SECTION 5: RANDOM VARIABLES

We often summarize the outcome from a random experiment by a simple number. In many of the examples of random experiments that we have considered, the ***sample space*** has been a description of possible outcomes. In some cases, descriptions of outcomes are sufficient, but in other cases, it is useful to associate a number with each outcome in the ***sample space.*** Because the particular outcome of the experiment is not known in advance, the resulting value of our variable is not known in advance. For this reason, the variable that associates a number with the outcome of a random experiment is referred to as a ***random variable***.

Definition: A **random variable** is a function that assigns a real number to each outcome in the sample space of a random experiment. A random variable is denoted by an uppercase letter such as X. After an experiment is conducted, the measured value of the random variable is denoted by a lowercase letter such as x=70 milliamperes.

Sometimes a measurement (such as current in a copper wire (bakır tel) or length of a machined part) can assume any value in an interval of real numbers (at last theoretically). Then arbitrary (keyfi) precision in the measurement is possible. Of course, in practice, we might round off to the nearest tenth or hundredth of a unit. The random variable that represents this measurement is said to be a ***continuous random variable***. The range of the random variable includes all values in an interval of real numbers; that is, the range can be thought of as continuum.

In other experiments, we might record a count such as the number of transmitted bits that are received in error. Then the measurement is limited to integers. Or we might record that a proportion such as 0.0042 of the 10000 transmitted bits were received in error. Then the measurement is fractional (kesirli/oransal), but it is still limited to discrete points on the real line. Whenever the measurement is limited to discrete points on the real line, the random variable is said to be a ***discrete random variable.***

Definition: A **discrete random variable** is random variable with a finite (or countably infinite) range. A **continuous random variable** is a random variable with an interval (either finite or infinite) of real numbers for its range.

In some cases, the random variable X is actually discrete but, because the range of possible values is so large, it might be more convenient to analyze X as a continuous random variable. For example, suppose that current measurements (akım ölçümleri) are read from a digital instrument that displays the current to the nearest one-hundredth of a milliampere. Because the possible measurements are limited, the random variable is discrete. However, it might be a more convenient, simple approximation to assume that the current measurements are values of a continuous random variable.

Examples of Random Variables:

Examples of continuous random variables: electrical current (elektrik akımı), length, pressure, temperature, time, voltage, weight
Examples of discrete random variables: number of scratches (çizikler) on a surface, proportion of defective parts among 1000 tested, number of transmitted bits received in error.

The following are examples of discrete random variables:

1. **1.** The number of seizures (kriz, nöbet) an epileptic patient has in a given week: x=0, 1, 2, … (it is not known that the number of seizures for any epileptic patient in a given week, that's why it is defined as a random event.)

2. **2.** The number of voters in a sample of 500 who favor impeachment (görevini kötüye kullanma suçlaması/ithamı) of the president: x=0, 1, 2, …, 500.

3. **3.** The shoe size of a tennis player: $x = \cdots 5, 5\frac{1}{2}, 6, 6\frac{1}{2}, 7, 7\frac{1}{2} \cdots$ (assume that we think the customers of sport store, and shoe size are different from a customer to another)

4. **4.** The number of customers waiting to be served in a restaurant at a particular time: x=0,1,2,…. (for any particular time, the number of customers waiting to be served in a restaurant differs.)

Note that several of the examples of discrete random variables begin with the words *The number of*… This wording is very common, since the discrete random variables most frequently observed are counts. The following are examples of continuous random variables:

1. **1.** The length of time (in seconds) between arrivals at a hospital clinic: $0 \le x \le \infty$ (infinity)
2. **2.** The length of time (in minutes) it takes a student to complete a one-hour exam: $0 \le x \le 60$
3. **3.** The amount (in ounces) of carbonated beverage (içecek, meşrubat) loaded into a 12-ounce can (teneke kutu) in a can-filling operation: $0 \le x \le 12$
4. **4.** The depth (in feet) at which a successful oil-drilling (petrol sondaj) venture (girişimi) first strikes (çıkarmak) oil: $0 \le x \le c$, where c is the maximum depth obtainable
5. **5.** The weight (in pounds) of a food item bought in a supermarket: $0 \le x \le 500$ [Note: Theoretically, there is no upper limit on x, but it is unlikely that it would exceed 500 pounds.]

5.1. DISCRETE RANDOM VARIABLES

Many physical systems can be modeled by the same or similar random experiments and random variables. The distribution of the random variables involved in each of these common systems can be analyzed, and the results of that analysis can be used in different applications and examples.

Example 1: A voice communication system (ses haberleşme sistemi) for a business contains 48 external lines (dış hatlar). 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. When the system is observed, if 10 lines are in use, x=10.

Example 2: In a semiconductor (yarı iletken) manufacturing (imalat) process, two wafers (silikon devre levhaları) from a lot (parti) are tested. Each wafer is classified as pass or fail. Assume that the probability that a wafer passes the test is 0.8 and that wafers are independent. The sample space for the experiment and associated probabilities are shown in Table 1. For example, because of the independence, the probability of the outcome that the first wafer tested passes and the second wafer tested fails, denoted as pf is $P(\text{pf}) = 0.8(0.2) = 0.16$. The random variable X is defined to be equal to the number of wafers that pass. The last column of the table shows the values of X that are assigned to each outcome in the experiment.

Table 1. Wafer Tests.

Outcome			
Wafer 1	Wafer 2	Probability	x
Pass	Pass	0.64	2
Fail	Pass	0.16	1
Pass	Fail	0.16	1
Fail	Fail	0.04	0

Example 3: Define the random variable X to be the number of contamination particles on a wafer in semiconductor manufacturing. Although wafers possess a number of characteristics, the random variable X summarizes the wafer only in terms of the number of particles. The possible values of X are integers from zero up to some large value that represents the maximum number of particles that can be found on one of the wafers. If this maximum number is very large, we might simply assume that the range of X is the set of integers from zero to infinity. (Note that more than one random variable can be defined on a sample space. In this example, we might define the random variable Y to be the number of chips (mikroçip) from a wafer that fail the final test.)

