2. Getting Started with Probability

Dave Goldsman

H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

3/2/20

Outline

Georgia Tech

Lesson 2.1 — Experiments, Sample Spaces, and Events

Next Few Lessons:

- Intro to Experiments, Sample Spaces, and Events
- Definition of Probability
- Basic Probability Results
- Finite Sample Spaces

Consider a “random” experiment whose outcome we can observe.

Examples:

- Toss a coin.
- Toss a coin 3 times.
- Ask 10 people if they prefer Coke or Pepsi.
- See how long a light bulb lasts.

Definition: A **sample space** associated with an experiment E is the set of *all* possible outcomes of E. It's usually denoted by S or Ω.

Examples of sample spaces:

- Coin toss: $S = \{H, T\}$.
- Toss a coin 3 times:

  $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

  or $S' = \{0, 1, 2, 3\}$ (number of heads).

  So a sample space doesn't have to be unique!
- Ask 10 people if they prefer Coke or Pepsi: $S = \{0, 1, \ldots, 10\}$.
- Light bulb life: $S = \{t | t \geq 0\}$.

Definition: An **event** is a set of possible outcomes. Thus, any subset of S is an event.

Example: Toss an 8-sided Dungeons and Dragons die. $S = \{1, 2, \ldots, 8\}$. If A is the event "an odd number occurs," then $A = \{1, 3, 5, 7\}$, i.e., when the die is tossed, we get 1 or 3 or 5 or 7.

Remark: The empty set $\emptyset$ is an event of S ("none of the possible outcomes of the experiment are observed").

S is an event of S ("something from the sample space happens").

Remark: If A is an event, then $\bar{A}$ is the complementary (opposite) event.

Example: Toss two coins. $A = \{HH\} \Rightarrow \bar{A} = \{HT, TH, TT\}$.

Remark: If A and B are events, then $A \cup B$ and $A \cap B$ are events.

Example: Toss 3 coins.

$$\begin{aligned} A &= \text{"exactly one } T \text{ was observed"} = \{HHT, HTH, THH\} \\ B &= \text{"no } T\text{'s observed"} = \{HHH\} \\ C &= \text{"first coin is } H\text{"} = \{HHH, HHT, HTH, HTT\} \end{aligned}$$

Then

$$\begin{aligned} A \cup B &= \text{"at most one } T \text{ observed"} \\ &= \{HHT, HTH, THH, HHH\} \\ A \cap C &= \{HHT, HTH\}. \quad \square \end{aligned}$$

Outline

Lesson 2.2 — What Is Probability?

Suppose A is some event from a sample space S. What's the probability that A will occur, i.e., $P(A)$?

Example: Toss a fair coin. $S = \{H, T\}$. What's the probability of H? Intuitively, $P(\{H\}) = P(H) = 1/2$.

What does this mean?

Frequentist view: If the experiment were repeated n times, where n is very large, we'd expect about 1/2 of the tosses to be H's.

$$\frac{\text{Total \# of } H\text{'s out of } n \text{ tosses}}{n} \approx 1/2.$$

Example: Toss a fair die. $S = \{1, 2, 3, 4, 5, 6\}$, where each individual outcome has probability 1/6. Then $P(1 \text{ or } 2) = 1/3$.

More-Formal Definition: The **probability** of a generic event $A \subseteq S$ is a function that adheres to the following *axioms*:

(1) $0 \leq P(A) \leq 1$ (probabilities are *always* between 0 and 1).

(2) $P(S) = 1$ (probability of *some* outcome is 1).
Example: Die. $P(S) = P(1, 2, 3, 4, 5, 6) = 1$.

(3) If A and B are *disjoint* events, i.e., $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.
Example: $P(1 \text{ or } 2) = P(1) + P(2) = 1/6 + 1/6 = 1/3$.

(4) Suppose $A_1, A_2, \ldots$ is a sequence of *disjoint* events (i.e., $A_i \cap A_j = \emptyset$ for $i \neq j$). Then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Georgia Tech

Example: Toss a coin until the first H appears.

$$S = \{H, TH, TTH, TTTH, \ldots\}.$$

Define the *disjoint* events

$$A_1 = \{H\}, A_2 = \{TH\}, A_3 = \{TTH\}, \ldots.$$

Then

$$1 = P(S) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{\infty} \frac{1}{2^i}.$$

(We'll eventually see why that last equality holds, though it may already be intuitively obvious.)

Outline

Lesson 2.3 — Basic Probability Results

Some Nifty Properties (see Meyer 1970)

Theorem: $P(\bar{A}) = 1 - P(A)$.

Proof: We have

$$\begin{aligned} 1 &= P(S) \quad \text{(by Axiom (2))} \\ &= P(A \cup \bar{A}) \\ &= P(A) + P(\bar{A}) \quad \text{(by Axiom (3) since } A \cap \bar{A} = \emptyset\text{).} \quad \Box \end{aligned}$$

Example: The probability that it'll rain tomorrow is 1 minus the probability that it won't rain.

Corollary: $P(\emptyset) = 0$. (You must observe *some* outcome from the sample space, i.e., you can't observe *no* outcome.)

Proof: By definition, $\emptyset = \bar{S}$; so the result follows from the Theorem and Axiom (2). $\Box$

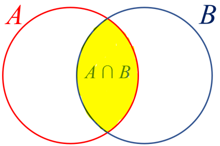
Remark: The converse is *false*: $P(A) = 0$ does *not* imply $A = \emptyset$.

Example: Pick a random number between 0 and 1. Later on, we'll show why any particular outcome actually has probability 0!

Theorem: For any two events A and B,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Proof: Easy Venn diagram argument. (Subtract $P(A \cap B)$ to avoid double-counting.)



Remark: Axiom (3) is a "special case" of this theorem with $A \cap B = \emptyset$.

Example: Suppose there's…

40% chance of colder weather,
10% chance of rain *and* colder weather,
80% chance of rain *or* colder weather.

