The Monty Hall Problem

"Let's make a Deal" was a popular TV game show in the 1960s. Monty Hall was the show's host. Thirty three contestants were selected from the studio audience before air time and became the day's possible traders. They swapped items brought with them or played pricing games for gifts in over-sized boxes or behind curtains, with the chance that the prize hidden might either be something more valuable or a worthless gag gift. The players on the trading floor began wearing costumes in 1964.

In one of the games on the show, a contestant was asked to choose one of three doors. Behind one of the doors was a valuable prize (like a car) and behind the other two were goats.
Monty would open a door (other than the one chosen by the contestant) and reveal a goat. He would then offer the contestant a chance to "Stay" or "Switch".

Question: Is it advantageous for the contestant to switch?

PARADE
AskMarilyn® BY MARILYN VOS SAVANT

This question was posed to Marilyn Vos Savant in her "Ask Marilyn" column in the newspaper "Parade". Her answer was that it would always be of advantage to switch. This provoked thousands of letters in response, nearly all arguing that she was wrong and that the doors are equally likely to win. A follow-up column affirming her answer only intensified the debate, which soon spread through the media, even reaching the front page of THE NEW YORK TIMES. Among the ranks of her opponents were hundreds of academics with Ph.D.s, some of them professional mathematicians scolding her for propagating innumeracy.

$$\boxed{\text{What do YOU think?}}$$

On the

Monty Hall Game website

you can find an applet that simulates playing this game. Play a couple times (at least 30 times with each strategy). Which strategy do you think is more promising? Why? Think about it...

Poll Question 4.1

Conditional Probability

A conditional probability of an event is the probability that the event occurs given that some other event has occurred. Conditional probabilities sometimes seem to be counterintuitive and may lead to surprising results.

Example: A rare disease occurs in 0.1% of the population. Fortunately, there is a screening test. If you have the disease, it will be detected with 99% probability. Similarly, if you do not have the disease, the test will reflect this 99% of the times.

A friend of yours gets tested for the disease and his test comes back positive. How concerned should he be (i.e. what is the probability that he actually has the disease)?

To investigate this question we will consider an (imaginary) population of 100,000 people.

		True Diagnosis		
		Positive	Negative	Total
Test Result	Positive	100*0.99= 99	999	1098
	Negative	1	98,901	98,902
	Total	100000*0.001 =100	99,900	100,000

Given that a test is positive, what is the probability of having the disease? Are you surprised about that?

DEFINITION: For any two events A and B with $P(B) > 0$, the conditional probability of A given that B has occurred is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Example: Consider the experiment of tossing a fair coin three times. Let B be the event that exactly two heads are tossed and let A denote the event that the first toss is a head. Find the conditional probability $P(A|B)$.

Remark: Effectively, our knowledge of B reduces the sample space from $\mathcal{S}$ to $\mathcal{S}_B = \{HHT, HTH, THH\}$.

THEOREM: **Multiplication Rule**

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

$$\text{or } P(A \cap B) = P(B|A)P(A)$$

Proof: Follows directly from the definition of conditional probability.

Example: Cards are drawn one by one from a standard 52-card deck without replacement. What is the probability that the fourth heart is drawn on the tenth draw?

Event A: get 3 hearts in first 9 draws
Event B

Remark: The general multiplication rule works also for more than two events, e.g. for three events A, B, C

$$P(A \cap B \cap C) = P(A|B \cap C)P(B|C)P(C).$$

THEOREM: **Law of Total Probability**

$$P(B) = P(B|A)P(A) + P(B|A')P(A')$$

Proof:

REMARK: A set $A_1, \ldots, A_k$ of mutually exclusive and exhaustive events is sometimes referred to as a PARTITION of the sample space $\mathcal{S}$. Recall that mutually exclusive means that no two sets have any common outcomes. Exhaustive means that one of the evens *must* occur, so that $A_1 \cup \cdots \cup A_k = \mathcal{S}$.

Poll Question 4.2

THEOREM: **Law of Total Probability** (general version)

Let $A_1, \ldots, A_k$ be mutually exclusive and exhaustive events. Then for any other event B

$$P(B) = P(B|A_1)P(A_1) + \cdots + P(B|A_k)P(A_k) = \sum_{i=1}^{k} P(B|A_i)P(A_i)$$

Remark: Effectively, the Law of Total Probability is a tool that allows you to tackle the computation of the probability of the event B in smaller steps. Look at everything that could possibly happen *before* our experiment (this is the partition $A_1, \ldots, A_k$). Then compute the probability of B assuming that you are in such a situation (these are the $P(B|A_i)$). Finally, weigh these probabilities with the probabilities that you are in the A_i situation (these are the $P(A_i)$).

Example: Suppose that we have two boxes. Box I contains one white and three black marbles and box II contains two white and two black marbles. We select first a box and then a marble from that box at random. What is the probability that the marble is black?

Remark: Sometimes we are given information about $P(B|A)$ but we are really interested in $P(A|B)$. In the above example that could mean that you are interested in: "If you draw a black marble what is the probability that it came from box I?"

THEOREM: **Bayes' Theorem**

Let $A_1, \ldots, A_k$ be a collection of mutually exclusive and exhaustive events with $P(A_i) > 0$ for $i = 1, \ldots, k$. Then for any other event B with $P(B) > 0$,

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{k} P(B|A_i)P(A_i)}, \qquad j = 1, \ldots, k$$

Example: (cont.)

