CS 237: Probability in Computing

Wayne Snyder
Computer Science Department
Boston University

Lecture 9:

- Counting Concluded
- Bose-Einstein combinatorics: One more case....
- Introduction to Random Variables

Counting: Combinations concluded

Reasoning with combinations is very powerful and sometimes the simplest way to analyze a problem, particularly when duplicates are involved....

Consider a series of related problems....

How many permutations are there of the word

PROBAbILiTY

(the two B's and the 2 I's are distinguishable).

Counting: Combinations concluded

Version 2:

How many permutations are there of the word

PROBABILITY

(the two B's and the two I's are INdistinguishable).

Counting: Combinations concluded

Version 3:

How many permutations are there of the word

PROBABILITY

(the B's and I's are INdistinguishable), if all the vowels must be in the order: O..A..I

Counting: Combinations concluded

Version 3:

How many permutations are there of the word

PROBABILITY

(the B's and I's are INdistinguishable), if all the vowels must be in the order: O..A..I.. I and all the consonants must be in the order: PRBBLTY

Counting: Bose-Einstein Combinatorics

	Selection Without Replacement	Selection With Replacement
Ordered Outcome **(Sequence or String)**	**Standard Problem 1(a):** How many permutations of all N letters ABC... (all different)? (If the problem simply says "permutations" this is what they mean.) **Formula:** P(N,N) = N! *Example: How many permutations of the word "COMPUTER"? 8!* **Standard Problem 1(b):** How many permutations of **all** N letters ABC... where some letters are repeated with multiplicities $m_1, m_2, ..., m_p$? **Formula:** $N! / (m_1!* ...\bullet m_p!.)$ *Example: How many different-looking permutations can be made of the word "STATISTICS" = 10! / (3! * 3! * 2!)* **Standard Problem 2:** How many permutations of K letters from N letters ABC... ? **Formula:** P(N,K) = N*(N-1)*(N-2)* ... *(N-K+1) = N! / (N-K)! *Example: How many 3-letter words can be made from letters in "COMPUTER"? P(8,3)*	**Standard Problem:** How many enumerations of K letters from N letters ABC....? **Formula:** N^K *Example: How many 10-letter words all in lower case?* 26^{10}
Unordered Outcome **(Set or Multiset)**	**Standard Problem 1(a):** How many combinations (sets) are there of size K from N objects... (all different)? **Formula:** C(N,K) = N! / ((N-K)! * K!) *Example: How many committees of 3 people can be chosen from 8 people? C(8,3)*	**Standard Problem:** How many ways to choose a multiset of K objects from a set of N objects, with replacement? **Formula:** C(N+K-1,K) *Example: At your favorite takeout place, there are 10 accompaniments, and you can choose any 3, with duplicates allowed (e.g., you can choose two servings of fries and one of mac and cheese). How many possibilities are there? C(10+3-1,3) = C(12,3) = 220*

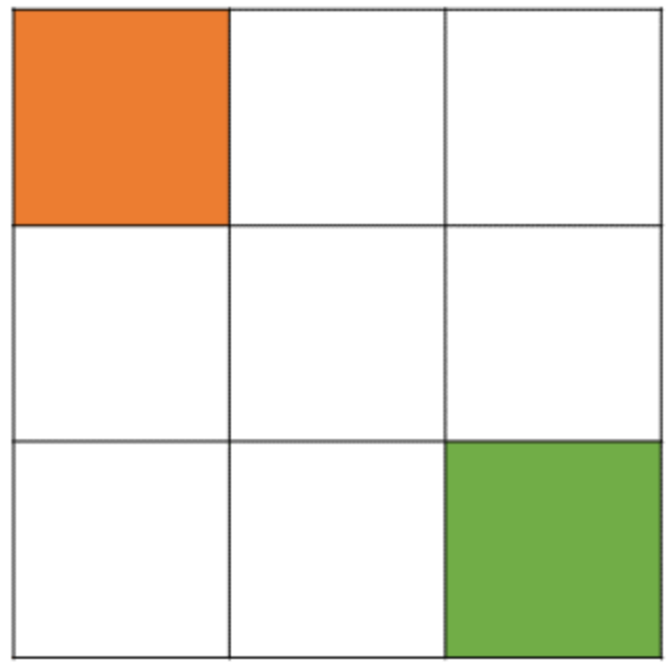
Counting: Bose-Einstein Combinatorics

EXAMPLE: At your favorite takeout place, there are 10 accompaniments, and you can choose any 3, with duplicates allowed (e.g., you can choose two servings of fries and one of mac and cheese). How many possibilities are there?