Then the chance of rain is

$$\begin{aligned} P(R) &= P(R \cup C) - P(C) + P(R \cap C) \\ &= 0.8 - 0.4 + 0.1 \;=\; 0.5. \quad \Box \end{aligned}$$

Theorem: For any three events A, B, and C,

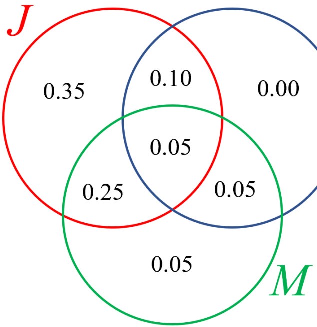
$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &\quad + P(A \cap B \cap C). \end{aligned}$$

The formal proof is a bit tedious. You can try an informal proof via Venn diagrams, but you'll need to be careful about double and triple counting events.

Example: 75% of Atlantans jog (J), 20% like ice cream (I), and 40% enjoy music (M). Also, 15% J *and* I, 30% J *and* M, 10% I *and* M, and 5% do all three. Find the probability that a random resident will engage in at least one of the three activities.

$$\begin{aligned}
&P(J \cup I \cup M) \\
&\quad = \quad P(J) + P(I) + P(M) \\
&\qquad\quad -P(J \cap I) - P(J \cap M) - P(I \cap M) \\
&\qquad\quad +P(J \cap I \cap M) \\
&\quad = \quad 0.75 + 0.20 + 0.40 - 0.15 - 0.30 - 0.10 + 0.05 \\
&\quad = \quad 0.85. \quad \Box
\end{aligned}$$

Now find the probability of precisely one activity. We can use a Venn diagram, starting from the center (since $P(J \cap I \cap M) = 0.05$) and building out.



$$\begin{aligned} & P(\text{only } J) + P(\text{only } I) + P(\text{only } M) \\ & \quad = \quad P(J \cap \bar{I} \cap \bar{M}) + P(\bar{J} \cap I \cap \bar{M}) + P(\bar{J} \cap \bar{I} \cap M) \\ & \quad = \quad 0.35 + 0 + 0.05 \; = \; 0.40. \quad \Box \end{aligned}$$

Theorem: Here is the general **principle of inclusion-exclusion**:

$$\begin{aligned} & P(A_1 \cup A_2 \cup \cdots \cup A_n) \\ &= \sum_{i=1}^{n} P(A_i) - \sum\sum_{i<j} P(A_i \cap A_j) \\ &\quad + \sum\sum\sum_{i<j<k} P(A_i \cap A_j \cap A_k) \\ &\quad - \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n). \end{aligned}$$

Remark: You "include" all of the "single" events, "exclude" the double events, include the triple events, etc.

The proof of this thing is quite tedious. In any case, the previous two theorems are special cases.

Outline

Lesson 2.4 — Finite Sample Spaces

Suppose S is **finite**, say $S = \{s_1, s_2, \dots, s_n\}$. Finite sample spaces often allow us to calculate the probabilities of certain events more efficiently.

To illustrate, let $A \subseteq S$ be any event. Then $P(A) = \sum_{s_i \in A} P(s_i)$.

Example: You have 2 red cards, a blue card, and a yellow card. Pick one card at random, where "**at random**" means that each of the 4 cards has the same probability (1/4) of being chosen.

The sample space $S = \{\text{red, blue, yellow}\} = \{s_1, s_2, s_3\}$.

$P(s_1) = 1/2$, $P(s_2) = 1/4$, $P(s_3) = 1/4$.

$P(\text{red or yellow}) = P(s_1) + P(s_3) = 3/4$. $\square$

Definition: A **simple sample space** (SSS) is a finite sample space in which all outcomes are *equally likely*.

Remark: In the above example, S is *not* simple since $P(s_1) \neq P(s_2)$.

Example: Toss 2 fair coins.

$S = \{HH, HT, TH, TT\}$ is a SSS (all probabilities are 1/4).

$S' = \{0, 1, 2\}$ (number of H's) is *not* a SSS. Why not? $\Box$

Theorem: For any event A in a SSS S,

$$P(A) \;=\; \frac{|A|}{|S|} \;=\; \frac{\text{\# elements in } A}{\text{\# elements in } S}.$$

Example: Toss a die. Let $A = \{1, 2, 4, 6\}$. Each outcome has probability 1/6, so $P(A) = 4/6$. $\Box$

Example: Roll a pair of dice. Possible results (each w.p. 1/36):

$$\begin{array}{cccc} 1,1 & 1,2 & \cdots & 1,6 \\ 2,1 & 2,2 & \cdots & 2,6 \\ & \vdots & & \\ 6,1 & 6,2 & \cdots & 6,6 \end{array}$$

Sum	2	3	4	5	6	7	8	9	10	11	12
Prob	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

E.g., $P(\text{Sum} = 4) = P((1,3)) + P((2,2)) + P((3,1)) = 3/36.$ $\Box$

With this material in mind, we can now move on to more-complicated counting problems. . . .

Outline

Lesson 2.5 — Counting Techniques: Baby Examples

Next Few Lessons: Count the elements in events from a SSS in order to calculate certain probabilities efficiently. We'll look at various helpful rules / techniques, including (i) some intuitive baby examples, (ii) **permutations,** and (iii) **combinations.**

Baby Example: Suppose that you can make choice A in n_A ways, and you can make choice B in n_B ways. If only one choice can be made, you have $n_A + n_B$ ways of doing so. For instance, go to Starbucks and have a muffin (blueberry or oatmeal) or a bagel (sesame, plain, salt, garlic), but not both. You have $2 + 4 = 6$ choices in total. $\Box$

Baby Example: $n_{\mathrm{AB}} = 3$ ways to go from City A to B (walk, car, bus), and $n_{\mathrm{BC}} = 4$ ways to go from B to C (car, bus, train, plane). Then you can go from A to C (via B) using $n_{\mathrm{AB}}\, n_{\mathrm{BC}} = 12$ itineraries. $\Box$

Baby Example: Roll 2 dice. How many outcomes? (Assume $(3,2) \neq (2,3)$.) Answer is $6 \times 6 = 36$. $\Box$

Example: Toss n dice. 6^n possible outcomes. $\Box$

Example: Flip 3 coins. $2 \times 2 \times 2 = 8$ possible outcomes,

$$\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

What's the probability that the 3rd coin is a head?

