STAT 109 Lecture Notes

3 Probability

Introduction

For the next three chapters, we will be focusing on probability, then at the end of the course we will be returning to some more advanced topics in statistics. Why are we breaking in the middle of statistics to focus on probability?

In this course, we have already learned how to carefully design experiments, studies, and surveys to collect meaningful data. We have learned how to summarize that data with charts, graphs, statistics, and verbal descriptions. We have learned that randomizing causes some differences in results from trial to trial and experiment to experiment. We want to be able to say when these differences are "just" due to randomization, and when they are "statistically significant."

One of our goals in this course is to talk about ways of quantifying these differences and their significance, as well as confidence and error in our data. Before we know what differences are significant, we need a baseline: we need to know what types of values to *expect*. Probability is one way (the only way?) of computing these expectations when dealing with randomization.

3.1 The Basics of Probability

Definition 1: An action that can be modeled by random numbers is called a random phenomenon. Each occasion upon which we observe a random phenomenon is a trial. The value of the random phenomenon at the end of each trial is called the trial's outcome.

Definition 2: A collection of outcomes of a random phenomenon is called an event. The collection of all possible outcomes is called the sample space.

Definition 3: For independent trials (meaning that the outcome of any one trial does not influence the outcome of any other trial), as the number of trials increases (to infinity!), the relative frequency of an event gets closer and closer to one single number, called the probability of the event. This guarantee (of stabilization toward a single frequency as the number of trials gets larger and larger) is called the Law of Large Numbers.

Note: The Law of Large Numbers only guarantees the frequency of an event to occur "in the long-run," meaning given an infinite amount of time and an infinite number of trials. It does NOT guarantee any results in the "short-term" (meaning in 5 trials, or 5 minutes, or 5 hours, or 5 days). A belief in "short-term" results is called *The Law of Averages.* **The Law of Averages is false.** For example, on a roulette wheel in a casino, let's say the number 6 has not been hit all day (several hours, with the wheel being spun every few minutes). This does NOT mean that the number 6 should be expected to hit very soon. The Law of Large Numbers guarantees that in an infinite number of spins, the number 6 should come up just as often as any other number. However, there is no guarantee on how many times the number 6 will come up in the next minute, or the next hour, or the next 100 spins...

Types of Probability

Example 3.1: You know it sounds crazy, but you believe that blue M&M's just taste better than the other colors. You are curious to find out how likely it is to pull out a blue M&M if you reach your hand into a standard bag of M&M's (ie, all the colors are mixed in, it's not a special order bag with only one color, etc.), but you can't find this information on the Internet anywhere! You want to design a study to investigate this question. How will you estimate the likelihood of drawing a blue M&M?

Definition 4: If you *repeatedly observe the outcome of a random phenomenon* and record the results, this is called empirical probability. Then the probability of event A occurring is denoted $P(A)$ and it is calculated in the following way:

$$P(A) = \frac{\text{number of times } A \text{ occurs}}{\text{total number of trials}}$$

Note: This calculation is only valid if the trials are independent.

Definition 5: Sometimes is is not as easy, as fun, or as accurate to actually repeat a random phenomenon many times. However, if all outcomes are equally likely to occur, then we can still compute the probability of event A:

$$P(A) = \frac{\text{number of outcomes that result in event } A}{\text{total number of possible outcomes}}$$

This calculation is called theoretical probability because we are not actually observing the events, but computing their likelihoods theoretically.

Note: This calculation is only valid of all outcomes are equally likely.

Definition 6: When we use the word "probability" informally to express a degree of uncertainty, but not to signify a measurement of relative frequency of an event in an infinite number of trials, we call this personal probability or subjective probability.

Example 3.2: For example, you may have heard yourself say in the past "The probability that I actually get to bed before midnight tonight is less than 10%." Here, you are expressing uncertainty about your ability to get to bed before midnight, but you have not actually run an infinite number of trials to model the situation (or even a finite number of trials!).

As another example, it is common in the Intelligence Studies program here at Mercyhurst for professors to use "words of estimated probability" and instructors of intelligence studies courses will encourage you to use these words in your work as well. Keep in mind that this is referring to **subjective probability**, not any type of formal or mathematical probability.

Example 3.3: Consider the random phenomenon of rolling one fair six-sided die.

- What is the sample space?

- If event A is rolling a 2, what is $P(A)$?

- If event B is rolling an odd number, what is $P(B)$?

- If event C is rolling a 5, what is $P(C)$?

Example 3.4: Your friend bets you you cannot flip a coin three times and get tails at least twice. What is the sample space? How many outcomes are in the sample space? What is the probability of you winning this bet?

