CSCI 2824: Discrete Structures

Lecture 28: Bayes Theorem and the Law of Total Probability

Rachel Cox

Department of Computer Science

Probability Theory

Conditional Probability: The probability that E occurs given that F occurred:

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

Multiplication Rule: $p(E \cap F) = p(E|F)p(F)$

Independence: Events A and B are independent if

$$p(A|B) = p(A)$$
$$p(B|A) = p(B)$$
$$p(A \cap B) = p(A)p(B)$$

Bayes' Theorem

From the idea of conditional probability, $p(E|F) = \frac{p(E\cap F)}{p(F)}$

$$\Longrightarrow p(E \cap F) = p(E|F)p(F) \qquad AND \qquad p(F \cap E) = p(F|E)p(E)$$

$$p(E \cap F) = p(F \cap E)$$

$$\Longrightarrow p(E|F)p(F) = p(F|E)P(E)$$

$$\Longrightarrow p(F|E) = \frac{p(E|F)p(F)}{p(E)}$$

Bayes’ Theorem

This formula is known as **Bayes’ Theorem.** $p(F \mid E) = \frac{p(E \mid F)\, p(F)}{p(E)}$

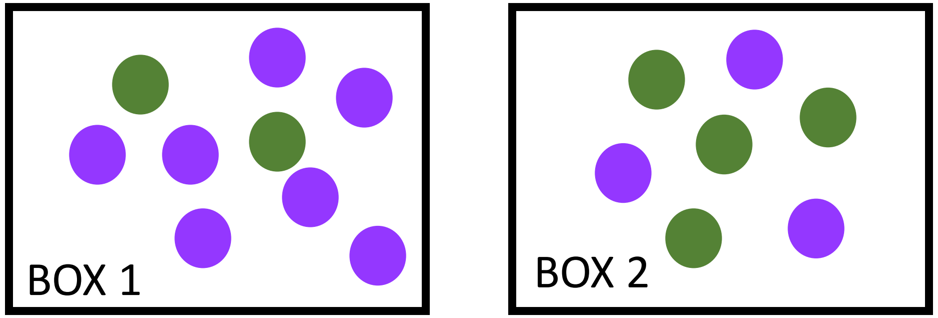
Thomas Bayes

Portrait purportedly of Bayes used in a 1936 book,[1] but it is doubtful whether the portrait is actually of him.[2] No earlier portrait or claimed portrait survives.

Bayes' Theorem

Example: Suppose we have two boxes filled with green and purple balls.

Box 1: 2 green balls, 7 purple balls
Box 2: 4 green balls, 3 purple balls



Suppose Anna selects a ball by first choosing one of the two boxes at random. She then selects one of the balls in this box at random. If Anna has selected a purple ball, what is the probability that she selected a ball from the first box?

$\boldsymbol{P}$ = event Anna picks a purple ball
$\boldsymbol{B_1}$= event Anna picks from Box 1
$\overline{\boldsymbol{B_1}}$ = event Anna picks from Box 2

Bayes' Theorem:

$$\boldsymbol{p(B_1|P) = \frac{p(P|B_1)p(B_1)}{p(P)}}$$

Bayes' Theorem

We need to calculate $p(P)$.

Let's define $\overline{B_1}$ as the event that Anna selects from Box 2.
Let $\overline{P}$ be the event that Anna has selected a green ball.

Note that: $P = (P \cap B_1) \cup (P \cap \overline{B_1})$ and that $\left((P \cap B_1) \cap (P \cap \overline{B_1}\,)\right) = \emptyset$

Law of Total Probability

Law of Total Probability:

$$p(P) = p(P|B_1)p(B_1) + p(P|\overline{B_1})p(\overline{B_1})$$

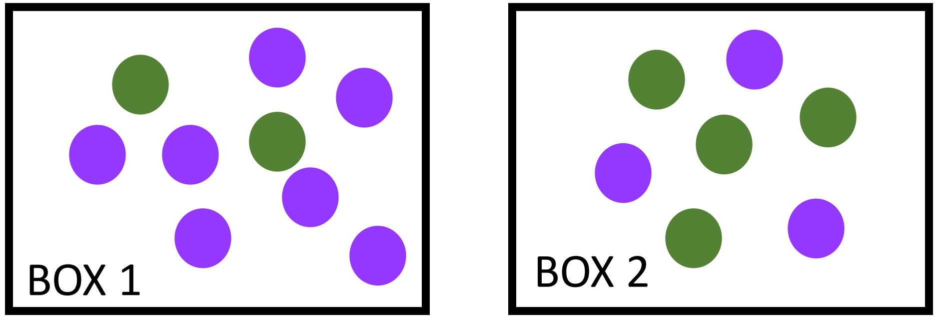
To generalize:

