SA402 – Dynamic and Stochastic Models — Fall 2013

Asst. Prof. Nelson Uhan

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 $\Pr\{X \leq x\}$
- 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:

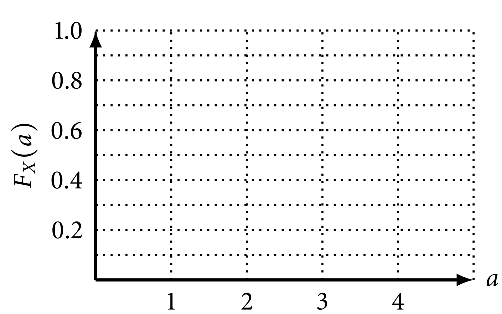
Example 1. From ESPN/AP on August 14, 2013:

> The Blue Jays traded utilityman Emilio Bonifacio to Kansas City for a player to be named later.

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 $\Pr\{X = 2\}$?



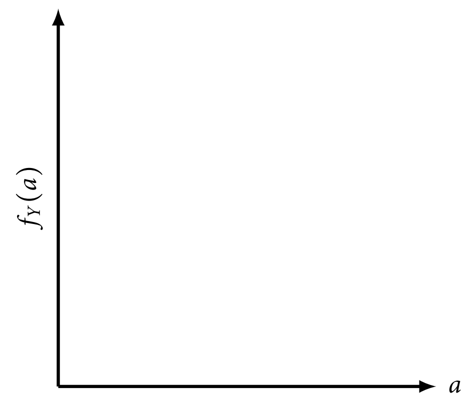
- Some properties of generic cdf $F(a)$:
  - Domain:
  - Range:
  - As a increases, $F(a)$
    - In other words, F is
  - F is **right-continuous**: if F has a discontinuity at a, then $F(a)$ is determined by the piece of the function on the right-hand side of the discontinuity
- A random variable is **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 $a_1, a_2, \ldots$
  - The **probability mass function (pmf)** p_X of X is:
  - The pmf and cdf of a discrete random variable are related:
- 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:

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

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
  - X is discrete:
  - 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 $a_1, a_2, \dots$. 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 Joint distributions and independence

- Let X and Y be random variables
- X and Y could be **dependent**
  - For example, X and Y are the times that the 1st and 2nd customers wait in the queue, respectively
- If we want to determine the probability of an event that depends both on X and Y, we need their **joint distribution**
- The **joint cdf** of X and Y is:

- Two random variables X and Y are **independent** if knowing the value of X does not change the probability distribution for the value of Y
- Mathematically speaking, X and Y are independent if

- Independence makes dealing with joint distributions easier!
- Idea generalizes to 3 or more variables

4 Exercises

Problem 1 (Nelson 3.1)**.** Calculate the following probabilities from the cdf given in Example 1. In each case, begin by writing the appropriate probability statement, for example $\Pr\{X = 4\}$; then calculate the probability.

a. The probability that player 4 will be traded.
b. The probability that player 2 will not be traded.
c. The probability that the index of the player traded will be less than 3.
d. The probability that the index of the player traded will be larger than 1.

Problem 2 (Nelson 3.2)**.** Let Y be an exponential random variable with parameter $\lambda = 2$ that models the time to deliver a pizza in hours. Calculate the following probabilities. In each case, begin by writing the appropriate probability statement, for example $\Pr\{Y \leq 1/2\}$; then calculate the probability.

a. The pizza-delivery company promises delivery within 40 minutes or the pizza is free. What is the probability that it will have to give a pizza away for free?
b. The probability that delivery takes longer than 1 hour.
c. The probability that delivery takes between 10 and 40 minutes.
d. The probability that delivery takes less than 5 minutes.

Problem 3 (Nelson 3.3)**.** A random variable X has the following cdf:

$$F_X(a) = \begin{cases} 0 & \text{if } a < 0, \\ \frac{a^2}{\delta^2} & \text{if } 0 \leq a \leq \delta, \\ 1 & \text{if } a > \delta \end{cases}$$

a. What is the density function of X?
b. What is the maximum possible value that X can take?
c. What is the expected value of X?

Problem 4 (Nelson 3.5)**.** A random variable Y has the following pdf:

$$f_Y(a) = \begin{cases} 0 & \text{if } a < 0, \\ \frac{3}{16}a^2 + \frac{1}{4} & \text{if } 0 \leq a \leq 2, \\ 0 & \text{if } a > 2 \end{cases}$$

a. What is the expected value of Y?
b. What is the cdf of Y?
c. For this random variable, which is more likely: a value near 1/2 or a value near 3/2?