6.01

Lecture 9: Probabilistic Reasoning

Adam Hartz (hz@mit.edu)

Overview and Perspective

Focus on **key concepts**

Overview and Perspective

Focus on **key concepts**

software engineering

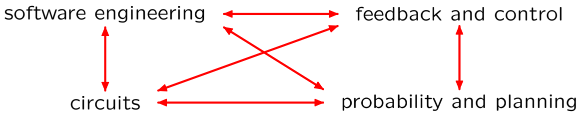
feedback and control

circuits

probability and planning

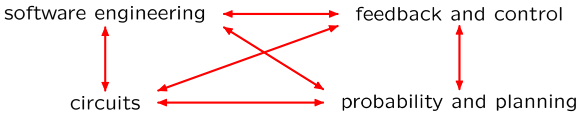
Overview and Perspective

Focus on **key concepts** with **explicit connections**

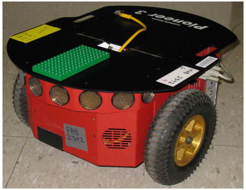


Overview and Perspective

Focus on **key concepts** with **explicit connections**



in an **authentic context** with overarching theme:
Modular Design of Complex Systems



Controlling complexity:

Primitives, Combinations, Abstractions, and Common Patterns

Module 3: Probabilistic Reasoning

Modeling uncertainty and desiging robust systems

Topics: Subjective Probability,
Markov Processes,
Bayesian Inference

Lab Exercises:

- Localization: Find location in known hallway

Probability Theory

Probability theory provides a framework for:

- Modeling and reasoning about uncertainty
  - Making precise statements about uncertain situations
  - Drawing reliable inferences from unreliable observations
- Designing systems that are robust to uncertainty

Check Yourself!

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5% (and a false negative rate of 0%). Patient 0 is found to have received a positive result on the test. What is the chance that Patient 0 actually has the disease, assuming you know nothing about the person's symptoms or signs?

Which of the following is closest to the right answer?

0. 1%
1. 2%
2. 5%
3. 10%
4. 50%
5. 95%

Check Yourself!

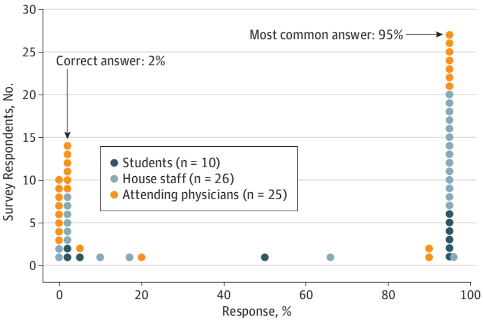
If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5% (and a false negative rate of 0%). Patient 0 is found to have received a positive result on the test. What is the chance that Patient 0 actually has the disease, assuming you know nothing about the person's symptoms or signs?

Which of the following is closest to the right answer?

0. 1%
1. 2%
2. 5%
3. 10%
4. 50%
5. 95%

Check Yourself!

Figure. Distribution of Responses to Survey Question Provided in the Article Text



Of 61 respondents, 14 provided the correct answer of 2%. The most common answer was 95%, provided by 27 of 61 respondents. The median answer was 66%, which is 33 times larger than the true answer.

`http://archinte.jamanetwork.com/data/Journals/INTEMED/0/ild140014.pdf`

Probability Review: Basics

Probability: Likelihood of an event occurring: $\Pr(A = a)$

Distribution: Function from elements a in domain A to probabilities: $\Pr(A) : a \to p$

Conditional Probability: Likelihood of an event occuring, after knowing that some other event occurred: $\Pr(A = a | B = b)$

Conditional Distribution: Function from elements b in domain B to distributions over A: $\Pr(A|B) : b \to (a \to p)$

Joint Probability: Probability of two events happening: $\Pr(A = a, B = b)$

Joint Distribution: Function from elements (a, b) in domain (A, B) to probabilities: $\Pr(A, B) : (a, b) \to p$

Probability: Events

Probabilities (representing a likelihood of occurrence) are assigned to **events**, which are possible outcomes of an experiment.

