Announcements

- HW2 due today at 5:00

STAT 430: Lecture 7

Discrete Random Variables

Chapter 2

Course page: imouzon.github.io/stat430

What is a "Random" Variable?

Random Variables

Meaning

Random Variables

So we've got this idea that we can define a probability on the outcomes of an experiment. For instance, if we toss a coin three times we can define the sample space as:

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

and (assuming the coin is fair) we can assign a probability of $1/8$ to each outcome.

Then we can talk about things like
$P(\{\text{"first flip is H"}\})$

But what if the scenario is not so simple? What if we are dealing with a situation where all the probabilities are not the same? Or if there are an infinite number of outcomes?

Many common real world scenarios can lead to such problems.

Random Variables

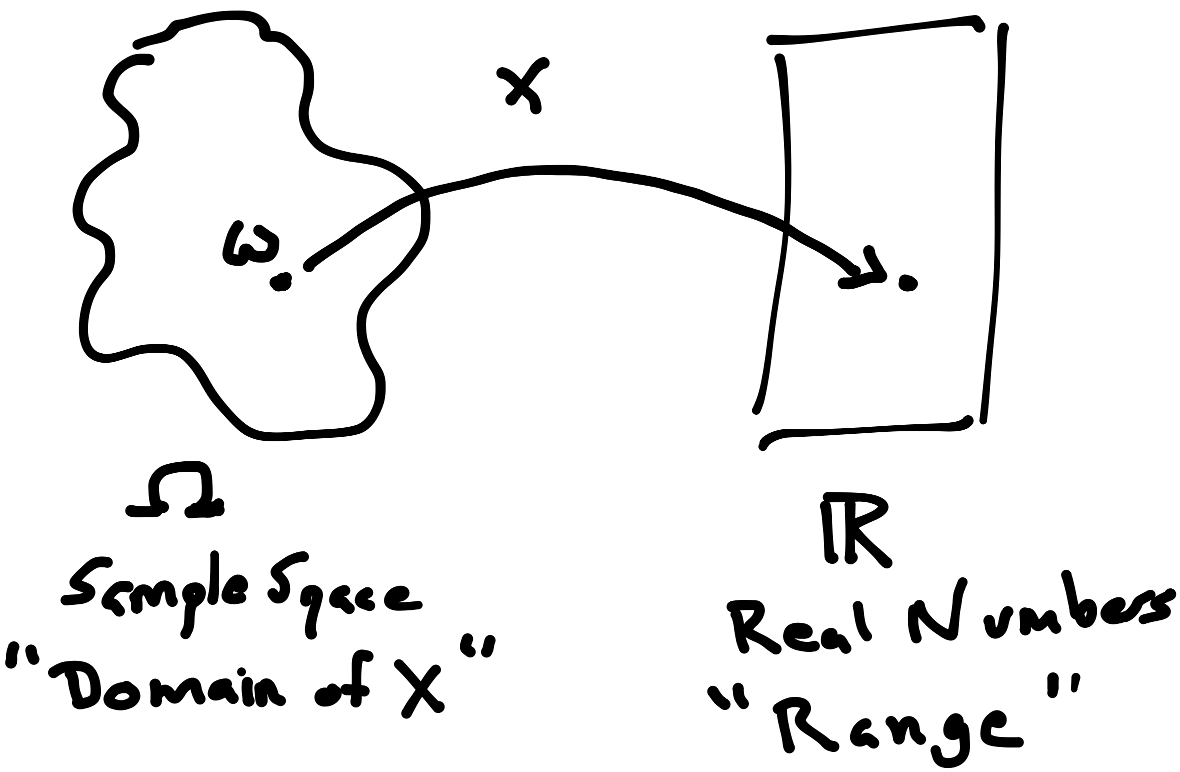
Meaning

Random Variables, cont.

Essentially a random variable is a *function* that takes elements of the sample space to the set of real numbers. Defining a function like this does two things for us

1. It moves elements ω from a sample space Ω that may be strange or hard to describe to the set of real numbers, which we have many tools to work with
2. Commonly occuring types of random variables can be explored more fully apart from the details of the scenario (coin flips, dice rolls, etc. can all be handled as a common type of random variable).

Illustration



$$X : \Omega \to \mathbb{R}$$

Random Variables

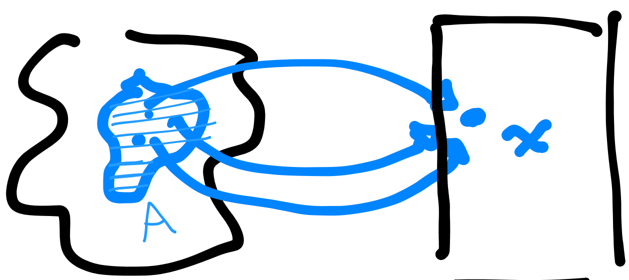
Meaning

Random Variables, cont.

