STAT 145
CHAPTER 12 - PROBABILITY - STUDENT VERSION

The **probability** of a random event, is the proportion of times the event will occur in a large number of repititions. For example, when flipping a coin, the probability of getting a heads is $\frac{1}{2}$, because if we flip the coin a large number of times, it should land on heads half of the time.

For a random phenomenon (such as flipping a coin) we want to construct a **probability model**. A probability model consists of two things:

(1) ________________
(2) ________________

In the case of flipping a coin, we can construct the following probability model.

Example 1: Flipping a coin

Outcome	Probability
Heads	1/2
Tails	1/2

Now we will look at some important definitions

- The **Sample Space**, which we write as S, is the ________________________ .
  -For above example: $S = \{\text{Heads}, \text{Tails}\}$
- An **Event** is an ___________________________ of a random phenomenon.
  -For above example, one possible event is: $A = \{\text{Heads}\}$
- For an event A, we can write the **Probability that A occurs** as $P(A)$.
  -For above example: $P(A) = 1/2$

The first important thing to realize, is that Probabilities are proportions, and therefore *must* be a number between 0 and 1.

$$0 \leq P(A) \leq 1$$

The Sample Space is the list of all things that can possibly happen. So the probability of S, is just the probability that something happens (which is clearly 1). This tells us that the probabilities in the probability model have to sum to 1.

$$P(S) = 1$$

Finally, we see that an event can be any combination of outcomes in the Sample Space. For Example 1, the possible events are

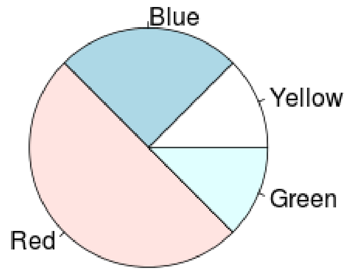
$$A = \{\text{Heads}\} \qquad B = \{\text{Tails}\} \qquad C = \{\text{Heads or Tails}\} \qquad D = \{\text{Niether Heads nor Tails}\}$$

For each of these events, the corresponding probabilities are

$$P(A) = \qquad P(B) = \qquad P(C) = \qquad P(D) =$$

Example 2: A spinner

FIGURE 1. A Simple 4 Color Spinner



If we have a needle on this spinner, and we spin it, this is a random phenomenom where the outcomes are the different colors that the needle can land on. Answer the following questions.

(1) Fill in the probability model

Outcome	Probability

(2) What is the Sample Space?

(3) An example of an event here, is $A = \{Blue\}$, come up with 4 other possible events.

(4) For the event A given above, $P(A) = 1/4$. What is the probability that each event from (3) occurs?

RANDOM VARIABLES

So far, we have looked at cases where the outcomes are categorical (Heads or tails, and color). From a math perspective, we know that we can do more interesting things if we can get outcomes whose values are quantitative. A **Random Variable** is a variable whose value is a ____________ outcome of a random phenomenon.

A common way to do this, is by counting the number of successes. For example, we return to Example 1, but this time, let's count the number of heads we get in 1 flip. (In 1 flip, we can either get 0 heads, or 1 heads).

Example 3: Flipping a coin - Number of heads

Outcome	Probability
1	1/2
0	1/2

This is similar to the probability model from before, but now that the outcomes are quantitative, we are dealing with a Random Variable, so we now refer to the model as a probability distribution. A **probability distribution** of a random variable X, tells us what values X can take, and how often X will take it.

Discrete Random Variables

Example 4 - Flipping a coin multiple times

Let's flip a coin 3 times. The sample space is going to be the list of all possible outcomes.

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

But if we flip a coin and let X be the number of times we get heads, what is the *sample space of X*? (What possible values can X take?)

$S = \{0, 1, 2, 3\}$

The probability distribution of X is given by something called the Binomial Distribution. It is an interesting and well-studied distribution, but we don't cover it in this class. It gives the probability of having x successes out of n trials, when the probability of success is p. In this example, n = 3, and p = .5. This distribution is covered in STAT 345. If you are interested, feel free to ask me for more information.

Outcome	Probability
0	1/8
1	1/4
2	1/4
3	1/8

(1) What is the probability that we get exactly 2 heads?

(2) What is the probability that we get at least 2 heads?

So far, we have only discussed Discrete RV's. A **Discrete Random Variable**, is a RV for which we can list out the possible outcomes. *The formal definition deals with the different sizes of infinity, a surprisingly simple concept. We call a RV discrete, if the outcomes are countable.* So far, we have looked only at cases where the RV can have a finite number of outcomes. But that isn't necessarily the case. Here is an example of a discrete random variable with infinite number of outcomes.

