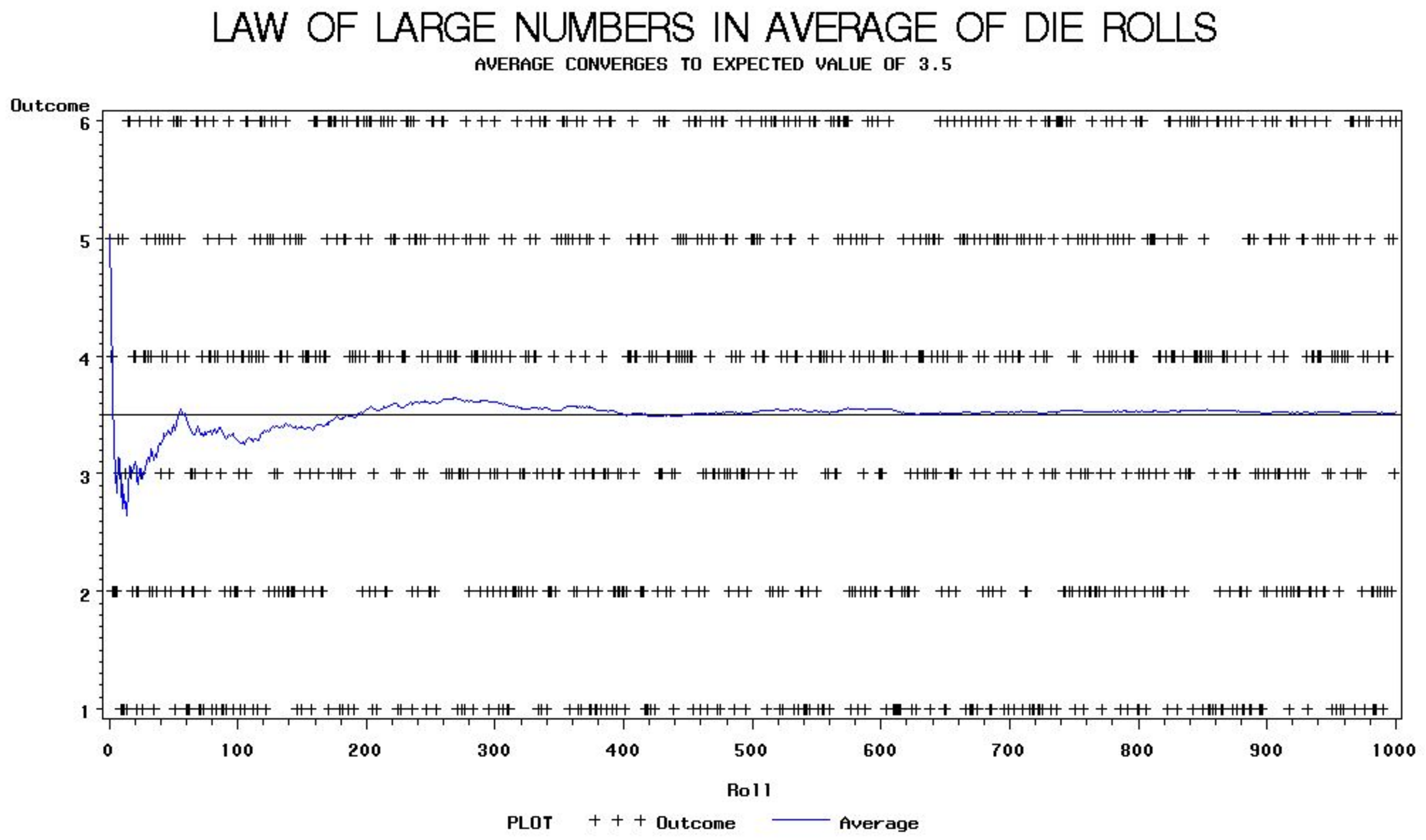
Example: throw a dice

- $P(\{1\})=P(\{2\})= \ldots =P(\{6\})=1/6$

Law of Large Numbers

- As the number of experiments increases the relative frequency of an event more closely approximates the theoretical probability of the event.
  - if the theoretical assumptions hold.
- Buffon's Needle for Computing π



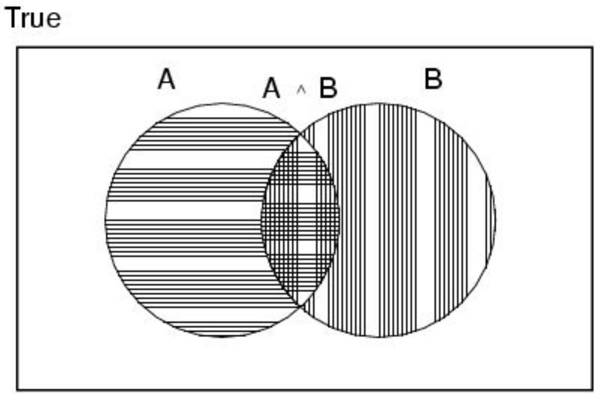

LAW OF LARGE NUMBERS IN AVERAGE OF DIE ROLLS
AVERAGE CONVERGES TO EXPECTED VALUE OF 3.5
Outcome
6
5
4
3
2
1
0
100
200
300
400
500
600
700
800
900
1000
Roll
PLOT
+ + + Outcome
Average