EXERCISES 1

Exercise 1.1: The random variable is the number of nonconforming solder (lehim) connections on a printed circuit board (anakart devresi) with 1000 connections. $x = \{0, 1, 2, \cdots, 1000\}$

Exercise 1.2: An electronic scale (elektronik tartı) that displays weights to the nearest pounds is used to weigh (tartmak) packages. The display shows only five digits. Any weight greater than the display can indicate is shown as 99999. The random variable is the displayed weight. $x = \{0, 1, 2, \cdots, 99999\}$

5.1.1. Probability Distributions and Probability Mass Functions for Discrete Random Variables

Random variables are so important in random experiments that sometimes we essentially ignore the original sample space of the experiment and focus on the probability distribution of the random variable. For example, in Example 1, our analysis might focus exclusively on the integers $\{0, 1, 2 \cdots, 48\}$ in the range of X. In Example 2, we might summarize the random experiment in terms of the three possible values of X, namely $\{0, 1, 2\}$. In this manner, a random variable can simplify the description and analysis of a random experiment.

The **probability distribution** of a random variable X is a description of the probabilities associated with the possible values of X. For a discrete random variable, the distribution is often specified by just a list of the possible values along with the probability of each. In some cases, it is convenient to express the probability in terms of a formula.

Example 4: Consider the experiment of tossing two coins, and let X be the number of heads observed. Find the probability associated with each value of the random variable X, assuming that the two coins are fair.

Solution: The random variable X can assume values 0, 1, 2. The probability associated with each of the four sample points is $1/4$. Then, identifying the probabilities of the sample points associated with each of these values of X, we have

$$P(X=0)=P(TT)=\frac{1}{4}$$

$$P(X=1)=P(TH)+P(HT)=\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$$

$$P(X=2)=P(HH)=\frac{1}{4}$$

Thus, we now know the values the random variable can assume $(0,1,2)$ and how the probability is distributed over those values $\left(1/4, 1/2, 1/4\right)$. This dual specification completely describes the random variable and referred to as the probability distribution, denoted by the symbol $p(x)$. In standard mathematical notation, the probability that a random variable X takes on a value x is denoted $P(X=x)=p(x)$. Thus, $P(X=0)=p(0), P(X=1)=p(1)$ etc. The probability distribution for the coin-toss example is shown in tabular form in Table 2 and in graphic form in Figure 1. Since the probability distribution for a discrete random variable is concentrated at specific points (values of x), the graph in Figure 1 represents the probabilities as the heights of vertical lines over the corresponding values of x.

Table 2. Probability Distribution for Coin-Toss Experiment: Tabular Form

x	$p(x)$
0	$1/4$
1	$1/2$
2	$1/4$

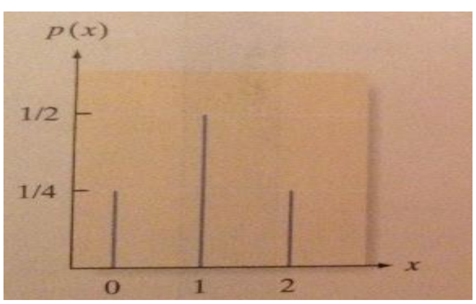

Figure 1. Probability distribution for coin-toss experiment: graphical form.

Example 5: There is a chance that a bit transmitted through a digital transmission channel is received in error. Let X equal the number of bits in error in the next four bits transmitted. The possible values for X are $\{0,1,2,3,4\}$. Based on a model for the errors that is presented in the following section, probabilities for these values will be determined. Suppose that the probabilities are

$$P(X=0)=0.6561 \quad P(X=1)=0.2916 \quad P(X=2)=0.0486 \quad P(X=3)=0.0036 \quad P(X=4)=0.0001$$

The probability distribution of X is specified by the possible values along with the probability of each. A graphical description of the probability distribution of X is shown in Figure 2.

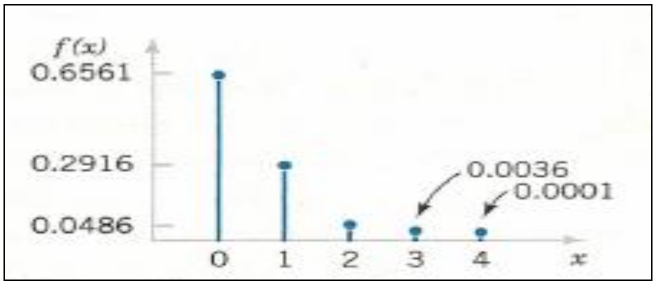

Figure 2. Probability distribution for bits in error.

<u>**Definition**</u>**:** For a discrete random variable X with possible values $x_1, x_2, \cdots, x_n$, a **probability mass function** is function such that

(1) $p(x_i) \geq 0$

(2) $\sum_{i=1}^{n} p(x_i) = 1$

(3) $p(x_i) = P(X = x_i)$

For example, in Example 5,
$p(0)=0.6561$, $p(1)=0.2916$, $p(2)=0.0486$, $p(3)=0.0036$ and $p(4)=0.0001$. Check that the sum of the probabilities in Example 5 is 1.

Example 6: Let the random variable X denote the number of semiconductor wafers that need to be analyzed in order to detect a large particle of contamination. Assume that the probability that a wafer contains a large particle is 0.01 and the wafers are independent. Determine the probability distribution of X. Let *p* denote a wafer in which a large particle is present, and let *a* denote a wafer in which it is absent. The sample space of the experiment is infinite, and it can be represented as all possible sequences that start with a string of *a*'s and end with *p*. That is,

$$s = \{p, ap, aap, aaap, aaaap, aaaaap, \text{and so forth}\}$$

Consider a few special cases. We have $P(X=1)=P(p)=0.01$. Also, using the independence assumption $P(X=2)=P(ap)=0.99(0.01)=0.0099$.