Example: Flip three coins in succession. Possible events:

- head, head, head
- head, tail, head
- one head and two tails
- first toss was a head

There are eight **atomic** (finest grain) events:
HHH, HHT, HTH, HTT, THH, THT, TTH, TTT

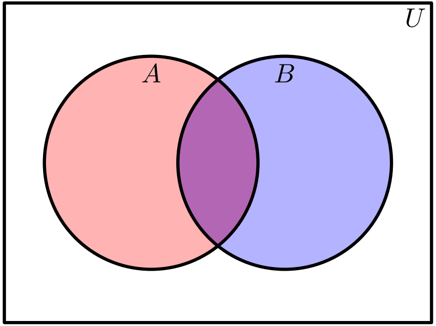
Atomic events are **mutually exclusive** (only one can happen).
Set of all atomic events is **collectively exhaustive** (cover all cases).
Set of all possible atomic events is called the **sample space** U.

Probability Theory: Axioms of Probability

A probability $\Pr(A)$ is assigned to each atomic event A.

The probabilities assigned to events must obey three axioms:

- $\Pr(A) \geq 0$ for all events A
- $\Pr(U) = 1$
- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$



Conditional Probability

Often times, the probability of an event happening changes depending on whether or not another event happened. The events are, generally, *dependent*.

Conditional probability:
$\Pr(A \mid B)$

This probability (pronounced "the probability of A given B") represents the probability of event A happening, *given that event B happened*.

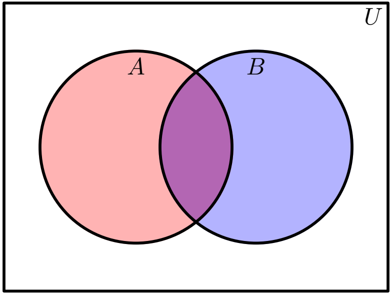
Conditional Probability

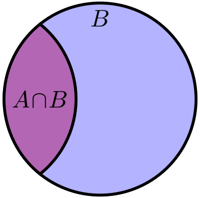
Here we *know* that B happened, so we can throw everything else away ("condition" on B).

Conditional Probability

Here we *know* that B happened, so we can throw everything else away ("condition" on B).

Conditioning on B restricts the sample space (which was U) to B:



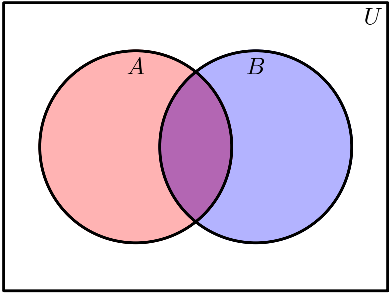


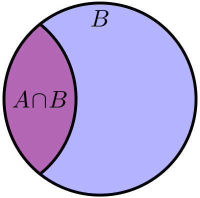
U has shrunk to B

Conditional Probability

Here we *know* that B happened, so we can throw everything else away ("condition" on B).

Conditioning on B restricts the sample space (which was U) to B:





U has shrunk to B

$$\Pr(A \mid B) = \frac{\Pr(A, B)}{\Pr(B)} \qquad \Pr(A, B) = \Pr(B) \times \Pr(A \mid B)$$

Check Yourself!

Ben Bitdiddle applied to both MIT and Harvard. Ben believes that his probability of being accepted at MIT is 0.10, at Harvard is 0.06, and at neither is 0.843.

What is the probability that Ben will get accepted to either MIT or Harvard but not both?

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

$\Pr(B_2 \mid \overline{B_1}) = 0.9$

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

$\Pr(B_2 \mid \overline{B_1}) = 0.9$
$\Pr(B_2 \mid B_1) = 0.8$

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

$\Pr(B_2 \mid \overline{B_1}) = 0.9$
$\Pr(B_2 \mid B_1) = 0.8$

What is the probability that the second song is by Ariana Grande?