If we can break the set for our event F up into $F = \bigcup_{i=1}^{N} F_i$, where $F_i \cap F_j = \emptyset$ for $i \neq j$, then the probability of some other event $E, P(E)$, is:

$$p(E) = p(E|F_1)p(F_1) + p(E|F_2)p(F_2) + \ldots + p(E|F_N)p(F_N) = \sum_{i=1}^{N} p(E|F_i)p(F_i)$$

Bayes' Theorem

Back to our example:



Putting it all together:

$$p(B_1|P) = \frac{p(P|B_1)p(B_1)}{p(P|B_1)p(B_1) + p(P|\overline{B_1})p(\overline{B_1})}$$

Bayes' Theorem

The crux of Bayesian reasoning is the following:

1. Without the observation that Anna picked a purple ball, you would have guessed that the probability of picking from Box 1 was 0.5

2. By assimilating this data, you were able to update this belief about the probability of this event. (to 0.645)

Bayes' Theorem

Example: Cancer Testing. Suppose we know that 1% of the people over the age of 40 have cancer. And assume that 90% of the people who have cancer will test positive for cancer, if tested. Finally suppose that 8% of people who do not have cancer will also test positive (false positives).

What is the probability that a person who tests positive for cancer actually has cancer?

Bayes' Theorem

From Bayes' Theorem:

$$p(C|\text{positive}) = \frac{p(\text{positive}|C)\, p(C)}{p(\text{positive})}$$

From the Law of Total Probability:

$$p(\text{positive}) = p(\text{positive}\ |C)\, p(C) + p(\text{positive}\ |\bar{C})p(\bar{C})$$

Bayes' Theorem

Example: Cancer Testing - continued.

What is the probability that a person who tests negative for cancer is actually cancer free?

$$p(no\ cancer\ |negative) = \frac{p(negative|no\ cancer)p(no\ cancer)}{p(negative|no\ cancer)p(no\ cancer) + p(negative|cancer)p(cancer)}$$

Bayes' Theorem

Generalized Bayes' Theorem:

We can generalize Bayes' theorem to the situation where there are more than two boxes of balls (going back to the ball-drawing example).

S'pose that E is an event from sample space S and F_1, F_2, …, F_n are mutually disjoint events such that $S = \cup_{k=1}^{n} F_k$. Then

$$p(F_i \mid E) = \frac{p(E \mid F_i)\, p(F_i)}{\sum_{k=1}^{n} p(E \mid F_k)\, p(F_k)}$$

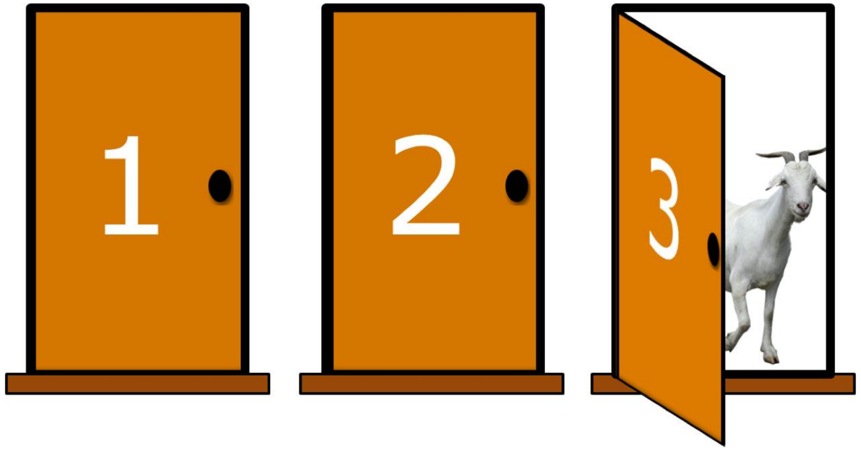
For example, the case where there are three bins of balls (B_k), and picking a red ball (R):

$$p(B_1 \mid R) = \frac{p(R \mid B_1)\, p(B_1)}{\underbrace{p(R \mid B_1)\, p(B_1)}_{\text{bin 1?}} + \underbrace{p(R \mid B_2)\, p(B_2)}_{\text{bin 2?}} + \underbrace{p(R \mid B_3)\, p(B_3)}_{\text{bin 3?}}}$$

