4.6 The Vocabulary and Axioms of Probability

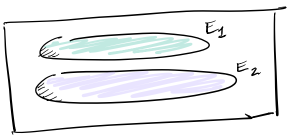
We will be using set theory to help us study and learn probability. Here is some of the notation and vocabulary of probability theory:

- An **experiment** is any process that allows researchers to obtain observations (outcomes).
- An **event** is a set that represents a collection of results or outcomes from an experiment. We will continue to denote these with capital letters: A, B, C.
- A **sample space** for an experiment is a set that consists of all possible outcomes. Again, this will be written Ω.
- P denotes probability.
- $P(A)$ denotes the probability of event A occurring.
- $\mathbb{P}$ is a function whose <u>domain</u> is the collection of events in Ω, and whose <u>range</u> is the set of real numbers $[0,1]$ (i.e., the closed interval $[0,1] := \{x \text{ in } \mathbb{R} \text{ such that } 0 \le x \le 1\}$).

Example 8. *Suppose that we toss a coin twice. The sample space for this experiment is $\Omega = \{HH, HT, TH, TT\}$. Let A be the event of tossing a head on the first toss. Thus, $A = \{HH, HT\}$. Suppose we let B be the event of tossing at least one head. Then $B = \{HH, HT, TH\}$ and we see that $A \subseteq B$.*

Example 9. *Suppose we select a student at random from a class of 30 adult students. Here Ω is the 30 students. Let A be the event that the student selected is a man. Let B be the event that the student selected is a woman. With the assumption that everyone in the class identifies themselves as either male or female, $A \cup B = \Omega$ and $A \cap B = \emptyset$.*

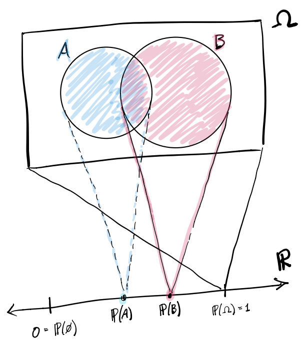
Probability Axioms
A1: $P(A) \ge 0$ for every event A.
A2: $P(\Omega) = 1$.
A3: If E_1 and E_2 are <u>disjoint</u>, then $P(E_1 \sqcup E_2) = P(E_1) + P(E_2)$.



"$E_1 \cap E_2 = \emptyset$."

$\boxed{\text{Notation}}$ If the events E_1, E_2 are disjoint, i.e., if $E_1 \cap E_2 = \emptyset$, we write "$E_1 \sqcup E_2$" to represent the <u>disjoint union</u> of the events E_1 and E_2, which is just a nice way to write $E_1 \cup E_2$ (the union of E_1 and E_2) while remembering that E_1 and E_2 have <u>no elements in common</u>.

4.7 Exercises



"Proof by picture"

1. Give a "proof by picture" that for any events A and B in the sample space Ω of an experiment, because $A \cup B = (A \cap B^c) \sqcup (A \cap B) \sqcup (A^c \cap B)$, the probability of $\mathbb{P}(A \cup B)$ is the sum $\mathbb{P}(A \cap B^c) + \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B)$. Argue that, moreover, because $A = (A \cap B) \sqcup (A \cap B^c)$ and $B = (A \cap B) \sqcup (A^c \cap B)$, substitution implies $\mathbb{P}(A \cap B^c) + \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) + \mathbb{P}(B)$.

2. Let Ω be the set of all students at a given university.

   Several subsets of Ω are represented below. Some of the sets are represented by symbols and some by the probability that a random student will belong to the set.

   Use the given symbols and probabilities to complete the missing items in the table. [1, 2.1.2]

Definition of set in words	Symbol	Probability
All social-science students	S	
All research students	R	
All social-science and/or research students		0.25
All research students who are not from social sciences		0.05
All social-science students who are not research students		0.10
All social-science research students		

3. Here is a rough description of an experiment carried out by a psychologist:

   Subjects were given the following judgment problem: "Think of a population of women with academic degrees in the social sciences. Consider a random woman from such a population. Rank order the following categories according to the probability that the woman will belong to them."

   (a) Employed at a university.
   (b) Married and unemployed.
   (c) Owns her own business.
   (d) Unemployed.

   The subjects' rankings, from the most probable possibility to the least probable one, matched the order in which the categories are written, that it, (a) was rated most probable and (d) least probable.

   Are these rankings compatible with the axioms of probability theory, or do they violate them in any way? Explain. (Note that you are not being asked about the state of employment of academic women. Rather, you are asked to evaluate the consistency of the judgments of the experimental subjects.)[1, 2.1.3]

4. Tom and Harriet cannot agree on whether to go to the baseball game (Tom's choice) or the movies (Harriet's choice). Flipping a coin and deciding for Harriet if 'heads' and for Tom if 'tails' appears too trivial. They discuss several chance procedures for making their decision.

   Let T be the event that 'Tom wins' and H the event that 'Harriet wins'. Find $P(T)$ and $P(H)$ for each of the suggestions described below, and determine whether each procedure is fair or biased.

   (a) Playing one round of "Rock, Paper, Scissors" to determine whose wish will be granted.
   (b) Flipping two coins. Harriet prevails if at least one outcome is heads, Tom– otherwise.
   (c) Giving each a box containing three notes numbered 1,2, and 3, and having each blindly draw one of the notes. If the sum of their draws is even– Tom wins, if it is odd – Harriet wins.
   (d) Rolling two dice and computing the absolute difference between the two numbers obtained (always a positive number). Tom wins if the outcome is either 1 or 2, Harriet–otherwise. [1, 2.1.4]

5. Suppose A and B are two *disjoint* events and that $P(A) = .22; P(B) = .33$. Calculate the following probabilities.

   (a) $P(A \cup B)$
   (b) $P(A \cap B)$
   (c) $P(\overline{A} \cup B)$
   (d) $P(\overline{A} \cap B)$
   (e) $P(\overline{A \cap B})$
   (f) $P(\overline{A} \cap \overline{B})$
   (g) $P(A - B)$. [1, 2.1.5]

6. Let A and B be events so that $A \subseteq B$. Also, let $P(A) = .30$, and $P(B) = .45$. Calculate the following probabilities.

   (a) $P(A \cup B)$
   (b) $P(\overline{A})$
   (c) $P(A \cap B)$
   (d) $P(\overline{A} \cap B)$
   (e) $P(A - B)$
   (f) $P(\overline{A \cup B})$
   (g) $P(\overline{A} - B)$
   (h) $P(\overline{A} - \overline{B})$.[1, 2.1.6]

7. Suppose A and B are events so that $P(A) = .40$ and $P(B) = .25$.
   (a) What is the very largest that $P(A \cup B)$ can be?
   (b) What is the very smallest that $P(A \cup B)$ can be?
   (c) What additional information do you need in order to compute $P(A \cup B)$ exactly?
8. Suppose that A and B are events so that $P(A) = .75$ and $P(B) = .60$. Answer questions (a) – (c) above with these probabilities.
9. Suppose that A and B are events so that $P(A) = 1$ and $P(B) = .35$. Answer questions (a) – (c) above with these probabilities.

4.8 Conditional Probability: Exercises

1. Suppose you have a class with 30 students enrolled. You know that 18 of the students are women and 12 of the students are men. In addition, 10 students wear glasses, 7 of whom are women.

   (a) How many men wear glasses?

   (b) How many women do not wear glasses?

   (c) How many students are either women or wear glasses (or both)?

   (d) Suppose you choose a student at random. What is the probability that you choose a woman?

   (e) What is the probability that the student chosen wears glasses?

   (f) What is the probability that the student chosen does not wear glasses and is not a woman?

   (g) What is the probability that the student chosen is a woman who wears glasses?

   (h) Suppose you know that you have chosen a woman. What is the probability that the chosen student wears glasses? Note: Be careful with this one. Your answer should be different than what you got in (g).

   (i) Suppose you know that the student you have chosen wears glasses. What is the probability that the chosen student also is a woman?

2. Consider the same class as in question 1. Suppose you choose a student at random from this class. Let W be the event that you choose a women, G be the event that you choose a person who wears glasses.

   (a) Use the notation of set theory and probability to describe each of the sets in (d)–(g) above. (For example, the probability that the student chosen is a man would be $P(\overline{W})$.

   (b) Use the notation of set theory and probability to write down a formula for the (h) and (i) from question 1 using the values $P(W), P(G), P(W \cap G)$ (not necessarily all three).

   Notation: $P(B|A)$ is the probability that event B will occur given that the event A has already occurred. We read "$B|A$" as "B given A."

3. Suppose that you have 10 slips of paper numbered 1-10. You put the slips of paper into a hat and draw one slip of paper at random. Let B be the event that you draw a six and let A be the event that you draw an even number. Find $P(B)$ and $P(B|A)$.

4. Find a general formula for $P(B|A)$ involving $P(A \cap B)$ and $P(A)$. Use the previous exercise to justify your answer.

4.9 Conditional Probability: Summary

As you saw in the previous exercises, the general formula for the *conditional probability* $P(B|A)$ is given by:

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

We say A and B are **independent** events if $P(B|A) = P(B)$. This means that the occurrence of A does not effect the probability of the occurrence of B. If A and B are not independent, then they are called **dependent** events.

Two useful formulas:

1. From the definition above, it is easy to see that
$$P(A \cap B) = P(A)P(B|A).$$
2. You should have seen in section 4.7 that:
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

4.10 Exercises

As always, you must FULLY justify your answer. If you are multiplying two things together, it is important to explain why. Often times, your justification will come from the formulas above.

1. You go to a casino to play a game of Texas Hold 'em. You are dealt two cards from an ordinary deck of cards.
   (a) What is the probability that the two cards are both aces?
   (b) What is the probability that the first card dealt is an ace and the second card dealt is a king of the same suit?
   (c) What is the probability that either both cards are aces or the first is an ace and the second is a king of the same suit?
   (d) What is the probability that the two cards are the same number or same face card? For example, two kings or two aces or two jacks or two 3's, etc.
   (e) What is the probability that both cards are face cards or both cards are spades?

2. Five people are standing for the first time on the edge of a cliff in Argentina "ready" to dive into the pool of water below. The five friends have been planning this trip for months but now that the big moment to jump has arrived, they are feeling pretty nervous. The cliff is *very* high. They need to come up with a method to determine who will go first. Unfortunately, standing in their bathing suits at the top of the cliff, shaking with fear, none of them have any ideas. Luckily, a nonpartisan bystander happens to be standing at the top of the cliff with them and offers a solution.

   This bystander's solution is to essentially "draw straws". The bystander is a smoker and takes out five cigarettes. Then, the bystander secretly breaks one of the cigarettes in half and then holds all five in his hand, displaying them to our fearful divers so that they are unable to tell which cigarette is the broken one.

   The five friends are then instructed to order themselves by height, shortest to tallest. Then, the shortest person is to select one of the cigarettes. If the cigarette selected is the broken one, our short friend will dive first. If not, the second shortest person draws. This process continues until one of the friends gets the broken cigarette.

   Is this procedure fair? (You need to figure out the probability that each friend will dive first.)

3. A population is distributed according to the four standard blood types as follows:

   A– 42%
   O – 33%
   B – 18%
   AB – 7%

   Assuming that people choose their mates independent of blood type, calculate the probability that a randomly sampled couple from this population will have the same blood type. [1, 2.3.7]

4. Assume that the weather can be described by one and only one of two states, *fair* or *rainy*, and there exists an unequivocal system for determining the state of the weather on any given day. Assume further that the probability that the weather is in a given state depends only on the preceding day's weather.

   For this problem, let's denote by $P(y|x)$ the probability of weather y on a given day, given that the previous day was x (both x and y assume the values *fair* or *rainy*).

The weather in a given region may described by the following conditional probabilities:
$P(rainy|fair) = 0.2$ $P(fair|fair) = 0.8$
$P(rainy|rainy) = 0.4$ $P(fair|rainy) = 0.6$
It rains Monday. A picnic is planned for Thursday. What is the probability that in three days there will be fair weather? [1, 2.3.13]

5. A doctor is called to see a sick child. The doctor knows (prior to the visit) that 90% of the children in that neighborhood are sick with the flu, denoted F, while 10% are sick with the measles, denoted M. Let us assume for simplicity's sake that M and F are complementary events.

   A well-known symptom of measles is a rash, denoted R. The probability of having a rash for a child sick with the measles is 0.95. However, occasionally children with the flu also develop a rash, with a conditional probability of 0.08.

   Upon examining the child, the doctor finds a rash. What is the probability that the child has the measles? [1, 2.4.1]

6. A man was arrested on suspicion of murder. Let us denote the event 'the man is guilty' by G. The investigating officer collected all the relevant information, added his impressions of the suspect, and arrived at the conclusion that the man's probability of guilt was 0.60.

   (a) As the investigation went on, it was learned (beyond any reasonable doubt) that the murderer's blood type was O. The relative frequency of blood type O in the population is 0.33 (that is the probability that a randomly selected person in the population has blood type O). The suspect's blood was tested and found to be O.

       Compute the 'posterior' probability of this suspect's guilt (from the officer's point of view) considering all the data.

   (b) Suppose both the murderer's and the suspect's blood types were found (with certainty) to be A. The relative frequency of blood type A in the population is 0.42. How would the posterior probability of guilt in that case compare with the same probability in part (a)? Would it be greater, smaller, or equal? Explain. [1, 2.4.2]

7. Suppose that we have two bags and each bag contains sixteen balls. Also, in each bag the sixteen balls are a combination of black balls and white balls. One bag contains three times as many white balls as black. The other bag contains three times as many black balls as white. Suppose we choose one of these bags at random.

(a) From this bag, we select five balls at random, replacing each ball after it has been selected. The result is that we find 4 white balls and one black. What is the probability that we were using the bag with mainly white balls?

(b) Now suppose instead we select five balls at random *without replacement*. The result is again that we find 4 white balls and one black. What is the probability that we were using the bag with mainly white balls?

4.11 Bayes' Rule: Summary

Bayes' Rule is basically the formula:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|\overline{A})P(\overline{A})}$$

We derived this formula via the following two equations:

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

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

4.12 Bayes' Rule: Another Exercise

1. Assume that people can be sorted unequivocally into two distinct sets according to their hair color: dark – denoted D, and blond – denoted B.

   A person's hair color is determined by two alleles of a gene, each transmitted at random by one parent. The allele for dark hair, denoted d, is dominant over that for blond hair, denoted b. Hence, of the three genotypes, dd, db, and bb, the first two would result phenotypically in D, only the third would be B. One can be certain that a blond person is homozygous (i.e., bb). However, upon observing a dark-haired person, one cannot know whether that person is genotypically homozygous (dd) or heterozygous (bd).

   Consider a couple with both mates dark haired and heterozygous for hair color. The genotypes and phenotypes of the potential offspring of the couple are given in the table below:

   Mother D

		d	b
		d	b
Father D	d	dd (D)	db (D)
	b	bd (D)	bb (B)

   The probability of such a couple giving birth to a dark-haired child is $3/4$ and to a blond child is $1/4$. The probability of a random D child of such parents being heterozygous is $2/3$.

   A couple is interested in knowing whether any of its future children could be blond. Both husband and wife are dark haired. Note that each of these prospective parents has two dark-haired parents and a blond brother.

Let H denote the genotypic event, or the hypothesis that the couple has the potential to produce a blond child. H is equal to the intersection of the events that husband and wife are heterozygotes (that is, each is a carrier of a recessive gene b). $\overline{H}$ is the event that the couple is incapable of producing a blond child (that is, at least one spouse is homozygous dd). If H is true, then the probability of that couple having a B child is $1/4$.

(a) Find the probability of H (before any children are born in that family). Denote this $P_0(H)$.

(b) A dark-haired baby is born to the couple. Denote this event D_1 (first-born child dark). Has the probability of H changed? Denote the probability of H after the birth of the first child by $P_1(H)$. What is the value of $P_1(H)$?

(c) How many dark haired children would the couple need to have in order to convince you that H is not true?

(d) After a few years the couple has another baby, which is blond. Denote this event B_2. What is the 'posterior' probability of H in light of this information? In line with the previous notations, what is $P_2(H)$? [1, 2.4.5]

Chapter 5

Probability Distributions

5.1 Definitions

- A **probability distribution** is an assignment of a probability to each *outcome* X of an experiment. We will typically use P to represent a probability distribution as this is essentially a special case of the P we've been using (which denotes the probability of events).

  In a probability distribution, each of the probabilities must be between 0 and 1 (inclusive) and the sum of all the probabilities must equal 1. We write these two conditions:

  1. $0 \leq P(X) \leq 1$ for all possible values of X,
  2. $\Sigma P(X) = 1$ where X assumes all possible values.

  Example 10. *Consider the following experiment: you flip a coin two times and record the number of heads that appeared. It should be clear that the set of all possible outcomes for this experiment is*
  $\{0\ heads, 1\ head, 2\ heads\}$.

  If the coin being flipped is a fair coin, then $P(2) = \frac{1}{4}$, $P(0) = \frac{1}{4}$, *and since there are two different ways of getting only one head (heads then tails and tails then heads) ,* $P(1) = \frac{1}{2}$.

- The **expected value** of an experiment is denoted $E(X)$, and, in some sense, represents the outcome X we "expect" to see. We compute $E(X)$ with the formula $E(X) = \Sigma(X * P(X))$, summing over all possible values of X.

  It is important to note that $E(X)$ might not actually be equal to a possible outcome of the experiment (see the following example). However, if the experiment is performed over and over again and the outcomes X from each experiment are averaged, this value will get closer and closer to $E(X)$.

