Lecture plan

3.1 Concept of a Random Variable

Definition 3.1: A **random variable** is a function that associates a real number with each element in the sample space.

Example 3.1: Two balls are drawn in succession without replacement from an urn containing 4 red balls and 3 black balls. The possible outcomes and the values y of the random variable Y, where Y is the number of red balls, are

Sample Space	y
RR	2
RB	1
BR	1
BB	0

Example 3.2: A stockroom clerk returns three safety helmets at random to three steel mill employees who had previously checked them. If Smith, Jones, and Brown, in that order, receive one of the three hats, list the sample points for the possible orders of returning the helmets, and find the value m of the random variable M that represents the number of correct matches.

***Solution*:** If S, J, and B stand for Smith's, Jones's, and Brown's helmets, respectively, then the possible arrangements in which the helmets may be returned and the number of correct matches are

Sample Space	m
SJB	3
SBJ	1
BJS	1
JSB	1
JBS	0
BSJ	0

Definition 3.2: If a sample space contains a finite number of possibilities or an unending sequence with as many elements as there are whole numbers, it is called a **discrete sample space**.

Definition 3.3: If a sample space contains an infinite number of possibilities equal to the number of points on a line segment, it is called a **continuous sample space**.

A random variable is called a **discrete random variable** if its set of possible outcomes is countable.

When a random variable can take on values on a continuous scale, it is called a **continuous random variable**.

Example 3.4: Statisticians use **sampling plans** to either accept or reject batches or lots of material. Suppose one of these sampling plans involves sampling independently 10 items from a lot of 100 items in which 12 are defective.

Let X be the random variable defined as the number of items found defective in the sample of 10. In this case, the random variable takes on the values $0, 1, 2, \ldots, 9, 10$.

Example 3.5: Suppose a sampling plan involves sampling items from a process until a defective is observed. The evaluation of the process will depend on how many consecutive items are observed. In that regard, let X be a random variable defined by the number of items observed before a defective is found. With N a nondefective and D a defective, sample spaces are $S = \{D\}$ given $X = 1$, $S = \{ND\}$ given $X = 2$, $S = \{NND\}$ given $X = 3$, and so on.

Example 3.3: Consider the simple condition in which components are arriving from the production line and they are stipulated to be defective or not defective. Define the random variable X by

$$X = \begin{cases} 1, & \text{if the component is defective,} \\ 0, & \text{if the component is not defective.} \end{cases}$$

Clearly the assignment of 1 or 0 is arbitrary though quite convenient. This will become clear in later chapters. The random variable for which 0 and 1 are chosen to describe the two possible values is called a **Bernoulli random variable**.

Further illustrations of random variables are revealed in the following examples.

Example 3.6: Interest centers around the proportion of people who respond to a certain mail order solicitation. Let X be that proportion. X is a random variable that takes on all values x for which $0 \leq x \leq 1$.

Example 3.7: Let X be the random variable defined by the waiting time, in hours, between successive speeders spotted by a radar unit. The random variable X takes on all values x for which $x \geq 0$.

3.2 Discrete Probability Distributions

m	0	1	3
$P(M = m)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

Definition 3.4: The set of ordered pairs $(x, f(x))$ is a **probability function**, **probability mass function**, or **probability distribution** of the discrete random variable X if, for each possible outcome x,

1. $f(x) \geq 0$,
2. $\sum_x f(x) = 1$,
3. $P(X = x) = f(x)$.

Example 3.8: A shipment of 20 similar laptop computers to a retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these computers, find the probability distribution for the number of defectives.

