ECS315 2012/1 Part III.1 Dr.Prapun

7 Random variables

In performing a chance experiment, one is often not interested in the particular outcome that occurs but in a specific numerical value associated with that outcome. In fact, for most applications, measurements and observations are expressed as numerical quantities.

Example 7.1. Take this course and observe your grades.

7.2. The advantage of working with numerical quantities is that we can perform mathematical operations on them.

In the mathematics of probability, averages are called expectations or expected values.

Definition 7.3. A real-valued function $X(\omega)$ defined for all points ω in a sample space Ω is called a ***random variable*** (r.v. or RV) [23].

- So, a random variable is a rule that assigns a numerical value to each possible outcome of a chance experiment.

- Intuitively, a random variable is a variable that takes on its values by chance.
- The convention is to use capital letters such as X, Y, Z to denote random variables.

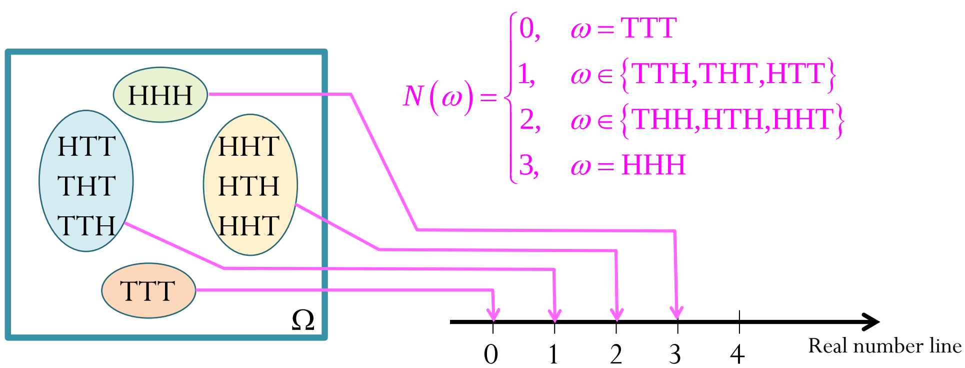
Example 7.4. Roll a fair dice: $\Omega = \{1, 2, 3, 4, 5, 6\}$.

[23] The term "random variable" is a misnomer. Technically, if you look at the definition carefully, a random variable is a deterministic function; that is, it is not random and it is not a variable. [Toby Berger][26, p 254]

- As a function, it is simply a rule that maps points/outcomes ω in Ω to real numbers.
- It is also a deterministic function; nothing is random about the mapping/assignment. The randomness in the observed values is due to the underlying randomness of the argument of the function X, namely the experiment outcomes ω.
- In other words, the randomness in the observed value of X is induced by the underlying random experiment, and hence we should be able to compute the probabilities of the observed values in terms of the probabilities of the underlying outcomes.

Example 7.5 (Three Coin Tosses)**.** Counting the number of heads in a sequence of three coin tosses.

$$\Omega = \{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}$$



Example 7.6 (Sum of Two Dice)**.** If S is the sum of the dots when rolling one fair dice twice, the random variable S assigns the numerical value $i+j$ to the outcome (i, j) of the chance experiment.

Example 7.7. Continue from Example 7.4,

(a) What is the probability that $X = 4$?

(b) What is the probability that $Y = 4$?

Definition 7.8. Shorthand Notation:

- $[X \in B] = \{\omega \in \Omega : X(\omega) \in B\}$
- $[a \le X < b] = [X \in [a, b)] = \{\omega \in \Omega : a \le X(\omega) < b\}$

- $[X > a] = \{\omega \in \Omega : X(\omega) > a\}$
- $[X = x] = \{\omega \in \Omega : X(\omega) = x\}$
    - ◦ We usually use the corresponding lowercase letter to denote
        (a) a possible value (realization) of the random variable
        (b) the value that the random variable takes on
        (c) the running values for the random variable

All of the above items are sets of outcomes. They are all events!