Large Number Reveals Untruth in Assumptions

Results of 1,000,000 throws of a die

Number	1	2	3	4	5	6
Fraction	.155	.159	.164	.169	.174	.179

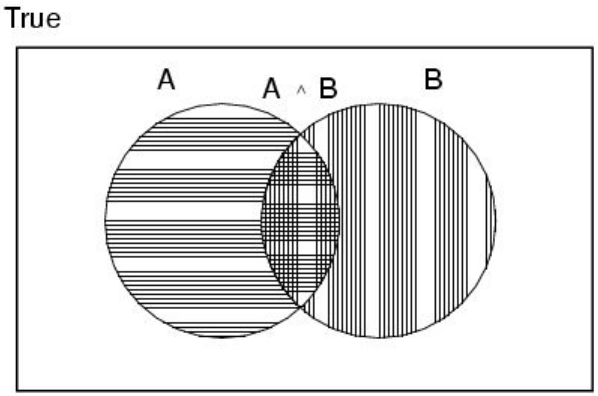
Axioms of probability



Axioms of probability

For any propositions *A*, *B*

- $0 \leq P(A) \leq 1$
- $P(true) = 1$ and $P(false) = 0$
- $P(A \lor B) = P(A) + P(B) - P(A \land B)$



Prior probability

Prior or unconditional probabilities of propositions

e.g., P(*Cavity* = true) = 0.1 (also written as P(cavity))

P(*Weather* = sunny) = 0.72 correspond to belief prior to arrival of any (new) evidence

Probability distribution gives values for all possible assignments:

P(*Weather*) = <0.72,0.1,0.08,0.1> (sums to 1)

P(Dice) = <0.167, 0.167, 0.167, 0.167, 0.167, 0.167>

Joint probability distribution

Joint probability distribution for a set of random variables gives the probability of every atomic event on those random variables

P(*Weather*,*Cavity*) = a 4 × 2 matrix of values:

Weather =	sunny	rainy	cloudy	snow
Cavity = true	0.144	0.02	0.016	0.02
Cavity = false	0.576	0.08	0.064	0.08

- Every question about a domain can be answered by the joint distribution

Properties of Probability

1. P(¬E) = 1– P(E)
2. If E1 and E2 are logically equivalent, then

   P(E1) = P(E2).

   – E1: Not all philosophers are more than six feet tall.
   – E2: Some philosopher is not more that six feet tall.

   Then P(E1) = P(E2).
3. P(E1, E2) ≤ P(E1).

Conditional Probability

- The probability of an event may change after knowing another event.

The probability of A given B is denoted by P(A|B).

A: the top card of a deck of poker cards is an eight of clubs

P(A) =

However, if we know

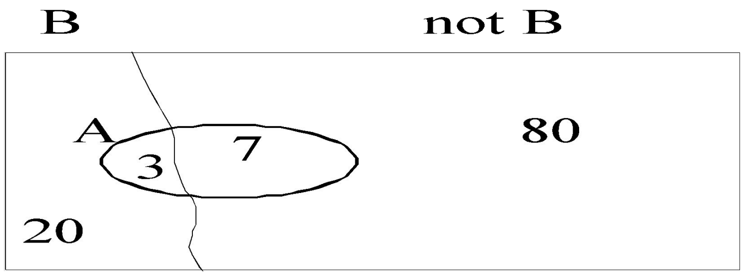
B: the top card is a club, then the probability of A given B is true is

P(A|B) =

How to Compute $P(A|B)$?

Business Students

Of 100 students completing a course, 20 were business majors. Ten students received As in the course, and three of these were business majors., suppose A is the event that a randomly selected student got an A in the course, B is the event that a randomly selected event is a business major. What is the probability of A? What is the probability of A after knowing B is true?



Probability of being hired

1000 people apply to a Boston company

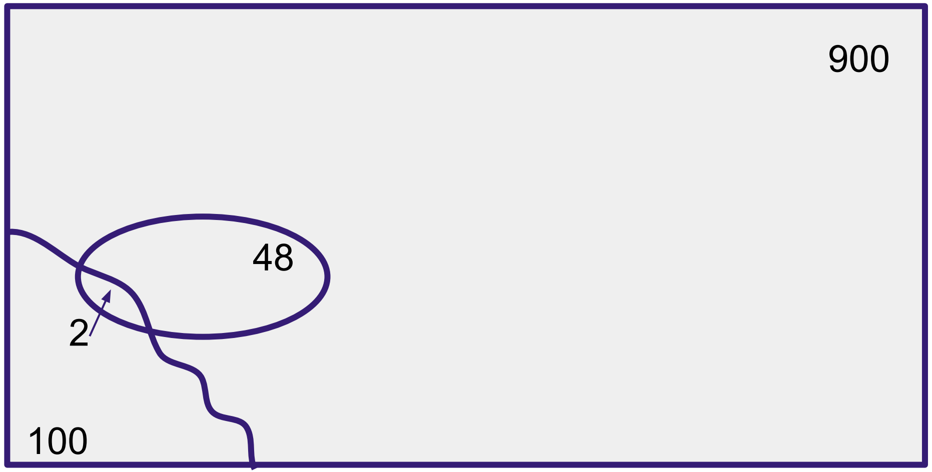
900 of them are Red Sox fans, 100 are not

