Outline

• Review of Conditional Probability
• Bayes’ Rule
• Random Variables
Definition of Conditional Probability

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

Given \( f(x) = \frac{1}{15} \) for all \( x \).
What is \( P(A \mid B) \)?
Examples of Conditional Probability

• A number between 1 and 10 is selected uniformly at random
• What is the probability that it is odd given that it is prime?

• What is the probability that a 5-card poker hand is a straight given that it is a flush?
Practice Problem

• A couple has two children
• What is the probability that the couple has two girls given that one of them is a girl?
Conditional Probability and Conditional Independence

• Recall our definition of the independence of two events A and B. We say A and B are independent events if and only if

\[ P(A \cap B) = P(A)P(B) \]

• Two events A and B are conditionally independent given a third event C if and only if

\[ P(A \cap B|C) = P(A|C)P(B|C) \]

• Equivalently:

\[ P(A|B \cap C) = P(A|C) \]
Conditional Independence Practice Problem

• We draw a 5-card hand from 52 cards
• We define the following events
  – $E_1$: The hand is a flush
  – $E_2$: The hand is a royal flush (10 through Ace same suit)
  – $E_3$: The hand is a straight
• Which if any of the following are true:
  – $E_1$ is conditionally independent of $E_3$ given $E_2$
  – $E_1$ is conditionally independent of $E_2$ given $E_3$
  – $E_2$ is conditionally independent of $E_3$ given $E_1$
Workspace
Workspace
Probabilities and Decision Trees (simplified version of Yahtzee)

• A person rolls three dice
• If all three dice are the same value the game stops and the player wins
• Otherwise the player chooses the number that appear most frequently on the rolled dice (if more than 1 value appears most frequently 1 value is selected arbitrarily) and sets all the dice with that value aside and rerolls the other dice
• After this second roll if all three dice are the same value the player wins
• What is the probability of the player winning?
Workspace
Workspace
Bayes’ Rule

• Motivation
  – Suppose we have two events A and B
  – Sometimes we wish to compute \( P(B|A) \) but it is much easier to compute \( P(A|B) \)
  – Bayes’ Rule allows us to write one in terms of the other

Thomas “The Rev” Bayes
Example from Last Time

• Recall my example last time of building a system to recognize whether someone is smiling given an image of their face

• We wish to compute $P(\text{smile} | \text{pixels})$ and $P(\text{nosmile} | \text{pixels})$ and we will predict whichever is highest

However, how do we calculate $P(\text{smile} | \text{pixels})$?
Solution

• Instead of computing $P(\text{smile} \mid \text{pixels})$ we computed $P(\text{pixels} \mid \text{smile})$

• This was very natural to compute since we are given a bunch of labeled images of smiles and nonsmiles
Solution Continued

• Bayes’ Rule allows us to use $P(\text{pixels} \mid \text{smile})$ and $P(\text{pixels} \mid \text{nosmile})$ to compute $P(\text{smile} \mid \text{pixels})$

• Recall our definition of conditional probability

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

• Doing some basic algebra:

$$P(A \mid B)P(B) = P(A \cap B)$$
Derivation of Bayes’

• From the definition of conditional probability we also have:
  \[ P(B|A) = \frac{P(A \cap B)}{P(A)} \]

• Doing some more algebra
  \[ P(B|A)P(A) = P(A \cap B) \]

• Now we equate our two expressions for \( P(A \text{ and } B) \)
  \[ P(A|B)P(B) = P(B|A)P(A) \]
Derivation of Bayes’

• More simple algebra yields the typical form of Bayes’ rule

\[ P(B|A) = \frac{P(A|B)P(B)}{P(A)} \]
Back to our Example of Expression Recognition

\[ P(smile|pixels) = \frac{P(pixels|smile)P(smile)}{P(pixels)} \]

\[ P(nosmile|pixels) = \frac{P(pixels|nosmile)P(nosmile)}{P(pixels)} \]
Applying Bayes’ Rule