A general formula is

$$P(X=x)=\underbrace{P(aa...ap)}_{(x-1)a's}=0.99^{x-1}(0.01),\ for\ x=1,2,3,\cdots$$

- Describing the probabilities associated with X in terms of this formula is the simplest method of describing the distribution of X in this example. Clearly, $f(x)\geq 0$. The fact that the sum of the probabilities is 1 left as an exercise. This is an example of a geometric random variable.

EXERCISES 2

Exercise 2.1: The sample space of a random experiment is $\{a,b,c,d,e,f\}$, and each outcome is equally likely. A random variable is defined as follows:

Outcome	a	b	c	d	e	f
x	0	0	1.5	1.5	2	3

Determine the probability mass function of X.

Solution:

$$\begin{aligned} p(x) &= \frac{1}{3}, \quad x=0,1.5 \\ &= \frac{1}{6}, \quad x=2,3 \\ &= 0,\ otherwise \end{aligned}$$

Exercise 2.2:

x	-2	-1	0	1	2
p(x)	1/8	2/8	2/8	2/8	1/8

a) $P(X\leq 2)=?$ **b)** $P(X>-2)=?$ **c)** $P(-1\leq X\leq 1)=?$ **d)** $P(X\leq -1\ or\ X=2)=?$

Solution:

a)
$$
\begin{aligned}
P(X \le 2) &= P(X=-2)+P(X=-1)+P(X=0)+P(X=1)+P(X=2)\\
&= \frac{1}{8}+\frac{2}{8}+\frac{2}{8}+\frac{2}{8}+\frac{1}{8}\\
&= 1
\end{aligned}
$$

b)
$$
\begin{aligned}
P(X > -2) &= P(X=-1)+P(X=0)+P(X=1)+P(X=2)\\
&= \frac{2}{8}+\frac{2}{8}+\frac{2}{8}+\frac{1}{8}\\
&= 7/8
\end{aligned}
$$

c)
$$
\begin{aligned}
P(-1 \le X \le 1) &= P(X=-1)+P(X=0)+P(X=1)\\
&= \frac{2}{8}+\frac{2}{8}+\frac{2}{8}\\
&= 3/4
\end{aligned}
$$

d)
$$
\begin{aligned}
P(X \le -1\, or\ X=2) &= P(X=-2)+P(X=-1)+P(X=2)\\
&= \frac{1}{8}+\frac{2}{8}+\frac{1}{8}\\
&= 1/2
\end{aligned}
$$

Exercise 2.3:
$$
\begin{aligned}
p(x) &= \frac{2x+1}{25}, \quad x = 0,1,2,3,4\\
&= 0, \qquad \textit{otherwise}
\end{aligned}
$$

a) $P(X=4)=?$ **b)** $P(X \le 1)=?$ **c)** $P(2 \le X < 4)=?$ **d)** $P(X > -10)=?$

Solution:

a) $P(X=4)=9/25$ **b)** $P(X \le 1)=P(X=0)+P(X=1)=(1/25)+(3/25)=4/25$

c) $P(2 \le X < 4)=P(X=2)+P(X=3)=(5/25)+(7/25)=12/25$

d) $P(X > -10)=P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)=1$

Exercise 2.4: Marketing estimates that a new instrument for the analysis of soil samples will be very successful, moderately successful, or unsuccessful, with probabilities 0.3, 0.6, and 0.1, respectively. The yearly revenue associated with a very successful, moderately successful, or unsuccessful product is \$10 million, \$5 million, and \$1 million, respectively. Let the random variable X denote the yearly revenue of the product. Determine the probability mass function of X.

Solution: $P(X = 10\ million) = 0.3 \quad P(X = 5\ million) = 0.6 \quad P(X = 1\ million) = 0.1$

$$\text{or} \quad \begin{aligned} p(x) &= 0.1,\ x = 1\ million \\ &= 0.6,\ x = 5\ million \\ &= 0.3,\ x = 10\ million \end{aligned}$$

Exercise 2.5: An optical inspection system is to distinguish among different part types. The probability of a correct classification of any part is 0.98. Suppose that three parts are inspected and that the classifications are independent. Let the random variable X denote the number of parts that are correctly classified. Determine the probability mass function of X.

Solution: $P(X = 0) = 0.02 \times 0.02 \times 0.02 = 8 \times 10^{-6} \quad P(X = 1) = (0.98 \times 0.02 \times 0.02) \times 3 = 0.0012$

$P(X = 2) = (0.98 \times 0.98 \times 0.02) \times 3 = 0.0576 \quad P(X = 3) = 0.98 \times 0.98 \times 0.98 = 0.9412$

5.1.2. Cumulative Distribution Function of a Discrete Random Variable

In Example 5, we might be interested in the probability of three or fewer bits being in error. This question can be expressed as $P(X \le 3)$. The event that $\{X \le 3\}$ is the union of the events $\{X = 0\}, \{X = 1\}, \{X = 2\}$, and $\{X = 3\}$. Clearly, these three events are mutually exclusive (ayrışık, karşılıklı dışarlayan).

Therefore, $$\begin{aligned} P(X \le 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0.6561 + 0.2916 + 0.0486 + 0.0036 = 0.9999 \end{aligned}$$

This approach can also be used to determine

$$P(X = 3) = P(X \le 3) - P(X \le 2) = 0.0036$$

Example 5 shows that it is sometimes useful to be able to provide **cumulative probabilities** such as $P(X \le x)$ and that such probabilities can be used to find the probability mass function of a random variable. <u>Therefore, using cumulative probabilities is an alternative method of describing the probability distribution of a random variable.</u>

In general, for any discrete random variable with possible values $x_1, x_2, \cdots, x_n$, the events $\{X = x_1\}, \{X = x_2\}, \cdots, \{X = x_n\}$ are mutually exclusive. Therefore, $P(X \le x) = \sum_{x_i \le x} p(x_i)$.