Of the 1000 applicants, they hire 48 Red Sox fans and 2 who are not Red Sox fans

Let A be the event of randomly selecting a non-Red Sox fan applicant

Let B be the event of randomly selecting a hired applicant

What is P(A)? P(B)? P(B|A)?



Inference by enumeration

- Start with the joint probability distribution:

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

- Sum out true events

Inference by enumeration

- Start with the joint probability distribution:

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

- Sum out true events
- P(*toothache*) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2

Inference by enumeration

- Start with the joint probability distribution:

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

- Can also compute conditional probabilities:

$$P(\neg \mathit{cavity} \mid \mathit{toothache}) = \frac{P(\neg \mathit{cavity} \land \mathit{toothache})}{P(\mathit{toothache})}$$

$$= \frac{0.016+0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

Normalization

- Denominator can be viewed as a normalization constant α

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

$\mathbf{P}(Cavity \mid toothache) = \alpha\, \mathbf{P}(Cavity, toothache)$

$= \alpha\, [\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg\, catch)]$

$= \alpha\, [\langle 0.108, 0.016\rangle + \langle 0.012, 0.064\rangle]$

$= \alpha\, \langle 0.12, 0.08\rangle = \langle 0.6, 0.4\rangle$

General idea: compute distribution on query variable by fixing evidence variables and summing over hidden variables

Probabilistic Reasoning

- Evidence
  - What we know about a situation.
- Hypothesis
  - What we want to conclude.
- Compute
  - P(Hypothesis | Evidence)

Credit Card Authorization

E is the data about the applicant's age, job, education, income, credit history, etc,

H is the hypothesis that the credit card will provide positive return (for ccard company).

The decision of whether to issue the credit card to the applicant is based on the probability P(H|E).

Medical Diagnosis

- E is a set of symptoms, such as, coughing, sneezing, headache, ...
- H is a disorder, e.g., common cold, cancer, swine flu.
- The diagnosis problem is to find an H (disorder) such that $P(H|E)$ is maximum.

The Three-Card Problem

Three cards are in a hat. One is red on both sides (the red-red card). One is white on both sides (the white-white card). One is red on one side and white on the other (the red-white card). A single card is drawn randomly and tossed into the air.

a. What is the probability that the red-red card was drawn? (RR)

b. What is the probability that the drawn cards lands with a white side up? (W-up)

c. What is the probability that the red-red card was not drawn, assuming that the drawn card lands with the a red side up. (not-RR|R-up)

Fair Bets

A bet is fair to an individual I if, according to the individual's probability assessment, the bet will break even in the long run.

The following three bets are fair to a naïve (typical?) individual:

fair would be:
a: win $4
a: lose $4

b: win $4?
b: lose $4?

c: win $4
c: lose $2

Bet (a): Win $4.20 if RR;
lose $2.10
otherwise. [since they believe P(RR)=1/3]

Bet (b): Win $2.00 if W-up;
lose $2.00
otherwise. [since they believe P(W-up)=1/2]

Bet (c): Win $4.00 if R-up and not-RR;
lose $4.00 if R-up and RR;
neither win nor lose if not-R-up.
[since they believe P(not-RR|R-up)=1/2]

Dutch Book

- The bets that this person accepted have an interesting property:

  No matter what card is drawn in the three-card problem, and no matter how it lands, you are guaranteed to lose money.

- This is called a Dutch Book

Verification

There are six possible outcomes

1. RR drawn, R-up (side 1)
2. RR drawn, R-up (side 2)
3. WR drawn, R-up
4. WR drawn, W-up
5. WW drawn, W-up (side 1)
6. WW drawn, W-up (side 2)

	1	2	3	4	5	6
a.	$4.20	$4.20	-$2.10	-$2.10	-$2.10	-$2.10
b.	-$2.00	-$2.00	-$2.00	$2.00	$2.00	$2.00
c.	-$4.00	-$4.00	$4.00	$0.00	$0.00	$0.00
Total	-$1.80	-$1.80	-$0.10	-$0.10	-$0.10	-$0.10

Another Example

A patient takes a lab test and the result comes back positive. The test has a false negative rate of 2% and false positive rate of 2%. Furthermore, 0.5% of the entire population have this cancer.

What is the probability of cancer if we know the test result is positive?