To avoid double use of brackets (round brackets over square brackets), we write $P[X \in B]$ when we means $P([X \in B])$. Hence,

$$P[X \in B] = P([X \in B]) = P(\{\omega \in \Omega : X(\omega) \in B\}).$$

Similarly,

$$P[X < x] = P([X < x]) = P(\{\omega \in \Omega : X(\omega) < x\}).$$

Example 7.9. Continue from Examples 7.4 and 7.7,

(a) $[X = 4] = \{\omega : X(\omega) = 4\}$

(b) $[Y = 4] = \{\omega : Y(\omega) = 4\} = \left\{\omega : (\omega - 3)^2 = 4\right\}$

Example 7.10. In Example 7.5 (Three Coin Tosses), if the coin is fair, then

$P[N < 2] =$

7.11. At a certain point in most probability courses, the sample space is rarely mentioned anymore and we work directly with random variables. The sample space often "disappears" along with the "(ω)" of $X(\omega)$ but they are really there in the background.

Definition 7.12. A set S is called a ***support*** of a random variable X if $P[X \in S] = 1$.

- To emphasize that S is a support of a particular variable X, we denote a support of X by S_X.
- Recall that a support of a probability measure P is any set $A \subset \Omega$ such that $P(A) = 1$.

Definition 7.13. The ***probability distribution*** is a description of the probabilities associated with the random variable.

7.14. There are three types of of random variables. The first type, which will be discussed in Section 8, is called ***discrete random variable***. To tell whether a random variable is discrete, one simple way is to consider the possible values of the random variable. If it is limited to only a finite or countably infinite number of possibilities, then it is discrete. We will later discuss ***continuous random variables*** whose possible values can be anywhere in some intervals of real numbers.

8 Discrete Random Variables

Intuitively, to tell whether a random variable is discrete, we simply consider the possible values of the random variable. If the random variable is limited to only a finite or countably infinite number of possibilities, then it is discrete.

Example 8.1. Voice Lines: A voice communication system for a business contains 48 external lines. At a particular time, the system is observed, and some of the lines are being used. Let the random variable X denote the number of lines in use. Then, X can assume any of the integer values 0 through 48. [15, Ex 3-1]

Definition 8.2. A random variable X is said to be a ***discrete random variable*** if there exists a countable number of distinct real numbers x_k such that

$$\sum_k P[X = x_k] = 1. \tag{11}$$

In other words, X is a discrete random variable if and only if X has a countable support.

Example 8.3. For the random variable N in Example 7.5 (Three Coin Tosses),

For the random variable S in Example 7.6 (Sum of Two Dice),

Definition 8.4. Important Special Case: An ***integer-valued random variable*** is a discrete random variable whose x_k in (11) above are all integers.

8.5. Recall, from 7.13, that the ***probability distribution*** of a random variable X is a description of the probabilities associated with X.

For a discrete random variable, the distribution is often characterized by just a list of the possible values $(x_1, x_2, x_3, \ldots)$ along with the probability of each:

$$(P[X = x_1], P[X = x_2], P[X = x_3], \ldots, \text{ respectively}).$$

In some cases, it is convenient to express the probability in terms of a formula. This is especially useful when dealing with a random variable that has an unbounded number of outcomes. It would be tedious to list all the possible values and the corresponding probabilities.

8.1 PMF: Probability Mass Function

Definition 8.6. When X is a discrete random variable satisfying (11), we define its ***probability mass function*** (pmf) by[24]

$$p_X(x) = P[X = x].$$

- Sometimes, when we only deal with one random variable or when it is clear which random variable the pmf is associated with, we write $p(x)$ or p_x instead of $p_X(x)$.
- The argument (x) of a pmf ranges over all real numbers. Hence, the pmf is defined for x that is not among the x_k in (11). In such case, the pmf is simply 0. This is usually expressed as "$p_X(x) = 0$, otherwise" when we specify a pmf for a particular r.v.

[24]Many references (including [15] and `MATLAB`) use $f_X(x)$ for pmf instead of $p_X(x)$. We will *NOT* use $f_X(x)$ for pmf. Later, we will define $f_X(x)$ as a probability density function which will be used primarily for another type of random variable (continuous r.v.)

Example 8.7. Continue from Example 7.5. N is the number of heads in a sequence of three coin tosses.

8.8. Graphical Description of the Probability Distribution: We can use ***stem plot*** to visualize p_X. To do this, we graph a pmf by marking on the horizontal axis each value with nonzero probability and drawing a vertical bar with length proportional to the probability.

8.9. Any pmf $p(\cdot)$ satisfies two properties:

(a) $p(\cdot) \geq 0$

(b) there exists numbers $x_1, x_2, x_3, \ldots$ such that $\sum_k p(x_k) = 1$ and $p(x) = 0$ for other x.

When you are asked to verify that a function is a pmf, check these two properties.