Definition 3.5: The **cumulative distribution function** $F(x)$ of a discrete random variable X with probability distribution $f(x)$ is

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad \text{for } -\infty < x < \infty.$$

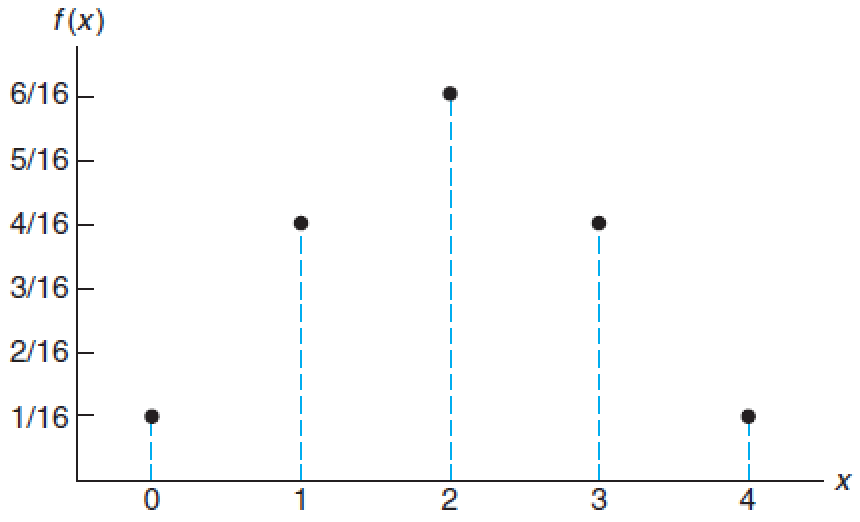


Figure 3.1: Probability mass function plot.

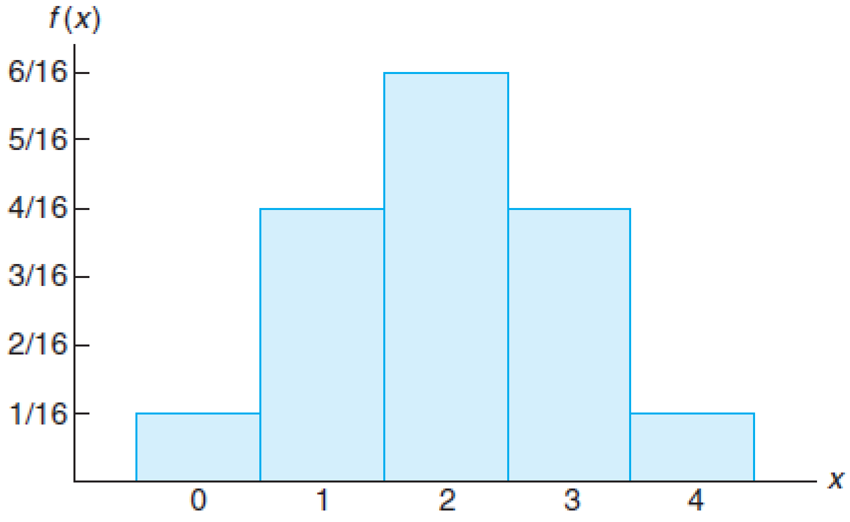


Figure 3.2: Probability histogram.

Example 3.10: Find the cumulative distribution function of the random variable X in Example 3.9. Using $F(x)$, verify that $f(2) = 3/8$.

***Solution*:** Direct calculations of the probability distribution of Example 3.9 give $f(0) = 1/16$, $f(1) = 1/4$, $f(2) = 3/8$, $f(3) = 1/4$, and $f(4) = 1/16$. Therefore,

$$
\begin{aligned}
F(0) &= f(0) = \frac{1}{16},\\
F(1) &= f(0) + f(1) = \frac{5}{16},\\
F(2) &= f(0) + f(1) + f(2) = \frac{11}{16},\\
F(3) &= f(0) + f(1) + f(2) + f(3) = \frac{15}{16},\\
F(4) &= f(0) + f(1) + f(2) + f(3) + f(4) = 1.
\end{aligned}
$$

Hence,

$$
F(x) = \begin{cases} 0, & \text{for } x < 0,\\ \frac{1}{16}, & \text{for } 0 \le x < 1,\\ \frac{5}{16}, & \text{for } 1 \le x < 2,\\ \frac{11}{16}, & \text{for } 2 \le x < 3,\\ \frac{15}{16}, & \text{for } 3 \le x < 4,\\ 1 & \text{for } x \ge 4. \end{cases}
$$