Answer (by looking at the list) $= 4/8$. $\Box$

Example: Select 2 cards from a deck ***without replacement*** and ***care about order*** (i.e., $(Q\spadesuit, 7\clubsuit) \neq (7\clubsuit, Q\spadesuit)$). How many ways can you do this? Answer: $52 \cdot 51 = 2652$. $\Box$

Example: Box of 10 sox — 2 red and 8 black. Pick 2 without replacement.

(a) Let A be the event that both are red.

$$P(A) \;=\; \frac{\text{\# ways to pick 2 reds}}{\text{\# ways to pick 2 sox}} \;=\; \frac{2 \cdot 1}{10 \cdot 9} \;=\; \frac{1}{45}. \quad \Box$$

(b) Let B be the event that both are black. $P(B) = \frac{8\cdot 7}{10\cdot 9} = \frac{28}{45}$. $\Box$

(c) Let C be one of each color. Since A and B are disjoint, we have

$$P(C) \;=\; 1 - P(\bar{C}) \;=\; 1 - P(A \cup B) \;=\; 1 - P(A) - P(B) \;=\; \frac{16}{45}. \quad \Box$$

Outline

Lesson 2.6 — Counting Techniques: Permutations

Definition: An arrangement of n symbols in a ***definite order*** is a **permutation** of the n symbols.

Example: How many ways to arrange the numbers 1,2,3?

Answer: 6 ways — 123, 132, 213, 231, 312, 321. $\square$

Example: How many ways to arrange $1, 2, \ldots, n$?

$$\text{(choose first)(choose second)} \cdots \text{(choose } n\text{th)}$$

$$n(n-1)(n-2)\cdots 2 \cdot 1 \;=\; n!. \quad \square$$

Example: A baseball manager has 9 players on his team. Find the number of possible batting orders. Answer: $9! = 362880$. $\square$

Definition: The number of r-tuples we can make from n different symbols (each used at most once) is called the **number of permutations of n things taken r-at-a-time**,

$$P_{n,r} \equiv \frac{n!}{(n-r)!}. \qquad (*)$$

Note that $0! = 1$ and $P_{n,n} = n!$.

Example: How many ways can you take two symbols from a, b, c, d?

Answer: $P_{4,2} = 4!/(4-2)! = 12$. Let's list them:

$$ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc. \quad \square$$

Proof (of $(*)$):

$$\begin{aligned}
P_{n,r} &= \text{(choose first)(choose second)} \cdots \text{(choose } r\text{th)} \\
&= n(n-1)(n-2)\cdots(n-r+1) \\
&= \frac{n(n-1)\cdots(n-r+1)(n-r)\cdots 2\cdot 1}{(n-r)\cdots 2\cdot 1} \\
&= \frac{n!}{(n-r)!}. \quad \square
\end{aligned}$$

Georgia Tech

Example: How many ways to fill the first 4 positions of a batting order?

$n = 9$ players, $r = 4$ positions.

$P_{9,4} = 9!/(9-4)! = 3024$ ways. $\square$

Example: How many of these 3024 ways has Smith batting first?

Method 1: First 4 positions: (Smith,?,?,?). This is equivalent to choosing 3 players from the remaining 8.

$P_{8,3} = 8!/(8-3)! = 336$ ways.

Method 2: It's clear that each of the 9 players is equally likely to bat first. Thus, $3024/9 = 336$. $\square$

Example: How many license plates of 6 digits can be made from the numbers $\{1, 2, \ldots, 9\} \ldots$?

(a) with no repetitions? (e.g., 123465 is OK, but 133354 isn't OK)
$P_{9,6} = 9!/3! = 60480.$

(b) allowing repetitions? (anything's OK) $9 \times 9 \times \cdots \times 9 = 9^6 = 531441.$

(c) containing repetitions? $531441 - 60480 = 470961.$ $\square$

Outline

Lesson 2.7 — Counting Techniques: Combinations

Suppose we only want to count the number of ways to choose r out of n objects ***without regard to order***. This is equivalent to counting the number of different *subsets* of these n objects that contain exactly r objects.

Example: How many subsets of $\{1, 2, 3\}$ contain exactly 2 elements? (Order isn't important.)

3 subsets — $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$

Definition: The number of subsets with r elements of a set with n elements is called the **number of combinations of n things taken r-at-a-time**.

Notation: $\binom{n}{r}$ or $C_{n,r}$ (read as "n choose r"). These are also called **binomial coefficients**. It turns out (see below) that $C_{n,r} = \frac{n!}{r!(n-r)!}$.

The difference between permutations and combinations:

Combinations — not concerned with order: $(a, b, c) = (b, a, c)$.

Permutations — concerned with order: $(a, b, c) \neq (b, a, c)$.

The number of permutations of n things taken r-at-a-time is always as least as large as the number of combinations. In fact,…

Choosing a permutation is the same as first choosing a combination *and* then putting the elements in order, i.e.,

$$\frac{n!}{(n-r)!} = \binom{n}{r} r!,$$

and so

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}.$$

In particular, the following results should all be intuitive:

$$\binom{n}{r} = \binom{n}{n-r}, \quad \binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{1} = \binom{n}{n-1} = n.$$

Binomial Theorem: We won't prove it here, but you may know this famous result.

$$(x+y)^n \;=\; \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

This is where Pascal's Triangle comes from!