8.10. Finding probability from pmf: for any subset B of $\mathbb{R}$, we can find

$$P[X \in B] = \sum_{x_k \in B} P[X = x_k] = \sum_{x_k \in B} p_X(x_k).$$

In particular, for integer-valued random variables,

$$P[X \in B] = \sum_{k \in B} P[X = k] = \sum_{k \in B} p_X(k).$$

Example 8.11. Suppose a random variable X has pmf

$$p_X(x) = \begin{cases} c/x, & x = 1, 2, 3, \\ 0, & \text{otherwise.} \end{cases}$$

(a) The value of the constant c is

(b) Sketch of pmf

(c) $P[X = 1]$

(d) $P[X \geq 2]$

(e) $P[X > 3]$

8.12. Any function $p(\cdot)$ on $\mathbb{R}$ which satisfies

(a) $p(\cdot) \geq 0$, and

(b) there exists numbers $x_1, x_2, x_3, \ldots$ such that $\sum_k p(x_k) = 1$ and $p(x) = 0$ for other x

is a pmf of some discrete random variable.

8.2 CDF: Cumulative Distribution Function

Definition 8.13. The (***cumulative***) ***distribution function*** (***cdf***) of a random variable X is the function $F_X(x)$ defined by

$$F_X(x) = P[X \le x].$$

- The argument (x) of a cdf ranges over all real numbers.
- From its definition, we know that $0 \le F_X \le 1$.
- Think of it as a function that collects the "probability mass" from $-\infty$ up to the point x.

8.14. In general, for any discrete random variable with possible values $x_1, x_2, \ldots$, the cdf of X is given by

$$F_X(x) = P[X \le x] = \sum_{x_k < x} p_X(x_i).$$

Example 8.15. Continue from Examples 7.5, 7.10, and 8.7 where N is defined as the number of heads in a sequence of three coin tosses. We have

$$p_N(0) = p_N(3) = \frac{1}{8} \text{ and } p_N(1) = p_N(2) = \frac{3}{8}.$$

(a) $F_N(0)$

(b) $F_N(1.5)$

(c) Sketch of cdf

8.16. Facts:

- For any discrete r.v. X, F_X is a right-continuous, ***staircase*** function of x with jumps at a countable set of points x_k.
- When you are given the cdf of a discrete random variable, you can derive its pmf from the locations and sizes of the jumps. If a jump happens at $x = c$, then $p_X(c)$ is the same as the amount of jump at c. At the location x where there is no jump, $p_X(x) = 0$.

Example 8.17. Consider a discrete random variable X whose cdf $F_X(x)$ is shown in Figure 3.

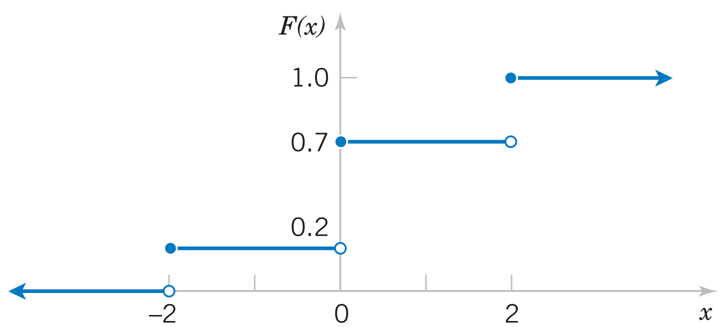


Figure 3: CDF for Example 8.17

Determine the pmf $p_X(x)$.

8.18. Characterizing[25] properties of cdf:

CDF1 F_X is non-decreasing (monotone increasing)

CDF2 F_X is right continuous (continuous from the right)

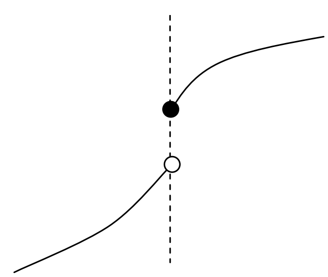


Figure 4: Right-continuous function at jump point

CDF3 $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X(x) = 1$.

8.19. F_X can be written as

$$F_X(x) = \sum_{x_k} p_X(x_k) u(x - x_k),$$

where $u(x) = 1_{[0,\infty)}(x)$ is the unit step function.

[25]These properties hold for any type of random variables. Moreover, for any function F that satisfies these three properties, there exists a random variable X whose CDF is F.