Now

$$
f(2) = F(2) - F(1) = \frac{11}{16} - \frac{5}{16} = \frac{3}{8}.
$$

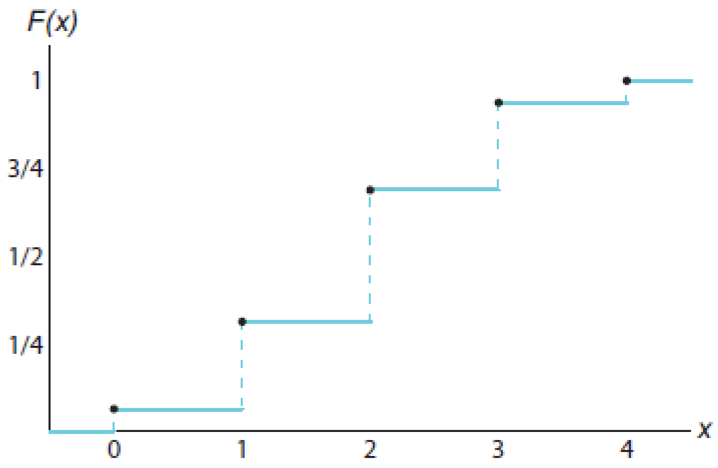


Figure 3.3: Discrete cumulative distribution function.

