

## Outline

1. Recap of probability
2. Simulation
3. Random Variables

## Independent vs. Disjoint Events

- Independent events **can** happen at the same time, but knowing that event “A” occurred **does not** change  $P(B)$  and vice versa
- Disjoint events cannot happen at the same time.
  - Knowing that event “A” occurred changed  $P(B) = 0$  and vice versa
- If A and B are independent,  $P(A \cap B) = P(A) P(B)$
- If A and B are disjoint,  $P(A \cap B) = 0$

### Example One

- Draw a tile from a bag of 100 scrabble tiles
- Event C = “the tile is a C”
- Event A = “the tile is an A”

$$P(C) = .02$$

$$P(A) = .09$$

Events “C” and “A” are disjoint

Events “C” and “A” are not independent

### Example Two

Draw one tile and set it outside

Event C = “first tile is a C”

Event A = “second tile is a A”

Event C and A are not disjoint

Events “C” and “A” are not independent

**Sampling without replacement**

### Example Three

Draw a tile, put it back in the bag and then draw another tile

Event C = “first tile is a C” Event A = “second tile is an A”

Event C and A are not disjoint Event C and A are independent

**Sampling with replacement**

## Simulation

Trying to imitate in the real world where the outcome is uncertain but is random

- Specify our model for an uncertain situation/random event
- “Randomly” generate an outcome for the model
- Repeat step two many, many times

### Why simulate?

- Once we set up the model, the math maybe too difficult
- Situation may be unique, or we only have ability to observe it once, due to physical/financial limitations
- For fun and/or profit

Report assumptions of the model!

## Random Variables (RVs)

Random variable is a variable whose numerical values describe outcomes of a random event

Typically we map outcomes in our sample space denoted as “S” to numerical values of the random variable.

Discrete Random Variable : probability mass function (PMF) places positive probability at specific numbers on the number line

- Only specific numbers
- Example: all outcomes are real, positive numbers

Continuous Random Variable : probability density function (PDF)

- Places positive probability along a possibly infinite interval of the number line.

## Writing the PMF of a Discrete Random Variable

Each unique key value  $X=x$  is mapped to a non unique value  $P(X=x)$

```
example_hash_map = {  
    key: value  
}
```

A hash table is another way to represent data mapping.

- value represents a random variable
- key represents a “realization” of value

### Example One

We can find  $P(Y=0)$

Once we have observed the random event either  $y = 0$  or  $y \neq 0$

Let  $X$  = the point value of the chosen tile

```
map = {  
    0: 0.02,  
    1: 0.68,  
    2: 0.07,  
    3: 0.08,  
    4: 0.10,  
    5: 0.01,  
    8: 0.02,  
    10: 0.02  
}
```

Sum of values in map == 0

## Example Two

Use PMF & probability rules to find:

- $P(X \leq 3)$ 
  - $P(X = 0, 1, 2 \text{ or } 3)$
  - $P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$
  - $= 0.85$
- $P(X > 1)$ 
  - $P(X = 2, 3, 4, 5, 8, \text{ or } 10)$
  - $1 - P(X \leq 1) = 1 - (X = 0 \text{ or } 1)$
  - $1 - [P(X = 0) + P(X = 1)] = 1 - [0.02 + 0.68] = 1 - .7 = 0.3$
- $P(X > 5)$ 
  - $P(X = \{0..5\})$
  - $0.04$
- $P(3 < x \leq 5)$ 
  - $0.11$
  - $P(X = 4, \text{ or } 5)$

## Expected Value (Mean) of a Random Variable

- Called expectation, mean, all the same thing
- On average, what value do we expect the random variable to be

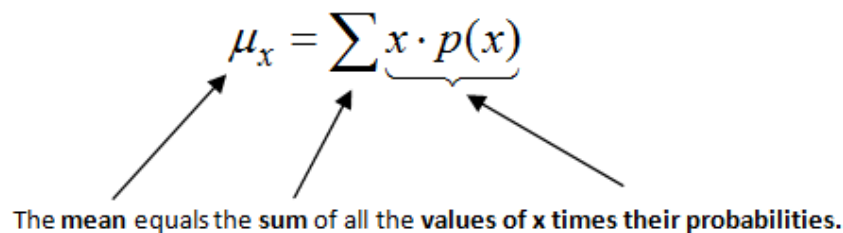
Recall idea of “weighted average”

## Mean of a Probability Distribution

$\mu$

Denotes the average of all events, which in turn gives us an expected value for a given function.

## Summation notation



The diagram shows the formula  $\mu_x = \sum x \cdot p(x)$ . Three arrows point from the text below to parts of the formula: one to  $\mu_x$ , one to the summation symbol  $\sum$ , and one to the term  $x \cdot p(x)$  which is underlined. The text below reads: "The mean equals the sum of all the values of x times their probabilities."

$$\mu_x = \sum x \cdot p(x)$$

The mean equals the sum of all the values of x times their probabilities.

Figure 1: discrete random variable formula

Expected value is a linear operator (can take in sum and give back a result in the form of a sum of the applied operators)

## For random variables X and Y, and constant C

- $E[X+Y] = E[X] + E[Y]$
- $E[cX] = cE[X]$
- ^ where “c” is a constant applied

This implies, for X, Y and arbitrary constants a,b  $E[aX + bY] = aE[X] + bE[Y]$

Consequences: a = 1 , b = -1

$$E[X - Y] = E[X] - E[Y]$$

$$\mu_{x-y} = \mu_x - \mu_y$$