

## CHAPTER 6

# Integration

The origin of integral calculus can be traced back to the ancient Greeks. They were motivated by the need to measure the length of a curve, the area of a surface, or the volume of a solid. Archimedes used techniques very similar to actual integration to determine the length of a segment of a curve. Democritus (410 B.C.) had the insight to consider that a cone was made up of infinitely many plane cross sections parallel to the base.

The theory of integration received very little stimulus after Archimedes's remarkable achievements. It was not until the beginning of the seventeenth century that the interest in Archimedes's ideas began to develop. Johann Kepler (1571–1630) was the first among European mathematicians to develop the ideas of infinitesimals in connection with integration. The use of the term “integral” is due to the Swiss mathematician Johann Bernoulli (1667–1748).

In the present chapter we shall study integration of real-valued functions of a single variable  $x$  according to the concepts put forth by the German mathematician Georg Friedrich Riemann (1826–1866). He was the first to establish a rigorous analytical foundation for integration, based on the older geometric approach.

### 6.1. SOME BASIC DEFINITIONS

Let  $f(x)$  be a function defined and bounded on a finite interval  $[a, b]$ . Suppose that this interval is partitioned into a finite number of subintervals by a set of points  $P = \{x_0, x_1, \dots, x_n\}$  such that  $a = x_0 < x_1 < x_2 < \dots < x_n = b$ . This set is called a partition of  $[a, b]$ . Let  $\Delta x_i = x_i - x_{i-1}$  ( $i = 1, 2, \dots, n$ ), and  $\Delta_p$  be the largest of  $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ . This value is called the norm of  $P$ . Consider the sum

$$S(P, f) = \sum_{i=1}^n f(t_i) \Delta x_i,$$

where  $t_i$  is a point in the subinterval  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ .

The function  $f(x)$  is said to be Riemann integrable on  $[a, b]$  if a number  $A$  exists with the following property: For any given  $\epsilon > 0$  there exists a number  $\delta > 0$  such that

$$|A - S(P, f)| < \epsilon$$

for any partition  $P$  of  $[a, b]$  with a norm  $\Delta_p < \delta$ , and for any choice of the point  $t_i$  in  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ . The number  $A$  is called the Riemann integral of  $f(x)$  on  $[a, b]$  and is denoted by  $\int_a^b f(x) dx$ . The integration symbol  $\int$  was first used by the German mathematician Gottfried Wilhelm Leibniz (1646–1716) to represent a sum (it was derived from the first letter of the Latin word *summa*, which means a sum).

## 6.2. THE EXISTENCE OF THE RIEMANN INTEGRAL

In order to investigate the existence of the Riemann integral, we shall need the following theorem:

**Theorem 6.2.1.** Let  $f(x)$  be a bounded function on a finite interval,  $[a, b]$ . For every partition  $P = \{x_0, x_1, \dots, x_n\}$  of  $[a, b]$ , let  $m_i$  and  $M_i$  be, respectively, the infimum and supremum of  $f(x)$  on  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ . If, for a given  $\epsilon > 0$ , there exists a  $\delta > 0$  such that

$$US_P(f) - LS_P(f) < \epsilon \quad (6.1)$$

whenever  $\Delta_p < \delta$ , where  $\Delta_p$  is the norm of  $P$ , and

$$LS_P(f) = \sum_{i=1}^n m_i \Delta x_i,$$

$$US_P(f) = \sum_{i=1}^n M_i \Delta x_i,$$

then  $f(x)$  is Riemann integrable on  $[a, b]$ . Conversely, if  $f(x)$  is Riemann integrable, then inequality (6.1) holds for any partition  $P$  such that  $\Delta_p < \delta$ . [The sums,  $LS_P(f)$  and  $US_P(f)$ , are called the lower sum and upper sum, respectively, of  $f(x)$  with respect to the partition  $P$ .]

In order to prove Theorem 6.2.1 we need the following lemmas:

**Lemma 6.2.1.** Let  $P$  and  $P'$  be two partitions of  $[a, b]$  such that  $P' \supset P$  ( $P'$  is called a refinement of  $P$  and is constructed by adding partition points between those that belong to  $P$ ). Then

$$US_{P'}(f) \leq US_P(f),$$

$$LS_{P'}(f) \geq LS_P(f).$$

*Proof.* Let  $P = \{x_0, x_1, \dots, x_n\}$ . By the nature of the partition  $P'$ , the  $i$ th subinterval  $\Delta x_i = x_i - x_{i-1}$  is divided into  $k_i$  parts  $\Delta_{x_i}^{(1)}, \Delta_{x_i}^{(2)}, \dots, \Delta_{x_i}^{(k_i)}$ , where  $k_i \geq 1$ ,  $i = 1, 2, \dots, n$ . If  $m_i^{(j)}$  and  $M_i^{(j)}$  denote, respectively, the infimum and supremum of  $f(x)$  on  $\Delta_{x_i}^{(j)}$ , then  $m_i \leq m_i^{(j)} \leq M_i^{(j)} \leq M_i$  for  $j = 1, 2, \dots, k_i$ ;  $i = 1, 2, \dots, n$ , where  $m_i$  and  $M_i$  are the infimum and supremum of  $f(x)$  on  $[x_{i-1}, x_i]$ , respectively. It follows that

$$LS_P(f) = \sum_{i=1}^n m_i \Delta x_i \leq \sum_{i=1}^n \sum_{j=1}^{k_i} m_i^{(j)} \Delta_{x_i}^{(j)} = LS_{P'}(f)$$

$$US_{P'}(f) = \sum_{i=1}^n \sum_{j=1}^{k_i} M_i^{(j)} \Delta_{x_i}^{(j)} \leq \sum_{i=1}^n M_i \Delta x_i = US_P(f). \quad \square$$

**Lemma 6.2.2.** Let  $P$  and  $P'$  be any two partitions of  $[a, b]$ . Then  $LS_P(f) \leq US_{P'}(f)$ .

*Proof.* Let  $P'' = P \cup P'$ . The partition  $P''$  is a refinement of both  $P$  and  $P'$ . Then, by Lemma 6.2.1,

$$LS_P(f) \leq LS_{P''}(f) \leq US_{P''}(f) \leq US_{P'}(f). \quad \square$$

*Proof of Theorem 6.2.1*

Let  $\epsilon > 0$  be given. Suppose that inequality (6.1) holds for any partition  $P$  whose norm  $\Delta_P$  is less than  $\delta$ . Let  $S(P, f) = \sum_{i=1}^n f(t_i) \Delta x_i$ , where  $t_i$  is a point in  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ . By the definition of  $LS_P(f)$  and  $US_P(f)$  we can write

$$LS_P(f) \leq S(P, f) \leq US_P(f). \quad (6.2)$$

Let  $m$  and  $M$  be the infimum and supremum, respectively, of  $f(x)$  on  $[a, b]$ ; then

$$m(b-a) \leq LS_P(f) \leq US_P(f) \leq M(b-a). \quad (6.3)$$

Let us consider two sets of lower and upper sums of  $f(x)$  with respect to partitions  $P, P', P'', \dots$  such that  $P \subset P' \subset P'' \subset \dots$ . Then, by Lemma 6.2.1, the set of upper sums is decreasing, and the set of lower sums is increasing. Furthermore, because of (6.3), the set of upper sums is bounded from below by  $m(b-a)$ , and the set of lower sums is bounded from above by  $M(b-a)$ . Hence, the infimum of  $US_P(f)$  and the supremum of  $LS_P(f)$  with respect to  $P$  do exist (see Theorem 1.5.1).

From Lemma 6.2.2 it is easy to deduce that

$$\sup_P LS_P(f) \leq \inf_P US_P(f).$$

Now, suppose that for the given  $\epsilon > 0$  there exists a  $\delta > 0$  such that

$$US_P(f) - LS_P(f) < \epsilon \quad (6.4)$$

for any partition whose norm  $\Delta_P$  is less than  $\delta$ . We have that

$$LS_P(f) \leq \sup_P LS_P(f) \leq \inf_P US_P(f) \leq US_P(f). \quad (6.5)$$

Hence,

$$\inf_P US_P(f) - \sup_P LS_P(f) < \epsilon.$$

Since  $\epsilon > 0$  is arbitrary, we conclude that if (6.1) is satisfied, then

$$\inf_P US_P(f) = \sup_P LS_P(f). \quad (6.6)$$

Furthermore, from (6.2), (6.4), and (6.5) we obtain

$$|S(P, f) - A| < \epsilon,$$

where  $A$  is the common value of  $\inf_P US_P(f)$  and  $\sup_P LS_P(f)$ . This proves that  $A$  is the Riemann integral of  $f(x)$  on  $[a, b]$ .

Let us now show that the converse of the theorem is true, that is, if  $f(x)$  is Riemann integrable on  $[a, b]$ , then inequality (6.1) holds.

If  $f(x)$  is Riemann integrable, then for a given  $\epsilon > 0$  there exists a  $\delta > 0$  such that

$$\left| \sum_{i=1}^n f(t_i) \Delta x_i - A \right| < \frac{\epsilon}{3} \quad (6.7)$$

and

$$\left| \sum_{i=1}^n f(t'_i) \Delta x_i - A \right| < \frac{\epsilon}{3} \quad (6.8)$$

for any partition  $P = \{x_0, x_1, \dots, x_n\}$  of  $[a, b]$  with a norm  $\Delta_P < \delta$ , and any choices of  $t_i, t'_i$  in  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ , where  $A = \int_a^b f(x) dx$ . From (6.7) and (6.8) we then obtain

$$\left| \sum_{i=1}^n [f(t_i) - f(t'_i)] \Delta x_i \right| < \frac{2\epsilon}{3}.$$

Now,  $M_i - m_i$  is the supremum of  $f(x) - f(x')$  for  $x, x'$  in  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ . It follows that for a given  $\eta > 0$  we can choose  $t_i, t'_i$  in  $[x_{i-1}, x_i]$  so that

$$f(t_i) - f(t'_i) > M_i - m_i - \eta, \quad i = 1, 2, \dots, n,$$

for otherwise  $M_i - m_i - \eta$  would be an upper bound for  $f(x) - f(x')$  for all  $x, x'$  in  $[x_{i-1}, x_i]$ , which is a contradiction. In particular, if  $\eta = \epsilon/[3(b-a)]$ , then we can find  $t_i, t'_i$  in  $[x_{i-1}, x_i]$  such that

$$\begin{aligned} US_P(f) - LS_P(f) &= \sum_{i=1}^n (M_i - m_i) \Delta x_i \\ &< \sum_{i=1}^n [f(t_i) - f(t'_i)] \Delta x_i + \eta(b-a) \\ &< \epsilon. \end{aligned}$$

This proves the validity of inequality (6.1).  $\square$

**Corollary 6.2.1.** Let  $f(x)$  be a bounded function on  $[a, b]$ . Then  $f(x)$  is Riemann integrable on  $[a, b]$  if and only if  $\inf_P US_P(f) = \sup_P LS_P(f)$ , where  $LS_P(f)$  and  $US_P(f)$  are, respectively, the lower and upper sums of  $f(x)$  with respect to a partition  $P$  of  $[a, b]$ .

*Proof.* See Exercise 6.1.  $\square$

EXAMPLE 6.2.1. Let  $f(x): [0, 1] \rightarrow \mathbb{R}$  be the function  $f(x) = x^2$ . Then,  $f(x)$  is Riemann integrable on  $[0, 1]$ . To show this, let  $P = \{x_0, x_1, \dots, x_n\}$  be any partition of  $[0, 1]$ , where  $x_0 = 0, x_n = 1$ . Then

$$\begin{aligned} LS_P(f) &= \sum_{i=1}^n x_{i-1}^2 \Delta x_i, \\ US_P(f) &= \sum_{i=1}^n x_i^2 \Delta x_i. \end{aligned}$$

Hence,

$$\begin{aligned} US_P(f) - LS_P(f) &= \sum_{i=1}^n (x_i^2 - x_{i-1}^2) \Delta x_i \\ &\leq \Delta_P \sum_{i=1}^n (x_i^2 - x_{i-1}^2), \end{aligned}$$

where  $\Delta_P$  is the norm of  $P$ . But

$$\sum_{i=1}^n (x_i^2 - x_{i-1}^2) = x_n^2 - x_0^2 = 1.$$

Thus

$$US_P(f) - LS_P(f) \leq \Delta_P.$$

It follows that for a given  $\epsilon > 0$  we can choose  $\delta = \epsilon$  such that for any partition  $P$  whose norm  $\Delta_p$  is less than  $\delta$ ,

$$US_P(f) - LS_P(f) < \epsilon.$$

By Theorem 6.2.1,  $f(x) = x^2$  is Riemann integrable on  $[0, 1]$ .