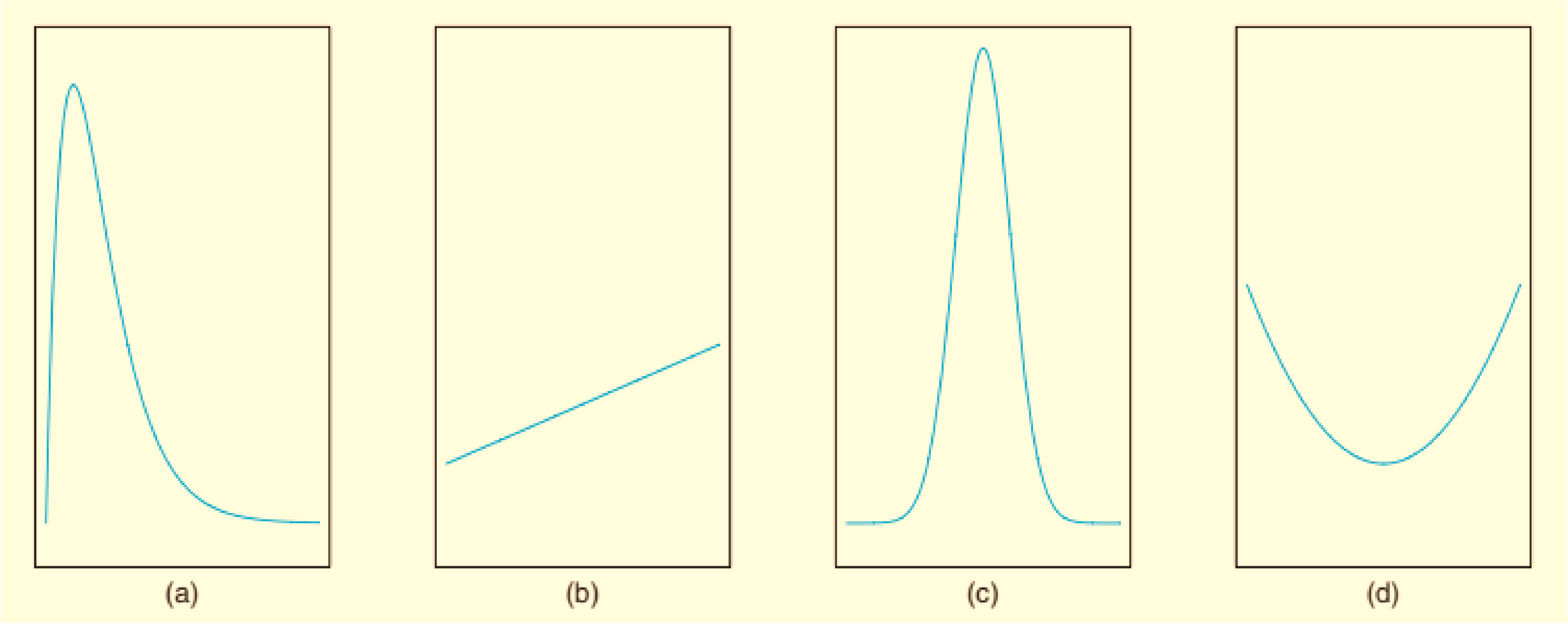
Continuous Probability Distributions

- A continuous random variable has a probability of 0 of assuming *exactly* any of its values.
- Because we are dealing with an interval rather than a point value of our random variable.
- It does not matter whether we include an endpoint of the interval or not. For example:

$$P(a < X \leq b) = P(a < X < b) + P(X = b) = P(a < X < b).$$

- Its probability distribution cannot be given in tabular form.
- *f*(*x*) is usually called the **probability density function**.
- Areas will be used to represent probabilities.

Some graphs of Density Functions



- **A probability density function is constructed so that the area under its curve bounded by the *x* axis is equal to 1 when computed over the range of *X*.**

3.3 Continuous Probability Distributions

$$P(a < X < b) = \int_a^b f(x)\, dx.$$

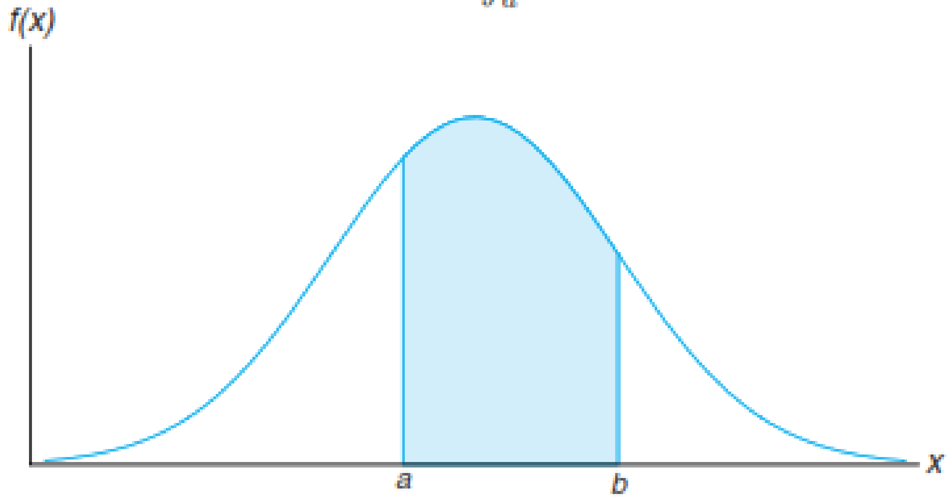


Figure 3.5: $P(a < X < b)$.

Definition 3.6: The function $f(x)$ is a **probability density function** (pdf) for the continuous random variable X, defined over the set of real numbers, if

1. $f(x) \geq 0$, for all $x \in R$.
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$.
3. $P(a < X < b) = \int_a^b f(x)\, dx$.

Definition 3.7: The **cumulative distribution function** $F(x)$ of a continuous random variable X with density function $f(x)$ is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt, \quad \text{for } -\infty < x < \infty.$$

As an immediate consequence of Definition 3.7, one can write the two results

$$P(a < X < b) = F(b) - F(a) \text{ and } f(x) = \frac{dF(x)}{dx},$$

if the derivative exists.

Example 3.11: Suppose that the error in the reaction temperature, in °C, for a controlled laboratory experiment is a continuous random variable X having the probability density function

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

.

(a) Verify that $f(x)$ is a density function.

(b) Find $P(0 < X \leq 1)$.

Example 3.12: For the density function of Example 3.11, find $F(x)$, and use it to evaluate $P(0 < X \leq 1)$.

