

STA 473 - Probability

Derek L. Sonderegger

2017-02-02

Contents

1	Introduction to Probability	5
1.1	Review of Set Theory (D&S 1.4)	5
1.2	Definition of Probability (D&S 1.5)	6
1.3	Finite Sample Spaces (D&S 1.6)	7
1.4	Ordered Counting (D&S 1.7)	8
1.5	Combinations (D&S 1.8)	9
1.6	Multinomial Coefficients (D&S 1.9)	10
2	Conditional Probability	11
2.1	Defining Conditional Probability (D&S 2.1)	11
2.2	Independence	13

Chapter 1

Introduction to Probability

1.1 Review of Set Theory (D&S 1.4)

1. Create a sample space \mathcal{S} where
 - a) Where the the number of outcomes is finite.
 - b) Define events (subsets of \mathcal{S}) that do not have a 1-to-1 correspondence with the outcomes.
2. Create a sample space \mathcal{S} where
 - a) Where the the number of outcomes is countably infinite.
 - b) Define a finite number of events (subsets of \mathcal{S}) that do not have a 1-to-1 correspondence with the outcomes and that the union of all your events is \mathcal{S} .
 - c) Define an infinite number of events (subsets of \mathcal{S}) that do not have a 1-to-1 correspondence with the outcomes and that the union of all your events is \mathcal{S} .
3. Create a sample space \mathcal{S} where
 - a) Where the the number of outcomes is uncountably infinite.
 - b) Define a finite number of events (subsets of \mathcal{S}) that do not have a 1-to-1 correspondence with the outcomes and that the union of all your events is \mathcal{S} .
 - c) Define an countably infinite number of events (subsets of \mathcal{S}) that do not have a 1-to-1 correspondence with the outcomes and that the union of all your events is \mathcal{S} .

It is time to define the set of events more carefully. The take-home idea is that if you add an event, say A , you also add some other events related to A . The rules are summarized below.

- S is an event. This is to say *something* will happen.
 - If A is an event, then A^c is also an event.
 - If A_i is a countable sequence of events, then $\cup_{i=1}^{\infty} A_i$ is also an event.
4. Prove that \emptyset is is an event.
 5. Prove two of the conclusions of theorem 1.4.4. I would expect a proof of $A \cup A^c = \mathcal{S}$ to look something like, “Let e be an arbitrary event in \mathcal{S} . Due to the nature of complements, either $e \in A$ or $e \in A^c$. Therefore $e \in A \cup A^c$ but because e was an arbitrary element of \mathcal{S} then $\mathcal{S} \subset A \cup A^c$. However because \mathcal{S} is the set of all possible events, then $A \cup A^c \subset \mathcal{S}$ and thus $A \cup A^c = \mathcal{S}$.”
 6. Chapter problem 1.4.1. Suppose $A \subset B$. Show that $B^c \subset A^c$. Do this in a similar fashion as problem 5.
 7. Chapter problem 1.4.2. Show this by Venn diagrams.
 8. Chapter problem 1.4.3. Prove DeMorgan’s Laws. Prove this via Venn diagrams.
 9. Chapter problem 1.4.6.
 10. Chapter problem 1.4.7

11. Chapter problem 1.4.13
12. Chapter problem 1.4.14

1.2 Definition of Probability (D&S 1.5)

Axiom 1 For every event A , the probability of the event, denoted $Pr(A)$ has the property $Pr(A) \geq 0$

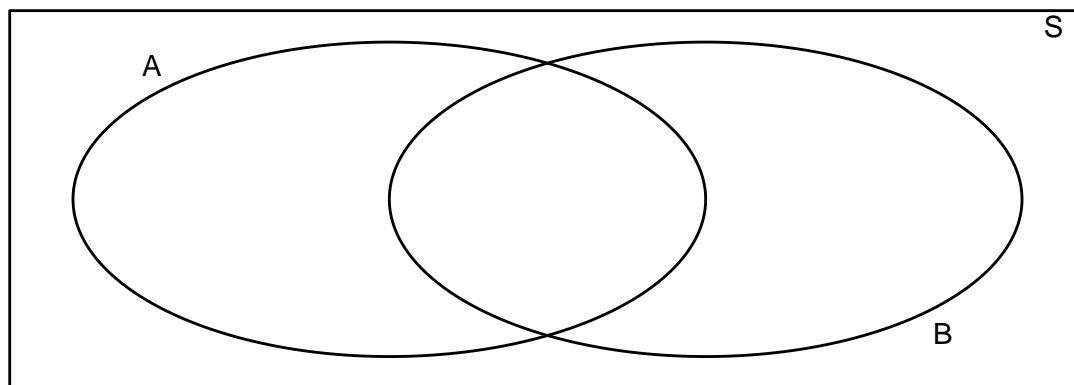
Axiom 2 If an event is sure to occur, then the event has probability 1. That is, $Pr(\mathcal{S}) = 1$.

Axiom 3 For every finite or countably infinite sequence of events A_1, A_2, \dots where $A_i \cap A_j = \emptyset \ \forall i, j$ (that is the sequence A_i is pairwise disjoint), then

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

1. Consider drawing a single card from a well shuffled deck of playing cards (4 suits, each with 13 cards Ace, two, ..., Queen, King). Consider the events H, S, C, D , which are drawing a Heart, Spade, Club, and Diamond. Explain why H and S are disjoint but H^c and S^c are not.
2. Prove $Pr(\emptyset) = 0$
3. Argue that *Axiom 3* should have been “For every countably infinite sequence of events A_i ” because you can pad any finite sequence with an infinite sequence of empty sets.
4. Prove $Pr(A^c) = 1 - Pr(A)$

Often we will draw Venn diagrams where the area of the event is its probability.



Many of the probability calculations can be most easily understood using a Venn diagram along with the algebraic proof.

5. Prove if $A \subset B$ then $Pr(A) \leq Pr(B)$. Show this formally and via Venn diagrams
6. Prove that $Pr(A) = Pr(A \cap B) + Pr(A \cap B^c)$ Show this formally and via Venn diagrams
7. Prove that $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$ Show this formally and via Venn diagrams. Notice our *Axiom 3* addresses the case where A and B are disjoint.
8. Consider events A and B where $Pr(A) = 1/3$ and $Pr(B) = 1/2$. Determine the value of $Pr(A \cap B^c)$ when
 - a) A and B are disjoint
 - b) $A \subset B$
 - c) $Pr(A \cap B) = 1/8$
9. Suppose that Adam has a probability of failing an exam of $Pr(A) = 0.5$ while Bob only has a probability of failing the exam of $Pr(B) = 0.2$. Suppose the probability of both students failing is 0.1

- a) What is the probability that at least one of these two students will fail?
 - b) What is the probability that neither student will fail?
 - c) What is the probability that exactly one student will fail?
10. A point (x, y) is to be selected from the unit square \mathcal{S} ($0 \leq x \leq 1$, $0 \leq y \leq 1$). Suppose that the probability that the point is selected from a specific region is equal to the area of the region. Find the probability the point selected is from each of the following regions:
- a) (x, y) such that $(x - 1/2)^2 + (y - 1/2)^2 \geq 1/4$
 - b) (x, y) such that $1/2 \leq x + y \leq 3/2$
 - c) (x, y) such that $y \leq 1 - x^2$
 - d) (x, y) such that $x = y$
11. Bonferroni's Inequality. Let A_1, A_2, \dots be an arbitrary infinite sequence of events. Define the sequence of events B_1, B_2, \dots as

$$B_1 = A_1$$

$$B_2 = A_1^c \cap A_2$$

$$B_3 = A_1^c \cap A_2^c \cap A_3$$

$$B_4 = A_1^c \cap A_2^c \cap A_3^c \cap A_4$$

- a) Prove that $B_i \subset A_i$, $B_i \cap B_j = \emptyset$ for $i \neq j$, and that $\bigcup_{i=1}^n A_i = \bigcup_{i=1}^n B_i$.
- b) Prove that

$$Pr\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n Pr(B_i)$$

- c) Prove that

$$Pr\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n Pr(A_i)$$

- d) Using the previous result (b), prove that for sets D_1, D_2, \dots, D_n that

$$Pr\left(\bigcap_{i=1}^n D_i\right) \geq 1 - \sum_{i=1}^n Pr(D_i^c)$$

1.3 Finite Sample Spaces (D&S 1.6)

When dealing with sample spaces with only a finite number of outcomes (say n outcomes s_i), it is often convenient to define each outcome as an event.

Let s_i be outcomes in the sample space \mathcal{S} . Let each of these outcomes have probability p_i .

For the axioms of probability to hold then:

$$p_i \geq 0$$

$$\sum_{i=1}^n p_i = 1$$

1. When fair dice, we assume that each side has equal probability of being rolled. For rolling a 6-sided die, what is the probability of rolling an even number?
2. When rolling two (or more) differently colored dice, we assume that the die do not affect the outcome of the other and that every pair of is equally likely. Alternatively you can think of rolling 1 die and then the other. So for rolling two six sided dice, there are 36 possible rolls, and notice, for example, $(2, 3)$ is a different roll than $(3, 2)$. What is the probability that the sum of the two rolls is even?
3. If a fair coin is flipped three times...

- a) What are the possible outcomes (enumerate these)?
- b) Explain why it is reasonable that each outcome is equally probable?
- c) What is the probability that all three faces will be the same?

1.4 Ordered Counting (D&S 1.7)

Often we situations where it is reasonable to believe that each outcome of an event is equally likely and therefore we can figure out the probability if we knew how many events there were. E.g. there are 36 different outcomes for rolling two fair 6-sided dice, so each outcome has a $1/36$ probability.

1. Prove/argue/justify that if the outcome of experiment is composed of 2 parts, where the first part has m outcomes x_1, \dots, x_m and the second part has n outcomes, y_1, \dots, y_n then there is a total of mn outcomes (x_i, y_j) . This is often called the Multiplication Rule for Counting.
2. I own 3 pair of pants that are “work appropriate.” I also own 6 different shirts and 5 pairs of shoes that are “work appropriate.” How many different outfits are possible?
3. How many way can the numbers 1, 2, 3, 4, and 5 be arranged?
4. **Ordered Sampling without Replacement** For distinct objects $1, 2, \dots, n$ prove that there are $P_{n,k} = \frac{n!}{(n-k)!}$ arrangements of k elements (where the order is important, i.e. $1, 2, 3$ is distinct from $2, 1, 3$) and $n! = n \cdot (n-1) \cdot (n-2) \cdots (2) \cdot (1)$ and by definition $0! = 1$. We call $P_{n,k}$ the number of *permutations* of k elements taken from a set of n distinct objects. *Hint: First consider base cases of $k = 1$ and then $k = 2$ and that the formula is appropriate. Then, to complete the induction argument, show that if we have a $P_{n,k}$ permutations of k objects, then increasing to $k + 1$ elements simply results in $(n - k) \cdot P_{n,k}$ arrangements due to the Multiplication Rule of counting.*
5. From $n = 17$ students, one student will get a candy bar, another will get a soda, and a third will receive some gummi bears. How many different ways could the treats be distributed to the students?
6. Suppose we are going to randomly select 3 elements from the digits $0, 1, 2, \dots, 9$ but we will select these *with replacement* (so we could get the 022). How many outcomes are there?
7. **Ordered Sampling with Replacement** Suppose that we have n distinct objects labeled $1, 2, \dots, n$ and we are going to sample k of these objects *with replacement*. Justify/derive a formula for the number of outcomes.
8. Consider the sequence of numbers 0000, 0001, 0002, \dots , 9998, 9999. How many of these numbers are composed of 4 different digits?

Often times I am interested in calculating the probability of a particular event and we can often do it in the following manner:

- First count the number of equally likely outcomes there are.
- Count the number of outcomes where the event of interest occurs.
- Then calculate

$$Pr(\text{Event}) = \frac{\text{Number of outcomes where event happens}}{\text{Total number of equally likely outcomes}}$$

9. As I work in the evenings, I often listen to music. Suppose that I have a playlist of $n = 300$ songs and I listen to them on shuffle where the software always selects from the list with equal probability when selecting which song to play next. If I listen to $k = 10$ songs, what is the probability that at least one of the songs will be duplicated? What about if I listen to $k = 30$ songs?
10. If 14 balls are randomly thrown into 25 boxes such that there is equal chance for a ball to land in any box, what is the probability that no box receives more than one ball?

1.5 Combinations (D&S 1.8)

Often we want to count the number of arrangements of k elements selected without replacement from n distinct objects but where the order doesn't matter. Another way of saying this is that we want to count the number of sets of size k taken from n distinct objects.

1. **Unordered Sampling without Replacement** For a set of k elements, prove that there are $k!$ permutations of those elements. Using this information, argue that the number of distinct sets of k objects taken from n elements is (which the book denotes as $C_{n,k}$ and many others denote $\binom{n}{k}$) is

$$C_{n,k} = \binom{n}{k} = \frac{P_{n,k}}{k!} = \frac{n!}{k!(n-k)!}$$

2. I have 3 identical cans of soda that I will distribute randomly to 17 students. I will select (with equal probabilities per student) 3 students. How many ways could I choose 3 students?
3. Suppose I have a character string composed of only 0s and 1s. The character string is 20 characters long and 8 of them are 0s. How many different strings are there?
4. **Unordered Sampling with Replacement** Suppose that I have $n = 7$ boxes into which I will randomly throw $k = 3$ balls.

```
n=7 boxes
+---+---+---+---+---+---+---+
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
+---+---+---+---+---+---+---+
```

Now suppose that we throw, at random, $k = 3$ balls into the boxes. We might end up with one ball in box 3 and two in box 6.

```
n=7 boxes
+---+---+---+---+---+---+---+
|   |   | 0 |   |   | 00 |   |
+---+---+---+---+---+---+---+
```

- a) Argue that throwing k balls randomly into n boxes is equivalent to selecting a set of k elements from n distinct objects with replacement. *Hint show that every set chosen with resampling can be represented via boxes/balls and that every boxes/balls combination represents a possible set of k elements from n distinct objects with replacement.*
- b) The guts of the boxes/balls diagram is the arrangement of box partitions and balls because the outer box walls don't matter because the balls get into a box.

```
|   |   | 0 |   |   | 00 |   |
```

which we can clean up a bit by remembering that the other walls have to be there and we'll represent the balls with 0 and the box partitions with a 1. `|110111001|`. This reduces the problem into how many binary strings can I produce with $n - 1$ 1s and k 0s.

5. Suppose that I will distribute my tree soda cans to 17 students by drawing names out of a hat, but replacing the student's name after it is draw. How many different outcomes could occur?
6. Suppose I draw 2 cards from a standard deck of 52 cards, what is the probability that I draw two cards of the same suit?
7. Ten teams are playing in a tournament. In the first round there, there will be five games played. How many possible arrangements are there? What is the probability that the Ashville Avalanche plays the Boston Behemoths (these are two of the ten teams playing)?
8. Suppose that I flip a fair coin 10 times. What is the probability I observe 3 heads?

1.6 Multinomial Coefficients (D&S 1.9)

1. From the Math/Stat department faculty, a committee of 5 members is to be selected. There are 8 Math, 4 Statistics, and 4 Math Ed Professors. What is the probability that committee is composed of 3 Math, 1 Stats, and 1 Math Ed professor.
2. Suppose that we are creating a string of beads from 9 red, 7 blue, and 10 yellow beads. How many different arrangements can be made?

Chapter 2

Conditional Probability

2.1 Defining Conditional Probability (D&S 2.1)

1. Out of $n = 17$ students in a class, I will chose one at random student to give gummi bears to. In this class there are 12 men and 5 women. Student A is very interested in her probability of being selected. Denote the event A as being that the student gets the gummi bears. Denote event W as a woman is selected.
 - a) What is the probability that student A is selected?
 - b) What is the probability that a woman is selected?
 - c) Suppose I restricted my selection to *only* the women? What is the probability that student A is selected. Can you write this probability as a function of your answers in parts (a) and (b)?
2. Consider the case where we take the sum of 2 six-sided dice. Define A as the event that the sum is greater than or equal to 9, and B being the event that the sum is greater than or equal to 6.
 - a) Create a table of the possible outcomes. Notice that each outcome is equally likely. What is the probability that A occurs? *Use correct notation and leave it as a fraction $\frac{??}{36}$*
 - b) Suppose that you are told that event B has occured. How many equally likely outcomes are there and what is the probability that A occurs? This will be denoted as $Pr(A|B)$. *(Leave this as a fraction).*
 - c) Notice that you the numerator in answer in part (a) is the same as the numerator as you had in part (b), but the denominators are different. What number do you need to multiple your part (a) answer by to get your part (b) answer. How does this relate to $P(B)$?
3. It seems that my children (Casey and Elise) get sick with annoying frequency. Suppose the probability that my son Casey gets sick is $Pr(C) = 0.05$ and furthermore the probability that **both** children get sick is $Pr(E \cap C) = 0.03$ If Casey is sick right now, what is the probability that Elise is also sick?

When we talk about events A and B , we defined them as possible outcomes, but it isn't until we define probability of the events $Pr(A)$ $Pr(B)$ that we care about the sample space \mathcal{S} of all possible events. What we are now trying to do is to claim that some addition knowledge allows us to refine the sample space to some smaller subset of \mathcal{S} , perhaps $B \subset \mathcal{S}$. So for events in B , we now need to re-scale all the probabilities to reflect that we now know that event B did happen.

If we previously know that $Pr(A \cap B) = 0.3$ and $Pr(B) = 0.6$, then half of the probability associated with B is overlapping with A . So if we just restrict ourselves to cases where B occurs, then the probability that A will occur is $1/2$.

Notice that our notation $Pr(A|B)$ is addressing the refinement of the sample space, but we've just defined our notation this way. We have not, and will not ever define $A|B$ because event A is event A , regardless of the sample space we use to figure out its probability.

We formally define the conditional probability of A given B as

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} \quad \text{Assuming } Pr(B) \neq 0$$

If $Pr(B) = 0$ then the conditional probability is undefined. Notice that this can be happily re-arranged to

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

I can also notice that I could condition on either event A or event B , so we could also have

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

4. If $B \subset A$ and $Pr(B) > 0$, what is $Pr(A|B)$
5. If A and B are disjoint and $Pr(B) > 0$, what is $Pr(A|B)$?
6. Suppose that events A_1, A_2, \dots, A_n are events such that $Pr(\bigcap_{i=1}^n A_i) > 0$. Show that

$$Pr\left(\bigcap_{i=1}^n A_i\right) = Pr\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right) Pr\left(A_{n-1} \mid \bigcap_{i=1}^{n-2} A_i\right) \dots Pr\left(A_3 \mid A_2 \cap A_1\right) Pr\left(A_2 \mid A_1\right) Pr(A_1)$$

7. For events A , B , and D such that $Pr(D) > 0$ show that:

- a) $Pr(A^C|D) = 1 - Pr(A|D)$.
- b) $Pr(A \cup B|D) = Pr(A|D) + Pr(B|D) - Pr(A \cap B|D)$

Suppose that we have K events B_k such that the B_1, B_2, \dots, B_K are disjoint and $\bigcup B_k = \mathcal{S}$. Then we call the events B_1, \dots, B_K a *partition* of the sample space \mathcal{S} . A partition of \mathcal{S} is often useful for calculating probabilities due to the disjoint nature of the B_k elements.

6. **Law of Total Probability** Prove that for for a partition B_1, \dots, B_K of \mathcal{S} , that

$$Pr(A) = \sum_{k=1}^K Pr(A \cap B_k) = \sum_{k=1}^K Pr(A|B_k)Pr(B_k)$$

Note: There is a conditional version of the Law of Total probability, which is proved in an analogous fashion:

$$Pr(A|C) = \sum_{k=1}^K Pr(A \cap B_k|C) = \sum_{k=1}^K Pr(A|B_k \cap C)Pr(B_k|C)$$

7. A child's bookshelf contains three shelves. On the shelves are $n_1 = 10, n_2 = 20$ and $n_3 = 30$ books. Within each set of books, there are some number of Dr Suess books $m_1 = 5, m_2 = 4, m_3 = 2$. The child will select a shelf at random (equal probability) and then from the shelf will select a book at random. What is the probability the child selects a Dr Suess book?
8. A camera with a motion detector was mounted facing a forest trail. 50% of the pictures were taken during the daytime, 15% were taken during twilight hours (dawn and dusk), and 35% were taken during the night. Of the pictures taken during the daytime, 80% were of hikers and 20% were of wild animals. Of the pictures taken at twilight 30% were of hikers and 70% were of wild animals. Finally, of the pictures taken during the night, 100% were wild animals.
 - a) What is the probability that a randomly selected photo is of a hiker and was taken at twilight?
 - b) What is the probability a photo was taken at night given that is of a wild animal?
9. I have kept track of the probabilities of how many cats will sit with me and/or my wife on the couch. Below is a table of probabilities.

	0 Cats	1 Cat	2 Cats	3 Cats
0 People	0.08	0.08	0.03	0.01
1 Person	0.1	0.25	0.125	0.025
2 People	0.03	0.09	0.12	0.06

- a) What is the probability that one human is sitting on the couch?
 - b) What is the probability that at least two cats are sitting on the couch?
 - c) Given that there are two cats sitting on the couch, what is the probability that there are two humans also on the couch?
10. My cat Kaylee occasionally likes to sit on people's laps while they are seated at the table. My wife is strongly opposed to this and will scold the cat when she catches her in the act. Suppose that that Kaylee will select my lap 60% of the time and the remaining 40% of the time she jumps into my wife's lap. If Kaylee jumps into my wife's lap, there is a 100% chance of being scolded, while if she jumps into mine, there is only a 20% chance of being scolded. Given that Kaylee was just scolded for being in a lap, what is the probability she was in my wife's lap?
 11. There are two brands of Mac & Cheese that my daughter will eat. When I go shopping I will pick from the two brands with a 70% probability of choosing the brand that I bought the previous time. The first time I went shopping, I chose from the two brands with equal probability. What is the probability that I chose brand A on the first and second trips, and brand B on the third and fourth trips?

2.2 Independence

Definition: Two events, A and B are independent if $Pr(A \cap B) = Pr(A)Pr(B)$. *Definition:* Events A_1, A_2, \dots, A_K are pairwise independent if A_i and A_j are independent for any i, j . *Definition:* Events A_1, A_2, \dots, A_K are mutually independent if for all subsets I of $1, 2, \dots, K$, $Pr(\bigcap_{i \in I} A_i) = \prod_{i \in I} Pr(A_i)$

1. Show that if $Pr(A) > 0$ and $Pr(B) > 0$, then A and B are independent if and only if $Pr(A|B) = Pr(A)$ and $Pr(B|A) = Pr(B)$
2. Give an example of three events that are pairwise independent but not mutually independent.

Convention If I say that a set of events are “independent”, then we intend to say “mutually independent” but are being lazy.

3. Show that if A and B are independent, then A and B^c are also independent.
4. Suppose that we flip a fair coin three times. Denote H_i as the event that I flip a head on the i th flip.
 - a) Find $Pr(H_1 \cap H_2 \cap H_3)$
 - b) Find $Pr(H_1 \cap H_2^c \cap H_3)$
 - c) Find $Pr(H_1^c \cap H_2 \cap H_3)$
 - d) How many ways can we have 2 heads?
 - e) What is the probability of 2 heads?
5. I will roll a 20-sided die three times. Define the event H_i as the event that I roll a 17 or greater on the i th roll.
 - a) Find $Pr(H_1 \cap H_2 \cap H_3)$
 - b) Find $Pr(H_1 \cap H_2^c \cap H_3)$
 - c) Find $Pr(H_1^c \cap H_2 \cap H_3)$
 - d) How many ways can we have 2 H events happen?
 - e) What is the probability of 2 H events happening?
6. I will roll my 20-sided die until I roll a 20.
 - a) What is the probability that I roll a 20 on my first roll?
 - b) What is the probability that the first 20 I roll is on the 5th roll?

Bibliography