<u>Definition:</u> The **cumulative distribution** function of a discrete random variable X, denoted as $F(x)$, is

$$F(x) = P(X \le x) = \sum_{x_i \le x} p(x_i)$$

For a discrete random variable X, $F(x)$ satisfies the following properties.

(1) $F(x) = P(X \le x) = \sum_{x_i \le x} p(x_i)$

(2) $0 \le F(x) \le 1$

(3) If $x \le y,$ then $F(x) \le F(y)$

(4) Since $F(x)$ is a probability, the value of the distribution function is always between 0 and 1. Moreover,

$$\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} P(X \le x) = 1$$

$$\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} P(X \le x) = 0$$

Like a probability mass function, a cumulative distribution function provides probabilities. Notice that even if the random variable X can only assume integer values, the cumulative distribution function can be defined at noninteger values.

In Example 5, $F(1.5) = P(X \le 1.5) = P(X = 0) + P(X = 1) = 0.6561 + 0.2916 = 0.9477.$ Properties (1) and (2) of a cumulative distribution function follow from the definition. Property (3) follows from the fact that if $x \le y$, the event that $\{X \le x\}$ is contained in the event $\{X \le y\}$.

*The next example shows how the cumulative distribution function can be used to determine the probability mass function of a discrete random variable.

Example 7: Determine the probability mass function of X from the following cumulative distribution function:

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \le x < 0 \\ 0.7 & 0 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

Figure 3 displays a plot of $F(x)$. From the plot, the only points that receive nonzero probability are -2, 0, and 2. The probability mass function at each point is the change in the cumulative distribution function at the point. Therefore,

$$p(-2) = 0.2 - 0 = 0.2 \qquad p(0) = 0.7 - 0.2 = 0.5 \qquad p(2) = 1.0 - 0.7 = 0.3$$

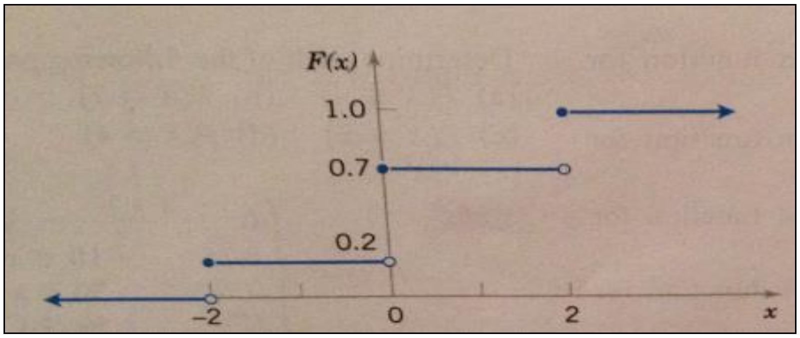

Figure 3. Cumulative distribution function for Example 7.

Example 8: Suppose that a day's production of 850 manufactured (üretilmiş) parts contains 50 parts that do not conform to customer requirements. Two parts are selected at random, without replacement, from the batch (yığın). Let the random variable X equal the number of nonconforming parts in the sample. What is the cumulative distribution function of X?

Solution: The question can be answered by first finding the probability mass function of X.

$$P(X=0)=\frac{800}{850}.\frac{799}{849}=0.886$$

$$P(X=1)=2.\frac{800}{850}.\frac{50}{849}=0.111, \text{ therefore,}$$

$$P(X=2)=\frac{50}{850}.\frac{49}{849}=0.003$$

$$F(0)=P(X\leq 0)=0.886$$

$$F(1)=P(X\leq 1)=0.886+0.111=0.997$$

$$F(2)=P(X\leq 2)=1$$

The cumulative distribution function for this example is graphed in Figure 4. Note that $F(x)$ is defined for all x from $-\infty<x<\infty$ and not only for 0, 1, and 2.

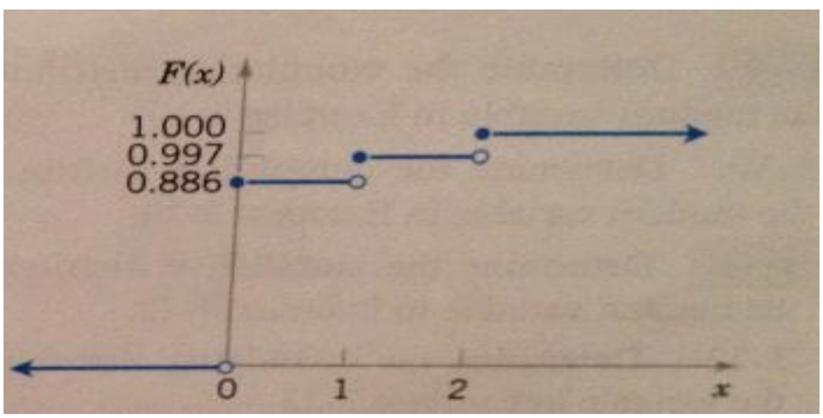

Figure 4. Cumulative distribution function for Example 8.