3.4 Joint Probability Distributions

Definition 3.8: The function $f(x, y)$ is a **joint probability distribution** or **probability mass function** of the discrete random variables X and Y if

1. $f(x, y) \geq 0$ for all (x, y),
2. $\sum\limits_{x} \sum\limits_{y} f(x, y) = 1$,
3. $P(X = x, Y = y) = f(x, y)$.

For any region A in the xy plane, $P[(X, Y) \in A] = \sum \sum\limits_{A} f(x, y)$.

Definition 3.9: The function $f(x, y)$ is a **joint density function** of the continuous random variables X and Y if

1. $f(x, y) \geq 0$, for all (x, y),
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$,
3. $P[(X, Y) \in A] = \int \int_A f(x, y)\, dx\, dy$, for any region A in the xy plane.

Definition 3.10: The **marginal distributions** of X alone and of Y alone are

$$g(x) = \sum_y f(x, y) \quad \text{and} \quad h(y) = \sum_x f(x, y)$$

for the discrete case, and

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \quad \text{and} \quad h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

for the continuous case.

Definition 3.11: Let X and Y be two random variables, discrete or continuous. The **conditional distribution** of the random variable Y given that $X = x$ is

$$f(y|x) = \frac{f(x,y)}{g(x)}, \text{ provided } g(x) > 0.$$

Similarly, the conditional distribution of X given that $Y = y$ is

$$f(x|y) = \frac{f(x,y)}{h(y)}, \text{ provided } h(y) > 0.$$

If we wish to find the probability that the discrete random variable X falls between a and b when it is known that the discrete variable $Y = y$, we evaluate

$$P(a < X < b \mid Y = y) = \sum_{a<x<b} f(x|y),$$

where the summation extends over all values of X between a and b. When X and Y are continuous, we evaluate

$$P(a < X < b \mid Y = y) = \int_a^b f(x|y)\, dx.$$

Example of Joint, Marginal and Conditional Probability

Example 3.14: Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, and 3 green pens. If X is the number of blue pens selected and Y is the number of red pens selected, find

(a) the joint probability function $f(x, y)$,

(b) $P[(X, Y) \in A]$, where A is the region $\{(x, y)|x + y \leq 1\}$.

Example 3.16: Show that the column and row totals of Table 3.1 give the marginal distribution of X alone and of Y alone.

Example 3.18: Referring to Example 3.14, find the conditional distribution of X, given that $Y = 1$, and use it to determine $P(X = 0 \mid Y = 1)$.

Example 3.15: A privately owned business operates both a drive-in facility and a walk-in facility. On a randomly selected day, let X and Y, respectively, be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \le x \le 1, 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Verify condition 2 of Definition 3.9.

(b) Find $P[(X,Y) \in A]$, where $A = \{(x,y) \mid 0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$.

Example 3.19: The joint density for the random variables (X, Y), where X is the unit temperature change and Y is the proportion of spectrum shift that a certain atomic particle produces, is

$$f(x, y) = \begin{cases} 10xy^2, & 0 < x < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the marginal densities $g(x)$, $h(y)$, and the conditional density $f(y|x)$.

(b) Find the probability that the spectrum shifts more than half of the total observations, given that the temperature is increased by 0.25 unit.

Definition 3.12: Let X and Y be two random variables, discrete or continuous, with joint probability distribution $f(x, y)$ and marginal distributions $g(x)$ and $h(y)$, respectively. The random variables X and Y are said to be **statistically independent** if and only if

$$f(x, y) = g(x)h(y)$$

for all (x, y) within their range.

Chapter 4

Mathematical Expectation

Definition 4.1: Let X be a random variable with probability distribution $f(x)$. The **mean**, or **expected value**, of X is

$$\mu = E(X) = \sum_x x f(x)$$

if X is discrete, and

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

if X is continuous.

Example 4.1: A lot containing 7 components is sampled by a quality inspector; the lot contains 4 good components and 3 defective components. A sample of 3 is taken by the inspector. Find the expected value of the number of good components in this sample.

Example 4.3: Let X be the random variable that denotes the life in hours of a certain electronic device. The probability density function is

$$f(x) = \begin{cases} \frac{20{,}000}{x^3}, & x > 100, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the expected life of this type of device.

