

Station 1 had 2 pumps

If this experiment is performed and results  $\omega = (2, 3)$ , then

$$X = 2 + 3 = 5$$

$$Y = 2 - 2 = -1$$

$$U = 3$$

2 had 3 pumps

**Example 24.** When a student calls a university help desk for technical support, he/she will either immediately be able to speak to someone ( $S$ , for success) or will be placed on hold ( $F$ , for failure). With  $S = \{S, F\}$ , define an random variable  $X$  by

$$X(S) = 1 \text{ and } X(F) = 0$$

The random variable  $X$  indicates \_\_\_\_\_

**Example 25.** Suppose a location in the United States is selected. Define the random variable  $Y$  by

$Y =$  the height above sea level at the selected location

Then the largest possible value of  $Y$  is 14,494 (Mt. Whitney), and the smallest possible value is 2282 (Death Valley). The set of all possible values of  $Y$  is the set of all numbers in the interval between 2282 and 14,494, that is,

infinitely  $[2282, 14494]$

and there are uncountably many values in this interval.

## Two Types of Random Variables

- Discrete. If the possible outcomes of a random variable can be listed out using a finite (or countably infinite) set of single numbers (Example 22, 23, 24), then the random variable is discrete.
- continuous. If the possible outcomes of a random variable can only be described using an interval or union of intervals of real numbers (Example 25), then the random variable is continuous.

## 3.2 Probability Distributions for Discrete Random Variables

### 3.2.1 The Probability Mass Function

The **probability distribution** of  $X$  says how the total probability of 1 is distributed among the various possible  $X$  values. The probability distribution of  $X$  lists all possible values of  $X$  and their corresponding probabilities.

**Definition 9.** For discrete random variables, the probability list of  $X$  is called probability mass function (pmf), which is defined for every number  $x$  by  $p(x) = P(X = x) = P(\text{all } \omega \in S : X(\omega) = x)$ .

The pmf returns the probability that the random variable  $X$  is equal to the value  $x$ .

To be a valid pmf, we need:

- (1)  $0 \leq p(x) \leq 1$  for all  $x$
- (2)  $\sum_x p(x) = 1$

**Example 26.** Suppose we toss a fair coin three times, and define the random variable  $X$  to be the number of heads that appear. Find the pmf of  $X$ .

*Solution.* The sample space is

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

So the possible values for the random variable  $X$  are in the set  $\{0, 1, 2, 3\}$ . The pmf tells us all possible values of  $X$  and their corresponding probabilities, i.e.  $p(x) = P(X = x)$ . Since we have a fair coin, so the \_\_\_\_\_ is assumed here. Therefore

$x$	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

**NOTE:** This is NOT a proper format of writing a pmf. Write it in a proper way,  $X$  should define on \_\_\_\_\_. So The pmf of  $X$  is given by

$$p(x) = \begin{cases} \frac{1}{8} & \text{if } x=0 \text{ or } 3 \\ \frac{3}{8} & \text{if } x=1 \text{ or } 2 \\ 0 & \text{o/w} \end{cases}$$

**Question:** Is this a valid pmf?

Check: values b/w 0 and 1 and  $\sum = 1$  ✓

The graph of the pmf of  $X$  could be



Event  $A = \bigcup_{x \in A} \{x\}$  ← countably many →  $P(A) = \sum_{x \in A} p(x)$

**Example 27.** (Exercise #13 on page 107 of the textbook) A mail-order computer business has six telephone lines. Let  $X$  denote the number of lines in use at a specified time. Suppose the pmf of  $X$  is as given in the accompanying table.

$x$	0	1	2	3	4	5	6
$p(x)$	.10	.15	.20	.25	.20	.06	.04

Check:  
 $0 \leq p(x) \leq 1$  ✓  
 $\sum = 1$  ✓

Calculate the probability of each of the following events.

- a. {at most three lines are in use} =  $\{0, 1, 2, 3\} : 0.1 + 0.15 + 0.2 + 0.25 = 0.70$
- b. {fewer than three lines are in use} =  $\{0, 1, 2\} : 0.45$
- c. {at least three lines are in use} =  $\{3, 4, 5, 6\} = \{0, 1, 2\}^c : 1 - 0.45$
- d. {between two and five lines, inclusive, are in use} =  $\{2, 3, 4, 5\} : \dots$
- e. {between two and four lines, inclusive, are not in use} =  $\{2, 3, 4\} : \dots$
- f. {at least four lines are not in use} =  $\{0, 1, 2\} : 0.45$

*Solution.* Before we calculate the probabilities, let's check whether it is a valid pmf.

**Example 28.** (Example 3.8 on textbook page 100) Six boxes of components are ready to be shipped by a certain supplier. The number of defective components in each box is as follows:

Box	1	2	3	4	5	6
Number of defectives	0	2	0	1	2	0

One of these boxes is to be randomly selected for shipment to a particular customer. Let  $X$  be the number of defectives in the selected box.

$x$	0	1	2
$p(x)$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$

check  $\sum = 1$  ✓  
 $0 \leq \leq 1$  ✓

Lec 10

### 3.2.2 The Cumulative Distribution Function

**Definition 10.** The cumulative distribution function (cdf)