Counting: Bose-Einstein Combinatorics

EXAMPLE 2:

Let's start out by looking at this 3x3 grid:

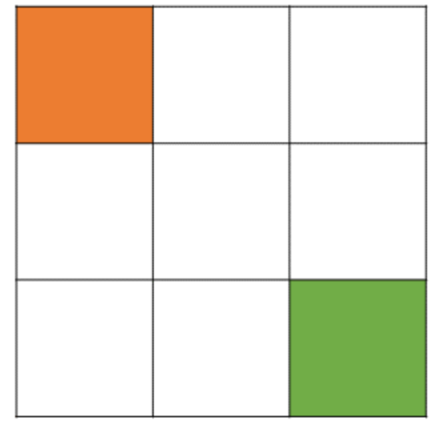


Suppose that we want to get from the orange to the green square. Each step we may move right or down. We will denote the **move to the right by R** and the **move down by D**. We are interested in the number of distinct paths we can take.

Counting: Bose-Einstein Combinatorics

Let's start out by looking at this 3x3 grid:

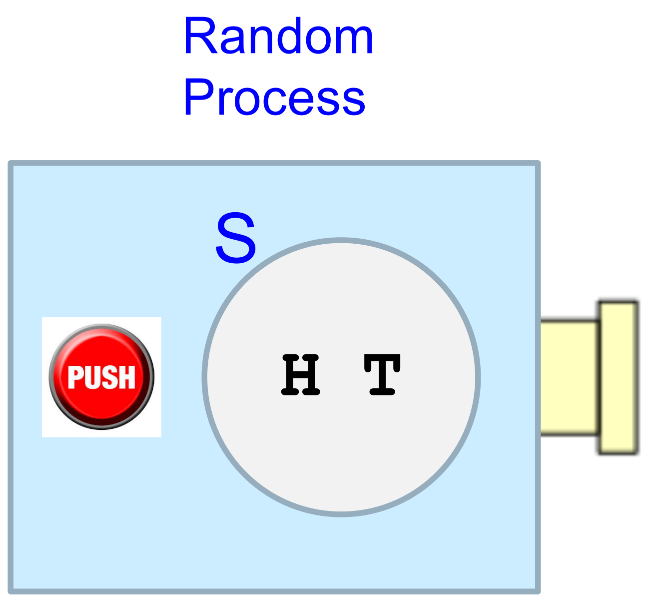
EXAMPLE 2:



Suppose that we want to get from the orange to the green square. Each step we may move right or down. We will denote the **move to the right by R** and the **move down by D**. We are interested in the number of distinct paths we can take.