EXERCISES 3

Exercise 3.1: Determine the cumulative distribution function for the random variable for

x	-2	-1	0	1	2
p(x)	1/8	2/8	2/8	2/8	1/8

also determine the following probabilities :

a) $P(X \le 1.25)$ **b)** $P(X \le 2.2)$ **c)** $P(-1.1 < X \le 1)$ **d)** $P(X > 0)$

Solution:

x	-2	-1	0	1	2
F(x)	1/8	3/8	5/8	7/8	1

The cumulative distribution function can be written as given below:

$$F(x)=\begin{cases} 0 & x<-2 \\ 1/8 & -2 \le x < -1 \\ 3/8 & -1 \le x < 0 \\ 5/8 & 0 \le x < 1 \\ 7/8 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

a) $P(X \le 1.25) = P(X \le 1) = F(1) = 7/8$

b) $P(X \le 2.2) = P(X \le 2) = F(2) = 1$

c) $P(-1.1 < X \le 1) = P(-1 \le X \le 1) = F(1) - F(-2) = (7/8) - (1/8) = 6/8 = 3/4$

d) $P(X > 0) = 1 - P(X \le 0) = 1 - F(0) = 1 - (5/8) = 3/8$

Exercise 3.2: $F(x)=\begin{cases} 0 & x<1 \\ 0.5 & 1 \le x < 3 \\ 1 & x \ge 3 \end{cases}$

a) $P(X \le 3)$ **b)** $P(X \le 2)$ **c)** $P(1 \le X \le 2)$ **d)** $P(X > 2)$

Solution:

a) $P(X \le 3) = F(3) = 1$ **b)** $P(X \le 2) = F(2) = 0.5$ **c)** $P(1 \le X \le 2) = F(2) - F(0) = 0.5 - 0 = 0.5$
d) $P(X > 2) = 1 - P(X \le 2) = 1 - F(2) = 1 - 0.5 = 0.5$

Exercise 3.3: $F(x) = \begin{cases} 0 & x < -10 \\ 0.25 & -10 \le x < 30 \\ 0.75 & 30 \le x < 50 \\ 1 & x \ge 50 \end{cases}$

a) $P(X \le 50)$ **b)** $P(X \le 40)$ **c)** $P(40 \le X \le 60)$ **d)** $P(X < 0)$ **e)** $P(0 \le X \le 10)$
f) $P(-10 < X < 10) = 0.25 - 0.25 = 0$

Solution:

a) $P(X \le 50) = F(50) = 1$ **b)** $P(X \le 40) = F(40) = 0.75$

c) $P(40 \le X \le 60) = F(60) - F(39) = 1 - 0.75 = 0.25$

d) $P(X < 0) = P(X \le -1) = F(-1) = 0.25$

e) $P(0 \le X \le 10) = F(10) - F(-1) = 0.25 - 0.25 = 0$

f) $P(-10 < X < 10) = F(9) - F(-10) = 0.25 - 0.25 = 0$

5.1.3. Mean and Variance of a Discrete Random Variable

Two numbers are often used to summarize a probability distribution for a random variable X. The ***mean*** is a measure of the center or middle of the probability distribution, and the ***variance*** is a measure of the dispersion, or variability in the distribution. These two measures do not uniquely identify a probability distribution. That is, two different distributions can have the same mean and variance. Still, these measures are simple, useful summaries of the probability distribution of X.

Definition: The **mean** or **expected value** of the discrete random variable X, denoted as μ or $E(X)$, i

$$\mu = E(X) = \sum_x x\, p(x)$$

The **variance** of X, denoted as σ^2 or $V(X)$, is

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_x (x-\mu)^2 p(x) = \sum_x x^2 p(x) - \mu^2 = E(X^2) - \mu^2, \text{ where } \mu = E(X)$$

The **standard deviation** of X is $\sigma = \sqrt{\sigma^2}$

The ***mean of a discrete random variable X*** is a weighted average of the possible values of X, with weights equal to the probabilities. If $p(x)$ is the probability mass function of a loading on a long, thin beam (kiriş, ışın), $E(X)$ is the point at which the beam (kiriş) balances. Consequently, $E(X)$ describes the "center" of the distribution of X in a manner similar to the balance point of a loading. See Figure 5.

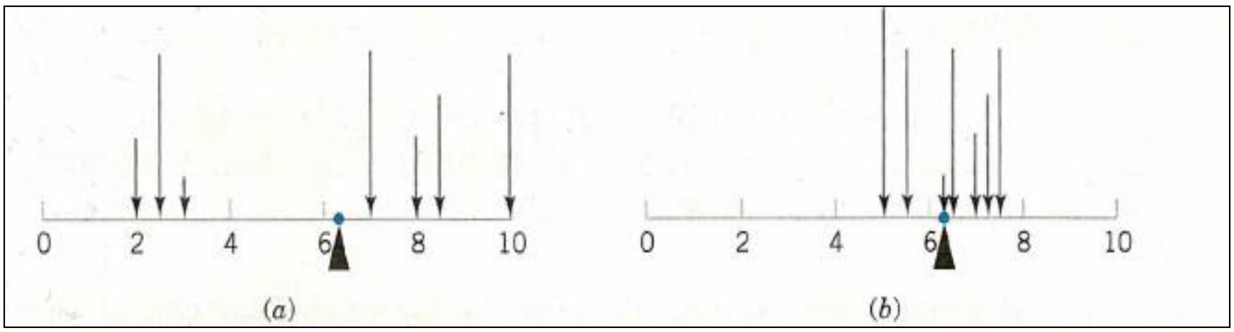

Figure 5. A probability distribution can be viewed as a loading with the mean equal to the balance point. Parts (a) and (b) illustrate equal means, but Part (a) illustrates a larger variance.

The ***variance of a random variable X*** is a measure of dispersion or scatter in the possible values for X. The variance of X uses weight $p(x)$ as the multiplier of each possible squared deviation $(x-\mu)^2$. Figure 5 illustrates probability distributions with equal means but different variances. Properties of summations and the definition of μ can be used to show the equality of the formulas for variance.

$$V(X) = \sum_x (x-\mu)^2 p(x) = \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x)$$

$$= \sum_x x^2 p(x) - 2\mu^2 + \mu^2 = \sum_x x^2 p(x) - \mu^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2$$

Either formula for $V(X)$ can be used. Figure 6 illustrates that two probability distributions can differ even though they have identical means and variances.