Notice that if we can talk about $A \subset \Omega$ then we can also talk about $\{\omega \in \Omega : X(\omega) = \underline{x}\}$.

Illustration

actual outcomes are unknown follow some defined prob.



Ω $\mathbb{R}$

$P(X = x)$

this probability is "induced" by the probability defined on Ω

$$P(A) = P(\{\omega \in \Omega : \omega \in A\})$$
$$= P(\{\omega \in \Omega : X(\omega) = x\})$$
$$= P(X = x)$$

If we can define a probability on Ω then that structure transfers to our real numbers (as X moves between sets)

Ex Flip a coin twice, record the number of times heads occurs as x

---

$$\Omega = \{HH, HT, TH, TT\}$$
$$\quad 1/4 \quad 1/4 \quad 1/4 \quad 1/4$$

define a random variable X (or $X(\omega)$ if we're being good)

$X(\{HH\}) = 2$

$X(\{HT\}) = 1$

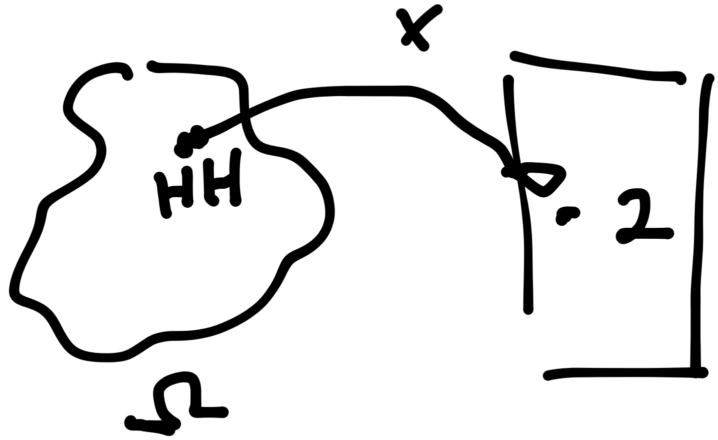
$X(\{TH\}) = 1$

$X(\{TT\}) = 0$

$$P(\{HH\}) = \frac{1}{4}$$

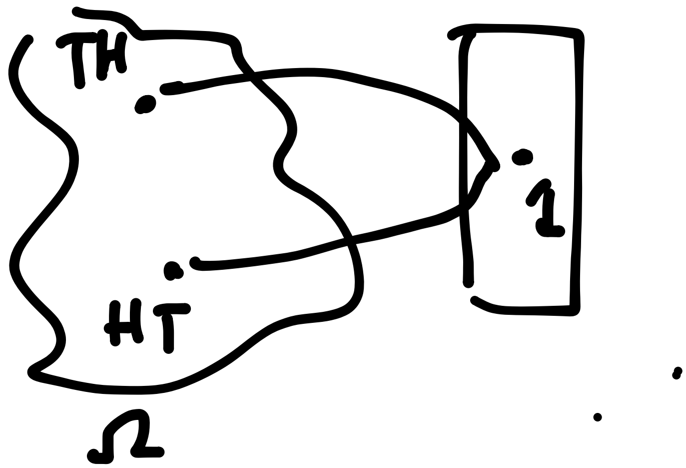
$\Downarrow$ induces

$$P(X=2) = \frac{1}{4}$$



$$P(\{TH\}) = \frac{1}{4}$$

$$P(\{HT\}) = \frac{1}{4}$$

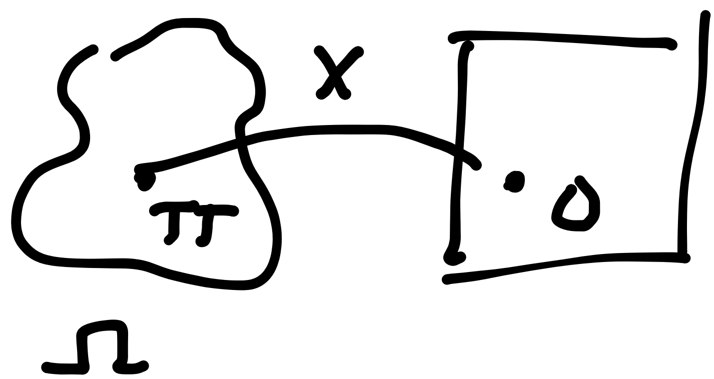


$\Downarrow$ induces

$$P(X=1) = P(\omega = \{TH\} \text{ or } \omega = \{HT\}) = \frac{2}{4}$$

$$P(\{TT\}) = \frac{1}{4}$$

$$P(X=0) = \frac{1}{4}$$



Random Variables

Meaning

Types

Random Variables, cont.