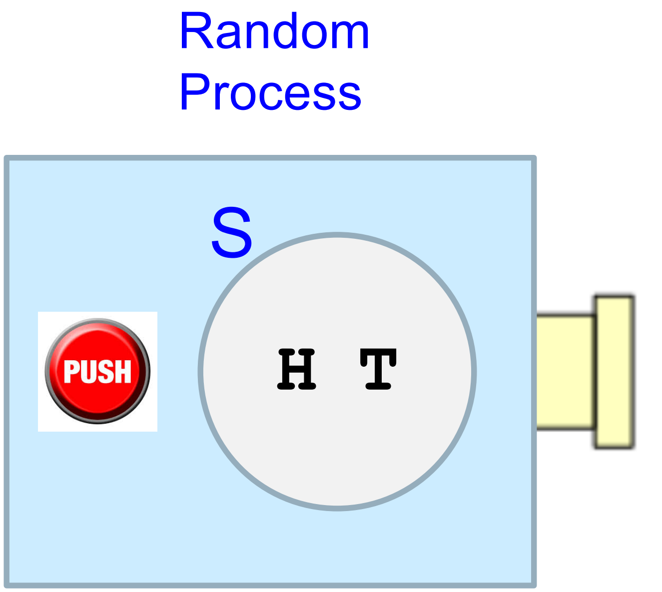
Random Experiments and Random Variables

A Random Experiment is a process that produces uncertain outcomes from a well-defined sample space.



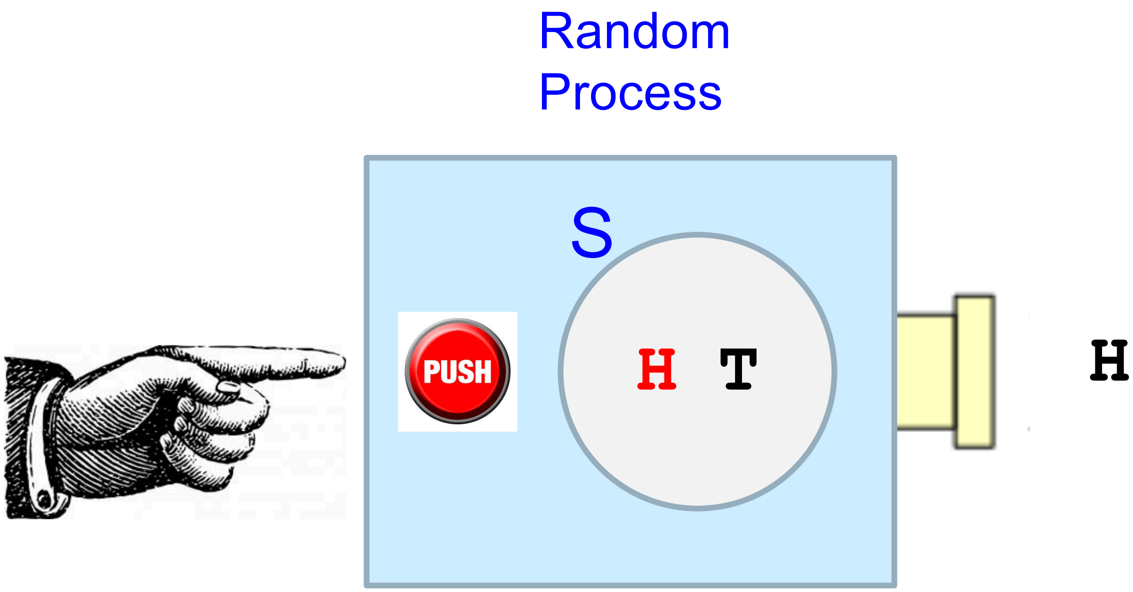
Random Experiments and Random Variables

A Random Experiment is a process that produces uncertain outcomes from a well-defined sample space.



Random Experiments and Random Variables

A Random Experiment is a process that produces uncertain outcomes from a well-defined sample space.



Random Variables

In order to formalize this notion, the notion of a Random Variable has been developed. A Random Variable X is a function from a sample space S into the reals:

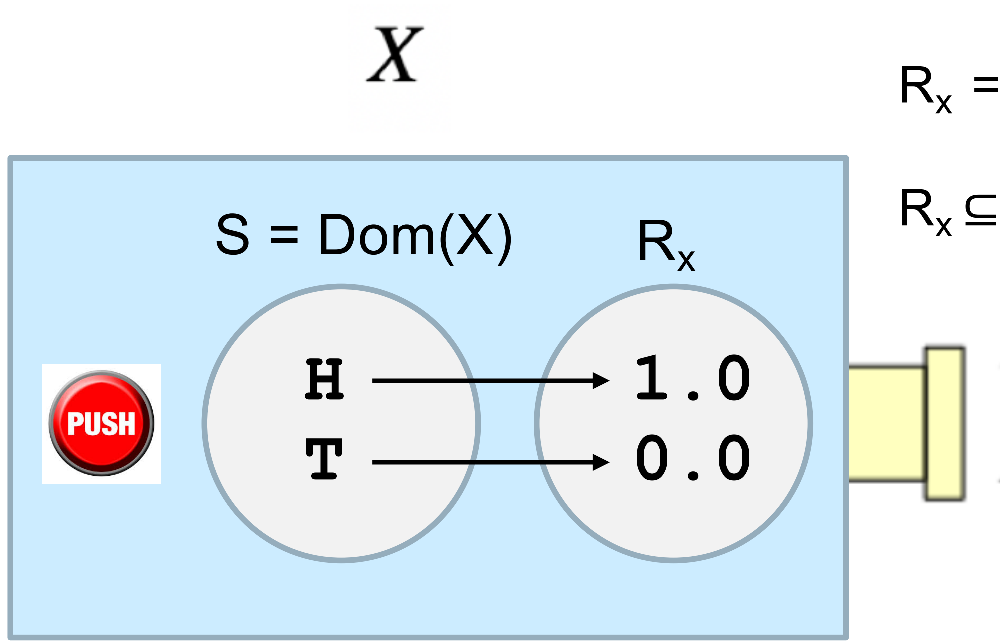
$$X : S \to \mathcal{R}$$

Now when an outcome is requested, the sample point is translated into a real number:

S = Domain(X)

$R_x =_{def} \text{Range}(X)$

$R_x \subseteq \mathcal{R}$



Random Variables

In order to formalize this notion, the notion of a Random Variable has been developed. A Random Variable X is a function from a sample space S into the reals:

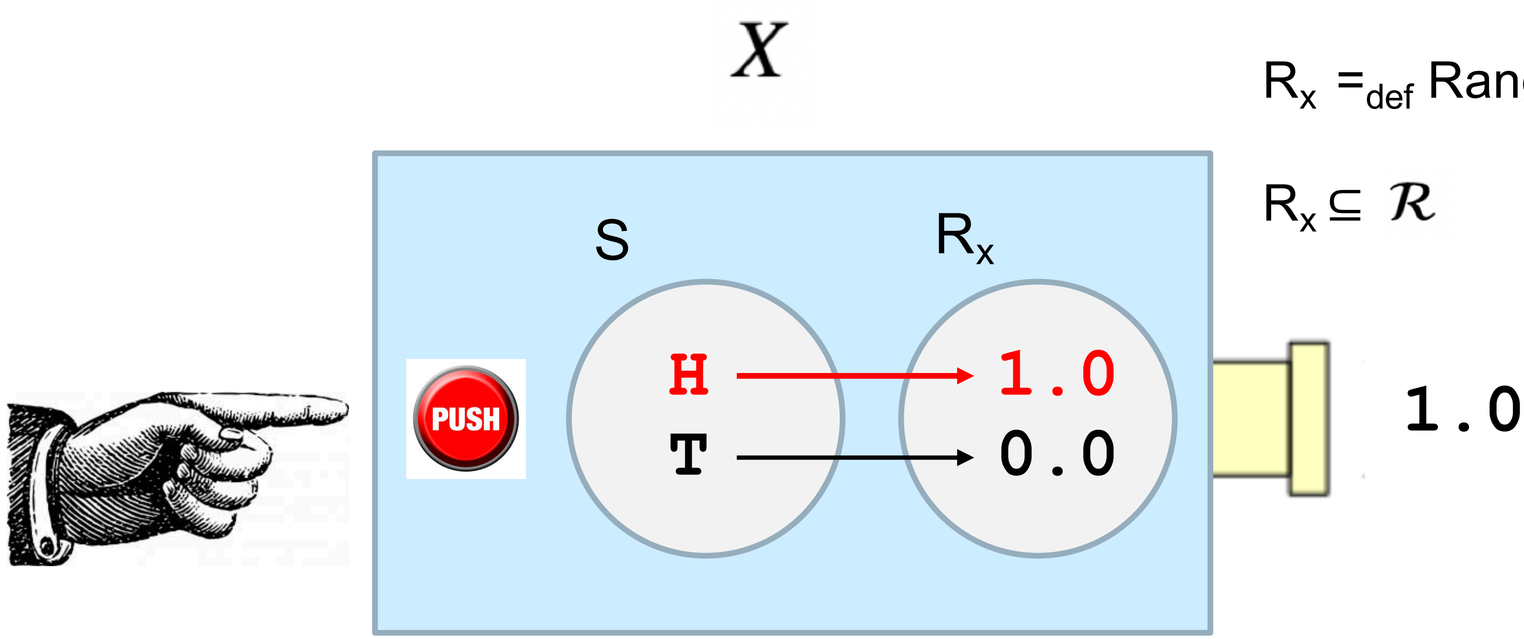
$$X : S \to \mathcal{R}$$

Now when an outcome is requested, the sample point is translated into a real number:

S = Domain(X)

$R_x =_{def} \text{Range}(X)$

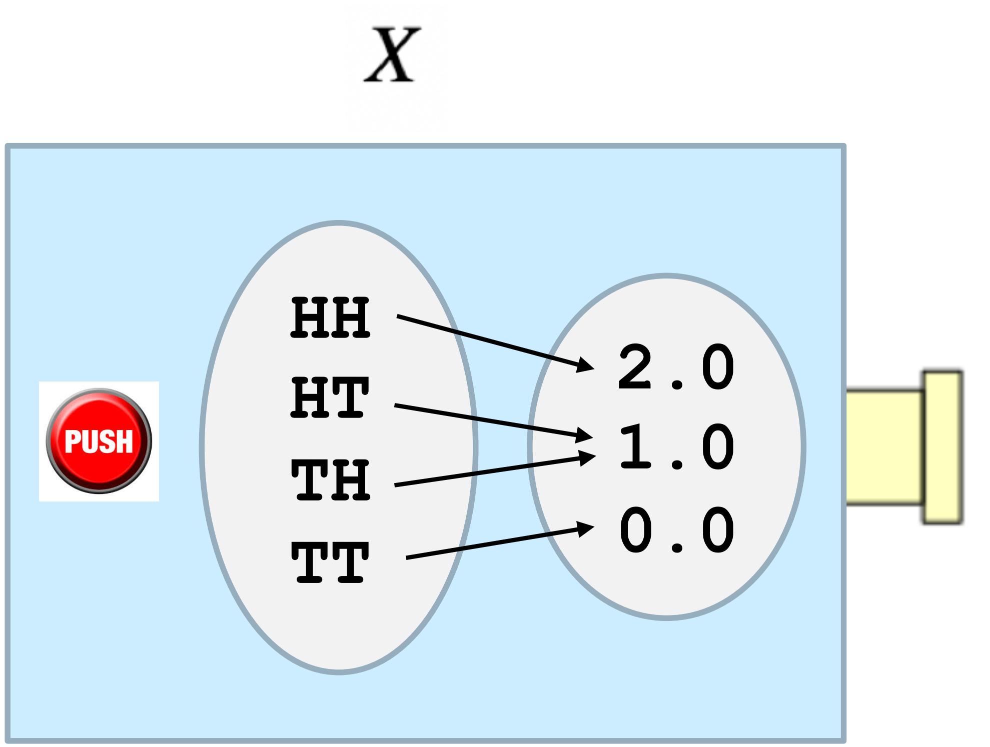
$R_x \subseteq \mathcal{R}$



Random Variables

This may seem awkward, but it helps to explain the difference between random experiments whose literal outcomes are not numbers, but which are translated into numbers for clarity.

