Lecture 12: Continuing Introduction to Probability

Mathematically Describing Randomness

and

Discrete Random Variables

Recap: Probability Axioms

What Are They?

and

How Do We Use Them?

Recap

Probability and Its Axioms

Probability: A special measurement used to describe the likelihood of a specific event. A probability of 0 means the event will not occur. A probability of 1 means the event will occur.

In order to valid for a sample space, S, the probability must follow these rules:

1. For any event A, $P(A) > 0$

2. $P(S) = 1$

3. If $A \cap B = \emptyset$ then $P(A \cup B) = P(A) + P(B)$.

We can translate these rules into words:

1. All probabilities are at least 0 (they can be 0).

2. The probability that something occurs is 1.

3. The probability of any event is the sum of the probabilities of its parts.

Recap

Example

Suppose that A and B are two events of a sample space S. Using the probability axioms, show that

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

Working with Sample Spaces

From Sample Spaces to Random Variables

Recap

Example: Red die, blue die

A fair red die and fair blue die are tossed at the same time.
The number of dots facing up on each die are recorded.
We can ask a lot of questions related to the outcome of the
toss:

1. What is the probability that the red die shows 4 dots
   facing up?
2. What is the probability that the blue die and the red
   die both have 4 dots facing up?
3. What is the probability that one of the die has 4 dots
   facing up?
4. What is the probability that neither of the die has 4
   dots facing up?
5. What is the most likely total number of dots facing up?
6. Assuming the red die shows a 6, what is the
   probability that the blue die shows a 4?
7. What is the probability that the total number of dots
   facing up is 5?

Recap

Example: Red die, blue die (cont)

Creating a Random Variable

With the sample space in hand, we might find many of our questions about the total are easy to handle by defining a new variable:

Let T be the total number of dots facing up on both die after performing the random experiment.

So for instance, if the outcome we observe is $(1, 3)$ (red is 1, blue is 3) then $T = 4$. In this way, the value T takes varies based on the outcome of a random experiment. Further, the probability of the outcomes of the experiment determine the probability of the value that T takes. We call such a variable a **Random Variable**.

Example: Red die, blue die (cont)

Creating a Random Variable

def: Random Variable: A variable which takes numeric values based on the outcome of a random experiment. We use capital letters for the variables and lower case letters when we need to generically refer to values it may take after the outcome of the random experiment is observed.

Since T takes values based on the outcome of our experiment and our outcomes have probabilities, then the value of T *inherits* the probability. We use a capital letter for the variable and a lower case letter for the specific value it takes after a random experiment.

t	2	3	4	5	6	7	8	9	10	11	12
#Outcomes	1	2	3	4	5	6	5	4	3	2	1
$P(T = t)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Recap

Example: Red die, blue die (cont)

Using Random Variables

Because a random variable inherits its probability from the sample space, it inherits the probability rules from the sample space too. For instance, we can write things like:

$$P(T < 4) = 1 - P(T \geq 4) = 1 - P(T > 5)$$

or

$$P(2 < T < 6) = P(T = 3) + P(T = 4) + P(T = 5)$$

Again, this all goes back to our *probability axioms* but now that T is a numeric random variable we don't have to go back and think about the events/sets involved. The rules stay the same.

Recap

Example: Red die, blue die (cont)

More Random Variables?

Suppose that we are interested in other aspects of the
random event. We could define other random variables:

- X_R as the number of dots facing up on the red die,
- X_B as the number of dots facing up on the blue die,
- $Y = X_R - X_B$, and
- $Z = \max\{X_R, X_B\}$.

These are all random variables, too, though they may not
be useful unless we are very concerned with the outcomes
they represent. Notice that a single random event will
result in all of these random variables taking a value at the
same time.

Example: Red die, blue die (cont)

More Random Variables?

Notice that a single random event will result in all of these random variables taking a value at the same time.

So if we role a $(3, 2)$ then we get:

- $T = 3 + 2 = 5$
- $X_R = 3$
- $X_B = 2$
- $Y = 1$
- $Z = 3$.

In other words, it is valid to ask questions involving multiple random variables, such as:

- What is $P(T \leq 9, Z = 6)$?
- What is $P(T \leq 9, Z = 6, Y = 1)$?