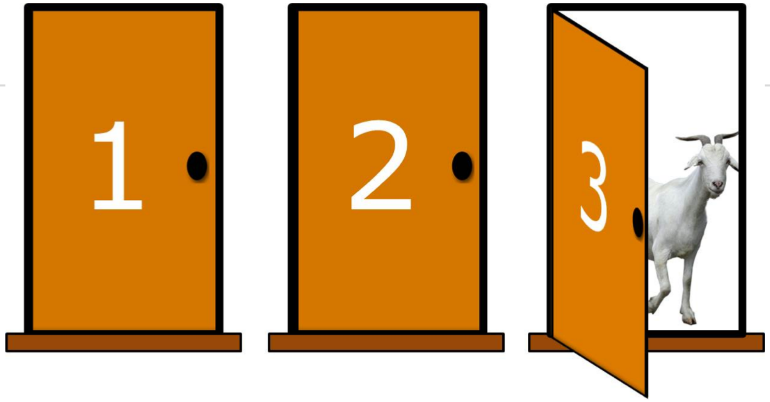
Example: Monty Hall Problem

You're on a game show, and you're given the choice of three doors. Behind one of the doors is a car; behind the others are goats. You pick a door - say, Door 1 - and the host, Monty, who knows what's behind all the doors, opens another door - say, Door 3 - which has a goat behind it. He then offers you the choice to switch to Door 2.

Do you switch?



Bayes' theorem



Example: Monty Hall Problem

Do you switch? **Answer:** Yes! Always!

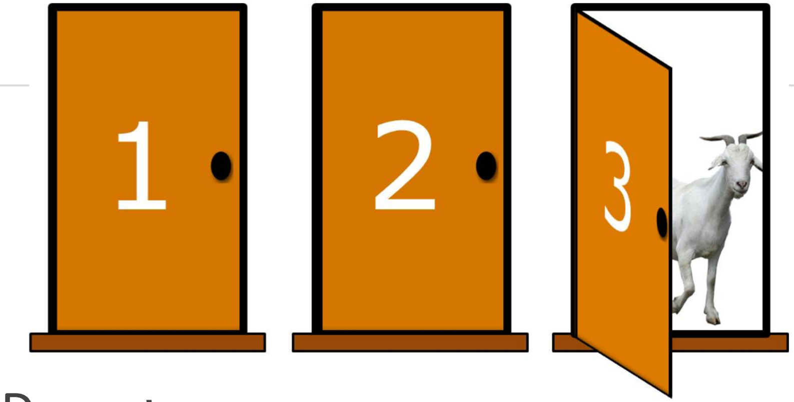
Why?

It turns out that if you don't switch doors, then your probability of winning is ⅓

But if you switch doors, then your probability of winning is ⅔

- There was originally ⅓ probability that your Door 1 was the winner, and ⅔ probability that Door 2 or Door 3 was the winner
- Monty then tells you that it isn't Door 3.
- **This doesn't change the fact that there's ⅔ *total* probability between Doors 2 & 3**

Bayes' theorem



Example: Monty Hall Problem

Let's work this out using **Bayes' Theorem:**

Assume you pick Door 1 and Monty shows you there's a goat behind Door 3

Let D_i be the event that the car is behind Door i

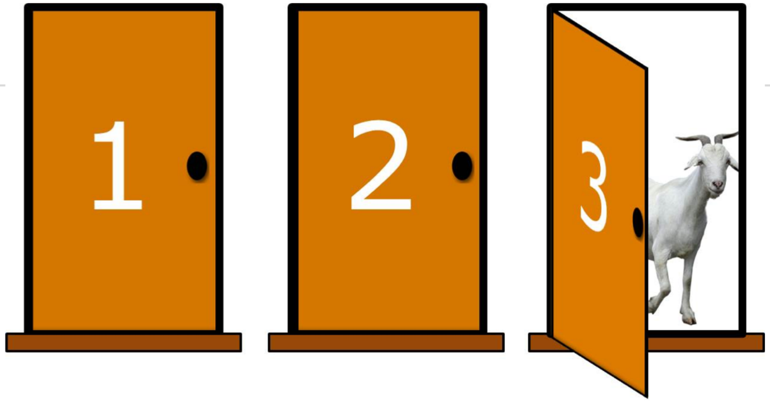
Let M_i be the event that Monty reveals a goat behind Door i

⇒ We want to calculate $p(D_1 \mid M_3)$ and $p(D_2 \mid M_3)$ and decide which door to go with, 1 or 2

Bayes' Theorem gives:

$$p(D_i \mid M_3) = \frac{p(M_3 \mid D_i)\; p(D_i)}{\sum_{k=1}^{3} p(M_3 \mid D_k)\; p(D_k)}$$