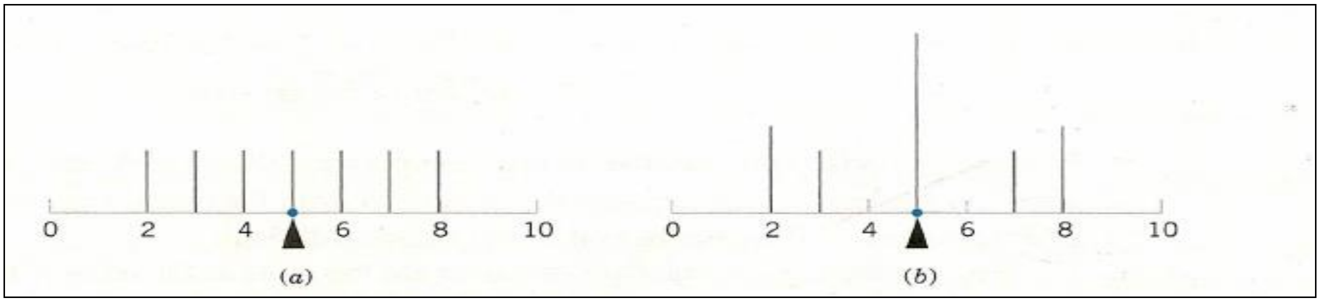

Figure 6. The probability distributions illustrated in Parts (a) and (b) differ even though they have equal means and variances.

Example 9: In Example 5, there is a chance that a bit transmitted through a digital transmission channel is received in error. Let X equal the number of bits in error in the next four bits transmitted. The possible values for X are $\{0,1,2,3,4\}$. Based on a model for the errors that is presented in the following section, probabilities for these values will be determined. Suppose that the probabilities are

$$P(X=0)=0.6561 \quad P(X=1)=0.2916 \quad P(X=2)=0.0486 \quad P(X=3)=0.0036 \quad P(X=4)=0.0001$$

Now
$$\begin{aligned}\mu = E(X) &= 0p(0)+1p(1)+2p(2)+3p(3)+4p(4)\\ &= 0(0.6561)+1(0.2916)+2(0.0486)+3(0.0036)+4(0.0001)\\ &= 0.4\end{aligned}$$

Although X never assumes the value 0.4, the weighted average of the possible values is 0.4. To calculate $V(X)$, a table is convenient.

x	$x-0.4$	$(x-0.4)^2$	$p(x)$	$p(x)(x-0.4)^2$
0	-0.4	0.16	0.6561	0.104976
1	0.6	0.36	0.2916	0.104976
2	1.6	2.56	0.0486	0.124416
3	2.6	6.76	0.0036	0.024336
4	3.6	12.96	0.001	0.001296

$V(X)=\sigma^2=\sum_{i=1}^{5} p(x_i)(x_i-0.4)^2=0.36$ (The alternative formula for variance could also be used to obtain the same result.)

Example 10: Two new product designs are to be compared on the basis of revenue (kazanç, hasılat) potential. Marketing feels that the revenue from design A can be predicted quite accurately to be $3 million. The revenue potential of design B is more difficult to assess. Marketing concludes that there is a probability of 0.3 that the revenue from design B will be $7 million, but there is a 0.7 probability that the revenue will be only $2 million. Which design do you prefer?

Let X denote the revenue from design A. Because there is no uncertainty in the revenue from design A, we can model the distribution of the random variable X as \$3 million with probability 1. Therefore, $E(X) = \$3$ million.

Let Y denote the revenue from design B. The expected value of Y in millions of dollars is $E(Y) = \$7(0.3) + \$2(0.7) = \$3.5$

Because $E(Y)$ exceeds $E(X)$, we might prefer design B. However, the variability of the result from design B is larger. That is, $\sigma^2 = (7-3.5)^2(0.3) + (2-3.5)^2(0.7) = 5.25\ \textit{million of dollars squared}$

Because the units of the variables in this example are millions of dollars, and because the variance of a random variable squares the deviations from the mean, the units of σ^2 are millions of dollars squared. These units make interpretation difficult. Because the units of standard deviation are the same as the units of the random variable, as the standard deviation σ is easier to interpret. In this example, we can summarize our results as "the average deviation of Y from its mean is \$2.29 million."

Example 11: The number of messages sent per hour over a computer network has the following distribution:

x=number of messages	10	11	12	13	14	15
$p(x)$	0.08	0.15	0.30	0.20	0.20	0.07

Determine the mean and standard deviation of the number of messages sent per hour.

$$E(X) = 10(0.08) + 11(0.15) + \cdots + 15(0.07) = 12.5$$
$$V(X) = 10^2(0.08) + 11^2(0.15) + \cdots + 15^2(0.07) - 12.5^2 = 1.85$$
$$\sigma = \sqrt{V(X)} = \sqrt{1.85} = 1.36$$

The variance of a random variable X can be considered to be the expected value of a specific function of X, namely, $g(X) = (X-\mu)^2$ (That means we can write $V(X) = E(X-\mu)^2$). In general, the expected value of any function $g(X)$ of a discrete random variable is defined in a similar manner.

<u>Expected Value of a Function of a Discrete Random Variable</u>

If X is a discrete random variable with probability mass function $f(x)$

$$E[g(X)] = \sum_x g(x)\,p(x) \qquad (1)$$

Example 12: In Example 9, X is the number of bits in error in the next four bits transmitted. What is the expected value of the square of the number of the number of bits in error? Now, $g(X)=X^2$. Therefore,

$$E\left[g(X)\right]=0^2\times 0.6561+1^2\times 0.2916+2^2\times 0.0486+3^2\times 0.0036+4^2\times 0.0001=0.52$$

In the previous example, the expected value of X^2 does not equal $E(X)$ squared (That means $E\left(X^2\right)\neq\left[E(X)\right]^2$. However, in the special case that $g(X)=aX+b$ for any constant a an b, $E\left[g(X)\right]=aE(X)+b$. This can be shown from the properties of sums in the definition in Eq. 1.