Example 5 - Shooting baskets

Let's say that I am a 50 percent free throw shooter. I will shoot freethrows until I finally make one, and X will be the number of total freethrows that I shoot.

Technically, there is no limit on the number of freethrows I will shoot. For example, the probability that it takes me 1000 shots to make a free throw, is pretty small, but not quite 0. In this case X has the following probability distribution

Outcome	Probability
1	1/2
2	1/4
3	1/8
4	1/16
$\vdots$	$\vdots$

This is called a geometric distribution and it is also well-know. In this case, p = .5

(1) What is the sample space of X?
(2) Do the probabilities for this distribution add up to 1?
(3) What is the probability that it takes me less than 3 shots to make a freethrow?

Continuous Random Variables

A Random Variable is called **continuous**, if we cannot list out all of its possible outcomes. This is true when the possible outcomes are all of the real numbers. Examples of this, is when a Random Variable can be a distance or a length of time. How do we assign probabilities to numbers if we cant even list them out? This leads to an important fact about continuous random variables.

$$P(\text{Continuous Random Variable} = \text{Any number}) = 0$$

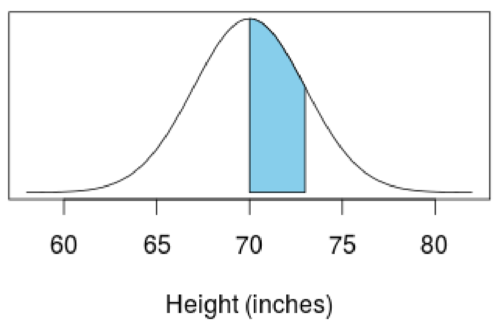
Continuous random variables use **density curves** to assign probabilities. The area under the curve, is the probability of the outcome occuring. Remember, this makes sense, because the area under a density curve must be equal to 1. A common example of density curve is the Normal Distribution.

Example 6 - Height of American Males - Normal Distribution

Assume the heights of American males is Normally distributed with mean, $\mu = 70$ and standard deviation $\sigma = 3$. If we pick an American male at random, and let X be his height, then $X \sim N(70, 3)$.

(1) What is the probability that the randomly chosen male is between 70 and 73 inches? In other words, find $P(70 < X < 73)$?

Figure 2. Normal with $\mu = 70$, $\sigma = 3$



The probability is given by the area under the curve between 70 and 73. To obtain the area, we can use the method from chapter 3 of converting to z-score and using the table.

Or in this case, we can use the 68-95-99.7 rule. 68% is contained between 67 and 73, so using symmetry, we know that 34% is between 70 and 73. So the answer is $P(70 < X < 73) = .34$.

(2) What is the probability that the randomly chosen male is taller than 75 inches? Find $P(X > 75)$.

First we calculate the z-score.

$$z = \frac{x - \mu}{\sigma} = \frac{75 - 70}{3} = 1.67$$

Looking up this z-score in the Standard Normal table gives 0.9525. But the table gives area's to the left. So to get the probability we are looking for we have to subtract from 1 like so

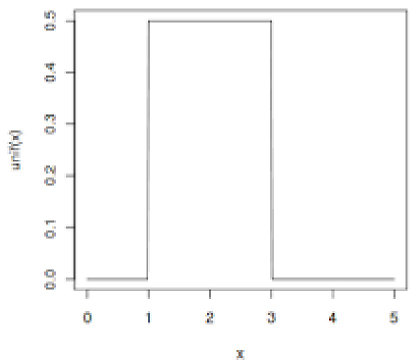
$$P(X > 75) = 1 - .9525 = .0475$$

In example 6, we saw that the Normal distribution was a valid density curve. But the same principle works with any valid density curve. So long as the curve is always greater than zero, and the area under the curve adds up to 1.

Example 7 - Random number

If we ask a computer to choose a number at random from 1 to 3, it will usually use a uniform distribution. This just means that it has an equal chance of picking any number. The density curve for this situation looks like this

FIGURE 3. Random number from 1 to 3



(1) Check that this is a valid density curve.

(2) What is the probability that the number we choose is between 1.5 and 2.5?

PROBABILITY RULES

Two events are called ____________ if the outcome of 1 has no effect on the outcome of the other.
Two events are called ____________ if they have no outcomes in common.

To determine if two events are disjoint, we ask the question: ______________________________
If they can, then the events are not disjoint. Let's look at an example.