Example 11. *Suppose we give a test with four multiple choice questions. Each question has three choices and only one correct answer. We then record the number of correctly answered questions. This experiment has possible outcomes $X = 0, 1, 2, 3,$ and 4.*

Now suppose that a student takes the test and arbitrarily guesses the answer to each question. A probability distribution table for this experiment is given below with an additional column to help us compute the expected value.

X	$P(X)$	$X * P(X)$
0	$(2/3)^4 = 16/81$	0
1	$4(1/3)(2/3)^3 = 32/81$	$32/81$
2	$6(1/3)^2(2/3)^2 = 24/81$	$48/81$
3	$4(1/3)^3(2/3) = 8/81$	$24/81$
4	$(1/3)^4 = 1/81$	$4/81$

So the expected value is $E(X) = 0 + 32/81 + 48/81 + 24/81 + 4/81 = 108/81 \approx 1.33$. *This means that if several people were to take this test by guessing the answers, we should expect the average number of correct answers for everyone to be about 1.33.*

Example 12. *Expected values comes up a lot in gambling. Here is a simple example. Suppose you choose one card from a deck of 52 playing cards. You make a deal with a friend that if the card drawn is a king, he will pay you \$50. If the card drawn is not a king, you will pay him \$10. We will compute your expected loss on this bet.*

Event	X	$P(X)$	$X * P(X)$
You win	*\$50*	*4/52*	*\$3.85*
You lose	*-\$10*	*48/52*	*-\$9.23*

$E(X) = -\$5.38$. *Thus, if you were to play this game over and over, in the long run, you should expect to lose \$5.38 per game. Not good for you.*

5.2 Exercises

1. Suppose you have a weighted coin that is twice as likely to land on heads as it is tails. You decide to flip this coin twice and record the number of times the coin lands on heads. Let X represent the possible outcomes of this experiment.

   (a) Find the probability distribution for this experiment. In other words, find $P(X)$ for all possible values of X.

   (b) Find the expected value of X, $E(X)$.

2. In a research study on animal behavior, mice are given a choice among four similar doors. One of them is the 'correct' door. If a mouse chooses the correct door, it is rewarded with food, and the experiment ends. If it chooses an incorrect door, the mouse is punished with a mild electric shock and then brought back to the starting point to choose again.

   Let X denote the number of trials in the experiment, that is the number of the trial on which the first correct choice occurs. Find the probability distribution of X in each of the two cases below. In other words, find $P(X)$ for every possible value of X. Note: You will need to use conditional probabilities.

   (a) On each trial, the mouse chooses, with equal probabilities, one of the doors that have not been chosen up to that moment. A door that has been tried is never chosen again (an intelligent mouse).

   (b) All doors are equally likely to be chosen on each trial (a dumb mouse). [1, 2.5.1]

3. An enthusiastic sports fan decides to express his support for his team by betting on it, and eventually making some profit out of that enjoyable activity.

   He knows from past statistics that the probability of his team winning in a given match is 0.75, and the probability of losing is 0.25 (let us ignore other possible outcomes).

   (a) He gets \$12 if his team wins. How much should he be willing to pay, if his team loses, in order to make and average profit of \$2 per bet?

   (b) He pays \$15 beforehand, and if his team wins he receives \$20. What is his expected gain from this bet? [1, 2.5.2]

4. An urn contains 7 balls: 3 red, 4 blue. Balls are randomly drawn from the urn, one after the other, *without* replacement.

Let X be the number of red balls drawn before drawing the first blue ball.

(a) Construct the probability-distribution table for this experiment.

(b) What is the expected value of X? [1, 2.5.4]

5. Consider an urn comprising of 3 red balls and 4 blue balls, as in the previous problem.

   Balls are randomly drawn from that urn, one after the other, *with* replacement.

   Let X be the number of red balls drawn before drawing the first blue ball.

   (a) Give the formula for the probability distribution of this experiment.

   (b) What is the expected value of X? [1, 2.5.5]

6. A medical clinic tests blood for a certain disease from which approximately one person in a hundred suffers. People come to the clinic in groups of 50. The director wonders whether he can increase the efficiency of the testing procedure by conducting pooled tests.

   Suppose that, instead of testing each individually, he would pool the 50 blood samples and test them all together. If the pooled test was negative, he could pronounce the whole group healthy. If not, he could then test each person's blood individually.

   What is the expected number of tests the director will have to perform if he pools the blood samples? [1, 2.5.9]