5.1.4. Expected Value of a Function of a Discrete Random Variable

Theorem 1. *Let X be a discrete random variable with probability mass function p(x) and g(X) be a realvalued function of X. Then the expected value of g(X) is given by*

$$E\left[g(X)\right]=\sum_x g(x)p(x)$$

5.1.5. Properties of Mathematical Expectation

5.1.5.1. Constants

Theorem 2. *Let X be a discrete random variable with probability function p(x) and c be a constant. Then* $E(c)=c$.

5.1.5.2. Constants Multiplied by Functions of Random Variables

Theorem 3. *Let X be a discrete random variable with probability function p(x), g(X) be a function of X, and let c be a constant. Then* $E\left[c\,g(X)\right]=c\,E\left[g(X)\right]$.

5.1.5.3. Sums of Functions of Random Variables

Theorem 4. *Let X be a discrete random variable with probability function p(x),* $g_1(X)$, $g_2(X)$, $g_3(X)$, …, $g_k(X)$ *be k functions of X. Then*

$$E\left[g_1(X)+g_2(X)+g_3(X)+\cdots+g_k(X)\right]=E\left[g_1(X)\right]+E\left[g_2(X)\right]+E\left[g_3(X)\right]+\cdots+E\left[g_k(X)\right]$$

EXERCISE 4

Exercise 4.1: The random variable X has the following probability distribution:

x	2	3	5	8
	0.2	0.4	0.3	0.1

Determine the following:

a) $P(X \le 3)$ **b)** $P(X > 2.5)$ **c)** $P(2.7 < X < 5.1)$ **d)** $E(X)$ **e)** $V(X)$

Solution:

a) $P(X \le 3) = P(X = 2) + P(X = 3) = 0.2 + 0.4 = 0.6$

b) $P(X > 2.5) = P(X = 3) + P(X = 5) + P(X = 8) = 0.4 + 0.3 + 0.1 = 0.8$

c) $P(2.7 < X < 5.1) = P(X = 3) + P(X = 5) = 0.4 + 0.3 = 0.7$

d)
$$\begin{aligned} E(X) &= 2P(X = 2) + 3P(X = 3) + 5P(X = 5) + 8P(X = 8) \\ &= 2(0.2) + 3(0.4) + 5(0.3) + 8(0.1) \\ &= 3.9 \end{aligned}$$

e)
$$\begin{aligned} E(X^2) &= 2^2 P(X = 2) + 3^2 P(X = 3) + 5^2 P(X = 5) + 8^2 P(X = 8) \\ &= 4(0.2) + 9(0.4) + 25(0.3) + 64(0.1) \\ &= 18.3 \end{aligned}$$

$$V(X) = E[X^2] - [E(X)]^2 = 18.3 - (3.9)^2 = 3.09$$

Exercise 4.2: A department supervisor is considering purchasing a photocopy machine. One consideration is how often the machine will need repairs. Let *X* denote the number of repairs during a year. Based on past performance, the distribution of *X* is shown below:

Number of repairs, x	0	1	2	3
$p(x)$	0.2	0.3	0.4	0.1

a) What is the expected number of repairs during a year?

b) What is the variance of the number of repairs during a year?

Solution:

a)

$$\begin{aligned} E(X) &= 0\times P(X=0)+1\times P(X=1)+2\times P(X=2)+3\times P(X=3) \\ &= 0\times(0.2)+1\times(0.3)+2\times(0.4)+3\times(0.1) \\ &= 1.4 \end{aligned}$$

b)

$$\begin{aligned} E(X^2) &= 0^2\times P(X=0)+1^2\times P(X=1)+2^2\times P(X=2)+3^2\times P(X=3) \\ &= 0\times(0.2)+1\times(0.3)+4\times(0.4)+9\times(0.1) \\ &= 2.8 \end{aligned}$$

$$V(X)=E\left[X^2\right]-\left[E(X)\right]^2=2.8-(1.4)^2=0.84$$

Exercise 4.3:X discrete random variable has probability mass function as follows:

$$\begin{aligned} p(x) &= kx, \qquad & x=1,2,3,4,5 \\ &= 0, \qquad & otherwise \end{aligned}$$

- **a)** Find constant k value.
- **b)** Find the expected value of X.
- **c)** Find the variance of X.
- **d)** Find the cumulative distribution function of X.
- **e)** Find probabilities: P(X>2), P(2<X≤4), P(2≤X<4).

Solution:

a) $\sum_{x=1}^{5} P(X=x)=\sum_{x=1}^{5} kx = k\frac{5.6}{2}=1 \qquad k=\frac{1}{15}$

b) $E(X)=\sum_{x=1}^{5} xP(X=x)=\sum_{x=1}^{5}\frac{x^2}{15}=\frac{1}{15}(1+4+9+16+25)=\frac{11}{3}$

c)

$$\begin{aligned} V(X) &= E(X^2)-\left[E(X)\right]^2=\sum_{x=1}^{5}\frac{x^3}{15}-\left(\frac{11}{3}\right)^2 \\ &= \frac{1}{15}(1+8+27+64+125)-\left(\frac{11}{3}\right)^2 \\ &= \frac{225}{15}-\left(\frac{121}{9}\right)=\frac{14}{9} \end{aligned}$$

d)

$$\begin{aligned} F(x) &= \sum_{t=1}^{x}\frac{t}{15}=\frac{x(x+1)}{30}, \qquad & x=1,2,3,4,5 \\ &= 0, \qquad & x<1 \\ &= 0, \qquad & x\geq 5 \end{aligned}$$

e) $P(X>2)=1-P(X\le 2)=1-\left(\frac{1}{15}+\frac{2}{15}\right)=\frac{12}{15}=\frac{4}{5}=1-F(2)$