Corollary: Surprising fact:

$$\sum_{i=0}^{n} \binom{n}{i} \;=\; 2^n.$$

Proof: By the Binomial Theorem,

$$2^n \;=\; (1+1)^n \;=\; \sum_{i=0}^{n} \binom{n}{i} 1^i 1^{n-i}. \quad \square$$

Example: An NBA team has 12 players. How many ways can the coach choose the starting 5? (Order doesn't matter.)

$$\binom{12}{5} = \frac{12!}{5!7!} = 792. \quad \Box$$

Example: Smith is one of the players on the team. How many of the 792 starting line-ups include him?

$$\binom{11}{4} = \frac{11!}{4!7!} = 330.$$

(Smith gets one of the five positions for free; there are now 4 left to be filled by the remaining 11 players.) $\Box$

Example: Given 7 red shoes and 5 blues, find the number of arrangements.

R B R R B B R R R B R B

I.e., how many ways can you put 7 reds in 12 slots?

Answer: $\binom{12}{7}$. $\square$

How many ways to put 5 blues in 12 slots? Same answer. $\square$

Outline

Georgia Tech

Lesson 2.8 — Hypergeometric, Binomial, and Multinomial Problems

Next Few Lessons — all involve interesting applications:

- Hypergeometric Distribution (sampling without replacement)
- Binomial Distribution (sampling with replacement)
- Multinomial Coefficients (generalizes binomial)
- Permutations vs. Combinations
- The Birthday Problem
- The Envelope Problem
- Poker Probabilities

Hypergeometric Distribution

Definition: You have a objects of type 1 and b objects of type 2. Select n objects ***without replacement*** from the $a+b$ objects. Then

$$
\begin{aligned}
& P(k \text{ type 1's were picked}) \\
&= \frac{(\#\text{ ways to choose } k \text{ type 1's out of } a)(\text{choose } n-k \text{ type 2's out of } b)}{\#\text{ ways to choose } n \text{ out of } a+b} \\
&= \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}}. \quad \square
\end{aligned}
$$

The number of type 1's chosen is said to have the **hypergeometric distribution**. We'll have a very thorough discussion on "distributions" later.

Example: 3 sox in a box, with $a = 2$ red, $b = 1$ blue. Pick $n = 3$ *without* replacement. It is trivial to see that

$$P(\text{exactly } k = 2 \text{ reds are picked}) = \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}} = \frac{\binom{2}{2}\binom{1}{1}}{\binom{3}{3}} = 1. \quad \Box$$

Example: 25 sox in a box. 15 red, 10 blue. Pick 7 without replacement. Then

$$P(\text{exactly 3 reds are picked}) = \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}} = \frac{\binom{15}{3}\binom{10}{4}}{\binom{25}{7}} = 0.1988. \quad \Box$$

Binomial Distribution

Definition: You again have a objects of type 1 and b objects of type 2. Now select n objects ***with replacement*** from the $a+b$ objects.

$$\begin{aligned} & P(k \text{ type 1's were picked}) \\ & \quad = \quad (\#\text{ ways to choose } k \text{ 1's and } n-k \text{ 2's}) \\ & \qquad\qquad \times P(\text{choose } k \text{ 1's in a row, then } n-k \text{ 2's in a row}) \\ & \quad = \quad \binom{n}{k}\left(\frac{a}{a+b}\right)^k\left(\frac{b}{a+b}\right)^{n-k}. \quad \square \end{aligned}$$

The number of type 1's chosen is said to have the **binomial distribution**, which will be discussed in great detail later.

Example: 3 sox in a box. $a = 2$ red, $b = 1$ blue. Pick $n = 3$ *with* replacement. We easily see that

$$P(\text{exactly } k = 2 \text{ reds are picked}) \;=\; \binom{3}{2}\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^{3-2} \;=\; 4/9. \quad \square$$

Example: 25 sox in a box. 15 red, 10 blue. Pick 7 with replacement.

$$P(\text{exactly 3 reds are picked}) \;=\; \binom{7}{3}\left(\frac{15}{25}\right)^3\left(\frac{10}{25}\right)^{7-3} \;=\; 0.1936. \quad \square$$

Make sure to compare these answers with the answers to the analogous hypergeometric examples.

Multinomial Coefficients

Example: n_1 blue sox, n_2 reds. The number of assortments is $\binom{n_1+n_2}{n_1}$ (binomial coefficients).

Generalization (for k types of objects): $n = \sum_{i=1}^{k} n_i$

The number of arrangements is

$$\binom{n}{n_1, n_2, \ldots, n_k} \equiv \frac{n!}{n_1! n_2! \cdots n_k!},$$

and this is known as a **multinomial coefficient**.

Example: How many ways can the letters in "Mississippi" be arranged?

$$\frac{\text{\# of permutations of 11 letters}}{(\#\ m\text{'s})!(\#\ p\text{'s})!(\#\ i\text{'s})!(\#\ s\text{'s})!} = \frac{11!}{1!2!4!4!} = 34{,}650. \quad \square$$

Outline

Lesson 2.9 — Permutations vs. Combinations

Which technique should you use?

It's all how you approach the problem!

Example: 4 red marbles, 2 whites. Put them in a row in random order. Find…

(a) P(2 end marbles are W).

(b) P(2 end marbles aren't both W).

(c) P(2 W's are side by side).

Method 1 (using permutations): Let the sample space

$$S = \{\text{every random ordering of the 6 marbles}\}.$$

(a) A: 2 end marbles are W — WRRRRW.

$$\begin{aligned} |A| &= (\text{\# ways to permute the 2 W's in the end slots}) \\ &\quad \times (\text{\# ways to permute the 4 R's in the middle slots}) \\ &= 2! \times 4! = 48. \end{aligned}$$

This implies that

$$P(A) = \frac{|A|}{|S|} = \frac{48}{720} = \frac{1}{15}.$$

(b) $P(\bar{A}) = 1 - P(A) = 14/15$.

(c) B: 2 W's side by side — WWRRRR or RWWRRR or … or RRRRWW.

$$\begin{aligned} |B| &= \text{(\# ways to select pair of slots for 2 W's)} \\ &\qquad \times \text{(\# ways to insert W's into pair of slots)} \\ &\qquad \times \text{(\# ways to insert R's into remaining slots)} \\ &= 5 \times 2! \times 4! \;=\; 240. \end{aligned}$$

Thus,

$$P(B) \;=\; \frac{|B|}{|S|} \;=\; \frac{240}{720} \;=\; \frac{1}{3}. \quad \square$$

But — The above method took too much time! Here's an easier way….