• Acme Grommet company has a quality control test for their product
• $P(\text{defective}) = 0.001$
• $P(\text{positive test} \mid \text{defective}) = 0.9$
• $P(\text{positive test} \mid \text{not defective}) = 0.05$
• What is $P(\text{defective} \mid \text{positive test})$?
• Is this a good test?
Workspace
Random Variables

• In order to define a random variable we combine what we have learned of probability theory with functions

• Recall that a probability space is defined by a sample set U and a probability function f

• A random variable is a function from U to the real numbers
Examples of Random Variables

• Sample space is the set of all 5 card hands from the standard deck
• The probability function is uniform over all hands (i.e. \( f(x) = \frac{1}{\binom{52}{5}} \))
• We define the random variable X to be the number of Queens in the 5-card hand
What does this function look like?
Questions

• What is the domain of the random variable?
• What is the range of the random variable?
Practice Defining a Random Variable

- A Fair Coin is Flipped twice
- The sample space is \( \{H, T\}^2 \)
- Define a random variable \( X \) in two line form that represents the number of heads that occurred
Distribution of a Random Variable

• The distribution function of a random variable capture how likely that random variable is to take on a particular value

• Given a sample space U, a probability function f, and a random variable X we define a new function \( f_X : \mathbb{R} \rightarrow \mathbb{R} \) that defines how probable it is for the random variable to take on any particular value.
Distribution of a Random Variable

\[ f_X(y) = \sum_{t \in U \text{ and } X(t) = y} f(x) \]

• We can think of this as just looping over all values in the sample space and adding up the probabilities of each the elementary events that the function X maps to the real number y
Example of a Distribution of a Random Variable

• Consider the sample space \( \{H, T\}^3 \) with \( f(x) = 1/8 \) (i.e. all elementary events are equally likely)

• We now define the random variable \( X \) to be the number of H’s in an elementary event
Distribution of random variable $X$

$$f_X(0) = \left(\frac{1}{2}\right)^3 \binom{3}{0}$$

$$f_X(1) = \left(\frac{1}{2}\right)^3 \binom{3}{1}$$

$$f_X(2) = \left(\frac{1}{2}\right)^3 \binom{3}{2}$$

$$f_X(3) = \left(\frac{1}{2}\right)^3 \binom{3}{3}$$

$$f_X(n) = 0 \quad \forall n \notin \{0, 1, 2, 3\}$$
Some Helpful Notation for Defining a Function

• Here is another notation we can use to define a function

\[ f_X(y) = \begin{cases} 
(\frac{1}{2})^3 \binom{3}{0}, & y = 0 \\
(\frac{1}{2})^3 \binom{3}{1}, & y = 1 \\
(\frac{1}{2})^3 \binom{3}{2}, & y = 2 \\
(\frac{1}{2})^3 \binom{3}{3}, & y = 3 \\
0, & \text{else}
\end{cases} \]
Sample problem

• What is the distribution function of the random variable $X$ that represents the number of queens in a hand of 5 cards?
Workspace
Expectation of a Random Variable

• The expectation of a random variable is defined as:

\[ E[X] = \sum_{i \in U} X(i)f(x) \]

- Sum over all elements of the sample space
- Value of the random variable at that elementary event
- Probability of that elementary event
Question

• We toss a biased coin 5 times (probability of heads is .8)
• What is the expected number of heads in the sequence of flips?
Workspace
Connection Between the Probability of an Event and Random Variables

• For a probability space \((U,f)\) we are given an event \(E\).

• We can define a random variable \(X\) in the following way

\[
X(i) = \begin{cases} 
1, & i \in E \\
0, & \text{else}
\end{cases}
\]
Problem

• Show that: \( \mathbb{E}[X] = P(E) \)
A Probability Paradox (not part of the class officially)

• Someone gives you two sealed envelopes which each contain some amount of money.
• You are told that one of the envelopes contains $x$ dollars and the other contains $2x$ dollars
• You are allowed to view how much money is contained in one of the envelopes. You are then given an opportunity to switch. Should you?