There are two main types of random variables: **Discrete** and **Continous**

Discrete Random Variable A function whose domain is the sample space and whose range is a countable subset of $\mathbb{R}$

Continous Random Variable A function whose domain is the sample space and whose range is a non-countable subset of $\mathbb{R}$

Discrete random variables are based on **counting** measurements, continuous random variables are based on a distance measurement.

Random Variables

Meaning

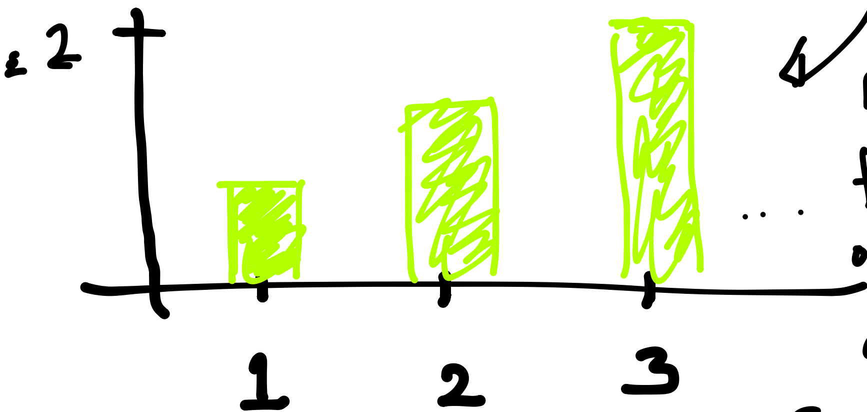
Types

Density/Mass Functions

Density and Mass Functions

Since we are dealing with real numbers when we deal with random variables, we can talk about their probabilities in terms of functions:

In discrete case



height of bar represents the amount of probability at that point

"probability mass function"

In continuous case

area under the curve between a and b represents the probability that $a \le X \le b$

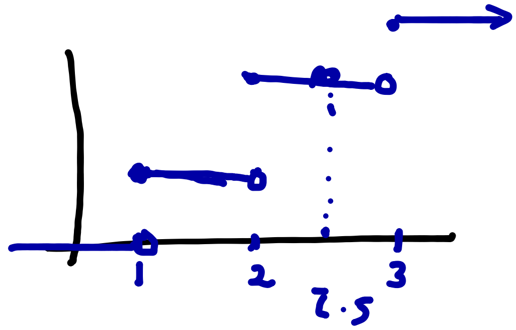
$P(a \le X \le b)$

"prob. density function"

Cumulative Probability Functions

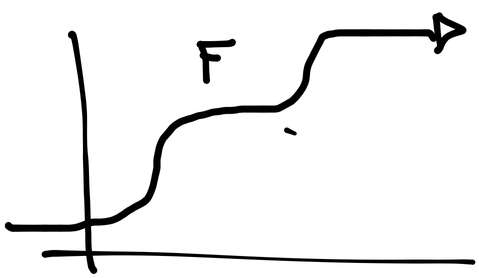
A way of adding up all the probabilities up to a certain value of x.

The Discrete Case



points on the CDF represent how much probability is at that value of X or below

Q: $P(X \leq 2.5) = F(2.5)$



for any x

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

Discrete Random Variables

The Common Types

Discrete Random Variables

Bernoulli

Bernoulli Experiments

Motivation: "I have an experiment that can be thought of as having two outcomes"

In a Bernoulli Experiment, we have the following:

- A single outcome of interest ("success" ~~and "failure")~~
- A probability associated with success, p

1 draw success = I drew A.

Examples:

- I flip a coin and hope I get a heads
- I draw a card from a deck and hope I get an Ace
- I run a marathon and hope to beat my previous record

1 marathon success = break my record.