Method 2 (using combinations): Which 2 positions do the W's occupy? Now let

$$S = \{\text{possible pairs of slots that the W's occupy}\}.$$

Clearly, $|S| = \binom{6}{2} = 15$.

(a) Since the W's must occupy the end slots in order for A to occur, $|A| = 1 \Rightarrow P(A) = |A|/|S| = 1/15$.

(b) $P(\bar{A}) = 14/15$.

(c) $|B| = 5 \Rightarrow P(B) = 5/15 = 1/3$. $\square$

(That was much nicer!)

Outline

Lesson 2.10 — The Birthday Problem

There are n people in a room. Find the probability that at least two have the same birthday. (Ignore Feb. 29, and assume that all 365 days have equal probability.)

The (simple) sample space is $S = \big\{(x_1, \dots, x_n) : x_i \in \{1, 2, \dots, 365\},\ \forall i\big\}$ (x_i is person i's birthday), and note that $|S| = (365)^n$.

Let A: All birthdays are different. Then

$$|A| \ = \ P_{365,n} \ = \ (365)(364)\cdots(365 - n + 1).$$

Thus, we have

$$P(A) = \frac{(365)(364)\cdots(365-n+1)}{(365)^n} = 1 \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365-n+1}{365}.$$

But we want

$$P(\bar{A}) = 1 - \left(1 \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365-n+1}{365}\right).$$

Notes: When $n = 366$, $P(\bar{A}) = 1$.

For $P(\bar{A})$ to be $> 1/2$, n must be ≥ 23 (a bit surprising?)

When $n = 50$, $P(\bar{A}) = 0.97$. $\square$

Outline

Lesson 2.11 — The Envelope Problem

A group of n people receives n envelopes with their names on them — but someone has completely mixed up the envelopes! Find the probability that at least one person will receive the proper envelope.

(FYI, there are lots of variations to this story that are mathematically equivalent.)

Let A_i: Person i receives the correct envelope.

We obviously want $P(A_1 \cup A_2 \cup \cdots \cup A_n)$.

By the general principle of inclusion-exclusion, we have....

$$
\begin{aligned}
& P(A_1 \cup A_2 \cup \cdots \cup A_n) \\
&= \sum_{i=1}^{n} P(A_i) - \sum\sum\nolimits_{i<j} P(A_i \cap A_j) \\
&\quad + \sum\sum\sum\nolimits_{i<j<k} P(A_i \cap A_j \cap A_k) \\
&\quad - \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n).
\end{aligned}
$$

Since all of the $P(A_i)$'s are the same, all of the $P(A_i \cap A_j)$'s are the same, etc., we have

$$
\begin{aligned}
& P(A_1 \cup A_2 \cup \cdots \cup A_n) \\
&= nP(A_1) - \binom{n}{2} P(A_1 \cap A_2) + \binom{n}{3} P(A_1 \cap A_2 \cap A_3) \\
&\quad - \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n).
\end{aligned}
$$

Finally, $P(A_1) = 1/n$, $P(A_1 \cap A_2) = 1/(n(n-1))$, etc. (why?) imply that

$$\begin{aligned} & P(A_1 \cup A_2 \cup \cdots \cup A_n) \\ &= \frac{n}{n} - \binom{n}{2}\frac{1}{n}\cdot\frac{1}{n-1} + \binom{n}{3}\frac{1}{n}\cdot\frac{1}{n-1}\cdot\frac{1}{n-2} - \cdots + (-1)^{n-1}\frac{1}{n!} \\ &= 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1}\frac{1}{n!} \doteq 1 - \frac{1}{e} \doteq 0.6321. \quad \Box \end{aligned}$$

Example: If there are just $n = 4$ envelopes, then

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} = 0.625,$$

which is right on the asymptotic money! $\Box$

Outline

Lesson 2.12 — Poker Problems

Draw 5 cards at random from a standard deck.

The number of possible hands is $|S| = \binom{52}{5} = 2{,}598{,}960$.

Terminology (not that I'm advocating gambling, but if you aren't familiar with poker, take a few minutes to learn the basics):

rank $= 2, 3, \ldots, Q, K, A$,
suit $= \clubsuit, \diamondsuit, \heartsuit, \spadesuit$.

We will calculate the probabilities of obtaining various special "hands"….

(a) **2 pairs** — e.g., $A\heartsuit, A\clubsuit, 3\heartsuit, 3\diamondsuit, 10\spadesuit$. (This hand does *not* include a full house, which will be explained below.)

Select 2 ranks (e.g., $A, 3$). Can do this $\binom{13}{2}$ ways.

Select 2 suits for first pair (e.g., $\heartsuit, \clubsuit$). $\binom{4}{2}$ ways.

Select 2 suits for second pair (e.g., $\heartsuit, \diamondsuit$). $\binom{4}{2}$ ways.

Select remaining card to complete the hand. 44 ways (because we are not allowing a full house).

$$|\text{2 pairs}| = \binom{13}{2}\binom{4}{2}\binom{4}{2}44 = 123{,}552$$

$$P(\text{2 pairs}) = \frac{123{,}552}{2{,}598{,}960} \doteq 0.0475. \quad \square$$

(b) **Full house** (1 pair, 3-of-a-kind) — e.g., $A\heartsuit, A\clubsuit, 3\heartsuit, 3\diamondsuit, 3\spadesuit$

Select 2 *ordered* ranks (e.g., $A, 3$) (because the triple and the pair are different). $P_{13,2}$ ways.

Select 2 suits for pair (e.g., $\heartsuit, \clubsuit$). $\binom{4}{2}$ ways.