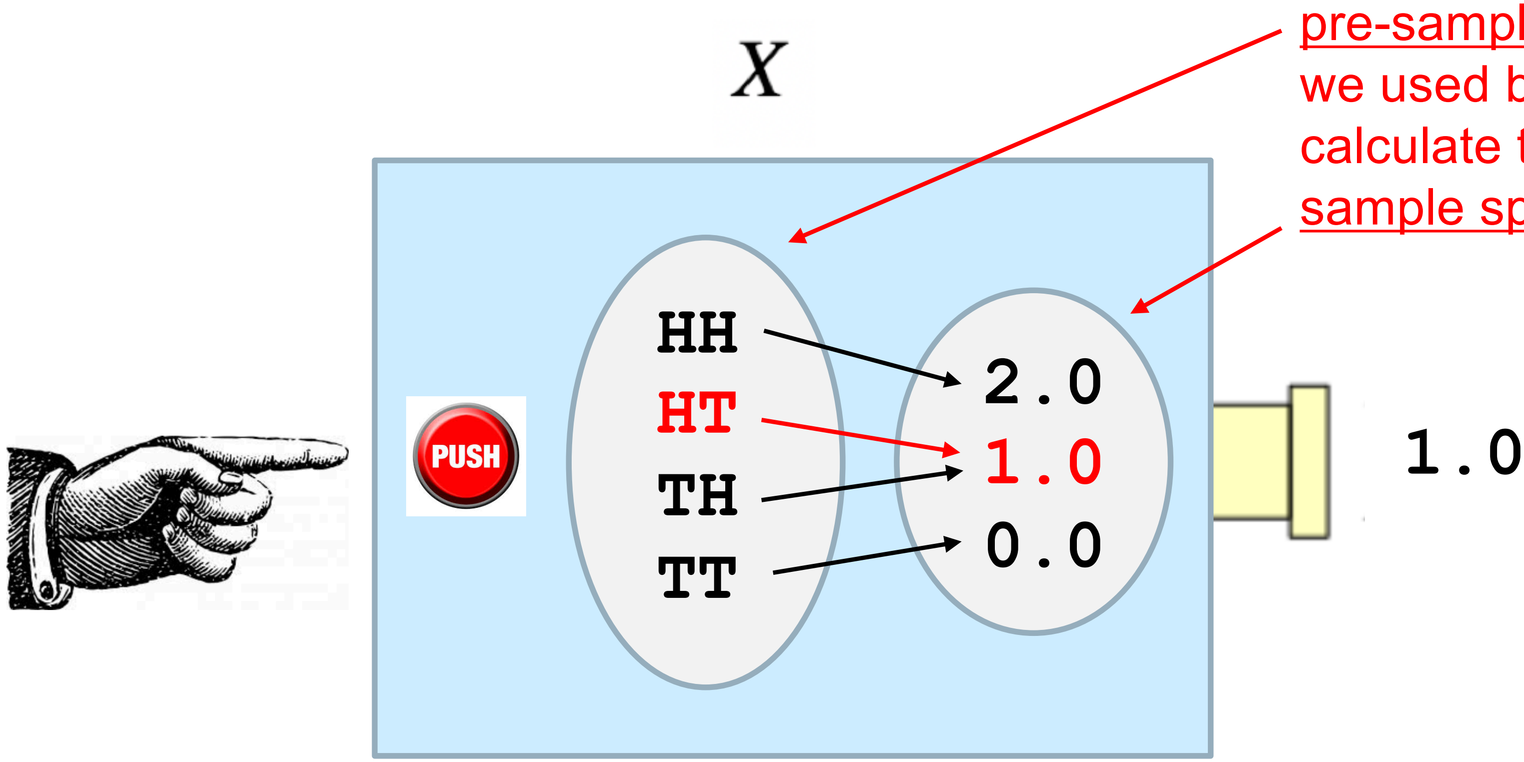
Example: X = "the number of heads which appear when two fair coins are flipped."



Random Variables

This may seem awkward, but it helps to explain the difference between random experiments whose literal outcomes are not numbers, but which are translated into numbers for clarity.