**EXAMPLE 6.2.2.** Consider the function  $f(x): [0, 1] \rightarrow R$  such that  $f(x) = 0$  if  $x$  a rational number and  $f(x) = 1$  if  $x$  is irrational. Since every subinterval of  $[0, 1]$  contains both rational and irrational numbers, then for any partition  $P = \{x_0, x_1, \dots, x_n\}$  of  $[0, 1]$  we have

$$\begin{aligned} US_P(f) &= \sum_{i=1}^n M_i \Delta x_i = \sum_{i=1}^n \Delta x_i = 1, \\ LS_P(f) &= \sum_{i=1}^n m_i \Delta x_i = \sum_{i=1}^n 0 \Delta x_i = 0. \end{aligned}$$

It follows that

$$\inf_P US_P(f) = 1 \quad \text{and} \quad \sup_P LS_P(f) = 0.$$

By Corollary 6.2.1,  $f(x)$  is not Riemann integrable on  $[0, 1]$ .

### 6.3. SOME CLASSES OF FUNCTIONS THAT ARE RIEMANN INTEGRABLE

There are certain classes of functions that are Riemann integrable. Identifying a given function as a member of such a class can facilitate the determination of its Riemann integrability. Some of these classes of functions include: (i) continuous functions; (ii) monotone functions; (iii) functions of bounded variation.

**Theorem 6.3.1.** If  $f(x)$  is continuous on  $[a, b]$ , then it is Riemann integrable there.

*Proof.* Since  $f(x)$  is continuous on a closed and bounded interval, then by Theorem 3.4.6 it must be uniformly continuous on  $[a, b]$ . Consequently, for a given  $\epsilon > 0$  there exists a  $\delta > 0$  that depends only on  $\epsilon$  such that for any  $x_1, x_2$  in  $[a, b]$  we have

$$|f(x_1) - f(x_2)| < \frac{\epsilon}{b - a}$$

if  $|x_1 - x_2| < \delta$ . Let  $P = \{x_0, x_1, \dots, x_n\}$  be a partition of  $P$  with a norm  $\Delta_p < \delta$ . Then

$$US_P(f) - LS_P(f) = \sum_{i=1}^n (M_i - m_i) \Delta x_i,$$

where  $m_i$  and  $M_i$  are, respectively, the infimum and supremum of  $f(x)$  on  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ . By Corollary 3.4.1 there exist points  $\xi_i, \eta_i$  in  $[x_{i-1}, x_i]$  such that  $m_i = f(\xi_i)$ ,  $M_i = f(\eta_i)$ ,  $i = 1, 2, \dots, n$ . Since  $|\eta_i - \xi_i| \leq \Delta_p < \delta$  for  $i = 1, 2, \dots, n$ , then

$$\begin{aligned} US_P(f) - LS_P(f) &= \sum_{i=1}^n [f(\eta_i) - f(\xi_i)] \Delta x_i \\ &< \frac{\epsilon}{b-a} \sum_{i=1}^n \Delta x_i = \epsilon. \end{aligned}$$

By Theorem 6.2.1 we conclude that  $f(x)$  is Riemann integrable on  $[a, b]$ .  $\square$

It should be noted that continuity is a sufficient condition for Riemann integrability, but is not a necessary one. A function  $f(x)$  can have discontinuities in  $[a, b]$  and still remains Riemann integrable on  $[a, b]$ . For example, consider the function

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

This function is discontinuous at  $x = 0$ . However, it is Riemann integrable on  $[-1, 1]$ . To show this, let  $\epsilon > 0$  be given, and let  $P = \{x_0, x_1, \dots, x_n\}$  be a partition of  $[-1, 1]$  such that  $\Delta_p < \epsilon/2$ . By the nature of this function,  $f(x_i) - f(x_{i-1}) \geq 0$ , and the infimum and supremum of  $f(x)$  on  $[x_{i-1}, x_i]$  are equal to  $f(x_{i-1})$  and  $f(x_i)$ , respectively,  $i = 1, 2, \dots, n$ . Hence,

$$\begin{aligned} US_P(f) - LS_P(f) &= \sum_{i=1}^n (M_i - m_i) \Delta x_i \\ &= \sum_{i=1}^n [f(x_i) - f(x_{i-1})] \Delta x_i \\ &< \Delta_p \sum_{i=1}^n [f(x_i) - f(x_{i-1})] = \Delta_p [f(1) - f(-1)] \\ &< \frac{\epsilon}{2} [f(1) - f(-1)] = \epsilon. \end{aligned}$$

The function  $f(x)$  is therefore Riemann integrable on  $[-1, 1]$  by Theorem 6.2.1.

On the basis of this example it is now easy to prove the following theorem:

**Theorem 6.3.2.** If  $f(x)$  is monotone increasing (or monotone decreasing) on  $[a, b]$ , then it is Riemann integrable there.

Theorem 6.3.2 can be used to construct a function that has a countable number of discontinuities in  $[a, b]$  and is also Riemann integrable (see Exercise 6.2).

### 6.3.1. Functions of Bounded Variation

Let  $f(x)$  be defined on  $[a, b]$ . This function is said to be of bounded variation on  $[a, b]$  if there exists a number  $M > 0$  such that for any partition  $P = \{x_0, x_1, \dots, x_n\}$  of  $[a, b]$  we have

$$\sum_{i=1}^n |\Delta f_i| \leq M,$$

where  $\Delta f_i = f(x_i) - f(x_{i-1})$ ,  $i = 1, 2, \dots, n$ .

Any function that is monotone increasing (or decreasing) on  $[a, b]$  is also of bounded variation there. To show this, let  $f(x)$  be monotone increasing on  $[a, b]$ . Then

$$\sum_{i=1}^n |\Delta f_i| = \sum_{i=1}^n [f(x_i) - f(x_{i-1})] = f(b) - f(a).$$

Hence, if  $M$  is any number greater than or equal to  $f(b) - f(a)$ , then  $\sum_{i=1}^n |\Delta f_i| \leq M$  for any partition  $P$  of  $[a, b]$ .

Another example of a function of bounded variation is given in the next theorem.

**Theorem 6.3.3.** If  $f(x)$  is continuous on  $[a, b]$  and its derivative  $f'(x)$  exists and is bounded on  $(a, b)$ , then  $f(x)$  is of bounded variation on  $[a, b]$ .

*Proof.* Let  $P = \{x_0, x_1, \dots, x_n\}$  be a partition of  $[a, b]$ . By applying the mean value theorem (Theorem 4.2.2) on each  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ , we obtain

$$\begin{aligned} \sum_{i=1}^n |\Delta f_i| &= \sum_{i=1}^n |f'(\xi_i) \Delta x_i| \\ &\leq K \sum_{i=1}^n (x_i - x_{i-1}) = K(b - a), \end{aligned}$$



where  $x_{i-1} < \xi_i < x_i$ ,  $i = 1, 2, \dots, n$ , and  $K > 0$  is such that  $|f'(x)| \leq K$  on  $(a, b)$ .  $\square$

It should be noted that any function of bounded variation on  $[a, b]$  is also bounded there. This is true because if  $a < x < b$ , then  $P = \{a, x, b\}$  is a partition of  $[a, b]$ . Hence,

$$|f(x) - f(a)| + |f(b) - f(x)| \leq M.$$

for some positive number  $M$ . This implies that  $|f(x)|$  is bounded on  $[a, b]$  since

$$\begin{aligned} |f(x)| &\leq \frac{1}{2} [|f(x) - f(a)| + |f(x) - f(b)| + |f(a) + f(b)|] \\ &\leq \frac{1}{2} [M + |f(a) + f(b)|]. \end{aligned}$$

The converse of this result, however, is not necessarily true, that is, if  $f(x)$  is bounded, then it may not be of bounded variation. For example, the function

$$f(x) = \begin{cases} x \cos\left(\frac{\pi}{2x}\right), & 0 < x \leq 1, \\ 0, & x = 0 \end{cases}$$

is bounded on  $[0, 1]$ , but is not of bounded variation there. It can be shown that for the partition

$$P = \left\{ 0, \frac{1}{2n}, \frac{1}{2n-1}, \dots, \frac{1}{3}, \frac{1}{2}, 1 \right\},$$

$\sum_{i=1}^{2n} |\Delta f_i| \rightarrow \infty$  as  $n \rightarrow \infty$  and hence cannot be bounded by a constant  $M$  for all  $n$  (see Exercise 6.4).

**Theorem 6.3.4.** If  $f(x)$  is of bounded variation on  $[a, b]$ , then it is Riemann integrable there.

*Proof.* Let  $\epsilon > 0$  be given, and let  $P = \{x_0, x_1, \dots, x_n\}$  be a partition of  $[a, b]$ . Then

$$US_P(f) - LS_P(f) = \sum_{i=1}^n (M_i - m_i) \Delta x_i, \quad (6.9)$$

where  $m_i$  and  $M_i$  are the infimum and supremum of  $f(x)$  on  $[x_{i-1}, x_i]$ , respectively,  $i = 1, 2, \dots, n$ . By the properties of  $m_i$  and  $M_i$ , there exist  $\xi_i$

and  $\eta_i$  in  $[x_{i-1}, x_i]$  such that for  $i = 1, 2, \dots, n$ ,

$$\begin{aligned} m_i &\leq f(\xi_i) < m_i + \epsilon', \\ M_i - \epsilon' &< f(\eta_i) \leq M_i, \end{aligned}$$

where  $\epsilon'$  is a small positive number to be determined later. It follows that

$$M_i - m_i - 2\epsilon' < f(\eta_i) - f(\xi_i) \leq M_i - m_i, \quad i = 1, 2, \dots, n.$$

Hence,

$$\begin{aligned} M_i - m_i &< 2\epsilon' + f(\eta_i) - f(\xi_i) \\ &\leq 2\epsilon' + |f(\xi_i) - f(\eta_i)|, \quad i = 1, 2, \dots, n. \end{aligned}$$

From formula (6.9) we obtain

$$US_P(f) - LS_P(f) < 2\epsilon' \sum_{i=1}^n \Delta x_i + \sum_{i=1}^n |f(\xi_i) - f(\eta_i)| \Delta x_i. \quad (6.10)$$

Now, if  $\Delta_p$  is the norm of  $P$ , then

$$\begin{aligned} \sum_{i=1}^n |f(\xi_i) - f(\eta_i)| \Delta x_i &\leq \Delta_p \sum_{i=1}^n |f(\xi_i) - f(\eta_i)| \\ &\leq \Delta_p \sum_{i=1}^m |f(z_i) - f(z_{i-1})|, \end{aligned} \quad (6.11)$$

where  $\{z_0, z_1, \dots, z_m\}$  is a partition  $Q$  of  $[a, b]$ , which consists of the points  $x_0, x_1, \dots, x_n$  as well as the points  $\xi_1, \eta_1, \xi_2, \eta_2, \dots, \xi_n, \eta_n$ , that is,  $Q$  is a refinement of  $P$  obtained by adding the  $\xi_i$ 's and  $\eta_i$ 's ( $i = 1, 2, \dots, n$ ). Since  $f(x)$  is of bounded variation on  $[a, b]$ , there exists a number  $M > 0$  such that

$$\sum_{i=1}^m |f(z_i) - f(z_{i-1})| \leq M. \quad (6.12)$$

From (6.10), (6.11), and (6.12) it follows that

$$US_P(f) - LS_P(f) < 2\epsilon'(b-a) + M\Delta_p. \quad (6.13)$$

Let us now select the partition  $P$  such that  $\Delta_p < \delta$ , where  $M\delta < \epsilon/2$ . If we also choose  $\epsilon'$  such that  $2\epsilon'(b-a) < \epsilon/2$ , then from (6.13) we obtain  $US_P(f) - LS_P(f) < \epsilon$ . The function  $f(x)$  is therefore Riemann integrable on  $[a, b]$  by Theorem 6.2.1.  $\square$

## 6.4. PROPERTIES OF THE RIEMANN INTEGRAL

The Riemann integral has several properties that are useful at both the theoretical and practical levels. Most of these properties are fairly simple and straightforward. We shall therefore not prove every one of them in this section.