Select 3 suits for 3-of-a-kind (e.g., $\heartsuit, \diamondsuit, \spadesuit$). $\binom{4}{3}$ ways.

$$|\text{full house}| \;=\; 13 \cdot 12 \binom{4}{2} \binom{4}{3} \;=\; 3744$$

$$P(\text{full house}) \;=\; \frac{3744}{2{,}598{,}960} \;\doteq\; 0.00144. \quad \square$$

(c) **Flush** (all 5 cards from same suit) (This includes all kinds of flushes.)

Select a suit. $\binom{4}{1}$ ways.

Select 5 cards from that suit. $\binom{13}{5}$ ways.

$$P(\text{flush}) = \frac{5148}{2{,}598{,}960} \doteq 0.00198. \quad \Box$$

(d) **Straight** (5 ranks in a row) (This includes all straights.)

Select a starting point for the straight $(A, 2, 3, \ldots, 10)$. $\binom{10}{1}$ ways.

Select a suit for each card in the straight. 4^5 ways.

$$P(\text{straight}) = \frac{10 \cdot 4^5}{2{,}598{,}960} \doteq 0.00394. \quad \Box$$

(e) Straight flush

Select a starting point for the straight. 10 ways.

Select a suit. 4 ways.

$$P(\text{straight flush}) = \frac{40}{2{,}598{,}960} \doteq 0.0000154. \quad \square$$

Remark: Can you do bridge problems? Yahtzee?

Outline

Lesson 2.13 — Conditional Probability

Next Few Lessons:

- Conditional Probability
- Independent Events
- Partition of a Sample Space and the Law of Total Probability
- Bayes Theorem (updating probabilities in a clever way)

Idea: Update probabilities as we obtain more information.

Example: If A is the event that a person weighs at least 150 pounds, then $P(A)$ certainly depends on the person's height, e.g., if B is the event that the person is at least 6 feet tall vs. B being the event that the person is < 5 feet tall.

Example: Die. $A = \{2, 4, 6\}$, $B = \{1, 2, 3, 4, 5\}$. So $P(A) = 1/2$, $P(B) = 5/6$.

Suppose we *know* that B occurs (so that there is no way that a "6" can come up). Then the probability that A occurs given that B occurs is

$$P(A|B) = \frac{2}{5} = \frac{|A \cap B|}{|B|}. \quad \Box$$

So the probability of A depends on the info that you have! The info that B occurs allows us to regard B as a new, restricted sample space. Assuming we have a simple sample space, then

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

Definition: If $P(B) > 0$, the **conditional probability of A given B** is $P(A|B) \equiv P(A \cap B)/P(B)$.

Remarks: If A and B are disjoint, then $P(A|B) = 0$. (If B occurs, there's no chance that A can also occur.)

What happens if $P(B) = 0$? Don't worry! In this case, makes no sense to consider $P(A|B)$.

Example: Toss 2 dice and take the sum.

A: odd sum $= \{3, 5, 7, 9, 11\}$
B: $\{2, 3\}$

$$\begin{aligned} P(A) &= P(3) + P(5) + \cdots + P(11) = \frac{2}{36} + \frac{4}{36} + \cdots + \frac{2}{36} = \frac{1}{2}, \\ P(B) &= P(2) + P(3) = \frac{1}{36} + \frac{2}{36} = \frac{1}{12}, \\ P(A|B) &= \frac{P(A \cap B)}{P(B)} = \frac{P(3)}{P(B)} = \frac{2/36}{1/12} = \frac{2}{3}. \end{aligned}$$

Thus, in light of the information provided by B, we see that $P(A) = 1/2$ increases to $P(A|B) = 2/3$. $\Box$

Example: 4 white sox, 8 red. Select 2 without replacement.

A: 1st sock is white
B: 2nd sock is white
C: Both are white ($= A \cap B$).

$$P(C) \;=\; P(A \cap B) \;=\; P(A)P(B|A) \;=\; \frac{4}{12} \cdot \frac{3}{11} \;=\; \frac{1}{11}.$$

It is easy to see that $B = (A \cap B) \cup (\bar{A} \cap B)$, where the two components of the union are disjoint. So

$$\begin{aligned} P(B) &= P(A \cap B) + P(\bar{A} \cap B) \\ &= P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) \\ &= \frac{4}{12} \cdot \frac{3}{11} + \frac{8}{12} \cdot \frac{4}{11} \;=\; \frac{1}{3}. \quad \Box \end{aligned}$$

Could you have gotten this result without thinking?

Example: A couple has two kids and at least one is a boy. What's the probability that BOTH are boys?

$S = \{GG, GB, BG, BB\}$ ('BG' means 'boy then girl')

C: Both are boys $= \{BB\}$.
D: At least 1 is a boy $= \{GB, BG, BB\}$.

$$P(C|D) = \frac{P(C \cap D)}{P(D)} = \frac{P(C)}{P(D)} = \frac{1/4}{3/4} = 1/3.$$

(My intuition was 1/2 — the *wrong* answer! The problem was that we didn't know whether D meant the first or second kid.) $\square$

As you get more information, you can make some surprising findings. . . .

Honors Example: A couple has two kids and at least one is a boy ***born on a Tuesday.*** What's the probability that BOTH are boys?

B_x [G_x] = Boy [Girl] born on day x, $x = 1, 2, \ldots, 7$ ($x = 3$ is Tuesday).

$S = \{(G_x, G_y), (G_x, B_y), (B_x, G_y), (B_x, B_y),\ x, y = 1, 2, \ldots, 7\}$
(so $|S| = 4 \times 49 = 196$).

C: Both are boys (with at least one born on a Tuesday)
$= \{(B_x, B_3),\ x = 1, 2, \ldots, 7\} \cup \{(B_3, B_y),\ y = 1, 2, \ldots, 7\}$.
Note that $|C| = 13$ (to avoid double counting (B_3, B_3)).

D: There is at least one boy born on a Tuesday
$= C \cup \{(G_x, B_3), (B_3, G_y),\ x, y = 1, 2, \ldots, 7\}$.
So $|D| = 27$ (list 'em out if you don't believe me). Then

$$P(C|D) = \frac{P(C \cap D)}{P(D)} = \frac{P(C)}{P(D)} = \frac{13/196}{27/196} = \mathbf{13/27}. \quad \square$$

Properties — analogous to Axioms of Probability.