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

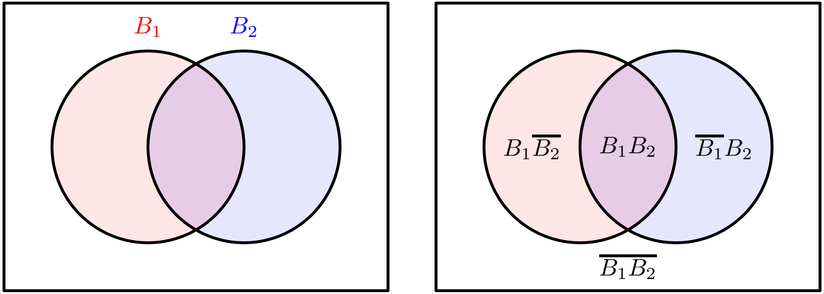
$\Pr(B_2 \mid \overline{B_1}) = 0.9$
$\Pr(B_2 \mid B_1) = 0.8$

What is the probability that the second song is by Ariana Grande?

$\Pr(B_2) = 0.56 + 0.27 = 0.83$

Symmetry

Decision trees are sequential, but set representation is symmetric.



We could compute the joint probability two ways:

$\Pr(B_1, B_2) = \Pr(B_1)\Pr(B_2 \mid B_1) = \Pr(B_2)\Pr(B_1 \mid B_2)$

Inverse Probability

We can compute the joint probability $\Pr(B_1, B_2)$ in two ways:

$$\Pr(B_1, B_2) = \Pr(B_1)\Pr(B_2 \mid B_1) = \Pr(B_2)\Pr(B_1 \mid B_2)$$

A slight manipulation gives us *Bayes' Theorem*:

$$\Pr(B_1 \mid B_2) = \frac{\Pr(B_1)\Pr(B_2 \mid B_1)}{\Pr(B_2)}$$

Allows for *anti-sequential* reasoning: infer causes from effects, or infer future events from past information.

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

$\Pr(B_2 \mid \overline{B_1}) = 0.9$

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

$\Pr(B_2 \mid \overline{B_1}) = 0.9$
$\Pr(B_2 \mid B_1) = 0.8$

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

$\Pr(B_2 \mid \overline{B_1}) = 0.9$
$\Pr(B_2 \mid B_1) = 0.8$

You tune in during the second song, and it turns out that it is *not* by Ariana Grande. What is the probability that the song before it *was* by Ariana Grande?

Check Yourself

Consider a sequence of two songs on the radio.

Let B_1 be the event that the first song is by Ariana Grande.
Let B_2 be the event that the second song is by Ariana Grande.

$\Pr(B_1) = 0.7$

$\Pr(B_2 \mid \overline{B_1}) = 0.9$
$\Pr(B_2 \mid B_1) = 0.8$

You tune in during the second song, and it turns out that it is *not* by Ariana Grande. What is the probability that the song before it *was* by Ariana Grande?

$\Pr(B_1 \mid \overline{B_2}) = \frac{0.14}{0.17} \approx 0.824$

Bayes' Theorem

"Inverse Probability:" infer causes from effects, or infer future events from past information

Basic idea: combine old *belief* with *evidence* to generate a new *belief*.

$$\Pr(H = h \mid E = e) = \frac{\Pr(E = e \mid H = h) \cdot \Pr(H = h)}{\Pr(E = e)}$$

Bayes' Theorem

"Inverse Probability:" infer causes from effects, or infer future events from past information

Basic idea: combine old *belief* with *evidence* to generate a new *belief*.

$$\Pr(H = h \mid E = e) = \frac{\Pr(E = e \mid H = h) \cdot \Pr(H = h)}{\Pr(E = e)}$$

$\Pr(H = h)$: how likely was the hypothesis h?

Bayes' Theorem

"Inverse Probability:" infer causes from effects, or infer future events from past information

Basic idea: combine old *belief* with *evidence* to generate a new *belief*.

$$\Pr(H = h \mid E = e) = \frac{\Pr(E = e \mid H = h) \cdot \Pr(H = h)}{\Pr(E = e)}$$

$\Pr(H = h)$: how likely was the hypothesis h?

$\Pr(E = e \mid H = h)$: how well is the evidence e supported by h?

Bayes' Theorem

"Inverse Probability:" infer causes from effects, or infer future events from past information

Basic idea: combine old *belief* with *evidence* to generate a new *belief*.

$$\Pr(H = h \mid E = e) = \frac{\Pr(E = e \mid H = h) \cdot \Pr(H = h)}{\Pr(E = e)}$$

$\Pr(H = h)$: how likely was the hypothesis h?

$\Pr(E = e \mid H = h)$: how well is the evidence e supported by h?