$P(2<X\le 4)=P(X=3)+P(X=4)=\frac{3}{15}+\frac{4}{15}=\frac{7}{15}=F(4)-F(2)$

$P(2\le X<4)=P(X=2)+P(X=3)=\frac{2}{15}+\frac{3}{15}=\frac{1}{3}=F(3)-F(1)$

Exercise 4.4: Suppose $\mathrm{p}(\mathrm{x})=\begin{cases}2\mathrm{ax}, & \mathrm{x}=1,2,3\\ \mathrm{a}(1+2\mathrm{x}), & \mathrm{x}=4,5,6,7\\ 0, & \text{otherwise}\end{cases}$

a) Find the constant a.
b) Determine the mean for X.
c) Determine the cumulative distribution function for X.
d) Determine the following probabilities

$P(X\ge 5)\quad P(X<3)\quad P(2\le X<5)\quad P(X=3)\quad P(X<1)\quad P(X>8)\quad P(X<8)$

Solution:

$$\sum_{x\in Rx} p(x)=1 \quad\Rightarrow\quad \sum_{x=1}^{7} p(x)=1$$

$$\sum_{x=1}^{3}2ax+\sum_{x=4}^{7}a(1+2x)=2a\sum_{x=1}^{3}x+\sum_{x=4}^{7}a+2a\sum_{x=4}^{7}x$$

$$=2a\left(\frac{3.4}{2}\right)+\left(\sum_{x=1}^{7}a-\sum_{x=1}^{3}a\right)+2a\left(\sum_{x=1}^{7}x-\sum_{x=1}^{3}x\right)$$

$$Note:\boxed{\begin{array}{l}\displaystyle\sum_{x=1}^{n}x=\frac{n(n+1)}{2}\\ \displaystyle\sum_{x=1}^{n}x^2=\frac{n(n+1)(2n+1)}{6}\end{array}}$$

$$=12a+(7a-3a)+2a\left(\frac{7.8}{2}-\frac{3.4}{2}\right)$$

$$=60a \quad\Rightarrow\quad 60a=1 \quad \boxed{a=1/60} \quad\Rightarrow\quad p(x)=\begin{cases}x/30, & x=1,2,3\\ (1+2x)/60, & x=4,5,6,7\\ 0, & otherwise\end{cases}$$

b) $E(X)=\sum_{x\in Rx} xp(x) \Rightarrow$

$$E(X)=\sum_{x=1}^{7} xp(x)=\sum_{x=1}^{3} x\frac{x}{30}+\sum_{x=4}^{7} x\left(\frac{1}{60}+\frac{x}{30}\right)$$
$$=\frac{1}{30}\sum_{x=1}^{3} x^2+\frac{1}{60}\sum_{x=4}^{7} x+\frac{1}{30}\sum_{x=4}^{7} x^2$$
$$=\frac{1}{30}\left(\frac{3.4.7}{6}\right)+\frac{1}{60}\sum_{x=4-3}^{7-3}(x+3)+\frac{1}{30}\sum_{x=4-3}^{7-3}(x+3)^2$$
$$=\frac{14}{30}+\frac{1}{60}\left(\sum_{x=1}^{4} x+\sum_{x=1}^{4} 3\right)+\frac{1}{30}\sum_{x=1}^{4}\left(x^2+6x+9\right)$$
$$=\frac{151}{30}$$

c)

$$F(x)=P(X\le x)=\begin{cases} 0, & x<1 \\ \sum_{t=1}^{x}\frac{t}{30}=\frac{x(x+1)}{60}, & x=1,2,3 \\ \sum_{x=1}^{3}\frac{x}{30}+\sum_{t=4}^{x}\left(\frac{1}{60}+\frac{t}{30}\right)=\frac{1}{30}\left(\frac{3.4}{2}\right)+\sum_{t=1}^{x-3}\frac{1}{60}+\frac{1}{30}\sum_{t=1}^{x-3}(t+3)=\frac{x^2+2x-3}{60}, & x=4,5,6,7 \\ 1, & x\ge 7 \end{cases}$$

$$F(x)=\begin{cases} 0, & x<1 \\ \frac{x(x+1)}{60}, & x=1,2,3 \\ \frac{x^2+2x-3}{60}, & x=4,5,6,7 \\ 1, & x\ge 7 \end{cases}$$

d)

$$P(X\ge 5)=\sum_{x=5}^{7} p(x)=P(X=5)+P(X=6)+P(X=7)=11/60+13/60+15/60=39/60=13/20$$

$$or \qquad P(X\ge 5)=1-P(X<5)=1-P(X\le 4)=1-F(4)=1-\frac{4^2+8-3}{60}=13/20$$

$$P(2\le X<5)=\sum_{x=2}^{4} p(x)=P(X=2)+P(X=3)+P(X=4)=2/30+3/30+9/60=19/60$$

$$or \quad P(2\le X<5)=P(1<X\le 4)=F(4)-F(1)=\frac{4^2+2.4-3}{60}-\frac{1.2}{60}=19/60$$

$$P(X<3)=\sum_{x=1}^{2} p(x)=\sum_{x=1}^{2}\frac{x}{30}=\frac{1}{30}\left(\frac{2.3}{2}\right)=1/10$$

$$or \quad P(X<3)=P(X\le 2)=F(2)=\frac{2.3}{60}=1/10$$

$$P(X=3)=3/30=1/10$$

$$or \quad P(X=3)=P(2<X<4)=P(2<X\le 3)=F(3)-F(2)=(3.4)/60-(2.3)/60=1/10$$

$$P(X<1)=0 \quad P(X>8)=0 \quad P(X<8)=P(X\le 8)=1$$