**Theorem 6.4.1.** If  $f(x)$  and  $g(x)$  are Riemann integrable on  $[a, b]$  and if  $c_1$  and  $c_2$  are constants, then  $c_1f(x) + c_2g(x)$  is Riemann integrable on  $[a, b]$ , and

$$\int_a^b [c_1f(x) + c_2g(x)] dx = c_1 \int_a^b f(x) dx + c_2 \int_a^b g(x) dx.$$

**Theorem 6.4.2.** If  $f(x)$  is Riemann integrable on  $[a, b]$ , and  $m \leq f(x) \leq M$  for all  $x$  in  $[a, b]$ , then

$$m(b-a) \leq \int_a^b f(x) dx \leq M(b-a).$$

**Theorem 6.4.3.** If  $f(x)$  and  $g(x)$  are Riemann integrable on  $[a, b]$ , and if  $f(x) \leq g(x)$  for all  $x$  in  $[a, b]$ , then

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

**Theorem 6.4.4.** If  $f(x)$  is Riemann integrable on  $[a, b]$  and if  $a < c < b$ , then

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

**Theorem 6.4.5.** If  $f(x)$  is Riemann integrable on  $[a, b]$ , then so is  $|f(x)|$  and

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

*Proof.* Let  $P = \{x_0, x_1, \dots, x_n\}$  be a partition of  $[a, b]$ . Let  $m_i$  and  $M_i$  be the infimum and supremum of  $f(x)$ , respectively, on  $[x_{i-1}, x_i]$ ; and let  $m'_i, M'_i$  be the same for  $|f(x)|$ . We claim that

$$M_i - m_i \geq M'_i - m'_i, \quad i = 1, 2, \dots, n.$$

It is obvious that  $M_i - m_i = M'_i - m'_i$  if  $f(x)$  is either nonnegative or nonpositive for all  $x$  in  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ . Let us therefore suppose that  $f(x)$  is

negative on  $D_i^-$  and nonnegative on  $D_i^+$ , where  $D_i^-$  and  $D_i^+$  are such that  $D_i^- \cup D_i^+ = D_i = [x_{i-1}, x_i]$  for  $i = 1, 2, \dots, n$ . We then have

$$\begin{aligned} M_i - m_i &= \sup_{D_i^+} f(x) - \inf_{D_i^-} f(x) \\ &= \sup_{D_i^+} |f(x)| - \inf_{D_i^-} (-|f(x)|) \\ &= \sup_{D_i^+} |f(x)| + \sup_{D_i^-} |f(x)| \\ &\geq \sup_{D_i} |f(x)| = M'_i, \end{aligned}$$

since  $\sup_{D_i} |f(x)| = \max\{\sup_{D_i^+} |f(x)|, \sup_{D_i^-} |f(x)|\}$ . Hence,  $M_i - m_i \geq M'_i \geq M'_i - m'_i$  for  $i = 1, 2, \dots, n$ , which proves our claim.

$$US_P(|f|) - LS_P(|f|) = \sum_{i=1}^n (M'_i - m'_i) \Delta x_i \leq \sum_{i=1}^n (M_i - m_i) \Delta x_i.$$

Hence,

$$US_P(|f|) - LS_P(|f|) \leq US_P(f) - LS_P(f). \quad (6.14)$$

Since  $f(x)$  is Riemann integrable, the right-hand side of inequality (6.14) can be made smaller than any given  $\epsilon > 0$  by a proper choice of the norm  $\Delta_P$  of  $P$ . It follows that  $|f(x)|$  is Riemann integrable on  $[a, b]$  by Theorem 6.2.1. Furthermore, since  $\mp f(x) \leq |f(x)|$  for all  $x$  in  $[a, b]$ , then  $\int_a^b \mp f(x) dx \leq \int_a^b |f(x)| dx$  by Theorem 6.4.3, that is,

$$\mp \int_a^b f(x) dx \leq \int_a^b |f(x)| dx.$$

Thus,  $|\int_a^b f(x) dx| \leq \int_a^b |f(x)| dx$ .  $\square$

**Corollary 6.4.1.** If  $f(x)$  is Riemann integrable on  $[a, b]$ , then so is  $f^2(x)$ .

*Proof.* Using the same notation as in the proof of Theorem 6.4.5, we have that  $m_i'^2$  and  $M_i'^2$  are, respectively, the infimum and supremum of  $f^2(x)$  on  $[x_{i-1}, x_i]$  for  $i = 1, 2, \dots, n$ . Now,

$$\begin{aligned} M_i'^2 - m_i'^2 &= (M'_i - m'_i)(M'_i + m'_i) \\ &\leq 2M'(M'_i - m'_i) \\ &\leq 2M'(M_i - m_i), \quad i = 1, 2, \dots, n, \end{aligned} \quad (6.15)$$

where  $M'$  is the supremum of  $|f(x)|$  on  $[a, b]$ . The Riemann integrability of  $f^2(x)$  now follows from inequality (6.15) by the Riemann integrability of  $f(x)$ .  $\square$

**Corollary 6.4.2.** If  $f(x)$  and  $g(x)$  are Riemann integrable on  $[a, b]$ , then so is their product  $f(x)g(x)$ .

*Proof.* This follows directly from the identity

$$4f(x)g(x) = [f(x) + g(x)]^2 - [f(x) - g(x)]^2, \quad (6.16)$$

and the fact that the squares of  $f(x) + g(x)$  and  $f(x) - g(x)$  are Riemann integrable on  $[a, b]$  by Theorem 6.4.1 and Corollary 6.4.1.  $\square$

**Theorem 6.4.6** (The Mean Value Theorem for Integrals). If  $f(x)$  is continuous on  $[a, b]$ , then there exists a point  $c \in [a, b]$  such that

$$\int_a^b f(x) dx = (b - a)f(c).$$

*Proof.* By Theorem 6.4.2 we have

$$m \leq \frac{1}{b - a} \int_a^b f(x) dx \leq M,$$

where  $m$  and  $M$  are, respectively, the infimum and supremum of  $f(x)$  on  $[a, b]$ . Since  $f(x)$  is continuous, then by Corollary 3.4.1 it must attain the values  $m$  and  $M$  at some points inside  $[a, b]$ . Furthermore, by the intermediate-value theorem (Theorem 3.4.4),  $f(x)$  assumes every value between  $m$  and  $M$ . Hence, there is a point  $c \in [a, b]$  such that

$$f(c) = \frac{1}{b - a} \int_a^b f(x) dx. \quad \square$$

**Definition 6.4.1.** Let  $f(x)$  be Riemann integrable on  $[a, b]$ . The function

$$F(x) = \int_a^x f(t) dt, \quad a \leq x \leq b,$$

is called an indefinite integral of  $f(x)$ .  $\square$

**Theorem 6.4.7.** If  $f(x)$  is Riemann integrable on  $[a, b]$ , then  $F(x) = \int_a^x f(t) dt$  is uniformly continuous on  $[a, b]$ .

*Proof.* Let  $x_1, x_2$  be in  $[a, b]$ ,  $x_1 < x_2$ . Then,

$$\begin{aligned} |F(x_2) - F(x_1)| &= \left| \int_a^{x_2} f(t) dt - \int_a^{x_1} f(t) dt \right| \\ &= \left| \int_{x_1}^{x_2} f(t) dt \right|, \quad \text{by Theorem 6.4.4} \\ &\leq \int_{x_1}^{x_2} |f(t)| dt, \quad \text{by Theorem 6.4.5} \\ &\leq M'(x_2 - x_1), \end{aligned}$$

where  $M'$  is the supremum of  $|f(x)|$  on  $[a, b]$ . Thus if  $\epsilon > 0$  is given, then  $|F(x_2) - F(x_1)| < \epsilon$  provided that  $|x_1 - x_2| < \epsilon/M'$ . This proves uniform continuity of  $F(x)$  on  $[a, b]$ .  $\square$

The next theorem presents a practical way for evaluating the Riemann integral on  $[a, b]$ .

**Theorem 6.4.8.** Suppose that  $f(x)$  is continuous on  $[a, b]$ . Let  $F(x) = \int_a^x f(t) dt$ . Then we have the following:

- i.  $dF(x)/dx = f(x)$ ,  $a \leq x \leq b$ .
- ii.  $\int_a^b f(x) dx = G(b) - G(a)$ , where  $G(x) = F(x) + c$ , and  $c$  is an arbitrary constant.

*Proof.* We have

$$\begin{aligned} \frac{dF(x)}{dx} &= \frac{d}{dx} \int_a^x f(t) dt = \lim_{h \rightarrow 0} \frac{1}{h} \left[ \int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f(t) dt, \quad \text{by Theorem 6.4.4} \\ &= \lim_{h \rightarrow 0} f(x + \theta h), \end{aligned}$$

by Theorem 6.4.6, where  $0 \leq \theta \leq 1$ . Hence,

$$\frac{dF(x)}{dx} = \lim_{h \rightarrow 0} f(x + \theta h) = f(x)$$

by the continuity of  $f(x)$ . This result indicates that an indefinite integral of  $f(x)$  is any function whose derivative is equal to  $f(x)$ . It is therefore unique up to a constant. Thus both  $F(x)$  and  $F(x) + c$ , where  $c$  is an arbitrary constant, are considered to be indefinite integrals.

To prove the second part of the theorem, let  $G(x)$  be defined on  $[a, b]$  as

$$G(x) = F(x) + c = \int_a^x f(t) dt + c,$$

that is,  $G(x)$  is an indefinite integral of  $f(x)$ . If  $x = a$ , then  $G(a) = c$ , since  $F(a) = 0$ . Also, if  $x = b$ , then  $G(b) = F(b) + c = \int_a^b f(t) dt + G(a)$ . It follows that

$$\int_a^b f(t) dt = G(b) - G(a). \quad \square$$

This result is known as the *fundamental theorem of calculus*. It is generally attributed to Isaac Barrow (1630–1677), who was the first to realize that differentiation and integration are inverse operations. One advantage of this theorem is that it provides a practical way to evaluate the integral of  $f(x)$  on  $[a, b]$ .

#### 6.4.1. Change of Variables in Riemann Integration

There are situations in which the variable  $x$  in a Riemann integral is a function of some other variable, say  $u$ . In this case, it may be of interest to determine how the integral can be expressed and evaluated under the given transformation. One advantage of this change of variable is the possibility of simplifying the actual evaluation of the integral, provided that the transformation is properly chosen.

**Theorem 6.4.9.** Let  $f(x)$  be continuous on  $[\alpha, \beta]$ , and let  $x = g(u)$  be a function whose derivative  $g'(u)$  exists and is continuous on  $[c, d]$ . Suppose that the range of  $g$  is contained inside  $[\alpha, \beta]$ . If  $a, b$  are points in  $[\alpha, \beta]$  such that  $a = g(c)$  and  $b = g(d)$ , then

$$\int_a^b f(x) dx = \int_c^d f[g(u)] g'(u) du.$$

*Proof.* Let  $F(x) = \int_a^x f(t) dt$ . By Theorem 6.4.8,  $F'(x) = f(x)$ . Let  $G(u)$  be defined as

$$G(u) = \int_c^u f[g(t)] g'(t) dt.$$

Since  $f$ ,  $g$ , and  $g'$  are continuous, then by Theorem 6.4.8 we have

$$\frac{dG(u)}{du} = f[g(u)] g'(u). \quad (6.17)$$

However, according to the chain rule (Theorem 4.1.3),

$$\begin{aligned}\frac{dF[g(u)]}{du} &= \frac{dF[g(u)]}{dg(u)} \frac{dg(u)}{du} \\ &= f[g(u)]g'(u).\end{aligned}\tag{6.18}$$

From formulas (6.17) and (6.18) we conclude that

$$G(u) - F[g(u)] = \lambda, \tag{6.19}$$

where  $\lambda$  is a constant. If  $a$  and  $b$  are points in  $[\alpha, \beta]$  such that  $a = g(c)$ ,  $b = g(d)$ , then when  $u = c$ , we have  $G(c) = 0$  and  $\lambda = -F[g(c)] = -F(a) = 0$ . Furthermore, when  $u = d$ ,  $G(d) = \int_c^d f[g(t)]g'(t) dt$ . From (6.19) we then obtain

$$\begin{aligned}G(d) &= \int_c^d f[g(t)]g'(t) dt = F[g(d)] + \lambda \\ &= F(b) \\ &= \int_a^b f(x) dx.\end{aligned}$$

For example, consider the integral  $\int_1^2 (2t^2 - 1)^{1/2} t dt$ . Let  $x = 2t^2 - 1$ . Then  $dx = 4t dt$ , and by Theorem 6.4.9,

$$\int_1^2 (2t^2 - 1)^{1/2} t dt = \frac{1}{4} \int_1^7 x^{1/2} dx.$$

An indefinite integral of  $x^{1/2}$  is given by  $\frac{2}{3}x^{3/2}$ . Hence,

$$\int_1^2 (2t^2 - 1)^{1/2} t dt = \frac{1}{4} \left( \frac{2}{3} \right) (7^{3/2} - 1) = \frac{1}{6} (7^{3/2} - 1). \quad \square$$

## 6.5. IMPROPER RIEMANN INTEGRALS

In our study of the Riemann integral we have only considered integrals of functions that are bounded on a finite interval  $[a, b]$ . We now extend the scope of Riemann integration to include situations where the integrand can become unbounded at one or more points inside the range of integration, which can also be infinite. In such situations, the Riemann integral is called an improper integral.

There are two kinds of improper integrals. If  $f(x)$  is Riemann integrable on  $[a, b]$  for any  $b > a$ , then  $\int_a^\infty f(x) dx$  is called an improper integral of the first kind, where the range of integration is infinite. If, however,  $f(x)$



becomes infinite at a finite number of points inside the range of integration, then the integral  $\int_a^b f(x) dx$  is said to be improper of the second kind.

**Definition 6.5.1.** Let  $F(z) = \int_a^z f(x) dx$ . Suppose that  $F(z)$  exists for any value of  $z$  greater than  $a$ . If  $F(z)$  has a finite limit  $L$  as  $z \rightarrow \infty$ , then the improper integral  $\int_a^\infty f(x) dx$  is said to converge to  $L$ . In this case,  $L$  represents the Riemann integral of  $f(x)$  on  $[a, \infty)$  and we write

$$L = \int_a^\infty f(x) dx.$$

On the other hand, if  $L = \pm\infty$ , then the improper integral  $\int_a^\infty f(x) dx$  is said to diverge. By the same token, we can define the integral  $\int_{-\infty}^a f(x) dx$  as the limit, if it exists, of  $\int_{-z}^a f(x) dx$  as  $z \rightarrow \infty$ . Also,  $\int_{-\infty}^\infty f(x) dx$  is defined as

$$\int_{-\infty}^\infty f(x) dx = \lim_{u \rightarrow \infty} \int_{-u}^a f(x) dx + \lim_{z \rightarrow \infty} \int_a^z f(x) dx,$$

where  $a$  is any finite number, provided that both limits exist.

The convergence of  $\int_a^\infty f(x) dx$  can be determined by using the Cauchy criterion in a manner similar to the one used in the study of convergence of sequences (see Section 5.1.1).

**Theorem 6.5.1.** The improper integral  $\int_a^\infty f(x) dx$  converges if and only if for a given  $\epsilon > 0$  there exists a  $z_0$  such that

$$\left| \int_{z_1}^{z_2} f(x) dx \right| < \epsilon, \quad (6.20)$$

whenever  $z_1$  and  $z_2$  exceed  $z_0$ .

*Proof.* If  $F(z) = \int_a^z f(x) dx$  has a limit  $L$  as  $z \rightarrow \infty$ , then for a given  $\epsilon > 0$  there exists  $z_0$  such that for  $z > z_0$ ,

$$|F(z) - L| < \frac{\epsilon}{2}.$$

Now, if both  $z_1$  and  $z_2$  exceed  $z_0$ , then

$$\begin{aligned} \left| \int_{z_1}^{z_2} f(x) dx \right| &= |F(z_2) - F(z_1)| \\ &\leq |F(z_2) - L| + |F(z_1) - L| < \epsilon. \end{aligned}$$

Vice versa, if condition (6.20) is satisfied, then we need to show that  $F(z)$  has a limit as  $z \rightarrow \infty$ . Let us therefore define the sequence  $\{g_n\}_{n=1}^\infty$ , where  $g_n$  is

given by

$$g_n = \int_a^{a+n} f(x) dx, \quad n = 1, 2, \dots$$

It follows that for any  $\epsilon > 0$ ,

$$|g_n - g_m| = \left| \int_{a+m}^{a+n} f(x) dx \right| < \epsilon,$$

if  $m$  and  $n$  are large enough. This implies that  $\{g_n\}_{n=1}^{\infty}$  is a Cauchy sequence; hence it converges by Theorem 5.1.6. Let  $g = \lim_{n \rightarrow \infty} g_n$ . To show that  $\lim_{z \rightarrow \infty} F(z) = g$ , let us write

$$\begin{aligned} |F(z) - g| &= |F(z) - g_n + g_n - g| \\ &\leq |F(z) - g_n| + |g_n - g|. \end{aligned} \quad (6.21)$$

Suppose  $\epsilon > 0$  is given. There exists an integer  $N_1$  such that  $|g_n - g| < \epsilon/2$  if  $n > N_1$ . Also, there exists an integer  $N_2$  such that

$$|F(z) - g_n| = \left| \int_{a+n}^z f(x) dx \right| < \frac{\epsilon}{2} \quad (6.22)$$

if  $z > a + n > N_2$ . Thus by choosing  $z > a + n$ , where  $n > \max(N_1, N_2 - a)$ , we get from inequalities (6.21) and (6.22)

$$|F(z) - g| < \epsilon.$$

This completes the proof.  $\square$

**Definition 6.5.2.** If the improper integral  $\int_a^{\infty} |f(x)| dx$  is convergent, then the integral  $\int_a^{\infty} f(x) dx$  is said to be absolutely convergent. If  $\int_a^{\infty} f(x) dx$  is convergent but not absolutely, then it is said to be conditionally convergent.  $\square$

It is easy to show that an improper integral is convergent if it converges absolutely.

As with the case of series of positive terms, there are comparison tests that can be used to test for convergence of improper integrals of the first kind of nonnegative functions. These tests are described in the following theorems.

**Theorem 6.5.2.** Let  $f(x)$  be a nonnegative function that is Riemann integrable on  $[a, b]$  for every  $b \geq a$ . Suppose that there exists a function  $g(x)$  such that  $f(x) \leq g(x)$  for  $x \geq a$ . If  $\int_a^{\infty} g(x) dx$  converges, then so does  $\int_a^{\infty} f(x) dx$ .

and we have

$$\int_a^\infty f(x) dx \leq \int_a^\infty g(x) dx.$$

*Proof.* See Exercise 6.7.  $\square$

**Theorem 6.5.3.** Let  $f(x)$  and  $g(x)$  be nonnegative functions that are Riemann integrable on  $[a, b]$  for every  $b \geq a$ . If

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = k,$$

where  $k$  is a positive constant, then  $\int_a^\infty f(x) dx$  and  $\int_a^\infty g(x) dx$  are either both convergent or both divergent.

*Proof.* See Exercise 6.8.  $\square$

EXAMPLE 6.5.1. Consider the integral  $\int_1^\infty e^{-x} x^2 dx$ . We have that  $e^x = 1 + \sum_{n=1}^\infty (x^n/n!)$ . Hence, for  $x \geq 1$ ,  $e^x > x^p/p!$ , where  $p$  is any positive integer. If  $p$  is chosen such that  $p - 2 \geq 2$ , then

$$e^{-x} x^2 < \frac{p!}{x^{p-2}} \leq \frac{p!}{x^2}.$$

However,  $\int_1^\infty (dx/x^2) = [-1/x]_1^\infty = 1$ . Therefore, by Theorem 6.5.2, the integral of  $e^{-x} x^2$  on  $[1, \infty)$  is convergent.

EXAMPLE 6.5.2. The integral  $\int_0^\infty [(\sin x)/(x+1)^2] dx$  is absolutely convergent, since

$$\frac{|\sin x|}{(x+1)^2} \leq \frac{1}{(x+1)^2}$$

and

$$\int_0^\infty \frac{dx}{(x+1)^2} = \left[ -\frac{1}{x+1} \right]_0^\infty = 1.$$

EXAMPLE 6.5.3. The integral  $\int_0^\infty (\sin x/x) dx$  is conditionally convergent. We first show that  $\int_0^\infty (\sin x/x) dx$  is convergent. We have that

$$\int_0^\infty \frac{\sin x}{x} dx = \int_0^1 \frac{\sin x}{x} dx + \int_1^\infty \frac{\sin x}{x} dx. \quad (6.23)$$

By Exercise 6.3,  $(\sin x)/x$  is Riemann integrable on  $[0, 1]$ , since it is continuous there except at  $x = 0$ , which is a discontinuity of the first kind (see Definition 3.4.2). As for the second integral in (6.23), we have for  $z_2 > z_1 > 1$ ,

$$\begin{aligned}\int_{z_1}^{z_2} \frac{\sin x}{x} dx &= \left[ -\frac{\cos x}{x} \right]_{z_1}^{z_2} - \int_{z_1}^{z_2} \frac{\cos x}{x^2} dx \\ &= \frac{\cos z_1}{z_1} - \frac{\cos z_2}{z_2} - \int_{z_1}^{z_2} \frac{\cos x}{x^2} dx.\end{aligned}$$

Thus

$$\left| \int_{z_1}^{z_2} \frac{\sin x}{x} dx \right| \leq \frac{1}{z_1} + \frac{1}{z_2} + \int_{z_1}^{z_2} \frac{dx}{x^2} = \frac{2}{z_1}.$$

Since  $2/z_1$  can be made arbitrarily small by choosing  $z_1$  large enough, then by Theorem 6.5.1,  $\int_1^\infty (\sin x/x) dx$  is convergent and so is  $\int_0^\infty (\sin x/x) dx$ .

It remains to show that  $\int_0^\infty (\sin x/x) dx$  is not absolutely convergent. This follows from the fact that (see Exercise 6.10)

$$\lim_{n \rightarrow \infty} \int_0^{n\pi} \left| \frac{\sin x}{x} \right| dx = \infty.$$

Convergence of improper integrals of the first kind can be used to determine convergence of series of positive terms (see Section 5.2.1). This is based on the next theorem.

**Theorem 6.5.4** (Maclaurin's Integral Test). Let  $\sum_{n=1}^\infty a_n$  be a series of positive terms such that  $a_{n+1} \leq a_n$  for  $n \geq 1$ . Let  $f(x)$  be a positive nonincreasing function defined on  $[1, \infty)$  such that  $f(n) = a_n$ ,  $n = 1, 2, \dots$ , and  $f(x) \rightarrow 0$  as  $x \rightarrow \infty$ . Then,  $\sum_{n=1}^\infty a_n$  converges if and only if the improper integral  $\int_1^\infty f(x) dx$  converges.

*Proof.* If  $n \geq 1$  and  $n \leq x \leq n+1$ , then

$$a_n = f(n) \geq f(x) \geq f(n+1) = a_{n+1}.$$

By Theorem 6.4.2 we have for  $n \geq 1$

$$a_n \geq \int_n^{n+1} f(x) dx \geq a_{n+1}. \quad (6.24)$$

If  $s_n = \sum_{k=1}^n a_k$  is the  $n$ th partial sum of the series, then from inequality (6.24) we obtain

$$s_n \geq \int_1^{n+1} f(x) dx \geq s_{n+1} - a_1. \quad (6.25)$$

If the series  $\sum_{n=1}^{\infty} a_n$  converges to the sum  $s$ , then  $s \geq s_n$  for all  $n$ . Consequently, the sequence whose  $n$ th term is  $F(n+1) = \int_1^{n+1} f(x) dx$  is monotone increasing and is bounded by  $s$ ; hence it must have a limit. Therefore, the integral  $\int_1^{\infty} f(x) dx$  converges.

Now, let us suppose that  $\int_1^{\infty} f(x) dx$  is convergent and is equal to  $L$ . Then from inequality (6.25) we obtain

$$s_{n+1} \leq a_1 + \int_1^{n+1} f(x) dx \leq a_1 + L, \quad n \geq 1, \quad (6.26)$$

since  $f(x)$  is positive. Inequality (6.26) indicates that the monotone increasing sequence  $\{s_n\}_{n=1}^{\infty}$  is bounded hence it has a limit, which is the sum of the series.  $\square$

Theorem 6.5.4 provides a test of convergence for a series of positive terms. Of course, the usefulness of this test depends on how easy it is to integrate the function  $f(x)$ .

As an example of using the integral test, consider the harmonic series  $\sum_{n=1}^{\infty} (1/n)$ . If  $f(x)$  is defined as  $f(x) = 1/x$ ,  $x \geq 1$ , then  $F(x) = \int_1^x f(t) dt = \log x$ . Since  $F(x)$  goes to infinity as  $x \rightarrow \infty$ , the harmonic series must therefore be divergent, as was shown in Chapter 5. On the other hand, the series  $\sum_{n=1}^{\infty} (1/n^2)$  is convergent, since  $F(x) = \int_1^x (dt/t^2) = 1 - 1/x$ , which converges to 1 as  $x \rightarrow \infty$ .

### 6.5.1. Improper Riemann Integrals of the Second Kind

Let us now consider integrals of the form  $\int_a^b f(x) dx$  where  $[a, b]$  is a finite interval and the integrand becomes infinite at a finite number of points inside  $[a, b]$ . Such integrals are called improper integrals of the second kind. Suppose, for example, that  $f(x) \rightarrow \infty$  as  $x \rightarrow a^+$ . Then  $\int_a^b f(x) dx$  is said to converge if the limit

$$\lim_{\epsilon \rightarrow 0^+} \int_{a+\epsilon}^b f(x) dx$$

exists and is finite. Similarly, if  $f(x) \rightarrow \infty$  as  $x \rightarrow b^-$ , then  $\int_a^b f(x) dx$  is convergent if the limit

$$\lim_{\epsilon \rightarrow 0^+} \int_a^{b-\epsilon} f(x) dx$$

exists. Furthermore, if  $f(x) \rightarrow \infty$  as  $x \rightarrow c$ , where  $a < c < b$ , then  $\int_a^b f(x) dx$  is the sum of  $\int_a^c f(x) dx$  and  $\int_c^b f(x) dx$  provided that both integrals converge. By definition, if  $f(x) \rightarrow \infty$  as  $x \rightarrow x_0$ , where  $x_0 \in [a, b]$ , then  $x_0$  is said to be a singularity of  $f(x)$ .

The following theorems can help in determining convergence of integrals of the second kind. They are similar to Theorems 6.5.1, 6.5.2, and 6.5.3. Their proofs will therefore be omitted.

**Theorem 6.5.5.** If  $f(x) \rightarrow \infty$  as  $x \rightarrow a^+$ , then  $\int_a^b f(x) dx$  converges if and only if for a given  $\epsilon > 0$  there exists a  $z_0$  such that

$$\left| \int_{z_1}^{z_2} f(x) dx \right| < \epsilon,$$

where  $z_1$  and  $z_2$  are any two numbers such that  $a < z_1 < z_2 < z_0 < b$ .

**Theorem 6.5.6.** Let  $f(x)$  be a nonnegative function such that  $\int_c^b f(x) dx$  exists for every  $c$  in  $(a, b]$ . If there exists a function  $g(x)$  such that  $f(x) \leq g(x)$  for all  $x$  in  $(a, b]$ , and if  $\int_c^b g(x) dx$  converges as  $c \rightarrow a^+$ , then so does  $\int_c^b f(x) dx$  and we have

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

**Theorem 6.5.7.** Let  $f(x)$  and  $g(x)$  be nonnegative functions that are Riemann integrable on  $[c, b]$  for every  $c$  such that  $a < c \leq b$ . If

$$\lim_{x \rightarrow a^+} \frac{f(x)}{g(x)} = k,$$

where  $k$  is a positive constant, then  $\int_a^b f(x) dx$  and  $\int_a^b g(x) dx$  are either both convergent or both divergent.

**Definition 6.5.3.** Let  $\int_a^b f(x) dx$  be an improper integral of the second kind. If  $\int_a^b |f(x)| dx$  converges, then  $\int_a^b f(x) dx$  is said to converge absolutely. If, however,  $\int_a^b f(x) dx$  is convergent, but not absolutely, then it is said to be conditionally convergent.  $\square$

**Theorem 6.5.8.** If  $\int_a^b |f(x)| dx$  converges, then so does  $\int_a^b f(x) dx$ .

**EXAMPLE 6.5.4.** Consider the integral  $\int_0^1 e^{-x} x^{n-1} dx$ , where  $n > 0$ . If  $0 < n < 1$ , then the integral is improper of the second kind, since  $x^{n-1} \rightarrow \infty$  as  $x \rightarrow 0^+$ . Thus,  $x = 0$  is a singularity of the integrand. Since

$$\lim_{x \rightarrow 0^+} \frac{e^{-x} x^{n-1}}{x^{n-1}} = 1,$$

then the behavior of  $\int_0^1 e^{-x} x^{n-1} dx$  with regard to convergence or divergence is the same as that of  $\int_0^1 x^{n-1} dx$ . But  $\int_0^1 x^{n-1} dx = (1/n)[x^n]_0^1 = 1/n$  is convergent, and so is  $\int_0^1 e^{-x} x^{n-1} dx$ .

EXAMPLE 6.5.5.  $\int_0^1 (\sin x/x^2) dx$ . The integrand has a singularity at  $x = 0$ . Let  $g(x) = 1/x$ . Then,  $(\sin x)/[x^2 g(x)] \rightarrow 1$  as  $x \rightarrow 0^+$ . But  $\int_0^1 (dx/x) = [\log x]_0^1$  is divergent, since  $\log x \rightarrow -\infty$  as  $x \rightarrow 0^+$ . Therefore,  $\int_0^1 (\sin x/x^2) dx$  is divergent.

EXAMPLE 6.5.6. Consider the integral  $\int_0^2 (x^2 - 3x + 1)/[x(x-1)^2] dx$ . Here, the integrand has two singularities, namely  $x = 0$  and  $x = 1$ , inside  $[0, 2]$ . We can therefore write

$$\begin{aligned} \int_0^2 \frac{x^2 - 3x + 1}{x(x-1)^2} dx &= \lim_{t \rightarrow 0^+} \int_t^{1/2} \frac{x^2 - 3x + 1}{x(x-1)^2} dx \\ &\quad + \lim_{u \rightarrow 1^-} \int_{1/2}^u \frac{x^2 - 3x + 1}{x(x-1)^2} dx \\ &\quad + \lim_{v \rightarrow 1^+} \int_v^2 \frac{x^2 - 3x + 1}{x(x-1)^2} dx. \end{aligned}$$

We note that

$$\frac{x^2 - 3x + 1}{x(x-1)^2} = \frac{1}{x} - \frac{1}{(x-1)^2}.$$

Hence,

$$\begin{aligned} \int_0^2 \frac{x^2 - 3x + 1}{x(x-1)^2} dx &= \lim_{t \rightarrow 0^+} \left[ \log x + \frac{1}{x-1} \right]_t^{1/2} \\ &\quad + \lim_{u \rightarrow 1^-} \left[ \log x + \frac{1}{x-1} \right]_{1/2}^u \\ &\quad + \lim_{v \rightarrow 1^+} \left[ \log x + \frac{1}{x-1} \right]_v^2. \end{aligned}$$

None of the above limits exists as a finite number. This integral is therefore divergent.

## 6.6. CONVERGENCE OF A SEQUENCE OF RIEMANN INTEGRALS

In the present section we confine our attention to the limiting behavior of integrals of a sequence of functions  $\{f_n(x)\}_{n=1}^\infty$ .

**Theorem 6.6.1.** Suppose that  $f_n(x)$  is Riemann integrable on  $[a, b]$  for  $n \geq 1$ . If  $f_n(x)$  converges uniformly to  $f(x)$  on  $[a, b]$  as  $n \rightarrow \infty$ , then  $f(x)$  is Riemann integrable on  $[a, b]$  and

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx.$$

*Proof.* Let us first show that  $f(x)$  is Riemann integrable on  $[a, b]$ . Let  $\epsilon > 0$  be given. Since  $f_n(x)$  converges uniformly to  $f(x)$ , then there exists an integer  $n_0$  that depends only on  $\epsilon$  such that

$$|f_n(x) - f(x)| < \frac{\epsilon}{3(b-a)} \quad (6.27)$$

if  $n > n_0$  for all  $x \in [a, b]$ . Let  $n_1 > n_0$ . Since  $f_{n_1}(x)$  is Riemann integrable on  $[a, b]$ , then by Theorem 6.2.1 there exists a  $\delta > 0$  such that

$$US_P(f_{n_1}) - LS_P(f_{n_1}) < \frac{\epsilon}{3} \quad (6.28)$$

for any partition  $P$  of  $[a, b]$  with a norm  $\Delta_P < \delta$ . Now, from inequality (6.27) we have

$$f(x) < f_{n_1}(x) + \frac{\epsilon}{3(b-a)},$$

$$f(x) > f_{n_1}(x) - \frac{\epsilon}{3(b-a)}.$$

We conclude that

$$US_P(f) \leq US_P(f_{n_1}) + \frac{\epsilon}{3}, \quad (6.29)$$

$$LS_P(f) \geq LS_P(f_{n_1}) - \frac{\epsilon}{3}. \quad (6.30)$$

From inequalities (6.28), (6.29), and (6.30) it follows that if  $\Delta_P < \delta$ , then

$$US_P(f) - LS_P(f) \leq US_P(f_{n_1}) - LS_P(f_{n_1}) + \frac{2\epsilon}{3} < \epsilon. \quad (6.31)$$

Inequality (6.31) shows that  $f(x)$  is Riemann integrable on  $[a, b]$ , again by Theorem 6.2.1.



Let us now show that

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx. \quad (6.32)$$

From inequality (6.27) we have for  $n > n_0$ ,

$$\begin{aligned} \left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| &\leq \int_a^b |f_n(x) - f(x)| dx \\ &< \frac{\epsilon}{3}, \end{aligned}$$

and the result follows, since  $\epsilon$  is an arbitrary positive number.  $\square$

## 6.7. SOME FUNDAMENTAL INEQUALITIES

In this section we consider certain well-known inequalities for the Riemann integral.

### 6.7.1. The Cauchy–Schwarz Inequality

**Theorem 6.7.1.** Suppose that  $f(x)$  and  $g(x)$  are such that  $f^2(x)$  and  $g^2(x)$  are Riemann integrable on  $[a, b]$ . Then

$$\left[ \int_a^b f(x)g(x) dx \right]^2 \leq \left[ \int_a^b f^2(x) dx \right] \left[ \int_a^b g^2(x) dx \right]. \quad (6.33)$$

The limits of integration may be finite or infinite.