$\Pr(E = e)$: normalizing factor

Bayes' Theorem

"Inverse Probability:" infer causes from effects, or infer future events from past information

Basic idea: combine old *belief* with *evidence* to generate a new *belief*.

$$\Pr(H=h \mid E=e) = \frac{\Pr(E=e \mid H=h) \cdot \Pr(H=h)}{\Pr(E=e)}$$

$\Pr(H=h)$: how likely was the hypothesis h?

$\Pr(E=e \mid H=h)$: how well is the evidence e supported by h?

$\Pr(E=e)$: normalizing factor

$\Pr(H=h \mid E=e)$: how likely is h after the evidence?

Bayes' Theorem

"Inverse Probability:" infer causes from effects, or infer future events from past information

Basic idea: combine old *belief* with *evidence* to generate a new *belief*.

$$\Pr(H = h \mid E = e) = \frac{\Pr(E = e \mid H = h) \cdot \Pr(H = h)}{\Pr(E = e)}$$

$\Pr(H = h)$: how likely was the hypothesis h?

$\Pr(E = e \mid H = h)$: how well is the evidence e supported by h?

$\Pr(E = e)$: normalizing factor

$\Pr(H = h \mid E = e)$: how likely is h after the evidence?

Dice Game

I have 6 dice:

- 4-sided
- 6-sided
- 8-sided
- 10-sided
- 12-sided
- 20-sided

I pick one at random and roll it, telling you what number I roll. Can you figure out which die I picked?

Check Yourself!

I pick a die and roll a "5". Obviously, this means the probability that I was rolling the 4-sided die, call it $p_4 = \Pr(\text{die} = 4 \mid \text{observe } 5)$, is 0.

What can be said about the updated probabilities of the other dice?

1. $p_6 < p_8 < p_{10} < p_{12} < p_{20}$
2. $p_6 = p_8 = p_{10} = p_{12} = p_{20}$
3. $p_6 > p_8 > p_{10} > p_{12} > p_{20}$
4. None of the above

Check Yourself!

I pick a die and roll a "5". Obviously, this means the probability that I was rolling the 4-sided die, call it $p_4 = \Pr(\text{die} = 4 \mid \text{observe } 5)$, is 0.

What can be said about the updated probabilities of the other dice?

1. $p_6 < p_8 < p_{10} < p_{12} < p_{20}$
2. $p_6 = p_8 = p_{10} = p_{12} = p_{20}$
3. $p_6 > p_8 > p_{10} > p_{12} > p_{20}$
4. None of the above

Check Yourself!

Say I roll a very large number n of 5's in a row, starting from uniform. What happens to the belief distribution over dice as $n \to \infty$?

1. It becomes uniform over all states but 4.
2. One state has probability $\to 1$.
3. None of the above

Check Yourself!

Say I roll a very large number n of 5's in a row, starting from uniform. What happens to the belief distribution over dice as $n \to \infty$?

1. It becomes uniform over all states but 4.
2. One state has probability $\to 1$.
3. None of the above

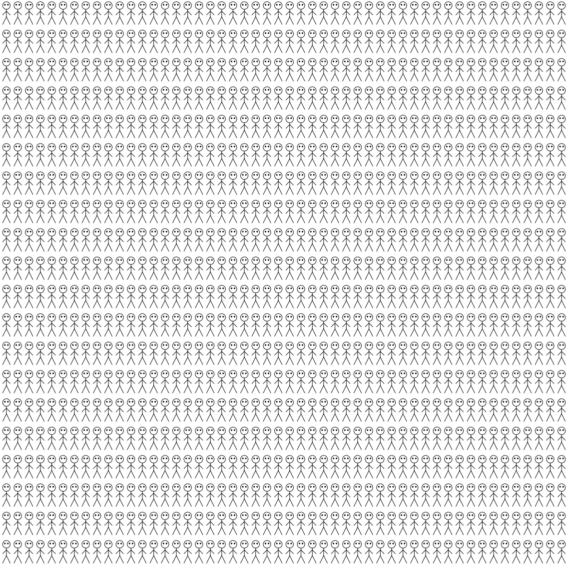
Check Yourself!