Example 3.5: You are playing a game where you roll a fair six-sided die and draw a random card from a standard 52-card deck. Your prize depends on what number comes up on the die and the suit of the card you select. What is the sample space? How many outcomes are in the sample space?

Example 3.6: You roll two six-sided dice simultaneously.

- What is the probability the sum of the face-up values is 4 or 5?

- Suppose you roll one die, and then the second die. What is the probability the second number you roll is greater than the first number?

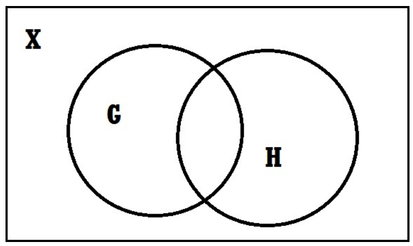
Basic Set Theory

Definition 7: A set is a collection of distinct elements. For example, in Example 2 of Chapter 12, we wrote our sample space as a set.

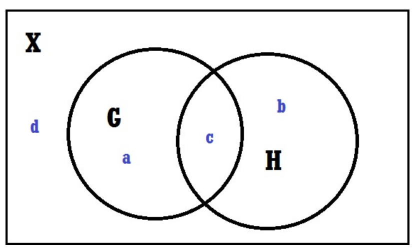
Definition 8: A Venn diagram is a diagram that models the relationships between sets visually.

Example 3.7: In the following examples, definitions, and diagrams, assume that X is some set (it will be like our sample space when we get back to probability), and that G and H are subsets of X, meaning that all elements of G are also in X and all elements of H are also in X. For example, if $X = \{a, b, c, d\}$ and $G = \{a, c\}$ and $H = \{b, c\}$ then G and H are subsets of X.

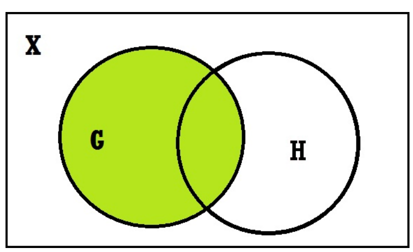
Below is a Venn diagram showing X, G, and H:



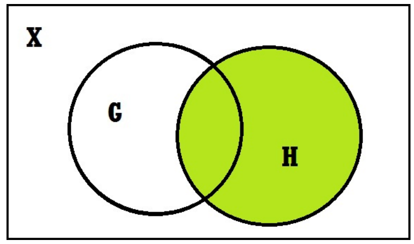
We can also draw or write in the individual elements of the sets (and later, the probabilities of each set) if it is helpful:



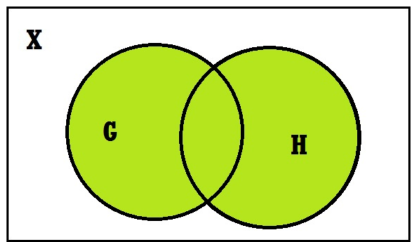
If we wish to highlight G, we can shade it in a different color:



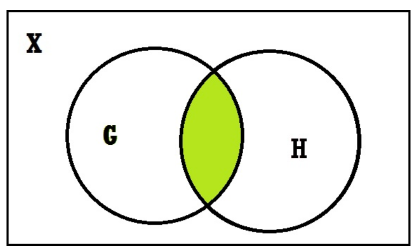
Or perhaps we want to pay more attention to H:



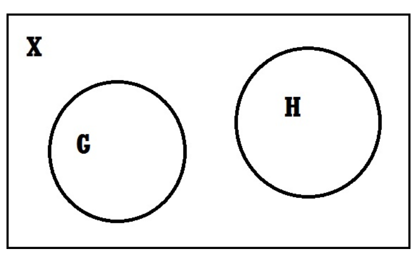
Definition 9: "$G \cup H$" means "G union H." This is the set of all elements that are EITHER in G OR in H (or in both). In our previous example with elements, $G \cup H = \{a, b, c\}$ because these are all of the elements in X that are either in G or in H. Below is a picture of $G \cup H$ in a Venn diagram:



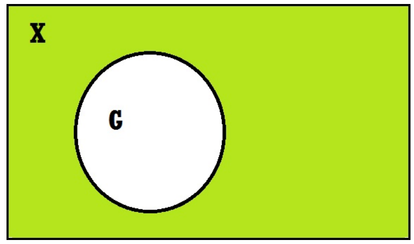
Definition 10: "$G \cap H$" means "G intersect H." This is the set of all elements that are in BOTH G AND in H. In our previous example with elements, $G \cap H = \{c\}$ because these are all of the elements in X that are in both G and in H. Below is a picture of $G \cap H$ in a Venn diagram:



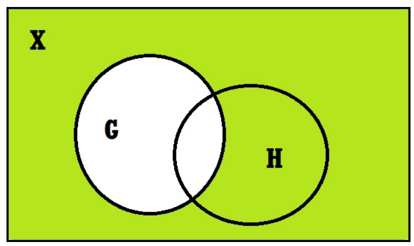
Definition 11: When a set has no members/elements, we say it is an "empty set" and we write the symbol $\emptyset$. If $G \cap H = \emptyset$ (ie, if there is no element that is in both G and in H), we say that G and H are disjoint, or mutually exclusive. Below is a picture of a Venn diagram where G and H are disjoint:



Definition 12: "G^C" means "G complement" or the complement of G. G^C is the set of all elements that are NOT in G. In our previous example with elements, $G^C = \{b, d\}$. Below is a picture of a Venn diagram with G^C highlighted, where only G is labelled in the diagram:



Below is a picture of a Venn diagram with G^C highlighted, where H is also labelled in the diagram:



Example 3.8: Recall from Example 2 that we are rolling a six-sided die, event A is rolling a 2, event B is rolling an odd number, and event C is rolling a 5. Draw a Venn diagram to visualize the sample space.

Example 3.9: Suppose a standard 52 card deck is shuffled and the top two cards are dealt.

1. How many ordered pairs of cards are possible? ("Ordered" means that 2 of clubs then 7 of hearts is different from 7 of hearts then 2 of clubs. That is, the order that the cards are dealt matters.)

2. What is the probability the first card is an ace?

3. What is the probability that both cards are aces?

4. What is the probability there is at least one ace?

Axioms of Probability: Here are a few important facts about theoretical and empirical probability experiments. Assume S stands for the whole sample space and let A and B be any events in S.

- $P(S) = 1$ and for any event A, $0 \leq P(A) \leq 1$.
- If events A and B are mutually exclusive (in other words: "no overlap," or "can't happen at the same time," or in terms of sets "disjoint" or "no intersection"), then $P(A \cup B) = P(A \text{ OR } B) = P(A) + P(B)$.
- Let A' denote the complement of A (A^C in set notation). Then $P(A') = 1 - P(A)$.

Example 3.10: A bag of jelly beans has been poured into a jar. You reach in and randomly select a jelly bean. Are the events "select a red jelly bean" and "select a yellow jelly bean" mutually exclusive?

Example 3.11: You stop a random person on the street. Are the events "the person loves cats" and "the person owns a dog" mutually exclusive?

Example 3.12: The table below shows number of employees for small businesses.

Number of employees	1-4	5-9	10-19	20-99	100 or more
Percent of businesses	42.9%	15.1%	9.6%	10.0%	22.4%

1. Find the probability a small business will have less than 5 employees.

2. Find the probability a small business will have at least 10 employees.

3. Find the probability a small business will have fewer than 20 employees.

Definition 13: Combinatorics is a field of mathematics concerned with three types of problems: existence problems, enumeration problems, and optimization problems. At its core, combinatorics is the field of mathematics that concerns counting (enumerating). This sounds pretty basic, and it is. Combinatorics problems are usually very easy to describe, but often very hard to solve. When you can't count things one-by-one, you need a clever way of enumerating them without counting them. These are the problems combinatorics tries to solve.

The Fundamental Counting Principle: If something can happen in n_1 ways, and no matter how the first thing happens, a second thing can happen in n_2 ways, and no matter how the first or second thing happens, a third thing can happen in n_3 ways, and .... then all of the things together can happen in

$$n_1 \times n_2 \times n_3 \times ...$$

ways.

Example 3.13: How many 5 letter words are possible with the 26-letter English alphabet?

Example 3.14: The way that telephone numbers are currently structured got its start in the 1950s and 1960s. Direct long distance dialing was possible with a ten digit number. At the time, the first three digits signified the area code, and all area codes could not begin with a zero or a one, the middle digit had to be a zero or a one, and the third digit could be anything. Then the area code was followed by a seven digit local number, the first two digits of which could not be zero or one, but the last five digits could be anything. **How many different telephone numbers are possible with this structure?**

Example 3.15: A T-shirt design company is going to release shirts from a certain material that they've ordered in a number of different designs. Each shirt will have a color (either blue, green, orange, pink, red, black, or gold), a design (solid, vertical stripes, horizontal stripes, or splattered), a sleeve length (long, short, or 3/4), and a style (V-neck, crew cut, or scoop neck). How many different-looking shirts can they design?