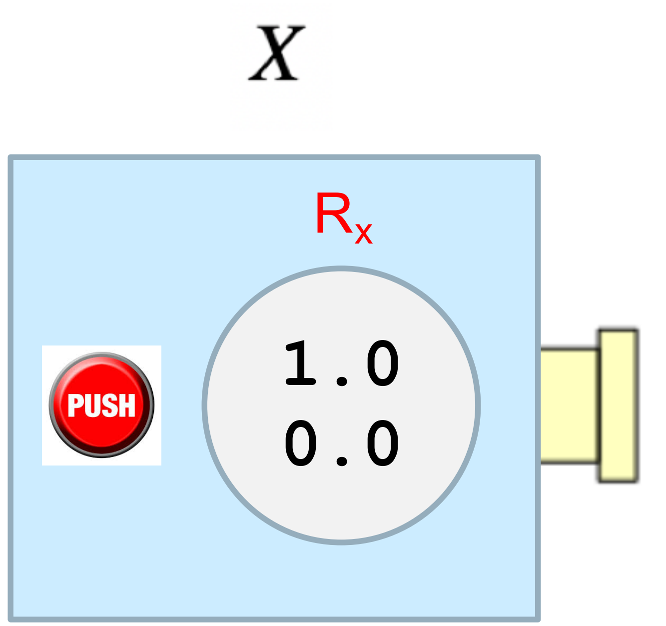
Example: X = “the number of heads which appear when two fair coins are flipped.



This is like the <u>pre-sample space</u> we used before to calculate the actual <u>sample space</u>.

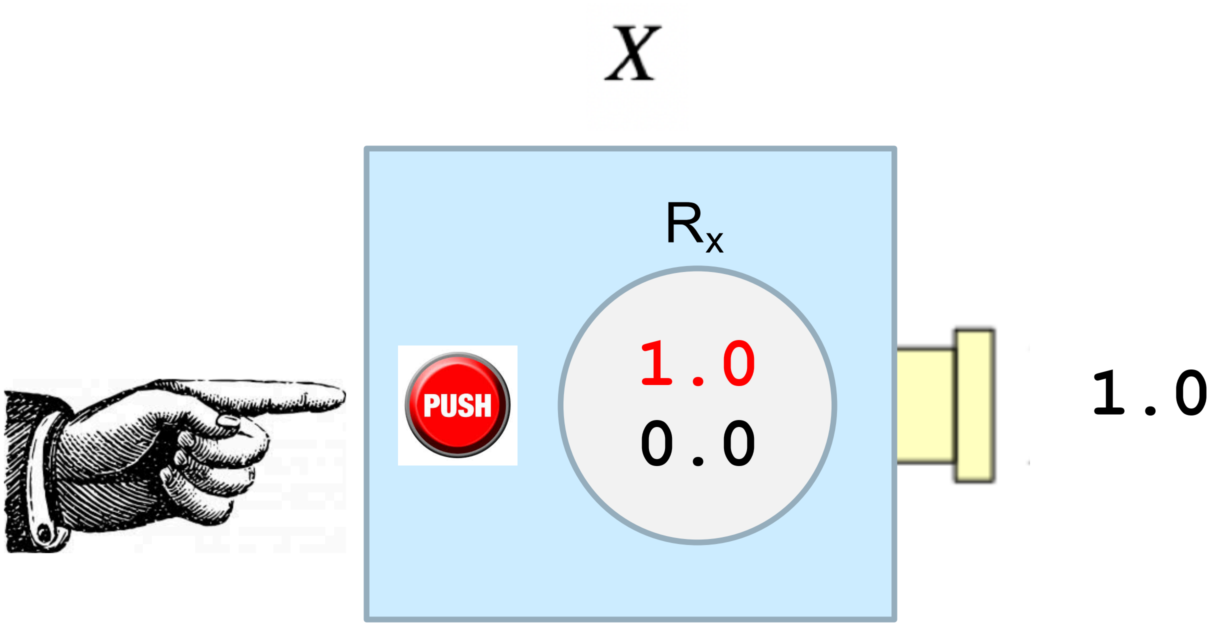
Random Variables

In general, in this class we will call the possible outputs R_x, since this the symbol used in your textbook, although you could just think of it as the sample space from which the outputs are drawn.



