

Lecture 13: Discrete Random Variables

Terminology, Use, and Common Distribution

Course page: imouzon.github.io/stat305

Last Lecture Leftovers

Conditional Probability

How Does Partial Information Effect Probabilities?

and

The Monty Hall Problem

Conditional Probability

What is it?

Conditional Probability

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

What is it?

How it works

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 occurred
- and it also means that the only possible outcomes that could have occurred are a **subset of the sample space**
- In other words, we know that **some event** has occurred!

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

What is it?

How it works

Notation

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 probabilities:

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

Which also means that we can write

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

Conditional Probability

What is it?

How it works

Notation

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

What is it?

How it works

Notation

Independence

Independence

A related concept is **independence**:

Independent Events

Events A and B are said to be **independent** if knowing event B has occurred 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

What is it?

How it works

Notation

Independence

Bayes' Theorem

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.

Discrete Random Variables

Meaning, Use, and Common Distributions

General Info

Reminder: RVs

General Info About Discrete RVs

Reminder: What is a Random Variable?

Random Variables, we have already defined, take real-numbered (\mathbb{R}) values based on outcomes of a random experiment.

- If we know the outcome, we know the value of the random variable (so that isn't the random part).
- However, before we perform the experiment we do not know the outcome - we can only make statements about what the outcome is likely to be (i.e., we make "probabilistic" statements).
- In the same way, we do not know the value of the random variable before the experiment, but we can make probability statements about what value the RV might take.

General Info

Reminder: RVs

Discrete?

Terms & Notation

Common Terms and Notation for Discrete RVs

Of course, we can't introduce a *sort of* new concept without introducing a whole lot of new terminology.

We use capital letters to refer to discrete random variables: X, Y, Z, \dots

We use lower case letters to refer to values the discrete RVs can take: x, x_1, y, z, \dots

While we can use $P(X = x)$ to refer to the probability that the discrete random variable takes the value x , we usually use what we call the **probability function**:

- For a discrete random variable X , the probability function $f(x)$ takes the value $P(X = x)$
- In otherwords, we just write $f(x)$ instead of $P(X = x)$.

General Info

Reminder: RVs

Discrete?

Terms & Notation

Common Terms and Notation for Discrete RVs

We also have another function related to the probabilities, called the **cumulative probability function**.

For a discrete random variable X taking values x_1, x_2, \dots the CDF or **cumulative probability function** of X , $F(x)$, is defined as

$$F(x) = \sum_{z \leq x} f(z)$$

Which in other words means that for any value x ,

$$f(x) = P(X = x)$$

and

$$F(x) = P(X \leq x)$$

General Info

Reminder: RVs

Discrete?

Terms & Notation

Common Terms and Notation for Discrete RVs (cont)

The values that X can take and the probabilities attached to those values are called the **probability distribution** of X (since we are talking about how the total probability 1 gets spread out on (or distributed to) the values that X can take).

Example

Suppose that the we roll a die and let T be the number of dots facing up. Define the probability distribution of T . Find $f(3)$ and $F(6)$.

General Info

Reminder: RVs

Discrete?

Terms & Notation

Mean of a Discrete Random Variable

For a discrete random variable, X , which can take values x_1, x_2, \dots we define **the mean of X** (also known as **the expected value of X**) as:

$$E(X) = \sum_{i=1}^n x_i \cdot f(x_i)$$

We often use the symbol μ instead of $E(X)$.

Also, just to be confusing, you will often see EX instead of $E(X)$. Use context clues.

Example:

Suppose that the we roll a die and let T be the number of dots facing up. Find the expected value of T .

General Info

Reminder: RVs

Discrete?

Terms & Notation

Variance of a Discrete Random Variable

For a discrete random variable, X , which can take values x_1, x_2, \dots and has mean μ we define **the variance of X** as:

$$Var(X) = \sum_{i=1}^n (x_i - \mu)^2 \cdot f(x_i)$$

There are other usefule ways to write this, most importantly:

$$Var(X) = \sum_{i=1}^n x_i^2 \cdot f(x_i) - \mu^2$$

Example:

Suppose that the we roll a die and let T be the number of dots facing up. What is the variance of T ?

General Info

Reminder: RVs

Discrete?

Terms & Notation

Summary

Discrete Random Variables

- Discrete RVs are RVs that will take one of a countable set of values.
- When working with a discrete random variable, it is common to need or use the RV's
 - probability distribution: the values the RV can take and their probabilities
 - probability function: a function where $f(x) = P(X = x)$
 - cumulative probability function: a function where $F(x) = P(X \leq x)$.
 - mean: a value for X defined by $EX = \sum_x x \cdot f(x)$
 - variance: a value for X defined by $VarX = \sum_x (x - \mu)^2 \cdot f(x)$

Common Distributions

Working with Off The Shelf Random Variables

Common Distributions

Why Are Some Distributions Worth Naming?

Even though you may create a random variable in a unique scenario, the way that its probability distribution behaves (mathematically) may have a lot in common with other random variables in other scenarios. For instance,

I roll a die until I see a 6 appear and then stop. I call X the number of times I have to roll the die in total.

I flip a coin until I see heads appear and then stop. I call Y the number of times I have to flip the coin in total.

I apply for home loans until I get accepted and then I stop. I call Z the number of times I have to apply for a loan in total.

General Info

Common Distributions

Background

Why Are Some Distributions Worth Naming? (cont)

In each of the above cases, we count the number of times we have to do some action until we see some specific result. The only thing that really changes from the random variables perspective is the likelihood that we see the specific result each time we try.

Mathematically, that's not a lot of difference. And if we can really understand the probability behavior of one of these scenarios then we can move our understanding to the different scenario pretty easily.

By recognizing the commonality between these scenarios, we have been able to identify many random variables that behave very similarly. We describe the similarity in the way the random variables behave by saying that they have a common/shared distribution.

We study the most useful ones by themselves.

General Info

Common Distributions

Background

Bernoulli

The Bernoulli Distribution

Origin: A random experiment is performed that results in one of two possible outcomes: successful or failure. The probability of a successful outcome is p .

Definition: X takes the value 1 if the outcome is a success. X takes the value 0 if the outcome is a failure.

probability function:

$$f(x) = \begin{cases} p & x = 1, \\ 1 - p & x = 0, \\ 0 & o. w. \end{cases}$$

which can also be written as

$$f(x) = \begin{cases} p^x(1 - p)^{1-x} & x = 0, 1 \\ 0 & o. w. \end{cases}$$

General Info

The Bernoulli Distribution

Common
Distributions

Expected value: $E(X) = p$

Background

Bernoulli

General Info

Common
Distributions

Background

Bernoulli

The Bernoulli Distribution

Variance: $Var(X) = (1 - p) \cdot p$