3.2 Conditional Probability

Definition 14: To find the probability of event B given that event A is true (or event A has already happened), we compute "the probability of B given A:":

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

$P(B|A)$ often comes from a conditional distribution, and it is called conditional probability. To help yourself remember the formula, remember that we already know event A is true, so we are only considering those outcomes (that is why $P(A)$ is in the denominator), then we want to know which of those events will result in event B (so the numerator are all the events that are BOTH A and B).

Example 3.16: Below is a contingency table showing the distribution of majors among men and women in the classes I taught in spring and summer 2016. Answer the questions below using the table.

Major	Females	Males
Natural Sciences	34	14
Business	2	9
Communications	10	6
Social Sciences	4	1
Arts & Humanities	2	0
Intelligence	6	24

1. What is the probability that a randomly chosen student is a natural science major, given that he is a male?

2. What is the probability that a randomly chosen student is a natural science major, given that she is a female?

3. What is the probability that a randomly chosen student is male, given that he a business major?

4. What is the probability that a randomly chosen student is female, given that she is an intelligence studies major?

Definition 15: Events A and B are independent if $P(B|A) = P(B)$. We already had an intuitive definition of independent events (that the probability of event A does not affect the probability of event B). Now we have an explicit way of computing independence. If you are not sure if A and B are independent, you can compute $P(B|A)$ and $P(B)$ and compare them.

Example 3.17: Fifty-six percent of all American workers have a workplace retirement plan, 68% have health insurance, and 49% have both benefits. Are having health insurance and a retirement plan independent events? Are the events disjoint?

Example 3.18: For a certain real estate company, 64% of homes-for-sale that they are currently running ads for have garages, 21% have swimming pools, and 17% have both a garage and a swimming pool. Are homes-for-sale having a garage and a swimming pool independent events? Are the events disjoint?

The Multiplication Rule. For two events A and B,

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

Note that when A and B are independent, this formula becomes $P(A \cap B) = P(A)P(B)$, but this is ONLY true when A and B are independent.

Example 3.19: A shipment of 250 laptops contains 3 defective units. Your company buys 3 units, and all three come from this shipment.

1. What is the probability you receive no defective units?

2. What is the probability you receive all defective units?

3. What is the probability you receive at least one good unit?

3.3 The General Addition Rule

For any two events A and B,

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

Example 3.20: A survey of students asked whether or not they ate breakfast on the morning of the survey. The results are in a table below.

	Ate breakfast	Did not eat breakfast	Total
Male	66	76	142
Female	125	74	199
Total	191	150	341

1. What is the probability that a randomly selected student ate breakfast?

2. What is the probability that a randomly selected student is a female who ate breakfast?

3. What is the probability that a randomly selected student is a female or the student ate breakfast?

4. What is the probability that a randomly selected student is female, given that the student ate breakfast?

5. What is the probability that a randomly selected student did not eat breakfast, given that the student is male?

Example 3.21: Consider a standard 52-card deck of playing cards that has been shuffled.

1. What is the probability that a randomly chosen card is an ace or it is red?

2. What is the probability that two randomly drawn cards (without replacement) are both aces?

3. What is the probability that five randomly drawn cards (without replacement) are all hearts?

4. What is the probability that a randomly chosen card is a heart, given that it is red?

5. Are "red card" and "spade" independent events? Are they disjoint events?

Example 3.22: A survey found that 58% of families eat turkey at their Thanksgiving meal, 44% eat ham, and 16% serve both turkey and ham.

1. What is the probability that a randomly chosen family serves either turkey or ham at Thanksgiving (or both)?

2. What is the probability that a randomly chosen family serves neither turkey nor ham at Thanksgiving?

3. What is the probability that a randomly chosen family serves turkey and no ham?

4. Given that a family serves turkey, what is the probability that it also serves ham?

Example 3.23: The table below shows the results of a survey in which 28,295 adults were asked whether they were sick on the previous day.

	Cold	Flu	Neither	Total
Smoker	526	153	4980	5659
Nonsmoker	1494	430	20712	22636
Total	2020	583	25692	28295

You randomly select one person. Find the probability that that person...

1. had a cold.

2. had a cold or the flu.

3. had neither illness, given that the person is a smoker.

4. had neither illness, given that the person is a nonsmoker.

5. is a smoker, given that the person had the flu.

6. had the flu or is a nonsmoker.

7. had a cold and is a smoker.

You randomly select two different people. Find the probability that...

1. both are smokers.

2. neither had the flu or a cold.

3. at least one had the flu.