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5% (and a false negative rate of 0%). Patient 0 is found to have received a positive result on the test. What is the chance that Patient 0 actually has the disease, assuming you know nothing about the person's symptoms or signs?

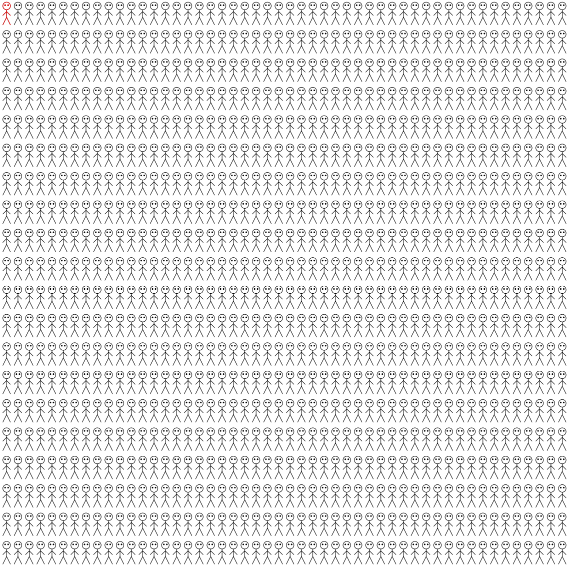
Which of the following is closest to the right answer?

0. 1%
1. 2%
2. 5%
3. 10%
4. 50%
5. 95%

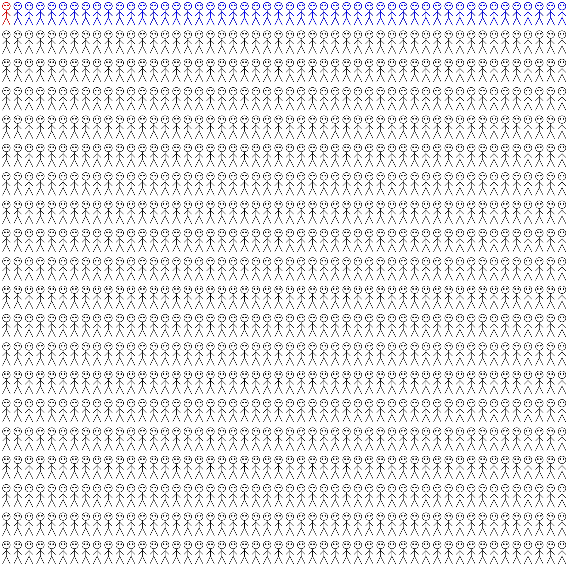
Check Yourself!



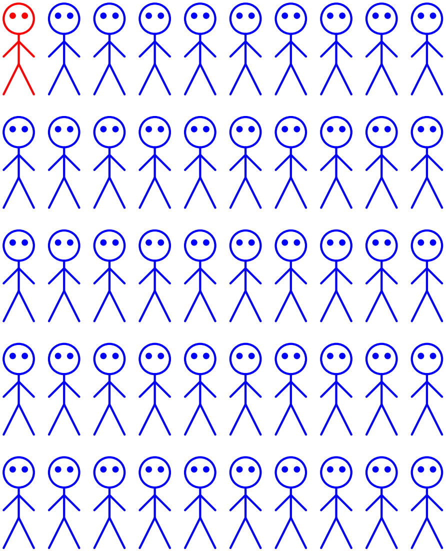
Check Yourself!



Check Yourself!



Check Yourself!



Check Yourself!

$$\Pr(\text{disease} \mid \text{test positive}) = \frac{\color{blue}{\Pr(\text{test positive} \mid \text{disease})} \cdot \color{red}{\Pr(\text{disease})}}{\Pr(\text{test positive})}$$

Check Yourself!

$$\Pr(\text{disease} \mid \text{test positive}) = \frac{\Pr(\text{test positive} \mid \text{disease}) \cdot \Pr(\text{disease})}{\Pr(\text{test positive})}$$

$$\Pr(\text{disease} \mid \text{test positive}) = \frac{1 \cdot (1/1000)}{1 \cdot (1/1000) + .05 \cdot (999/1000)}$$

Check Yourself!