Theorem 4.1: Let X be a random variable with probability distribution $f(x)$. The expected value of the random variable $g(X)$ is

$$\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x)$$

if X is discrete, and

$$\mu_{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

if X is continuous.

Example 4.4: Suppose that the number of cars X that pass through a car wash between 4:00 P.M. and 5:00 P.M. on any sunny Friday has the following probability distribution:

x	4	5	6	7	8	9
$P(X = x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Let $g(X) = 2X - 1$ represent the amount of money, in dollars, paid to the attendant by the manager. Find the attendant's expected earnings for this particular time period.

Definition 4.2: Let X and Y be random variables with joint probability distribution $f(x, y)$. The mean, or expected value, of the random variable $g(X, Y)$ is

$$\mu_{g(X,Y)} = E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

if X and Y are discrete, and

$$\mu_{g(X,Y)} = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y)\ dx\ dy$$

if X and Y are continuous.

Generalization of Definition 4.2 for the calculation of mathematical expectations of functions of several random variables is straightforward.

Example 4.6: Let X and Y be the random variables with joint probability distribution indicated in Table 3.1 on page 96. Find the expected value of $g(X,Y) = XY$. The table is reprinted here for convenience.

	$f(x,y)$	x = 0	1	2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Example 4.7: Find $E(Y/X)$ for the density function

$$f(x,y) = \begin{cases} \frac{x(1+3y^2)}{4}, & 0 < x < 2,\ 0 < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Definition 4.3: Let X be a random variable with probability distribution $f(x)$ and mean μ. The variance of X is

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x), \qquad \text{if } X \text{ is discrete, and}$$

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx, \qquad \text{if } X \text{ is continuous.}$$

The positive square root of the variance, σ, is called the **standard deviation** of X.

Theorem 4.2: The variance of a random variable X is

$$\sigma^2 = E(X^2) - \mu^2.$$

Example 4.8: Let the random variable X represent the number of automobiles that are used for official business purposes on any given workday. The probability distribution for company A [Figure 4.1(a)] is

x	1	2	3
$f(x)$	0.3	0.4	0.3

and that for company B [Figure 4.1(b)] is

x	0	1	2	3	4
$f(x)$	0.2	0.1	0.3	0.3	0.1

Show that the variance of the probability distribution for company B is greater than that for company A.

Example 4.10: The weekly demand for a drinking-water product, in thousands of liters, from a local chain of efficiency stores is a continuous random variable X having the probability density

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the mean and variance of X.

Theorem 4.3: Let X be a random variable with probability distribution $f(x)$. The variance of the random variable $g(X)$ is

$$\sigma^2_{g(X)} = E\{[g(X) - \mu_{g(X)}]^2\} = \sum_x [g(x) - \mu_{g(X)}]^2 f(x)$$

if X is discrete, and

$$\sigma^2_{g(X)} = E\{[g(X) - \mu_{g(X)}]^2\} = \int_{-\infty}^{\infty} [g(x) - \mu_{g(X)}]^2 f(x)\, dx$$

if X is continuous.

Theorem 4.4: The covariance of two random variables X and Y with means μ_X and μ_Y, respectively, is given by

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y.$$

Definition 4.4: Let X and Y be random variables with joint probability distribution $f(x, y)$. The covariance of X and Y is

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_y) f(x, y)$$

if X and Y are discrete, and

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_y) f(x, y)\, dx\, dy$$

if X and Y are continuous.