Again, suppose you have the two boxes from the previous example. If you draw a marble and the marble is black, what is the probability that it came from box I?

Poll Question 4.3

Example: According to the Arizona Chapter of the American Lung Association, 7% of the population has lung disease. Of those people having lung disease, 90% are smokers; and of those not having lung disease, 74.7% are non-smokers. What are the chances that a smoker has lung disease?

Start solving these kinds of problems by defining events and translating the given information into probabilities or conditional probabilities.

Step 1: Define meaningful abbreviations for events that you will need for the problem.

Note: On your homework or on quizzes/exams make sure to write down your definitions. Do not expect that the grader will guess what your abbreviations meant.

Poll Question 4.4

Step 2: Translate all given information into probabilities or conditional probabilities involving your previously defined events.

Step 3: Write down the probability you want to find in terms of your events and look for theorems that connect what you want to find with what you have. Solve.

The Monty Hall Problem - Revisited

Recall: In the show, a contestant was asked to choose one of three doors. Behind one of the doors was a valuable prize and behind the other two was junk (i.e. goats). Monty would open a door (other than the one chosen by the contestant) and reveal a goat. He would then offer the contestant a chance to "Switch".

There are two more or less plausible different theories. Of course, at least one of them must be wrong...

Theory 1: When you first pick your door, the probability that you got the prize is $\frac{1}{3}$. When Monty opens one door, there are now two doors and since we have no information about where the price is, the probability that we got the right door *changes* to $\frac{1}{2}$. In this case it makes no difference if you switch, so you might as well STICK.

Theory 2: When you first pick your door, the probability that you got the prize is $\frac{1}{3}$. This probability *does not* change when Monty opens another door. Since there is only one other door (besides yours) the probability that the price is behind the other door is now $\frac{2}{3}$. Hence you should SWITCH.

Which strategy looks more promising? Let's take a closer look by computing the conditional probabilities that you will win if you switch or that you will win if you stick.

Suppose the prize is behind door one. Consider everything that could possible happen:

You Pick	Door 1		Door 2		Door 3	
Monty Opens	$2(\frac{1}{2})$	$3(\frac{1}{2})$	$3(1)$		$3(1)$	
Switch	X	X	✓		✓	
Stick	✓	✓	X		X	

Poll Question 4.5

Therefore, P(Win|Switch) = $\frac{2}{3}$ and P(Win|Stick) = $\frac{1}{3}$

Does it matter that we assumed the prize to be behind door one?

NO

Independence

For some events, it is fairly obvious that they influence each other. Some examples are

- The crude oil price and the price of gasoline at the pump;
- Being over 80 years of age and developing complications from COVID-19;

For other events it is not so obvious. How about

- Receiving a COVID-19 vaccine and frequency of subsequent infections;
- Having a STEM degree and having a job offer upon graduation.

If our knowledge of an event A does not change the probability that the event B will occur, we say that A and B are independent.

DEFINITION: Two events A and B with $P(A) > 0$ and $P(B) > 0$ are independent if and only if

$$P(A|B) = P(A)$$

or $$P(B|A) = P(B)$$

or $$\boxed{P(A \cap B) = P(A)P(B)}$$ ← often easiest to check!

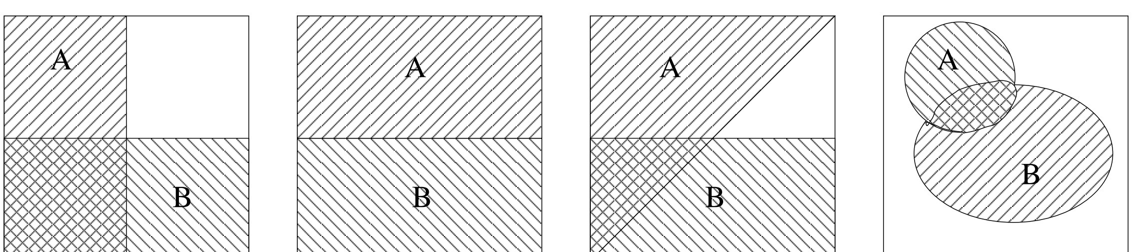
Example: Consider the experiment of drawing one card from a standard 52-card deck. Are the events A = card is an Ace and B = card is a spade independent?

if(P(A\capB) =P(A)*P(B))
P(A)=4/52
P(B)=13/52
P(A\capB)=1/52

Independence is not to be confused with mutually exclusive! It *cannot* usually be seen (or proved) in Venn diagrams.

mutually exclusive != independence

Example: For the following Venn diagrams decide whether or not the events A and B are independent and whether or not they are mutually exclusive.



mutually exclusive ✗ ✓ ✗ ✗

Poll Question 4.6 - 4.9

independence ✓ ✗ ✗ ?

Fact: If two events A and B are mutually exclusive ($A \cap B = \emptyset$), then they cannot be independent as long as $P(A) > 0$ and $P(B) > 0$

Proof:

Fact: If A and B are independent, then so are

$$A \text{ and } B', \quad A' \text{ and } B, \quad A' \text{ and } B'.$$

Proof: