

MATH 2510-001: ELEMENTARY PROBABILITY AND MEASURE

COLTON GRAINGER

SEPTEMBER 16, 2019

This was prepared for University of Colorado's MATH 2510-001 (Introduction to Statistics) in Fall 2019. Please send questions, comments, complaints, and corrections to colton.grainger@colorado.edu. Any mistakes are my own. I appreciate Joe Timmer (2018) as the original author of these lecture notes, as well as the textbook authors Blitzstein and Bwang (2019) and Sameniego (2014) for the main definitions and examples.

CONTENTS

Why is probability relevant to statistics?

Probability is about assigning a likelihood to a particular outcome of an unknown experiment. The most relevant situation is assigning probabilities to sampling from a **known** population.

Statistics is about using the results of sampling to infer information about an **unknown property** of the population.

Index of Terms and Keywords

did something, 3
discrete problems, 4
domain, 4

element, 3
equal to 1, 6
experiment, 3

finite, 4
function, 4

Kentucky Derby, 7

Naive Set Theory, 3

odds, 6
odds against an event, 7

probability, 6

random experiment, 3
randomness, 3
range, 4
real-valued function, 4
rolling a single six-sided fair die, 5
rolling two six-sided fair dice, 5

sample space, 3
Sicherman Dice, 5
simple event, 3
some of the possible outcomes, 3

universal set, 3

very, very primitive concept, 3

Imagine a situation in which you can [do anything](#). Ahead of time, say, you don't know what's going to happen, but then afterwards, you know exactly what happened, because after you *did something*, [something happened](#).

Definition 0.1. We call this *very, very primitive concept* an [experiment](#). I'd like you to interpret this definition in the broadest possible sense.

If that's too vague, you might try to complete the following definition.

Definition 0.2 (Random Experiment). A *random experiment* is an experiment whose outcome cannot be [predicted with certainty](#). All other experiments are said to be deterministic.

I didn't want us to use the word [random](#) until we defined it, and unfortunately, a solid definition of *randomness* will really take some work. But don't worry, we're quickly getting closer to the "right way"¹ of thinking about random events. Let's focus instead on this notion of an [event](#). (The earlier you start thinking about events, the better.)

Definition 0.3 (Event). An event is a collection of *some of the possible outcomes* of an experiment.

Now, having the notion of an event, we might as well collect *all* the possible outcomes of an experiment into a single object:

Definition 0.4 (Sample Space). A *sample space* is the [set](#) of all possible outcomes of an experiment. The sample space plays the role of the *universal set* in problems involving the corresponding experiment, and it will be denoted by Ω .

Note. One of the big breakthroughs in probability that made it possible to actually treat this as a mathematical subject instead of just something more like astrology was the idea of using sets. ◀

Hence, the whole point of the mathematical review we've done over the last 3 weeks has been to become comfortable with the notions of [unions](#), [intersections](#), and [complements](#), which are the basic operations of set theory. (See section 0.5 of math2510.coltongrainger.com/guide).

Note. Before realizing questions of probability in terms of set theory, people tried to just solve probability problems by just kind of writing down some stuff that sounded intuitive, or reasoning by analogy and various heuristics. Most of those heuristics unfortunately turned out to be completely wrong. Roughly around the same time gamblers were asking Isaac Newton (probably one of the top three most famous mathematician-physicists of all time) for gambling advice (also begging for help from Fermat and Pascal).

Why? Because at the time, no one knew how to "do probabilities", so if you were like a degenerate gambler and you really needed to know the odds, then you had to go to someone of the stature of Isaac Newton, Fermat, or Pascal to get an answer. You know, that was 300 years ago, and so one of the cool things is that after a few weeks of [MATH 2510], you'll be able to easily do calculations that 300 years ago, you'd have to consult Isaac Newton with. (Joe Blitzstein, <https://youtu.be/KbB0FjPg0mw>) ◀

Definition 0.5 (Simple Event). A single outcome of a experiment, i.e., a single [element](#) of Ω , is called a *simple event*.

¹When mathematicians say that a definition is "done right", they usually mean that its a definition presented in such a way folks can unanimously agree on it without spending more time squabbling over details than actually using the definition. But, if you're interested in squabbling over details, I highly recommend the book [Naive Set Theory](#) by Paul Halmos.

Note. We have thus shown: an event is simply a subset of Ω . While simple events can be viewed as “compound” events of size one, we will typically reserve the phrase “event” for subsets of Ω with more than one element. ◀

Ok, here’s a very naive definition of probability. I’m going to stop using boldface text and all the fancy symbols because we are just going to hash this one out, it’s not formal anymore.²

$$\mathbb{P}(A) = \frac{\text{Number of outcomes in } A}{\text{Number of all possible outcomes}} = \frac{|A|}{|\Omega|}.$$

You can only use this definition when you have *strong justification* for doing so. Fortunately, I will give you some strong justification now:

Note (It’s all finite!). I assure you, we will deal exclusively with “*discrete problems*”, that is, with problems in which the sample space is finite. We care about the size of the sample space, that is, how many elements are contained in it, because *developing a precise description of the sample space of a experiment is always the first step in formulating a probability model for that experiment*. ◀

So I’m imagining we have some experiment that we’re considering, we have this sample space Ω , we have a subset $A \subset \Omega$ we’re interested in. We want to know what’s the chance that this particular event A will occur?

