Lesson 2. Probability Review

1 Random variables

- A **random variable** is a variable that takes on its values by chance
  - One perspective: a random variable represents unknown future results
- Notation convention:
  - Uppercase letters (e.g. X, Y, Z) to denote random variables
  - Lowercase letters (e.g. x, y, z) to denote real numbers
- $\{X \leq x\}$ is the event that the random variable X is less than or equal to the real number x
- The probability this event occurs is written as $\Pr\{X \leq x\}$

2 Discrete random variables

- A random variable is **discrete** if it can take on only a finite or countably infinite number of values
- Let X be a discrete random variable that takes on values $a_1, a_2, \ldots$
- The **probability mass function (pmf)** p_X of X is:

Example 1. From ESPN/AP on July 7, 2016:

> The Los Angeles Lakers have swooped in ahead of the Brooklyn Nets to acquire Calderon, according to league sources Wednesday. Calderon and two future second-round picks are headed to Los Angeles in exchange for a player to be named later, the sources said.

Let X be a random variable that represents the player to be named later (using integer ID numbers 1, 2, 3, 4 instead of four names). Suppose the pmf for X is:

$$p_X(a) = \begin{cases} 0.1 & \text{if } a = 1, \\ 0.3 & \text{if } a = 2, \\ 0.5 & \text{if } a = 3, \\ 0.1 & \text{if } a = 4, \\ 0 & \text{otherwise} \end{cases}$$

Why is X a discrete random variable? What is $\Pr\{X = 2\}$? What is $\Pr\{1 < X \leq 3\}$?

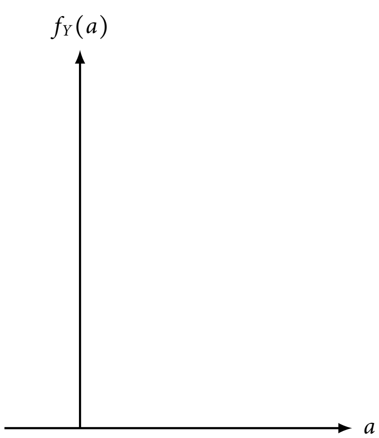
3 Continuous random variables

- A random variable is **continuous** if it can take on a continuum of values
- Let X be a continuous random variable
- The **probability density function (pdf)** f_X of X gives us the relative likelihood of the values of X
- We can use the pdf f_X of X to compute the probability that X lies within some interval:

Example 2. Let Y be a random variable that represents the time to deliver a pizza, in hours. In particular, suppose Y is a **exponentially distributed** random variable with parameter $\lambda = 2$. The pdf of Y is

$$f_Y(a) = \begin{cases} 0 & \text{if } a < 0, \\ \lambda e^{-\lambda a} & \text{if } a \geq 0 \end{cases} \qquad \text{where } \lambda = 2.$$

Plot the pdf of Y. Why is Y a continuous random variable? Which values are more or less likely? What is $\Pr\{1 < Y \leq 3\}$? What is $\Pr\{Y = 2\}$?



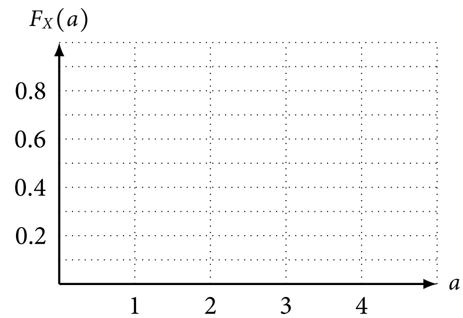
4 The cumulative distribution function

- The **cumulative distribution function (cdf)** F_X of random variable X is:
- We can use the cdf of X to find the probability that X is between a and b:

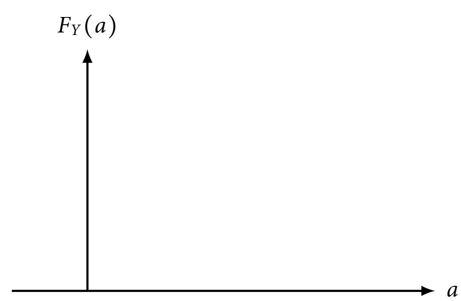
- X is a discrete random variable taking values $a_1, a_2, \dots$ with pmf p_X
- The cdf and pmf of X are related:

- X is a continuous random variable with pdf f_X
- The cdf and pdf of X are related:

Example 3. What is the cdf of X defined in Example 1? Plot the cdf.



Example 4. What is the cdf of Y defined in Example 2? Plot the cdf.



- Some properties of a generic cdf $F(a)$:
  - Domain:
  - Range:
  - As a increases, $F(a)$
  - F is **right-continuous**: if F has a discontinuity at a, then $F(a)$ is determined by the piece of the function on the right-hand side of the discontinuity

- In general, the cdf of a discrete random variable is

- In general, the cdf of a continuous random variable is

5 Expected value and variance

- The **expected value** of a random variable is its weighted average

- X is a discrete random variable taking values $a_1, a_2, \ldots$ with pmf p_X
- The expected value of X is
- The expected value of a function g of X is

- X is a continuous random variable with pdf f_X
- The expected value of X is
- The expected value of a function g of X is

- The **variance** of X is

Example 5. Find the expected value and the variance of X, as defined in Example 1.

- Some useful properties: let X, Y be random variables, and a, b be constants
  - $E[aX + b] = aE[X] + b$
  - $E[X + Y] = E[X] + E[Y]$
  - $\text{Var}(a + bX) = b^2\text{Var}(X)$
  - In general, $E[g(X)] \neq g(E[X])$

6 Exercises

Problem 1. Suppose X is a discrete random variable with the following cdf:

$$F_X(a) = \begin{cases} 0 & \text{if } a < 2, \\ 0.4 & \text{if } 2 \leq a < 4, \\ 0.9 & \text{if } 4 \leq a < 5, \\ 1 & \text{if } a \geq 5. \end{cases}$$

a. What is the pmf of X?
b. What is the expected value of X?
c. What is the variance of X?
d. Professor I. M. Wright peeks over your shoulder and declares,

> "The probability that $X = 3$ is 0.4, since $F_X(3) = 0.4$."

Is Professor Wright correct? Briefly explain.

Problem 2. Suppose X is a random variable with pdf

$$f_X(a) = \begin{cases} 0 & \text{if } a < 1, \\ \frac{1}{4}a - \frac{1}{4} & \text{if } 1 \leq a < 3, \\ \frac{1}{2} & \text{if } 3 \leq a < 4, \\ 0 & \text{if } a \geq 4. \end{cases}$$

a. What is the probability that $2 < X \leq 3$?
b. What is the expected value of X?
c. What is the probability that $X \leq 6$?
d. Professor I. M. Wright peeks over your shoulder and declares,

> "Since the maximum value of $f_X(a)$ is attained when $a = 3$, the maximum value that X can take is 3."

Is Professor Wright correct? Briefly explain.