SA402 – Dynamic and Stochastic Models Fall 2016
Assoc. Prof. Nelson Uhan

Lesson 2. Probability Review

Course standards covered in this lesson: B1 – Computing with probability distributions, B2 – Interpreting probability distributions.

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 $\Pr\{X \leq x\}$
- A random variable X is characterized by its **probability distribution**, which can be described by its **cumulative distribution function (cdf)**, F_X:

- We can use the cdf of X to find the probability that X is between a and b:

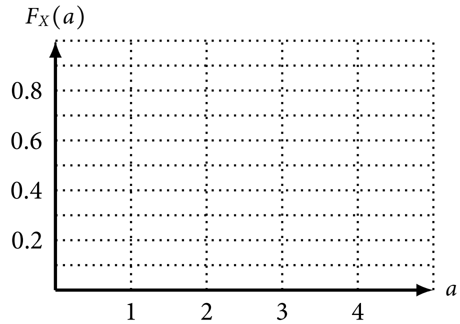
Example 1. From ESPN/AP on July 7, 2016:

> 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 cdf for X is:

$$F_X(a) = \begin{cases} 0 & \text{if } a < 1, \\ 0.1 & \text{if } 1 \leq a < 2, \\ 0.4 & \text{if } 2 \leq a < 3, \\ 0.9 & \text{if } 3 \leq a < 4, \\ 1 & \text{if } a \geq 4 \end{cases}$$

Plot the cdf for X. What is $\Pr\{X \leq 3\}$? What is $\Pr\{X = 2\}$?



- 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
- 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 $a_1, a_2, \dots$
  - The **probability mass function (pmf)** p_X of X is:
  - The pmf and cdf of a discrete random variable are related:

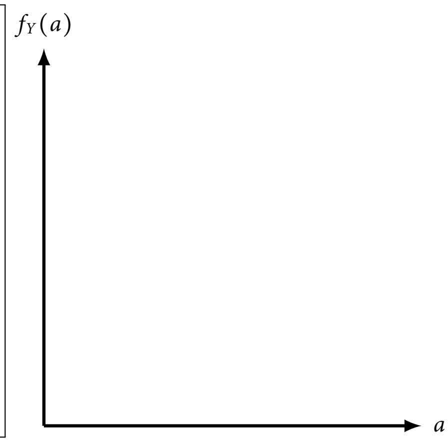
Example 2. Find the pmf of X defined in Example 1.

- A random variable is **continuous** if it can take on a continuum of values
  - The **probability density function (pdf)** f_X of a continuous random variable X is:
  - We can get the cdf from the pdf:
  - Therefore, the probability that X is between a and b can be computed as:

Example 3. Let Y be a **exponentially distributed** random variable with parameter λ. In other words, the cdf of Y is

$$F_Y(a) = \begin{cases} 0 & \text{if } a < 0, \\ 1 - e^{-\lambda a} & \text{if } a \geq 0. \end{cases}$$

a. Find the pdf of Y.
b. Let $\lambda = 2$. Plot the pdf of Y.
c. For this random variable, which values are more likely or less likely?
d. What is $\Pr\{Y = 3\}$?



2 Expected value

- The **expected value** of a random variable is its weighted average
- If X is a discrete random variable taking values $a_1, a_2, \dots$ with pmf p_X, then the expected value of X is
- If X is a continuous random variable with pdf f_X, then the expected value of X is

- We can similarly take the expected value of a function g of a random variable X
  - When X is discrete:
  - When X is continuous:
- The **variance** of X is

Example 4. Find the expected value and the variance of X, as defined in Example 1.

Example 5. The indicator function $\mathcal{I}(\cdot)$ takes on the value 1 if its argument is true, and 0 otherwise. Let X be a discrete random variable that takes on values $a_1, a_2, \ldots$. Find $E[\mathcal{I}(X \leq b)]$.

- Example 5 works similarly for continuous random variables
- Probabilities can be expressed as the expected value of an indicator function
- 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, $E[g(X)] \neq g(E[X])$

3 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.