Bayes' theorem



Example: Monty Hall Problem

Bayes' Theorem:

$$p(D_i \mid M_3) = \frac{p(M_3 \mid D_i)\, p(D_i)}{\sum_{k=1}^{3} p(M_3 \mid D_k)\, p(D_k)}$$

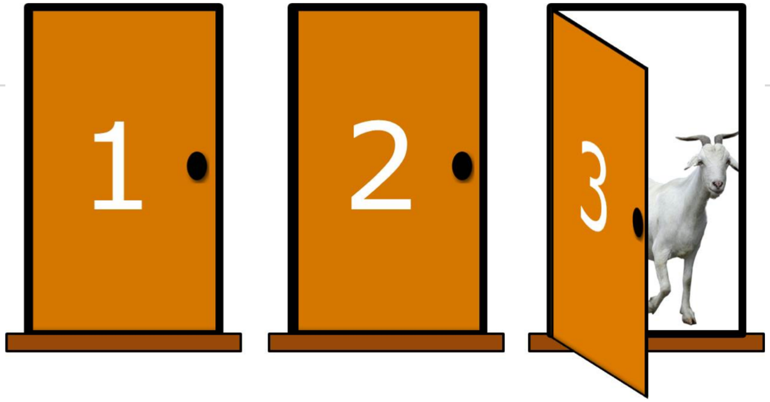
We can assume that initially, there's an equal probability the car is behind any given door:

$$p(D_1) = p(D_2) = p(D_3) = ⅓$$

What should $p(M_3 \mid D_1)$, $p(M_3 \mid D_2)$ and $p(M_3 \mid D_3)$ be?

$\Rightarrow$ $p(M_3 \mid D_3) = 0$, because Monty would never show you which door has the car

Bayes' theorem



Example: Monty Hall Problem

Bayes' Theorem:

$$p(D_i \mid M_3) = \frac{p(M_3 \mid D_i)\, p(D_i)}{\sum_{k=1}^{3} p(M_3 \mid D_k)\, p(D_k)}$$

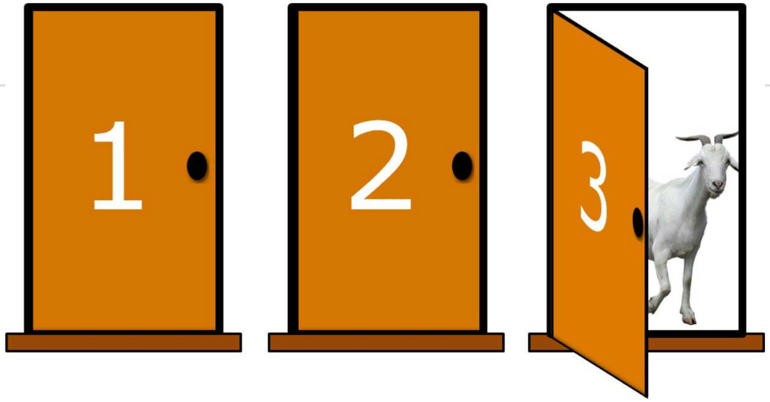
We can assume that initially, there's an equal probability the car is behind any given door:

$$p(D_1) = p(D_2) = p(D_3) = ⅓$$

What should $p(M_3 \mid D_1)$, $p(M_3 \mid D_2)$ and $p(M_3 \mid D_3)$ be?

$\Rightarrow$ $p(M_3 \mid D_2) = 1$, because you guessed Door 1, so Monty can only show Doors 2 or 3 and the car is behind Door 2, so he can only reveal Door 3

Bayes' theorem



Example: Monty Hall Problem

Bayes' Theorem: $$p(D_i \mid M_3) = \frac{p(M_3 \mid D_i)\, p(D_i)}{\sum_{k=1}^{3} p(M_3 \mid D_k)\, p(D_k)}$$

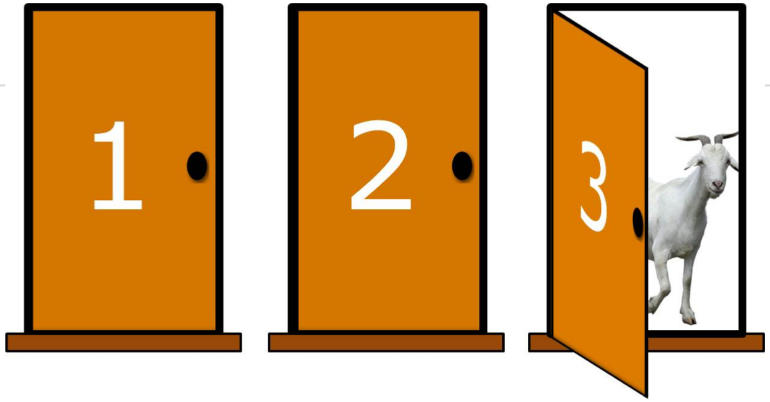
We can assume that initially, there's an equal probability the car is behind any given door:

$$p(D_1) = p(D_2) = p(D_3) = ⅓$$

What should $p(M_3 \mid D_1)$, $p(M_3 \mid D_2)$ and $p(M_3 \mid D_3)$ be?