Recap

Example: Deck of Cards

Goal 1: Find the probability that if you are dealt two cards that you will have a pair

Goal 2: Find the probability that if you are dealt five cards you will have a full house (3 of one rank, 2 of another rank)

Problems

- Sample space very large, can we avoid writing it out?
- How do we account for the order the cards are dealt (ignore ordering vs use ordering)?
- How do we make

Recap

Summary

Understanding Random Experiments

- We talked about how random experiments result in outcomes
- That the set of all outcomes is called a **sample space**
- That we can group outcomes together into **events**
- That the likelihood of the outcomes can be measured using probability

Understanding Probability

- We talked about the rules that a probability must follow to be a valid way of measuring likelihood.
- the simple set of rules (called the **probability axioms**) can be used to show many more complicated rules that must also be true.

Recap

Summary (cont)

Random Variables

- We gave a definition of random variables.
- We can create random variables to make answering questions about our outcomes easier.
- We can create multiple random variables for the same experiment
- The random variables we create can be used together to answer certain questions.

Conditional Probability

How Does Partial Information Effect Probabilities?

and

The Monty Hall Problem

What is Conditional Probability?

Most of what we've discussed up until this point has assumed that we have a random experiment and the outcome we observe is revealed to us all at once.

But what if we could know *partial information* about the outcome?

For instance, if we roll a pair of dice (one red, one blue) and you are told the total, could you better guess what number was on the red die?

In terms of probability, that information changes the conditions in which a specific outcome's likelihood is being measured - it means that we are now dealing with something called a **conditional probability**.

Conditional Probability

How Does Conditional Probability Work?

Here's how it works:

- By learning some detail about the actual outcome, we know that there are only some outcomes that have that detail and some that do not have that detail

- Since we have learned that the actual observed outcome *does* have that detail, then the outcomes that did not have that detail could not have been the outcome that occured

- and it also means that the only possible outcomes that could have occured are a **subset of the sample space**

- In other words, we know that **some event** has occured!

Conditional Probability

What is it?

How it works

How Does Conditional Probability Work?

Drawing Showing How It Works

Conditional Probability

What is it?

How it works

How Does Conditional Probability Work?

Example: Suppose I deal you two cards. You are interested in predicting if the second card I dealt you is Red.

1. What is the probability that the second card is red?

2. What is the probability that the first card is red?

3. What is the probability that the second card is red *given* the first card is red?

4. What is the probability that the second card is red *given* the first card is black?

5. What is the probability that the second card is red *given* the second card is red?

6. What is the probability that the second card is red *given* the second card is black?

Conditional Probability

Conditional Probability Notation

Event Notation

Since the information can be described in terms of events, we use the following notation:

- $P(A|B)$: probability the outcome is in event A given that the outcome is in event B

Since we know that the event is in B, then for it to be in A we must have in $A \cap B$. We can actually get values for conditional probabilities if we know the original probabilites:

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

Which also means that we can write

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

Conditional Probability Notation

Random Variable Notation

Since random variables are just sets of outcomes too, the we can use a similar notation when dealing with random variables:

- $P(X = x | Y = y)$: the probability the random variable X takes the specific values x given that the random variable Y took the value y.

And similarly we can find the conditional probability using:

$$P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

Which also means that we can write

$$P(X = x | Y = y) \cdot P(Y = y) = P(X = x, Y = y)$$

Conditional Probability

Independence

A related concept is **independence**:

Independent Events

Events A and B are said to be **independent** if knowing event B has occured does not impact the probability that A will occur. In other words,

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

Independent Random Variables

Random variables X and Y are said to be **independent** if knowing the values taken by Y has not impact on the probabilities associated with values taken by X. In other words, for any value of x and y,

$$P(X = x|Y = y) = P(X = x)$$

Conditional Probability

Bayes Theorem

Wildly important and useful way of connecting conditional probabilities

Bayes Theorem

For events A and B,

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

For Random Variables X and Y,

$$P(Y = y | X = x) = \frac{P(X = x | Y = y) \cdot P(Y = y)}{P(X = x)}$$

Example

I flip a coin. If I flip heads, I roll a six-sided die. If I flip tails, I roll a 10 sided die. I tell you the number on the die. You tell me the flip of the coin.