Note (Measure it!). The idea is to measure the size of A , then to measure the size of Ω , then to compare the two. We will let $\mathbb{P}: \Omega \rightarrow [0, 1]$ be the *real-valued function* that measures the size of each event $A \subset \Omega$, just by counting up how many elements contained in A . (Remember that sets do not have repeated elements!) ◀

This naive definition gives us the probability of an event A , written as “big \mathbb{P} of A ”, where A is an event. “Big \mathbb{P} ” (for the P in *Probability*) is a *function* with *domain* Ω and *range* $\mathbb{P}(\Omega) \subset [0, 1]$.

$$\mathbb{P}: \Omega \rightarrow [0, 1].$$

So we want “Big \mathbb{P} of A ”, i.e., we want the real number $\mathbb{P}(A) \in [0, 1]$, we want to know all about it. That’s the question we’re going to be considering throughout this course, OK?

How do we get that?

Well, that’s a hard question, that’s what this entire course is about. But the naive definition would be to just say, that’s just number of possible outcomes in the denominator, and then the number of favorable outcomes—and by favorable, I mean “favorable to A ”—divided by a number of possible outcomes. So the denominator is just the size of the sample space, it’s the number of possible outcomes. And the numerator is just how many of those outcomes did A occur? OK?

So for example, if we flip a coin—flip coin twice, there are four possible outcomes, right? Either the coin lands heads on the first toss and heads on the second toss, or heads and then tails, or tails and then heads, or tails and then tails. So we have these four different outcomes, OK? Now suppose we want to know, what’s the probability that both tosses are tails? Then according to this, it would be one quarter, right? Because we have one—this would be the favorable outcome, there are four of them, $1/4$, that’s it.

So that’s sort of like the high school definition of probability as well, to just count how many possibilities there are, how many of them did the thing you want happen, and that’s it.

But notice, though, I didn’t say anything about, is it a fair coin? Is it, you know—and well, that’s the question, is what does it actually mean for a coin to be fair?

This is where we need to start working *together*.

²An astute reader will notice that I am now directly copying the transcript of Blitzstein’s lecture 01. I encourage you to watch it yourself. I cannot claim the rest of this as my own writing.

Exercise 0.6 (Rolling fair dice). Consider the experiment of *rolling a single six-sided fair die* and counting the number of dots on the top of the die.³

- (1) What is the sample space of all possible outcomes? Are the outcomes equally likely?

The sample space is the set $\{1, 2, 3, 4, 5, 6\}$. The outcomes are all equally likely to occur.

- (2) Assign a probability to each of the outcomes in the sample space you just found. Do your probabilities add up to 1? Should they add up to 1? Explain.

The probability is $\frac{1}{6}$ for each of the six outcomes in the sample space. They do and they should add up to 1, since they represent all possible distinct outcomes.

- (3) What is the probability of rolling a number less than 5 on a single throw?

The event of “less than 5” consists of the outcomes $\{1, 2, 3, 4\}$. Therefore the probability of the event is $\frac{4}{6} = \frac{2}{3}$.

- (4) What is the probability of rolling an **odd** number?

The probability is 0.5.

Exercise 0.7 (Summing fair dice). Now consider the experiment of *rolling two six-sided fair dice* and recording the **sum** of the number of dots on the top of each die.

- (1) Using the chart below for assistance, find the sample space of all outcomes. Are the outcomes equally likely?

	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

The sample space is the set $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. These outcomes are NOT equally likely to occur. For example, there is only one way to roll a sum of 2, but there are three ways to roll a sum of 4.

- (2) Assign a probability to each of the outcomes in the sample space you just found. Do your probabilities add up to 1? Should they add up to 1? Explain.

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

They do and they should add up to 1, since they represent all possible distinct outcomes.

- (3) What is the most likely outcome of this experiment?

Rolling a “7” is the most likely outcome

- (4) What is the probability of rolling a sum less than 5 on a single throw?

The event of “less than 5” consists of the outcomes $\{2, 3, 4\}$. Therefore the probability of the event is $\frac{6}{36} = \frac{1}{6}$.

- (5) What is the probability of rolling an **odd** sum? Are you certain?

The probability of rolling an odd number is 0.5. Yes, as for each outcome of the first die, exactly half of the outcomes of the second die will yield an odd sum.

Exercise 0.8 (Sicherman Dice). *Sicherman Dice* There is in fact, one other way to mimic the above distribution with a pair of six-sided die without allowing blank faces or sides with “negative” pip counts. One die is labeled $\{1, 2, 2, 3, 3, 4\}$ and the other is $\{1, 3, 4, 5, 6, 8\}$.

- (1) Using the chart below for assistance, find the sample space of all outcomes. Are the outcomes equally likely?

³The rest of these exercises are from Joseph Timmer (2018).