(1′) $0 \leq P(A|B) \leq 1$.

(2′) $P(S|B) = 1$.

(3′) $A_1 \cap A_2 = \emptyset \Rightarrow P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$.

(4′) If $A_1, A_2, \ldots$ are all disjoint, then $P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) = \sum_{i=1}^{\infty} P(A_i|B)$.

Outline

Lesson 2.14 — Independence Day

Any unrelated events are independent.

A: It rains on Mars tomorrow.
B: Coin lands on H.

Definition: A and B are **independent** iff $P(A \cap B) = P(A)P(B)$.

Example: If $P(\text{rains on Mars}) = 0.2$ and $P(H) = 0.5$, then
$P(R \cap H) = P(R)P(H) = 0.1$. $\square$

Remark: If $P(A) = 0$, then A is independent of any other event.

Remark: Events don't have to be physically unrelated to be independent.

Example: Die. $A = \{2, 4, 6\}$, $B = \{1, 2, 3, 4\}$, $A \cap B = \{2, 4\}$, so $P(A) = 1/2$, $P(B) = 2/3$, $P(A \cap B) = 1/3$.

$P(A)P(B) = 1/3 = P(A \cap B) \Rightarrow A, B$ independent. $\square$

More natural interpretation of independence...

Theorem: Suppose $P(B) > 0$. Then A and B are independent $\Leftrightarrow$ $P(A|B) = P(A)$.

Proof: A and B independent $\Leftrightarrow P(A \cap B) = P(A)P(B) \Leftrightarrow$ $P(A \cap B)/P(B) = P(A)$. $\square$

Remark: So if A and B are independent, the probability of A doesn't depend on whether or not B occurs.

Bonus Theorem: A and B independent $\Leftrightarrow$ A and $\bar{B}$ independent .

Proof: Only need to prove in $\Rightarrow$ direction (then $\Leftarrow$ follows trivially). To this end, note that

$$P(A) \;=\; P(A\cap\bar{B})+P(A\cap B),$$

so that

$$\begin{aligned} P(A\cap\bar{B}) &= P(A)-P(A\cap B) \\ &= P(A)-P(A)P(B) \quad (A, B \text{ indep}) \\ &= P(A)[1-P(B)] \\ &= P(A)P(\bar{B}). \quad \Box \end{aligned}$$

Georgia Tech

Don't confuse independence with disjointness!

Theorem: If $P(A) > 0$ and $P(B) > 0$, then A and B can't be both independent and disjoint at the same time.

Proof: Suppose A and B are disjoint ($A \cap B = \emptyset$). Then $P(A \cap B) = 0 < P(A)P(B)$. Thus, A and B are not independent. Similarly, independent implies not disjoint. $\square$

Remark: In fact, independence and disjointness are almost opposites. If A and B are disjoint and A occurs, then you have *information* that B cannot occur — so A and B can't be independent!

Extension to more than two events.

Definition: A, B, C are independent iff

(a) $P(A \cap B \cap C) = P(A)P(B)P(C)$, *and*

(b) All *pairs* are independent:

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

Georgia Tech

Careful! Note that condition (a) by itself isn't enough.

Example: $S = \{1, 2, \ldots, 8\}$ (each element w.p. 1/8).
$A = \{1, 2, 3, 4\}$, $B = \{1, 5, 6, 7\}$, $C = \{1, 2, 3, 8\}$.

(a) $A \cap B \cap C = \{1\}$. $P(A \cap B \cap C) = P(A)P(B)P(C) = 1/8$, so (a) is satisfied. However, (b) *is not*. . . .

(b) $A \cap B = \{1\}$. $P(A \cap B) = 1/8 \neq 1/4 = P(A)P(B)$. ※

Stay vigilant, because (b) by itself isn't enough either!

Example: $S = \{1, 2, 3, 4\}$ (each element w.p. 1/4).
$A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{1, 4\}$.

(b) $P(A \cap B) = 1/4 = P(A)P(B)$. Same deal with A, C and B, C. So (b) is OK. But (a) *is not*....

(a) $P(A \cap B \cap C) = 1/4 \neq 1/8 = P(A)P(B)P(C)$. ※

General Definition: $A_1, \ldots, A_k$ are independent iff $P(A_1 \cap \cdots \cap A_k) = P(A_1) \cdots P(A_k)$ *and* all subsets of $\{A_1, \ldots, A_k\}$ are independent.

Independent Trials: Perform n trials of an experiment such that the outcome of one trial is independent of the outcomes of the other trials.

Example: Flip 3 coins independently.

(a) $P(\text{1st coin is } H) = 1/2$. Don't worry about the other two coins since they're independent of the first.

(b) $P(\text{1st coin } H, \text{3rd } T) = P(\text{1st coin } H)P(\text{3rd } T) = 1/4$. $\Box$

Remark: For independent trials, you just multiply the individual probabilities.

Example: Flip a coin infinitely many times (each flip is independent of the others).

$$\begin{aligned} p_n &\equiv P(\text{1st } H \text{ on } n\text{th trial}) \\ &= P(\underbrace{TT\cdots T}_{n-1}H) \\ &= \underbrace{P(T)P(T)\cdots P(T)}_{n-1}P(H) \;=\; 1/2^n. \end{aligned}$$

$$P(H \text{ eventually}) \;=\; \sum_{n=1}^{\infty} p_n \;=\; \sum_{n=1}^{\infty} 2^{-n} \;=\; 1. \quad \square$$

Georgia Tech

Outline

Lesson 2.15 — Partitions and the Law of Total Probability

Partition of a Sample Space — split the sample space into disjoint, yet all-encompassing subsets.

Definition: The events $A_1, A_2, \ldots, A_n$ form a **partition** of sample space S if

- $A_1, A_2, \ldots, A_n$ are disjoint,
- $\bigcup_{i=1}^{n} A_i = S$,
- $P(A_i) > 0$ for all i.