*Proof.* Let  $c_1$  and  $c_2$  be constants, not both zero. Without loss of generality, let us assume that  $c_2 \neq 0$ . Then

$$\int_a^b [c_1 f(x) + c_2 g(x)]^2 dx \geq 0.$$

Thus the quadratic form

$$c_1^2 \int_a^b f^2(x) dx + 2c_1 c_2 \int_a^b f(x)g(x) dx + c_2^2 \int_a^b g^2(x) dx$$

is nonnegative for all  $c_1$  and  $c_2$ . It follows that its discriminant, namely,

$$c_2^2 \left[ \int_a^b f(x)g(x) dx \right]^2 - c_2^2 \left[ \int_a^b f^2(x) dx \right] \left[ \int_a^b g^2(x) dx \right]$$

must be nonpositive, that is,

$$\left[ \int_a^b f(x)g(x) dx \right]^2 \leq \left[ \int_a^b f^2(x) dx \right] \left[ \int_a^b g^2(x) dx \right].$$

It is easy to see that if  $f(x)$  and  $g(x)$  are linearly related [that is, there exist constants  $\tau_1$  and  $\tau_2$ , not both zero, such that  $\tau_1 f(x) + \tau_2 g(x) = 0$ ], then inequality (6.33) becomes an equality.  $\square$

### 6.7.2. Hölder's Inequality

This is a generalization of the Cauchy–Schwarz inequality due to Otto Hölder (1859–1937). To prove Hölder's inequality we need the following lemmas:

**Lemma 6.7.1.** Let  $a_1, a_1, \dots, a_n$ ;  $\lambda_1, \lambda_2, \dots, \lambda_n$  be nonnegative numbers such that  $\sum_{i=1}^n \lambda_i = 1$ . Then

$$\prod_{i=1}^n a_i^{\lambda_i} \leq \sum_{i=1}^n \lambda_i a_i. \quad (6.34)$$

The right-hand side of inequality (6.34) is a weighted arithmetic mean of the  $a_i$ 's, and the left-hand side is a weighted geometric mean.

*Proof.* This lemma is an extension of a result given in Section 3.7 concerning the properties of convex functions (see Exercise 6.19).  $\square$

**Lemma 6.7.2.** Suppose that  $f_1(x), f_2(x), \dots, f_n(x)$  are nonnegative and Riemann integrable on  $[a, b]$ . If  $\lambda_1, \lambda_2, \dots, \lambda_n$  are nonnegative numbers such that  $\sum_{i=1}^n \lambda_i = 1$ , then

$$\int_a^b \left[ \prod_{i=1}^n f_i^{\lambda_i}(x) \right] dx \leq \prod_{i=1}^n \left[ \int_a^b f_i(x) dx \right]^{\lambda_i}. \quad (6.35)$$

*Proof.* Without loss of generality, let us assume that  $\int_a^b f_i(x) dx > 0$  for  $i = 1, 2, \dots, n$  [inequality (6.35) is obviously true if at least one  $f_i(x)$  is

identically equal to zero]. By Lemma 6.7.1 we have

$$\begin{aligned} & \frac{\int_a^b [\prod_{i=1}^n f_i^{\lambda_i}(x)] dx}{\prod_{i=1}^n [\int_a^b f_i(x) dx]^{\lambda_i}} \\ &= \int_a^b \left[ \frac{f_1(x)}{\int_a^b f_1(x) dx} \right]^{\lambda_1} \left[ \frac{f_2(x)}{\int_a^b f_2(x) dx} \right]^{\lambda_2} \cdots \left[ \frac{f_n(x)}{\int_a^b f_n(x) dx} \right]^{\lambda_n} dx \\ &\leq \int_a^b \sum_{i=1}^n \frac{\lambda_i f_i(x)}{\int_a^b f_i(x) dx} dx = \sum_{i=1}^n \lambda_i = 1. \end{aligned}$$

Hence, inequality (6.35) follows.  $\square$

**Theorem 6.7.2** (Hölder's Inequality). Let  $p$  and  $q$  be two positive numbers such that  $1/p + 1/q = 1$ . If  $|f(x)|^p$  and  $|g(x)|^q$  are Riemann integrable on  $[a, b]$ , then

$$\left| \int_a^b f(x)g(x) dx \right| \leq \left[ \int_a^b |f(x)|^p dx \right]^{1/p} \left[ \int_a^b |g(x)|^q dx \right]^{1/q}.$$

*Proof.* Define the functions

$$u(x) = |f(x)|^p, \quad v(x) = |g(x)|^q.$$

Then, by Lemma 6.7.2,

$$\int_a^b u(x)^{1/p} v(x)^{1/q} dx \leq \left[ \int_a^b u(x) dx \right]^{1/p} \left[ \int_a^b v(x) dx \right]^{1/q},$$

that is,

$$\int_a^b |f(x)| |g(x)| dx \leq \left[ \int_a^b |f(x)|^p dx \right]^{1/p} \left[ \int_a^b |g(x)|^q dx \right]^{1/q}. \quad (6.36)$$

The theorem follows from inequality (6.36) and the fact that

$$\left| \int_a^b f(x)g(x) dx \right| \leq \int_a^b |f(x)| |g(x)| dx.$$

We note that the Cauchy-Schwarz inequality can be deduced from Theorem 6.7.2 by taking  $p = q = 2$ .  $\square$

### 6.7.3. Minkowski's Inequality

The following inequality is due to Hermann Minkowski (1864–1909).

**Theorem 6.7.3.** Suppose that  $f(x)$  and  $g(x)$  are functions such that  $|f(x)|^p$  and  $|g(x)|^p$  are Riemann integrable on  $[a, b]$ , where  $1 \leq p < \infty$ . Then

$$\left[ \int_a^b |f(x) + g(x)|^p dx \right]^{1/p} \leq \left[ \int_a^b |f(x)|^p dx \right]^{1/p} + \left[ \int_a^b |g(x)|^p dx \right]^{1/p}.$$

*Proof.* The theorem is obviously true if  $p = 1$  by the triangle inequality. We therefore assume that  $p > 1$ . Let  $q$  be a positive number such that  $1/p + 1/q = 1$ . Hence,  $p = p(1/p + 1/q) = 1 + p/q$ . Let us now write

$$\begin{aligned} |f(x) + g(x)|^p &= |f(x) + g(x)| |f(x) + g(x)|^{p/q} \\ &\leq |f(x)| |f(x) + g(x)|^{p/q} + |g(x)| |f(x) + g(x)|^{p/q}. \end{aligned} \quad (6.37)$$

By applying Hölder's inequality to the two terms on the right-hand side of inequality (6.37) we obtain

$$\begin{aligned} \int_a^b |f(x)| |f(x) + g(x)|^{p/q} dx \\ \leq \left[ \int_a^b |f(x)|^p dx \right]^{1/p} \left[ \int_a^b |f(x) + g(x)|^p dx \right]^{1/q}, \end{aligned} \quad (6.38)$$

$$\begin{aligned} \int_a^b |g(x)| |f(x) + g(x)|^{p/q} dx \\ \leq \left[ \int_a^b |g(x)|^p dx \right]^{1/p} \left[ \int_a^b |f(x) + g(x)|^p dx \right]^{1/q}. \end{aligned} \quad (6.39)$$

From inequalities (6.37), (6.38), and (6.39) we conclude that

$$\begin{aligned} \int_a^b |f(x) + g(x)|^p dx &\leq \left[ \int_a^b |f(x) + g(x)|^p dx \right]^{1/q} \\ &\quad \times \left\{ \left[ \int_a^b |f(x)|^p dx \right]^{1/p} + \left[ \int_a^b |g(x)|^p dx \right]^{1/p} \right\}. \end{aligned} \quad (6.40)$$

Since  $1 - 1/q = 1/p$ , inequality (6.40) can be written as

$$\left[ \int_a^b |f(x) + g(x)|^p dx \right]^{1/p} \leq \left[ \int_a^b |f(x)|^p dx \right]^{1/p} + \left[ \int_a^b |g(x)|^p dx \right]^{1/p}.$$

Minkowski's inequality can be extended to integrals involving more than two functions. It can be shown (see Exercise 6.20) that if  $|f_i(x)|^p$  is Riemann integrable on  $[a, b]$  for  $i = 1, 2, \dots, n$ , then

$$\left[ \int_a^b \left| \sum_{i=1}^n f_i(x) \right|^p dx \right]^{1/p} \leq \sum_{i=1}^n \left[ \int_a^b |f_i(x)|^p dx \right]^{1/p}. \quad \square$$

#### 6.7.4. Jensen's Inequality

**Theorem 6.7.4.** Let  $X$  be a random variable with a finite expected value,  $\mu = E(X)$ . If  $\phi(x)$  is a twice differentiable convex function, then

$$E[\phi(X)] \geq \phi[E(X)].$$

*Proof.* Since  $\phi(x)$  is convex and  $\phi''(x)$  exists, then we must have  $\phi''(x) \geq 0$ . By applying the mean value theorem (Theorem 4.2.2) around  $\mu$  we obtain

$$\phi(X) = \phi(\mu) + (X - \mu)\phi'(c),$$

where  $c$  is between  $\mu$  and  $X$ . If  $X - \mu > 0$ , then  $c > \mu$  and hence  $\phi'(c) \geq \phi'(\mu)$ , since  $\phi''(x)$  is nonnegative. Thus,

$$\phi(X) - \phi(\mu) = (X - \mu)\phi'(c) \geq (X - \mu)\phi'(\mu). \quad (6.41)$$

On the other hand, if  $X - \mu < 0$ , then  $c < \mu$  and  $\phi'(c) \leq \phi'(\mu)$ . Hence,

$$\phi(X) - \phi(\mu) = (X - \mu)\phi'(c) \geq (X - \mu)\phi'(\mu). \quad (6.42)$$

From inequalities (6.41) and (6.42) we conclude that

$$E[\phi(X) - \phi(\mu)] \geq \phi'(\mu)E(X - \mu) = 0,$$

which implies that

$$E[\phi(X)] \geq \phi(\mu),$$

since  $E[\phi(\mu)] = \phi(\mu)$ .  $\square$

### 6.8. RIEMANN-STIELTJES INTEGRAL

In this section we consider a more general integral, namely the Riemann–Stieltjes integral. The concept on which this integral is based can be attributed to a combination of ideas by Georg Friedrich Riemann (1826–1866) and the Dutch mathematician Thomas Joannes Stieltjes (1856–1894).

The Riemann–Stieltjes integral involves two functions  $f(x)$  and  $g(x)$ , both defined on the interval  $[a, b]$ , and is denoted by  $\int_a^b f(x) dg(x)$ . In particular, if  $g(x) = x$  we obtain the Riemann integral  $\int_a^b f(x) dx$ . Thus the Riemann integral is a special case of the Riemann–Stieltjes integral.

The definition of the Riemann–Stieltjes integral of  $f(x)$  with respect to  $g(x)$  on  $[a, b]$  is similar to that of the Riemann integral. If  $f(x)$  is bounded on  $[a, b]$ , if  $g(x)$  is monotone increasing on  $[a, b]$ , and if  $P = \{x_0, x_1, \dots, x_n\}$  is a partition of  $[a, b]$ , then as in Section 6.2, we define the sums

$$LS_P(f, g) = \sum_{i=1}^n m_i \Delta g_i,$$

$$US_P(f, g) = \sum_{i=1}^n M_i \Delta g_i,$$

where  $m_i$  and  $M_i$  are, respectively, the infimum and supremum of  $f(x)$  on  $[x_{i-1}, x_i]$ ,  $\Delta g_i = g(x_i) - g(x_{i-1})$ ,  $i = 1, 2, \dots, n$ . If for a given  $\epsilon > 0$  there exists a  $\delta > 0$  such that

$$US_P(f, g) - LS_P(f, g) < \epsilon \quad (6.43)$$

whenever  $\Delta_p < \delta$ , where  $\Delta_p$  is the norm of  $P$ , then  $f(x)$  is said to be Riemann–Stieltjes integrable with respect to  $g(x)$  on  $[a, b]$ . In this case,

$$\int_a^b f(x) dg(x) = \inf_P US_P(f, g) = \sup_P LS_P(f, g).$$

Condition (6.43) is both necessary and sufficient for the existence of the Riemann–Stieltjes integral.

Equivalently, suppose that for a given partition  $P = \{x_0, x_1, \dots, x_n\}$  we define the sum

$$S(P, f, g) = \sum_{i=1}^n f(t_i) \Delta g_i, \quad (6.44)$$

where  $t_i$  is a point in the interval  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ . Then  $f(x)$  is Riemann–Stieltjes integrable with respect to  $g(x)$  on  $[a, b]$  if for any  $\epsilon > 0$

there exists a  $\delta > 0$  such that

$$\left| S(P, f, g) - \int_a^b f(x) dg(x) \right| < \epsilon \quad (6.45)$$

for any partition  $P$  of  $[a, b]$  with a norm  $\Delta_p < \delta$ , and for any choice of the point  $t_i$  in  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ .

Theorems concerning the Riemann–Stieltjes integral are very similar to those seen earlier concerning the Riemann integral. In particular, we have the following theorems:

**Theorem 6.8.1.** If  $f(x)$  is continuous on  $[a, b]$ , then  $f(x)$  is Riemann–Stieltjes integrable on  $[a, b]$ .

*Proof.* See Exercise 6.21.  $\square$

**Theorem 6.8.2.** If  $f(x)$  is monotone increasing (or monotone decreasing) on  $[a, b]$ , and  $g(x)$  is continuous on  $[a, b]$ , then  $f(x)$  is Riemann–Stieltjes integrable with respect to  $g(x)$  on  $[a, b]$ .

*Proof.* See Exercise 6.22.  $\square$

The next theorem shows that under certain conditions, the Riemann–Stieltjes integral reduces to the Riemann integral.

**Theorem 6.8.3.** Suppose that  $f(x)$  is Riemann–Stieltjes integrable with respect to  $g(x)$  on  $[a, b]$ , where  $g(x)$  has a continuous derivative  $g'(x)$  on  $[a, b]$ . Then

$$\int_a^b f(x) dg(x) = \int_a^b f(x) g'(x) dx.$$

*Proof.* Let  $P = \{x_0, x_1, \dots, x_n\}$  be a partition of  $[a, b]$ . Consider the sum

$$S(P, h) = \sum_{i=1}^n h(t_i) \Delta x_i, \quad (6.46)$$

where  $h(x) = f(x)g'(x)$  and  $x_{i-1} \leq t_i \leq x_i$ ,  $i = 1, 2, \dots, n$ . Let us also consider the sum

$$S(P, f, g) = \sum_{i=1}^n f(t_i) \Delta g_i, \quad (6.47)$$

If we apply the mean value theorem (Theorem 4.2.2) to  $g(x)$ , we obtain

$$\Delta g_i = g(x_i) - g(x_{i-1}) = g'(z_i) \Delta x_i, \quad i = 1, 2, \dots, n, \quad (6.48)$$

where  $x_{i-1} < z_i < x_i$ ,  $i = 1, 2, \dots, n$ . From (6.46), (6.47), and (6.48) we can then write

$$S(P, f, g) - S(P, h) = \sum_{i=1}^n f(t_i) [g'(z_i) - g'(t_i)] \Delta x_i. \quad (6.49)$$

Since  $f(x)$  is bounded on  $[a, b]$  and  $g'(x)$  is uniformly continuous on  $[a, b]$  by Theorem 3.4.6, then for a given  $\epsilon > 0$  there exists a  $\delta_1 > 0$ , which depends only on  $\epsilon$ , such that

$$|g'(z_i) - g'(t_i)| < \frac{\epsilon}{2M(b-a)},$$

if  $|z_i - t_i| < \delta_1$ , where  $M > 0$  is such that  $|f(x)| \leq M$  on  $[a, b]$ . From (6.49) it follows that if the partition  $P$  has a norm  $\Delta_p < \delta_1$ , then

$$|S(P, f, g) - S(P, h)| < \frac{\epsilon}{2}. \quad (6.50)$$

Now, since  $f(x)$  is Riemann–Stieltjes integrable with respect to  $g(x)$  on  $[a, b]$ , then by definition, for the given  $\epsilon > 0$  there exists a  $\delta_2 > 0$  such that

$$\left| S(P, f, g) - \int_a^b f(x) dg(x) \right| < \frac{\epsilon}{2}, \quad (6.51)$$

if the norm  $\Delta_p$  of  $P$  is less than  $\delta_2$ . We conclude from (6.50) and (6.51) that if the norm of  $P$  is less than  $\min(\delta_1, \delta_2)$ , then

$$\left| S(P, h) - \int_a^b f(x) dg(x) \right| < \epsilon.$$

Since  $\epsilon$  is arbitrary, this inequality implies that  $\int_a^b f(x) dg(x)$  is in fact the Riemann integral  $\int_a^b h(x) dx = \int_a^b f(x)g'(x) dx$ .  $\square$

Using Theorem 6.8.3, it is easy to see that if, for example,  $f(x) = 1$  and  $g(x) = x^2$ , then  $\int_a^b f(x) dg(x) = \int_a^b f(x)g'(x) dx = \int_a^b 2x dx = b^2 - a^2$ .

It should be noted that Theorems 6.8.1 and 6.8.2 provide sufficient conditions for the existence of  $\int_a^b f(x) dg(x)$ . It is possible, however, for the Riemann–Stieltjes integral to exist even if  $g(x)$  is a discontinuous function. For example, consider the function  $g(x) = \gamma I(x - c)$ , where  $\gamma$  is a nonzero



constant,  $a < c < b$ , and  $I(x - c)$  is such that

$$I(x - c) = \begin{cases} 0, & x < c, \\ 1, & x \geq c. \end{cases}$$

The quantity  $\gamma$  represents what is called a jump at  $x = c$ . If  $f(x)$  is bounded on  $[a, b]$  and is continuous at  $x = c$ , then

$$\int_a^b f(x) dg(x) = \gamma f(c). \quad (6.52)$$

To show the validity of formula (6.52), let  $P = \{x_0, x_1, \dots, x_n\}$  be any partition of  $[a, b]$ . Then,  $\Delta g_i = g(x_i) - g(x_{i-1})$  will be zero as long as  $x_i < c$  or  $x_{i-1} \geq c$ . Suppose, therefore, that there exists a  $k$ ,  $1 \leq k \leq n$ , such that  $x_{k-1} < c \leq x_k$ . In this case, the sum  $S(P, f, g)$  in formula (6.44) takes the form

$$S(P, f, g) = \sum_{i=1}^n f(t_i) \Delta g_i = \gamma f(t_k).$$

It follows that

$$|S(P, f, g) - \gamma f(c)| = |\gamma| |f(t_k) - f(c)|. \quad (6.53)$$

Now, let  $\epsilon > 0$  be given. Since  $f(x)$  is continuous at  $x = c$ , then there exists a  $\delta > 0$  such that

$$|f(t_k) - f(c)| < \frac{\epsilon}{|\gamma|},$$

if  $|t_k - c| < \delta$ . Thus if the norm  $\Delta_p$  of  $P$  is chosen so that  $\Delta_p < \delta$ , then

$$|S(P, f, g) - \gamma f(c)| < \epsilon. \quad (6.54)$$

Equality (6.52) follows from comparing inequalities (6.45) and (6.54).

It is now easy to show that if

$$g(x) = \begin{cases} \lambda, & a \leq x < b, \\ \lambda', & x = b, \end{cases}$$

and if  $f(x)$  is continuous at  $x = b$ , then

$$\int_a^b f(x) dg(x) = (\lambda' - \lambda)f(b). \quad (6.55)$$

The previous examples represent special cases of a class of functions defined on  $[a, b]$  called step functions. These functions are constant on  $[a, b]$

except for a finite number of jump discontinuities. We can generalize formula (6.55) to this class of functions as can be seen in the next theorem.

**Theorem 6.8.4.** Let  $g(x)$  be a step function defined on  $[a, b]$  with jump discontinuities at  $x = c_1, c_2, \dots, c_n$ , and  $a < c_1 < c_2 < \dots < c_n = b$ , such that

$$g(x) = \begin{cases} \lambda_1, & a \leq x < c_1, \\ \lambda_2, & c_1 \leq x < c_2, \\ \vdots & \vdots \\ \lambda_n, & c_{n-1} \leq x < c_n, \\ \lambda_{n+1}, & x = c_n. \end{cases}$$

If  $f(x)$  is bounded on  $[a, b]$  and is continuous at  $x = c_1, c_2, \dots, c_n$ , then

$$\int_a^b f(x) dg(x) = \sum_{i=1}^n (\lambda_{i+1} - \lambda_i) f(c_i). \quad (6.56)$$

*Proof.* The proof can be easily obtained by first writing the integral in formula (6.56) as

$$\begin{aligned} \int_a^b f(x) dg(x) &= \int_a^{c_1} f(x) dg(x) + \int_{c_1}^{c_2} f(x) dg(x) \\ &\quad + \dots + \int_{c_{n-1}}^{c_n} f(x) dg(x). \end{aligned} \quad (6.57)$$

If we now apply formula (6.55) to each integral in (6.57) we obtain

$$\begin{aligned} \int_a^{c_1} f(x) dg(x) &= (\lambda_2 - \lambda_1) f(c_1), \\ \int_{c_1}^{c_2} f(x) dg(x) &= (\lambda_3 - \lambda_2) f(c_2), \\ &\vdots \\ \int_{c_{n-1}}^{c_n} f(x) dg(x) &= (\lambda_{n+1} - \lambda_n) f(c_n). \end{aligned}$$

By adding up all these integrals we obtain formula (6.56).  $\square$

**EXAMPLE 6.8.1.** One example of a step function is the greatest-integer function  $[x]$ , which is defined as the greatest integer less than or equal to  $x$ . If  $f(x)$  is bounded on  $[0, n]$  and is continuous at  $x = 1, 2, \dots, n$ , where  $n$  is a

positive integer, then by Theorem 6.8.4 we can write

$$\int_0^n f(x) d[x] = \sum_{i=1}^n f(i). \quad (6.58)$$

It follows that every finite sum of the form  $\sum_{i=1}^n a_i$  can be expressed as a Riemann–Stieltjes integral with respect to  $[x]$  of a function  $f(x)$  continuous on  $[0, n]$  such that  $f(i) = a_i$ ,  $i = 1, 2, \dots, n$ . The Riemann–Stieltjes integral has therefore the distinct advantage of making finite sums expressible as integrals.

## 6.9. APPLICATIONS IN STATISTICS

Riemann integration plays an important role in statistical distribution theory. Perhaps the most prevalent use of the Riemann integral is in the study of the distributions of continuous random variables.

We recall from Section 4.5.1 that a continuous random variable  $X$  with a cumulative distribution function  $F(x) = P(X \leq x)$  is absolutely continuous if  $F(x)$  is differentiable. In this case, there exists a function  $f(x)$  called the density function of  $X$  such that  $F'(x) = f(x)$ , that is,

$$F(x) = \int_{-\infty}^x f(t) dt. \quad (6.59)$$

In general, if  $X$  is a continuous random variable, it need not be absolutely continuous. It is true, however, that most common distributions that are continuous are also absolutely continuous.

The probability distribution of an absolutely continuous random variable is completely determined by its density function. For example, from (6.59) it follows that

$$P(a < X < b) = F(b) - F(a) = \int_a^b f(x) dx. \quad (6.60)$$

Note that the value of this probability remains unchanged if one or both of the end points of the interval  $[a, b]$  are included. This is true because the probability assigned to these individual points is zero when  $X$  has a continuous distribution. The mean  $\mu$  and variance  $\sigma^2$  of  $X$  are given by

$$\begin{aligned} \mu &= E(X) = \int_{-\infty}^{\infty} xf(x) dx, \\ \sigma^2 &= \text{Var}(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx. \end{aligned}$$

In general, the  $k$ th central moment of  $X$ , denoted by  $\mu_k$  ( $k = 1, 2, \dots$ ), is defined as

$$\mu_k = E[(X - \mu)^k] = \int_{-\infty}^{\infty} (x - \mu)^k f(x) dx. \quad (6.61)$$

We note that  $\sigma^2 = \mu_2$ . Similarly, the  $k$ th noncentral moment of  $X$ , denoted by  $\mu'_k$  ( $k = 1, 2, \dots$ ), is defined as

$$\mu'_k = E(X^k). \quad (6.62)$$

The first noncentral moment of  $X$  is its mean  $\mu$ , while its first central moment is equal to zero.

We note that if the domain of the density function  $f(x)$  is infinite, then  $\mu_k$  and  $\mu'_k$  are improper integrals of the first kind. Therefore, they may or may not exist. If  $\int_{-\infty}^{\infty} |x|^k f(x) dx$  exists, then so does  $\mu'_k$  (see Definition 6.5.2 concerning absolute convergence of improper integrals). The latter integral is called the  $k$ th absolute moment and is denoted by  $\nu_k$ . If  $\nu_k$  exists, then the noncentral moments of order  $j$  for  $j \leq k$  exist also. This follows because of the inequality

$$|x^j| \leq |x|^k + 1 \quad \text{if } j \leq k, \quad (6.63)$$

which is true because  $|x|^j \leq |x|^k$  if  $|x| \geq 1$  and  $|x|^j < 1$  if  $|x| < 1$ . Hence,  $|x^j| \leq |x|^k + 1$  for all  $x$ . Consequently, from (6.63) we obtain the inequality

$$\nu_j \leq \nu_k + 1, \quad j \leq k,$$

which implies that  $\mu'_j$  exists for  $j \leq k$ . Since the central moment  $\mu_j$  in formula (6.61) is expressible in terms of noncentral moments of order  $j$  or smaller, the existence of  $\nu_j$  also implies the existence of  $\mu_j$ .