	1	2	2	3	3	4
1						
3						
4						
5						
6						
8						

The sample space is the set $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. These outcomes are NOT equally likely to occur. For example, there is only one way to roll a sum of 2, but there are three ways to roll a sum of 4.

- (2) Assign a probability to each of the outcomes in the sample space you just found. Does this match with the normal pair of die?

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

Amazingly enough, they do!

- (3) What is the most likely outcome of this experiment?

Rolling a “7” is the most likely outcome

- (4) What is the probability of rolling an **odd** sum? Are you certain?

The probability of rolling an odd number is 0.5. Yes, as for each outcome of the first die, exactly half of the outcomes of the second die will yield an odd sum.

Exercise 0.9 (Boolean Algebra). What is the value of $\mathbb{P}(A)$ if event A is certain to occur? What is the value of $\mathbb{P}(A^c)$, in that case?

When the event A is certain to occur, $\mathbb{P}(A)$ is equal to 1. In that case, $\mathbb{P}(A^c) = 0$, since the complement must be impossible to occur.

Exercise 0.10. Although often used interchangeably in common language, there is a difference between the *probability* of an event A and the *odds* in favor of an event A . Whereas the probability of an event A measures the likelihood of A occurring out of all possible outcomes of an experiment, the odds in favor of an event A measures the ratio of the likelihood of A occurring compared to the likelihood that it does not.

For example, the probability that a flipped fair coin lands heads up is $\frac{1}{2}$ because there is one way for it to land heads and 2 total possible outcomes of the coin flip. However, the odds in favor of the coin landing heads is 1 to 1, because there is one way for it to land heads and 1 way for it not to land heads.

- (1) If you roll a (fair) six-sided die, what is the probability of rolling a 4?

When a six-sided die is rolled, there is ONE way to roll a 4 and SIX total outcomes. Therefore, the probability of rolling a 4 is $\frac{1}{6}$.

- (2) If you roll a (fair) six-sided die, what are the odds in favor of rolling a 4?

When a six-sided die is rolled, there is ONE way to roll a 4 and FIVE ways to not roll a 4. Therefore, the odds in favor of rolling a 4 are 1 to 5.

- (3) If you pick a single random card from a deck of 52, what is the probability of drawing an ace?

In a standard deck of 52 cards, there are 4 aces and 52 total cards. So, the probability of drawing an ace is $\frac{4}{52} = \frac{1}{13}$.

- (4) If you pick a single random card from a deck of 52, what are the odds in favor of drawing an ace?

In a standard deck of 52 cards, there are 4 aces and 48 cards that are not the aces. So, the odds in favor of drawing an ace are 4 to 48 or 1 to 12.

Exercise 0.11 (Odds against an event). The *odds against an event* are the reciprocal of the odds in favor of the event.

- (1) In a recent *Kentucky Derby*, the betting odds for (or equivalently, the odds *against*) the favorite horse, Point Given, winning were 9 to 5, then what was the probability that Point Given would win the race?

Since the odds against the horse winning are 9 to 5, the probability of the horse NOT winning is $\frac{9}{14}$. This means that the probability that the horse will win is $\frac{5}{14} \approx 35.7\%$.

- (2) In the same race, the betting odds for the horse Monarchos were 6 to 1. What was the probability that Monarchos would win the race?

The probability is $\frac{1}{7} \approx 14.3\%$.

- (3) If the betting odds for Invisible Ink were 30 to 1, what was the probability that this horse would lose the race?

The probability was $\frac{30}{31} \approx 96.8\%$.

Exercise 0.12 (Roulette). In casinos, **payouts** are quite different from odds, even though the same notation is used. Often when people quote the “odds of an event”, they are actually stating payouts. The game of Roulette is a fine example of this.

American Roulette is played by spinning a marble around a rotating wheel with slots for each number 1–36 and an additional slots marked 0 and 00. In number ranges from 1 to 10 and 19 to 28, odd numbers are red and even are black. In ranges from 11 to 18 and 29 to 36, odd numbers are black and even are red. Both 0 and 00 are colored green. Assume the wheel is unbiased and all outcomes are equally likely.

- (1) The payout for betting on any single number is 35 to 1, meaning you **earn** \$35 for every \$1 wagered. What would the probability of a single number be **if these were the odds against** the single number being selected? How does this compare to the **actual** odds against and the probability of any given number being the winner?

If the payout was the actual odds against, the probability of landing any single number would be $\frac{1}{36}$. In truth, the probability of any single number is $\frac{1}{38}$ and the odds against are 37 to 1.

- (2) The payout for betting on red (or black) is 1 to 1. If these were the odds against red, what would be the probability of landing on red? How does this compare to the **actual** odds against and the probability of the outcome being red slot on the wheel?

If this were the odds against, the probability of landing on red would be 0.5. In truth, the probability of landing on red is $\frac{18}{38} = \frac{9}{19}$ and the odds against are 10 to 9.

- (3) The payout for betting on even (or odd) is also 1 to 1. The slots 0 and 00 are not considered even, nor odd, in this game. How does this distinction affect the game?

The distinction gives similar probabilities to the previous question. The purpose of this consideration, and coloring them green, gives the house the edges. Note the payouts are the same as odds against if these spaces didn't exist.