Random Variables

In general, in this class we will call the possible outputs R_x, since this the symbol used in your textbook, although you could just think of it as the sample space from which the outputs are drawn.



Discrete vs Continuous Random Variables

A random variable X is called discrete if R_x is finite or countably infinite:

Example of finite random variable:

X = "the number of dots showing after rolling two dice"

$R_X = \{ 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 \}$

Example of countably infinite random variable:

Y = "the number of flips of a coin until a head appears"

$R_Y = \{ 1, 2, 3, ... \}$

A random variable is called continuous if Rx is uncountable. Example:

Z = "the distance of a thrown dart from the center of a circular target of 1 meter radius"

$R_Z = [0.0 .. 1.0)$

Discrete vs Continuous Random Variables

A random variable X is called discrete if R_x is finite or countably infinite:

Example of finite random variable:

X = "the number of dots showing after rolling two dice"

R_X = { 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 }

Example of countably infinite random variable:

Y = "the number of flips of a coin until a head appears"

R_Y = { 1, 2, 3, ... }

For several weeks we will only consider discrete random variables!

A random variable is called continuous if Rx is uncountable. Example:

Z = "the distance of a thrown dart from the center of a circular target of 1 meter radius"

R_Z = [0.0 .. 1.0)

Discrete Random Variables: Probability Distributions

To specify a random variable precisely, you simply need to give the range R_X and the PMF f:

Examples:

X = "The number of dots showing on a thrown die"

$$R_X = \{1, 2, 3, 4, 5, 6\}$$

$$f_X = \{ \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6} \}$$

For simplicity, we simply list the values in Range(f_X) corresponding to the listing of R_X.

Y = "The number of tosses of a fair coin until a head appears"

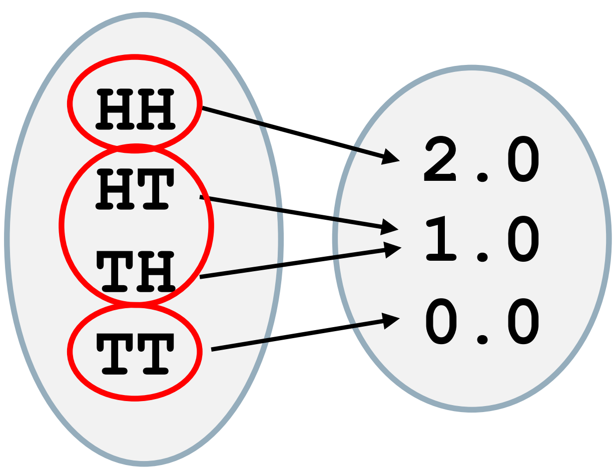
$$R_Y = \{ 1, 2, 3, \ \ldots \}$$

$$f_Y = \{ \tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \ \ldots \}$$

Discrete RandomVariables: Probability Distributions

How does this relate to our first definition of a probability space, events, probability function, etc., etc. ??

Probability Space	Random Variable X
Sample Space	R_X
Event	Subset of real numbers (with some restrictions).
Probability Function	Probability Distribution f_X



For continuous random variables there are additional conditions about events having to be the countable sum of intervals on the real number line.

Discrete Random Variables: Probability Distributions

The Probability Mass Function (PMF) of a discrete random variable X is a function from the range of X into $\mathcal{R}$:

$$f_X : R_X \to [0..1]$$

such that (i) $\forall a \in R_x \quad f_X(a) \geq 0$

(ii) $$\sum_{a \in R_x} f_X(a) = 1.0$$

If there is no possibility of confusion we will write f instead of f_X.

Discrete Random Variables: Probability Distributions

We will emphasize (starting next week) the distributions of random variables, using graphical representations (as in HW 01) to help our intuitions.

Example:

Y = "The number of tosses of a fair coin until a head appears"

$$R_Y = \{ 1, 2, 3, \ \ldots \ \}$$

$$f_Y = \{ \tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \ \ldots \}$$