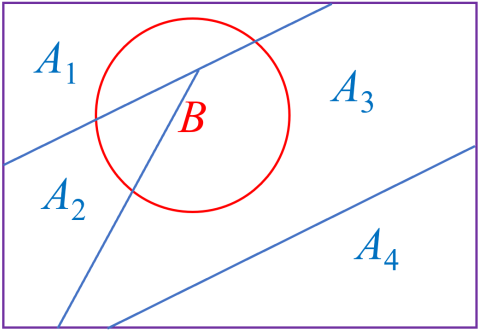
Remark: When an experiment is performed, *exactly one* of the A_i's occurs.

Example: A and $\bar{A}$ form a partition.

Example: "Vowels" and "consonants" form a partition of the letters (if you pretend that only a,e,i,o,u are vowels).

Suppose $A_1, A_2, \ldots, A_n$ form a partition of S, and B is some arbitrary event. Then

$$B = \bigcup_{i=1}^{n} (A_i \cap B).$$



Georgia Tech

So if $A_1, A_2, \ldots, A_n$ is a partition, we can decompose B into pieces of the A_i's. . . .

$$\begin{aligned} P(B) &= P\left(\bigcup_{i=1}^{n}(A_i \cap B)\right) \\ &= \sum_{i=1}^{n} P(A_i \cap B) \quad \text{(since } A_1, A_2, \ldots, A_n \text{ are disjoint)} \\ &= \sum_{i=1}^{n} P(A_i)P(B|A_i) \quad \text{(definition of conditional probability).} \end{aligned}$$

This is the **Law of Total Probability**.

Example: $P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$, which we saw in the previous lesson.

Example: Suppose we have 10 Georgia Tech students and 20 University of Georgia students taking a test. GT students have a 95% chance of passing the test, but UGA students (assuming that they don't cheat) only have a 50% chance of passing. Pick a student at random, and determine the probability that he/she passes.

By the Law of Total Probability,

$$\begin{aligned} P(\text{passes}) &= P(\text{GT})P(\text{passes}\,|\,\text{GT}) + P(\text{UGA})P(\text{passes}\,|\,\text{UGA}) \\ &= (1/3)(0.95) + (2/3)(0.5) \;=\; 0.65. \quad \Box \end{aligned}$$

Outline

Lesson 2.16 — Bayes Theorem

And here is an immediate consequence of the Law of Total Probability.

Bayes Theorem: If $A_1, A_2, \ldots, A_n$ form a partition of S, and B is any event, then

$$P(A_j|B) \;=\; \frac{P(A_j \cap B)}{P(B)} \;=\; \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}.$$

The $P(A_j)$'s are **prior** probabilities ("before B").

The $P(A_j|B)$'s are **posterior** probabilities ("after B").

The $P(A_j|B)$'s add up to 1. That's why the funny-looking denominator.

Example: Two political candidates are at a debate. Candidate A is asked 60% of the questions, and candidate B is asked just 40% (for some reason). Candidate A is likely to make a stupid answer 20% of the time, and B makes a dumb answer a whopping 50% of the time. One of the candidates is asked a question and makes a dumb answer. What's the probability that it was A?

[Answer should be $< 60\%$ since dumb answers favor Candidate B.]

Let D denote the event that the person makes a dumb answer. The relevant partition is $\{A, B\}$.

$$\begin{aligned} P(A|D) &= \frac{P(A)P(D|A)}{P(A)P(D|A) + P(B)P(D|B)} \\ &= \frac{(0.6)(0.2)}{(0.6)(0.2) + (0.4)(0.5)} = 0.375. \end{aligned}$$

Notice how the posterior probabilities depend strongly on the priors and the information we receive. $\square$

Example: In a certain city with good police,

$P(\text{Any defendant brought to trial is guilty}) \;=\; 0.99.$

In any trial,
$P(\text{Jury sets defendant free if he is innocent}) \;=\; 0.95.$
$P(\text{Jury convicts if defendant is guilty}) \;=\; 0.95.$

Find $P(\text{Defendant is innocent}|\text{Jury sets free})$.

Events: $I =$ "innocent", $G =$ "guilty" $= \bar{I}$, $F =$ "sets him free". Since the partition is $\{I, G\}$, Bayes $\Rightarrow$

$$\begin{aligned} P(I|F) &= \frac{P(I)P(F|I)}{P(I)P(F|I) + P(G)P(F|G)} \\ &= \frac{(0.01)(0.95)}{(0.01)(0.95) + (0.99)(0.05)} \;=\; 0.161. \quad \Box \end{aligned}$$

Example: You are a contestant at a game show. Behind one of three doors is a car; behind the other two are goats. You pick Door 1. Monty Hall opens Door 2 and reveals a goat. Monty offers you a chance to switch to Door 3. What should you do? (Note that, if the car is actually behind your Door 1, Monty just randomly shows you either Door 2 or 3.)

By Bayes, we have

$$
\begin{aligned}
& P(\text{Car behind 1} \,|\, \text{Monty shows you Door 2}) \\
&= \frac{P(\text{Monty shows you Door 2} \,|\, \text{Car behind 1}) P(\text{Car behind 1})}{\sum_{i=1}^{3} P(\text{Monty shows you Door 2} \,|\, \text{Car behind } i) P(\text{Car behind } i)} \\
&= \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} \;=\; 1/3.
\end{aligned}
$$

Georgia Tech

On the other hand,

$$P(\text{Car behind 3}\,|\,\text{Monty shows you door 2})$$
$$= \frac{P(\text{Monty shows you door 2}\,|\,\text{Car behind 3})P(\text{Car behind 3})}{\sum_{i=1}^{3} P(\text{Monty shows you door 2}\,|\,\text{Car behind } i)P(\text{Car behind } i)}$$
$$= \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} \;=\; 2/3.$$

Thus, the prudent action is to switch to door 3! $\square$

If you don't quite believe this, you aren't alone. But think what you would do if there were 1000 doors and Monty revealed 998 of them — of course you would switch from your door to the remaining one!