$\Rightarrow$ $p(M_3 \mid D_1) = ½$, because you guessed Door 1, so Monty can show either Doors 2 or 3 and will with equal probability

Bayes' theorem



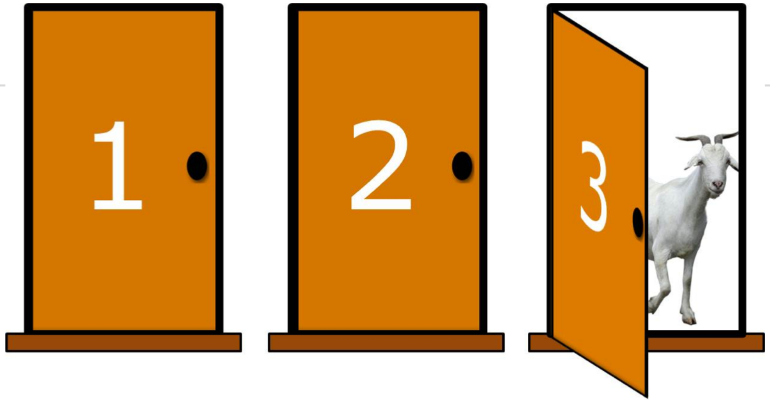
Example: Monty Hall Problem

Bayes' Theorem:

$$p(D_i \mid M_3) = \frac{p(M_3 \mid D_i)\, p(D_i)}{\sum_{k=1}^{3} p(M_3 \mid D_k)\, p(D_k)}$$

Putting all this together, we find:

Bayes’ theorem



Example: Monty Hall Problem

Bayes’ Theorem:
$$p(D_i \mid M_3) = \frac{p(M_3 \mid D_i)\; p(D_i)}{\sum_{k=1}^{3} p(M_3 \mid D_k)\; p(D_k)}$$

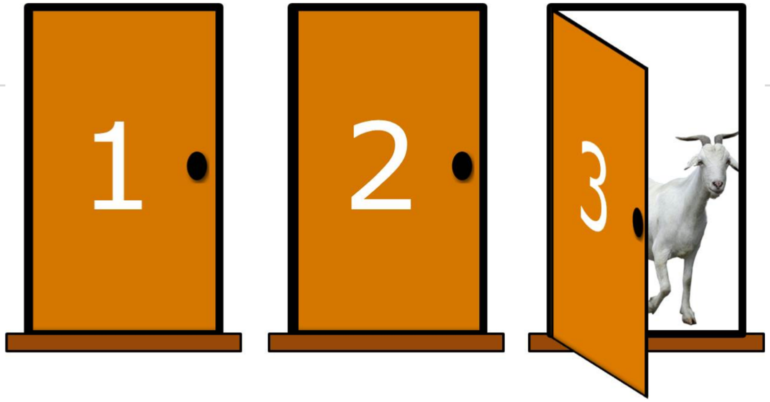
Putting all this together, we find:

$$p(D_1 \mid M_3) = \frac{p(M_3 \mid D_1)\; p(D_1)}{p(M_3 \mid D_1)\; p(D_1) + p(M_3 \mid D_2)\; p(D_2) + p(M_3 \mid D_3)\; p(D_3)}$$

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

$$\Rightarrow \;\; p(D_1 \mid M_3) = \frac{1}{3}$$

Bayes' theorem



Example: Monty Hall Problem

Bayes' Theorem:
$$p(D_i \mid M_3) = \frac{p(M_3 \mid D_i)\ p(D_i)}{\sum_{k=1}^{3} p(M_3 \mid D_k)\ p(D_k)}$$

Similarly…

$$p(D_2 \mid M_3) = \frac{p(M_3 \mid D_2)\ p(D_2)}{p(M_3 \mid D_1)\ p(D_1) + p(M_3 \mid D_2)\ p(D_2) + p(M_3 \mid D_3)\ p(D_3)}$$
$$= \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}}$$

$$\Rightarrow\ p(D_2 \mid M_3) = \frac{2}{3}$$

Next: Applications of Bayes' Theorem!