Theorem 4.4: The covariance of two random variables X and Y with means μ_X and μ_Y, respectively, is given by

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y.$$

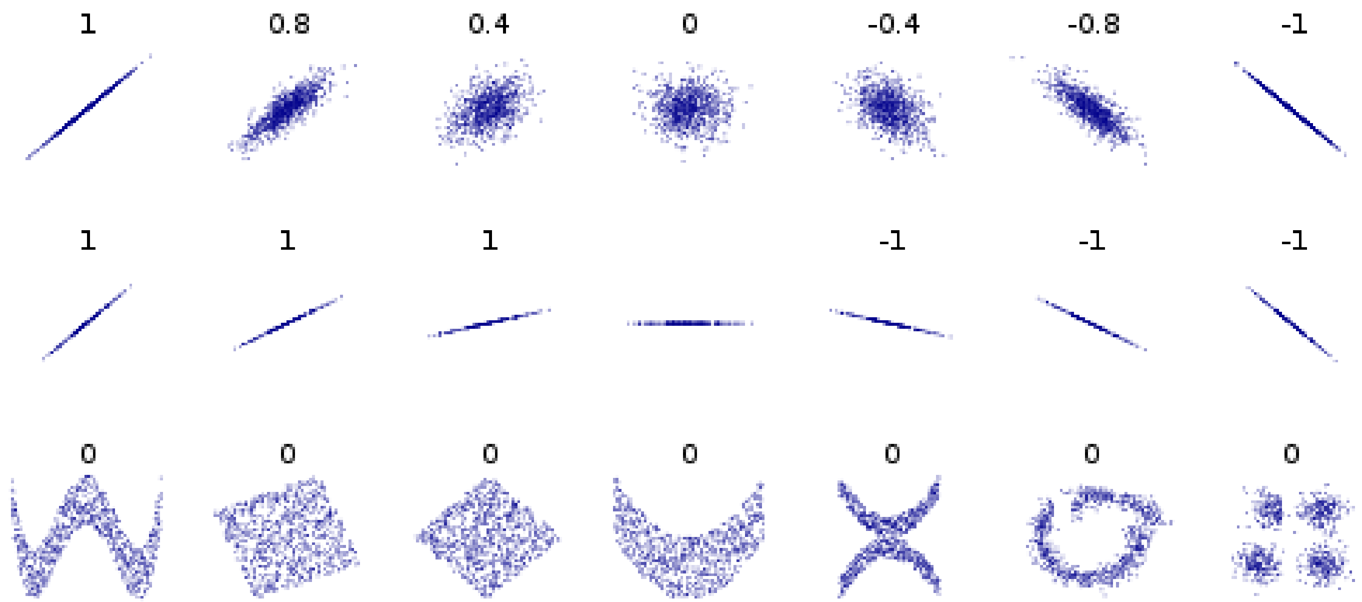
Important points related to variance and Covariance

- The variance or standard deviation has meaning only when we compare two or more distributions that have the same units of measurement.
- It would not be meaningful to compare the variance of a distribution of heights to the variance of a distribution of aptitude scores.

- The covariance between two random variables is a measure of the nature of the association between the two.
- The sign of the covariance indicates whether the relationship between two dependent random variables is positive or negative.
- When X and Y are statistically independent, it can be shown that the covariance is zero. The converse, however, is not generally true.

- The covariance only describes the linear relationship between two random variables. Therefore, if a covariance between X and Y is zero, X and Y may have a nonlinear relationship, which means that they are not necessarily independent.

Example of Linear and Non Linear Relationship



Although the covariance between two random variables does provide information regarding the nature of the relationship, the magnitude of σ_{XY} *does not indicate anything regarding the strength of the relationship*, since σ_{XY} is not scale-free. Its magnitude will depend on the units used to measure both X and Y. There is a scale-free version of the covariance called the **correlation coefficient** that is used widely in statistics.

Definition 4.5: Let X and Y be random variables with covariance σ_{XY} and standard deviations σ_X and σ_Y, respectively. The correlation coefficient of X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

It should be clear to the reader that ρ_{XY} is free of the units of X and Y. The correlation coefficient satisfies the inequality $-1 \leq \rho_{XY} \leq 1$. It assumes a value of zero when $\sigma_{XY} = 0$. Where there is an exact linear dependency, say $Y \equiv a + bX$,

$\rho_{XY} = 1$ if $b > 0$ and $\rho_{XY} = -1$ if $b < 0$. (See Exercise 4.48.) The correlation coefficient is the subject of more discussion in Chapter 12, where we deal with linear regression.