$$\Pr(\text{disease} \mid \text{test positive}) = \frac{\Pr(\text{test positive} \mid \text{disease}) \cdot \Pr(\text{disease})}{\Pr(\text{test positive})}$$

$$\Pr(\text{disease} \mid \text{test positive}) = \frac{1 \cdot (1/1000)}{1 \cdot (1/1000) + .05 \cdot (999/1000)}$$

$$\Pr(\text{disease} \mid \text{test positive}) \approx \frac{1}{1+50}$$

Check Yourself!

$$\Pr(\text{disease} \mid \text{test positive}) = \frac{\Pr(\text{test positive} \mid \text{disease}) \cdot \Pr(\text{disease})}{\Pr(\text{test positive})}$$

$$\Pr(\text{disease} \mid \text{test positive}) = \frac{1 \cdot (1/1000)}{1 \cdot (1/1000) + .05 \cdot (999/1000)}$$

$$\Pr(\text{disease} \mid \text{test positive}) \approx \frac{1}{1 + 50}$$

$$\Pr(\text{disease} \mid \text{test positive}) \approx 0.02$$

Something Strange

These probabilities are not about repeated events! There is only one Patient 0, and he/she either has the disease or doesn't (can't do repreated trials).

What is going on?

Check Yourself!

There are two people: Pat and Cameron.

What is the probability that Pat is older than Cameron?

Subjective Probability

In this view, probabilities represent not frequencies of occurrance, but our *belief* about the likelihood of occurrance, and our uncertainty about the results.

Same math! Different *interpretation*!

Game 2

I have one cup with 4 dice in it. Each die is either red or white. Your goal is to guess how many red dice are in the cup.

Check Yourself!

What is a good initial belief about the number of red dice in the cup?

Thinking About the Game Quantitatively

Which dice could be in the cup?

- 4 white
- 3 white + 1 red
- 2 white + 2 red
- 1 white + 3 red
- 4 red

How likely are these?

Thinking About the Game Quantitatively

Which dice could be in the cup?

- 4 white
- 3 white + 1 red
- 2 white + 2 red
- 1 white + 3 red
- 4 red

How likely are these?
Assume equally likely (for lack of a better assumption).

n =# of red	0	1	2	3	4
$\Pr(N = n)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

Thinking About the Game Quantitatively

Which dice could be in the cup?

- 4 white
- 3 white + 1 red
- 2 white + 2 red
- 1 white + 3 red
- 4 red

How likely are these?
Assume equally likely (for lack of a better assumption).

$n =$# of red	0	1	2	3	4
$\Pr(N = n)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

Incorporating New Information

Assume that, before the bet, Adam pulls a random die, tells you its color, and returns it.

To update the belief based on this information, which of the following must be applied?

1. Bayes' Theorem
2. Total Probability
3. Something Else

Incorporating New Information

Assume that, before the bet, Adam pulls a random die, tells you its color, and returns it.

To update the belief based on this information, which of the following must be applied?

1. Bayes' Theorem
2. Total Probability
3. Something Else

$$\Pr(N = n \mid E = e) = \frac{\Pr(E = e \mid N = n)\,\Pr(N = n)}{\Pr(E = e)}$$

Incorporating New Information

Assume that, before the bet, Adam pulls a random die, tells you it is **red**, and returns it.

What should your belief be now? We need to update the belief!

Incorporating New Information

Assume that, before the bet, Adam pulls a random die, tells you it is **red**, and returns it.

What should your belief be now? We need to update the belief!

$n = \#$ of red	0	1	2	3	4
$\Pr(N = n)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

"Prior" Belief

Incorporating New Information

Assume that, before the bet, Adam pulls a random die, tells you it is **red**, and returns it.

What should your belief be now? We need to update the belief!

$n = \#$ of red	0	1	2	3	4
$\Pr(N = n)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

"Prior" Belief

$\Pr(N = n)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$
$\Pr(E = \text{red} \mid N = n)$	0	1/4	1/2	3/4	1
$\Pr(E = \text{red},\ N = n)$	0	1/20	2/20	3/20	4/20
$\Pr(N = n \mid E = \text{red})$	0	1/10	2/10	3/10	4/10

"Posterior" Belief

Incorporating More New Information

After telling you about the red die, Adam pulls another random die, tells you it is **red**, and returns it.

We need to update the state belief again! Previous "posterior" belief is now the "prior" belief.