Example 8 - Rolling a die

Event A	Event B	Disjoint?
Roll a 1, 2 or 3	Roll a 4, 5 or 6	Yes
Roll a 1	Roll an odd	No
Roll < 3	Roll > 3	
Roll ≤ 3	Roll ≥ 3	
Roll a 3	S	

PROBABILITY RULE 1 - If A and B are disjoint, then $P(A \text{ or } B) = P(A) + P(B)$.

Lets look at 2 cases of example 8 to help us understand this. In the first row, events A and B are disjoint, so we can use this rule. we can see that P(A) and P(B) = 1/2. So according to our rule, P(A or B) = 1/2 + 1/2 = 1. Lets check and see if this is correct
$P(A \text{ or } B) = P(\{1, 2, 3, 4, 5, 6\}) = P(S) = 1$. But for the second case, this rule does not apply! Since the events are not disjoint, the rule is not true. Check it yourself to make sure.

PROBABILITY RULE 2 - P(A does not occur) = 1 - P(A)

Hopefully this makes sense to us. We already said that the probability something occurs must be 1. So the probability that A doesn't occur + the probability that A does occur must equal 1. Think about the connection to disjoint events.

To illustrate the next probability rule, we will look at an example.

Example 9 - Two 4-sided dice Lets say we have two, 4 sided dice. The probability distribution for each die looks like this

Outcome	Probability
1	1/4
2	1/4
3	1/4
4	1/4

If I roll both of them and record their sum, there will be 16 possible combinations.

Die 1	Die 2	Sum	Prob	Die 1	Die 2	Sum	Prob
1	1	2	1/16	3	1	4	1/16
1	2	3	1/16	3	2	5	1/16
1	3	4	1/16	3	3	6	1/16
1	4	5	1/16	3	4	7	1/16
2	1	3	1/16	4	1	5	1/16
2	2	4	1/16	4	2	6	1/16
2	3	5	1/16	4	3	7	1/16
2	4	6	1/16	4	4	8	1/16

(1) What is the sample space for the sum of the two dice rolls?

(2) What is the probability of the sum being 5?

If we use probability rule 1, we will get P(Sum=5) = 1/16 + 1/16 + 1/16 + 1/16 = 4/16 = 1/4. This is correct, but notice this, there were 16 possible combinations, and 4 ways we can get the sum to be 5. Therefore P(Sum=5) = 4/16. This leads us to our last probability rule.

PROBABILITY RULE 3 - If all outcomes are equally likely, we can say that $P(A) = \frac{\text{Number of ways A can occur}}{\text{Total number of possibilities}}$
This is a field of math called combinatorics. A fancy name for counting. This is the method you would use to find the probability of different poker hands.

One Last Thing

Personal Probability is a number between 0 and 1 that expresses a persons judgement of how likely the outcome is.

Example 9 - Personal Probability

What is the probability that Donald Trump is elected president?

Practice Problems

Problem 1 - Use the probability rules.

I have 4 favorites breakfasts: Waffles (W), Pancakes (P), Bacon (B) and Oatmeal (O). I eat one of these breakfasts every morning, and never anything else.

1. The probability that I will eat Waffles or Pancakes for breakfast is .75.
2. The probability that I don't eat bacon is .80.

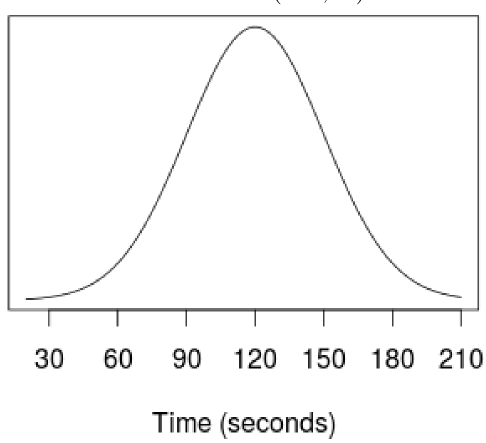
Fill in the probability model for what I eat for breakfast.

Outcome	Probability
W	.45
P	
B	
O	

Problem 2

The time it takes to complete a maze is Normally Distributed with mean $\mu = 120$ seconds and standard deviation $\sigma = 30$ seconds.

Figure 4. N(120,30)



(1) What is the probability that it takes exactly 150 seconds to finish the maze? (Hint: Check the beggining of the Continuous Random Variables section).

(2) What is the probability it takes less than 60 seconds to finish the maze?

(3) What is the probability it takes less than 55 seconds to finish the maze?

(4) What is the probability it takes more than 55 seconds to finish the maze?