- single flip
- call "H" success

Bernoulli Random Variables

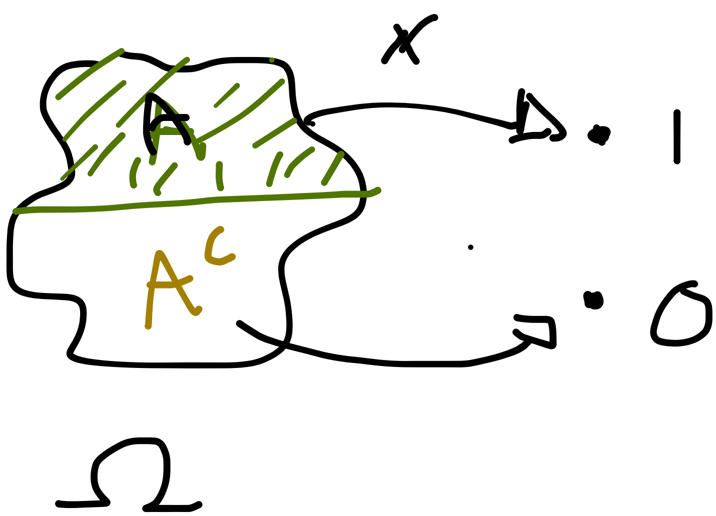
If we have a Bernoulli experiment, then we can define a **Bernoulli Random Variable** by writing:

Let A be an event composed of outcomes ω of the sample space Ω with $P(A) = p$. Then a random variable defined as

$$X(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

is a **Bernoulli Random Variable**. In this case, we would refer to A as the set of all "successful" outcomes. Further, $P(X = 1) = p$ and we can call the probability of success p

i.e. $X = \begin{cases} 1 & \text{if "success"} \\ 0 & \text{otherwise} \end{cases}$



$P(A) = p \longrightarrow P(X=1) = p$

Discrete Random Variables

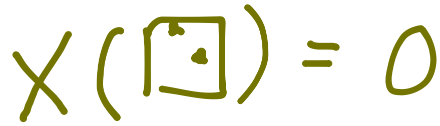
Bernoulli

Example: Rolling dice

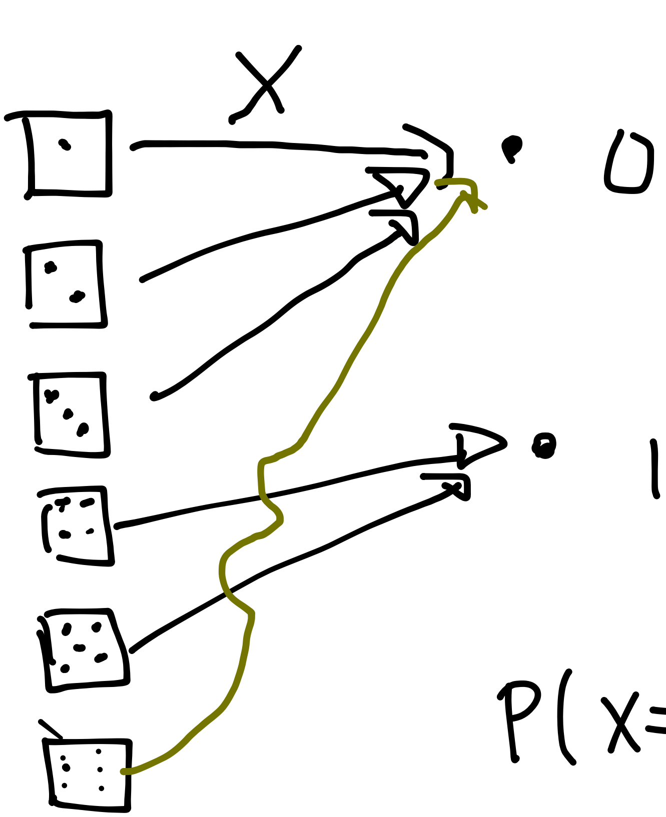
Scenario: we are rolling a fair six-sided die.

a. Our goal is to roll higher than 3 but lower than 6.

b. Our goal is to roll a pair of dice and get a double



a)

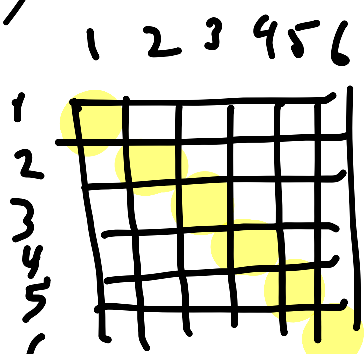


$$P(X=1) = 2/6 = 1/3$$

$$P(X=0) = 4/6 = 2/3$$

$$P(X=19) = 0$$

b)



$$X(\text{"doubles"}) = 1$$

$$X(\text{"not doubles"}) = 0$$

$$P(X=1) = 6/36$$