Incorporating More New Information

After telling you about the red die, Adam pulls another random die, tells you it is **red**, and returns it.

We need to update the state belief again! Previous "posterior" belief is now the "prior" belief.

$n = \#$ of red	0	1	2	3	4
$\Pr(N = n)$	0	1/10	2/10	3/10	4/10

"Prior" Belief

Incorporating More New Information

After telling you about the red die, Adam pulls another random die, tells you it is **red**, and returns it.

We need to update the state belief again! Previous "posterior" belief is now the "prior" belief.

$n =$# of red	0	1	2	3	4
$\Pr(N = n)$	0	1/10	2/10	3/10	4/10

"Prior" Belief

$n =$# of red	0	1	2	3	4
$\Pr(E = \text{red} \mid N = n)$	0	1/4	1/2	3/4	1
$\Pr(E = \text{red},\ N = n)$	0	1/40	4/40	9/40	16/40
$\Pr(N = n \mid E = \text{red})$	0	1/30	4/30	9/30	16/30

"Posterior" Belief

Bayesian Estimation

Using observations to improve on initial guess.

Bayesian Estimation

Using observations to improve on initial guess.

We started with no information:

$n = \#$ of red	0	1	2	3	4
$\Pr(N = n)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

Bayesian Estimation

Using observations to improve on initial guess.

We started with no information:

$n = \#$ of red	0	1	2	3	4
$\Pr(N = n)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

Then we "observed" a red die:

$n = \#$ of red	0	1	2	3	4
$\Pr(N = n)$	0	$1/10$	$2/10$	$3/10$	$4/10$

Bayesian Estimation

Using observations to improve on initial guess.

We started with no information:

$n =$# of red	0	1	2	3	4
$\Pr(N = n)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

Then we "observed" a red die:

$n =$# of red	0	1	2	3	4
$\Pr(N = n)$	0	$1/10$	$2/10$	$3/10$	$4/10$

Then we "observed" a red die:

$n =$# of red	0	1	2	3	4
$\Pr(N = n)$	0	$1/30$	$4/30$	$9/30$	$16/30$

Alternate Observations

Assume that now, Adam doesn't tell you the color of the die. Instead, one of the following people does:

- Pat is sneaky and wants to cheat you. Pat always says:
  - "red" if a white die was drawn
  - "white" if a red die was drawn

- Cameron can't tell the difference between white and red; and so always chooses to tell you a color at random.

We are aware of these predispositions!

Check Yourself!

Pat always says:

- "red" if a white die was drawn
- "white" if a red die was drawn

How does our belief state change when Pat tells us that a white die was drawn?

1. Same as if Adam (honest) told us red was drawn.
2. Same as if Adam (honest) told us white was drawn.
3. It does not change.
4. None of the above.

Check Yourself!

Pat always says:

- "red" if a white die was drawn
- "white" if a red die was drawn

How does our belief state change when Pat tells us that a white die was drawn?

1. Same as if Adam (honest) told us red was drawn.
2. Same as if Adam (honest) told us white was drawn.
3. It does not change.
4. None of the above.

Check Yourself!

Cameron says any of the colors with probability $1/2$, regardless of what was actually drawn.

How does our belief state change when Cameron tells us that a white die was drawn?

1. Same as if Adam (honest) told us red was drawn.
2. Same as if Adam (honest) told us white was drawn.
3. It does not change.
4. None of the above.

Check Yourself!

Cameron says any of the colors with probability $1/2$, regardless of what was actually drawn.

How does our belief state change when Cameron tells us that a white die was drawn?

1. Same as if Adam (honest) told us red was drawn.
2. Same as if Adam (honest) told us white was drawn.
3. It does not change.
4. None of the above.

Check Yourself!

What if Adam (honest) is reporting, but we use the observation model associated with Pat (systematically says different color)?

Labs

Bayesian estimation of robot location.
Model the location of the robot as a Markov process
Estimate the location of the robot from sonar observations

- SL09: Practice with Theoretical "Robot in Hallway"
- DL10: "Robot in Hallway" in Python
- Lec11: Bayesian State Estimation, Probabilistic Modeling
- SL11: Localization and Parking 1
- DL11: Localization and Parking 2