EXAMPLE 6.9.1. Consider a random variable  $X$  with the density function

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty.$$

Such a random variable has the so-called Cauchy distribution. Its mean  $\mu$  does not exist. This follows from the fact that in order of  $\mu$  to exist, the two limits in the following formula must exist:

$$\mu = \frac{1}{\pi} \lim_{a \rightarrow \infty} \int_{-a}^0 \frac{x dx}{1+x^2} + \frac{1}{\pi} \lim_{b \rightarrow \infty} \int_0^b \frac{x dx}{1+x^2}. \quad (6.64)$$

But,  $\int_{-a}^0 x dx / (1 + x^2) = -\frac{1}{2} \log(1 + a^2) \rightarrow -\infty$  as  $a \rightarrow \infty$ , and  $\int_0^b x dx / (1 + x^2) = \frac{1}{2} \log(1 + b^2) \rightarrow \infty$  as  $b \rightarrow \infty$ . The integral  $\int_{-\infty}^{\infty} x dx / (1 + x^2)$  is therefore divergent, and hence  $\mu$  does not exist.

It should be noted that it would be incorrect to state that

$$\mu = \frac{1}{\pi} \lim_{a \rightarrow \infty} \int_{-a}^a \frac{x dx}{1 + x^2}, \quad (6.65)$$

which is equal to zero. This is because the limits in (6.64) must exist for any  $a$  and  $b$  that tend to infinity. The limit in formula (6.65) requires that  $a = b$ . Such a limit is therefore considered as a subsequential limit.

The higher-order moments of the Cauchy distribution do not exist either. It is easy to verify that

$$\mu'_k = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x^k dx}{1 + x^2} \quad (6.66)$$

is divergent for  $k \geq 1$ .

**EXAMPLE 6.9.2.** Consider a random variable  $X$  that has the logistic distribution with the density function

$$f(x) = \frac{e^x}{(1 + e^x)^2}, \quad -\infty < x < \infty.$$

The mean of  $X$  is

$$\begin{aligned} \mu &= \int_{-\infty}^{\infty} \frac{x e^x}{(1 + e^x)^2} dx \\ &= \int_0^1 \log\left(\frac{u}{1-u}\right) du, \end{aligned} \quad (6.67)$$

where  $u = e^x / (1 + e^x)$ . We recognize the integral in (6.67) as being an improper integral of the second kind with singularities at  $u = 0$  and  $u = 1$ . We therefore write

$$\mu = \lim_{a \rightarrow 0^+} \int_a^{1/2} \log\left(\frac{u}{1-u}\right) du + \lim_{b \rightarrow 1^-} \int_{1/2}^b \log\left(\frac{u}{1-u}\right) du. \quad (6.68)$$

Thus,

$$\begin{aligned}
 \mu &= \lim_{a \rightarrow 0^+} [u \log u + (1-u) \log(1-u)]_a^{1/2} \\
 &\quad + \lim_{b \rightarrow 1^-} [u \log u + (1-u) \log(1-u)]_{1/2}^b \\
 &= \lim_{b \rightarrow 1^-} [b \log b + (1-b) \log(1-b)] \\
 &\quad - \lim_{a \rightarrow 0^+} [a \log a + (1-a) \log(1-a)]. \tag{6.69}
 \end{aligned}$$

By applying l'Hospital's rule (Theorem 4.2.6) we find that

$$\lim_{b \rightarrow 1^-} (1-b) \log(1-b) = \lim_{a \rightarrow 0^+} a \log a = 0.$$

We thus have

$$\mu = \lim_{b \rightarrow 1^-} (b \log b) - \lim_{a \rightarrow 0^+} [(1-a) \log(1-a)] = 0.$$

The variance  $\sigma^2$  of  $X$  can be shown to be equal to  $\pi^2/3$  (see Exercise 6.24).

### 6.9.1. The Existence of the First Negative Moment of a Continuous Distribution

Let  $X$  be a continuous random variable with a density function  $f(x)$ . By definition, the first negative moment of  $X$  is  $E(X^{-1})$ . The existence of such a moment will be explored in this section.

The need to evaluate a first negative moment can arise in many practical applications. Here are some examples.

**EXAMPLE 6.9.3.** Let  $\mathcal{P}$  be a population with a mean  $\mu$  and a variance  $\sigma^2$ . The coefficient of variation is a measure of variation in the population per unit mean and is equal to  $\sigma/|\mu|$ , assuming that  $\mu \neq 0$ . An estimate of this ratio is  $s/|\bar{y}|$ , where  $s$  and  $\bar{y}$  are, respectively, the sample standard deviation and sample mean of a sample randomly chosen from  $\mathcal{P}$ . If the population is normally distributed, then  $\bar{y}$  is also normally distributed and is statistically independent of  $s$ . In this case,  $E(s/|\bar{y}|) = E(s)E(1/|\bar{y}|)$ . The question now is whether  $E(1/|\bar{y}|)$  exists or not.

**EXAMPLE 6.9.4 (Calibration or Inverse Regression).** Consider the simple linear regression model

$$E(y) = \beta_0 + \beta_1 x. \tag{6.70}$$

In most regression situations, the interest is in predicting the response  $y$  for

a given value of  $x$ . For this purpose we use the prediction equation  $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ , where  $\hat{\beta}_0$  and  $\hat{\beta}_1$  are the least-squares estimators of  $\beta_0$  and  $\beta_1$ , respectively. These are obtained from the data set  $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$  that results from running  $n$  experiments in which  $y$  is measured for specified settings of  $x$ . There are other situations, however, where the interest is in predicting the value of  $x$ , say  $x_0$ , that corresponds to an observed value of  $y$ , say  $y_0$ . This is an inverse regression problem known as the calibration problem (see Graybill, 1976, Section 8.5; Montgomery and Peck, 1982, Section 9.7).

For example, in calibrating a new type of thermometer,  $n$  readings,  $y_1, y_2, \dots, y_n$ , are taken at predetermined known temperature values,  $x_1, x_2, \dots, x_n$  (these values are known by using a standard temperature gauge). Suppose that the relationship between the  $x_i$ 's and the  $y_i$ 's is well represented by the model in (6.70). If a new reading  $y_0$  is observed using the new thermometer, then it is of interest to estimate the correct temperature  $x_0$  (that is, the temperature on the standard gauge corresponding to the observed temperature reading  $y_0$ ).

In another calibration problem, the date of delivery of a pregnant woman can be estimated by the size  $y$  of the head of her unborn child, which can be determined by a special electronic device (sonogram). If the relationship between  $y$  and the number of days  $x$  left until delivery is well represented by model (6.70), then for a measured value of  $y$ , say  $y_0$ , it is possible to estimate  $x_0$ , the corresponding value of  $x$ .

In general, from model (6.70) we have  $E(y_0) = \beta_0 + \beta_1 x_0$ . If  $\beta_1 \neq 0$ , we can solve for  $x_0$  and obtain

$$x_0 = \frac{E(y_0) - \beta_0}{\beta_1}.$$

Hence, to estimate  $x_0$  we use

$$\hat{x}_0 = \frac{y_0 - \hat{\beta}_0}{\hat{\beta}_1} = \bar{x} + \frac{y_0 - \bar{y}}{\hat{\beta}_1},$$

since  $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ , where  $\bar{x} = (1/n)\sum_{i=1}^n x_i$ ,  $\bar{y} = (1/n)\sum_{i=1}^n y_i$ . If the response  $y$  is normally distributed with a variance  $\sigma^2$ , then  $\bar{y}$  and  $\hat{\beta}_1$  are statistically independent. Since  $y_0$  is also statistically independent of  $\hat{\beta}_1$  ( $y_0$  does not belong to the data set used to estimate  $\beta_1$ ), then the expected value of  $\hat{x}_0$  is given by

$$\begin{aligned} E(\hat{x}_0) &= \bar{x} + E\left(\frac{y_0 - \bar{y}}{\hat{\beta}_1}\right) \\ &= \bar{x} + E(y_0 - \bar{y})E\left(\frac{1}{\hat{\beta}_1}\right). \end{aligned}$$

Here again it is of interest to know if  $E(1/\hat{\beta}_1)$  exists.

Now, suppose that the density function  $f(x)$  of the continuous random variable  $X$  is defined on  $(0, \infty)$ . Let us also assume that  $f(x)$  is continuous. The expected value of  $X^{-1}$  is

$$E(X^{-1}) = \int_0^{\infty} \frac{f(x)}{x} dx. \quad (6.71)$$

This is an improper integral with a singularity at  $x = 0$ . In particular, if  $f(0) > 0$ , then  $E(X^{-1})$  does not exist, because

$$\lim_{x \rightarrow 0^+} \frac{f(x)/x}{1/x} = f(0) > 0.$$

By Theorem 6.5.7, the integrals  $\int_0^{\infty} (f(x)/x) dx$  and  $\int_0^{\infty} (dx/x)$  are of the same kind. Since the latter is divergent, then so is the former. Note that if  $f(x)$  is defined on  $(-\infty, \infty)$  and  $f(0) > 0$ , then  $E(X^{-1})$  does not exist either. In this case,

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{f(x)}{x} dx &= \int_{-\infty}^0 \frac{f(x)}{x} dx + \int_0^{\infty} \frac{f(x)}{x} dx \\ &= - \int_{-\infty}^0 \frac{f(x)}{|x|} dx + \int_0^{\infty} \frac{f(x)}{x} dx. \end{aligned}$$

Both integrals on the right-hand side are divergent.

A sufficient condition for the existence of  $E(X^{-1})$  is given by the following theorem [see Piegorsch and Casella (1985)]:

**Theorem 6.9.1.** Let  $f(x)$  be a continuous density function for a random variable  $X$  defined on  $(0, \infty)$ . If

$$\lim_{x \rightarrow 0^+} \frac{f(x)}{x^{\alpha}} < \infty \quad \text{for some } \alpha > 0, \quad (6.72)$$

then  $E(X^{-1})$  exists.

*Proof.* Since the limit of  $f(x)/x^{\alpha}$  is finite as  $x \rightarrow 0^+$ , there exist finite constants  $M$  and  $\delta > 0$  such that  $f(x)/x^{\alpha} < M$  if  $0 < x < \delta$ . Hence,

$$\int_0^{\delta} \frac{f(x)}{x} dx < M \int_0^{\delta} x^{\alpha-1} dx = \frac{M\delta^{\alpha}}{\alpha}.$$



Thus,

$$\begin{aligned} E(X^{-1}) &= \int_0^\infty \frac{f(x)}{x} dx = \int_0^\delta \frac{f(x)}{x} dx + \int_\delta^\infty \frac{f(x)}{x} dx \\ &< \frac{M\delta^\alpha}{\alpha} + \frac{1}{\delta} \int_\delta^\infty f(x) dx \leq \frac{M\delta^\alpha}{\alpha} + \frac{1}{\delta} < \infty. \quad \square \end{aligned}$$

It should be noted that the condition of Theorem 6.9.1 is not a necessary one. Piegorsch and Casella (1985) give an example of a family of density functions that all violate condition (6.72), with some members having a finite first negative moment and others not having one (see Exercise 6.25).

**Corollary 6.9.1.** Let  $f(x)$  be a continuous density function for a random variable  $X$  defined on  $(0, \infty)$  such that  $f(0) = 0$ . If  $f'(0)$  exists and is finite, then  $E(X^{-1})$  exists.

*Proof.* We have that

$$f'(0) = \lim_{x \rightarrow 0^+} \frac{f(x) - f(0)}{x} = \lim_{x \rightarrow 0^+} \frac{f(x)}{x}.$$

By applying Theorem 6.9.1 with  $\alpha = 1$  we conclude that  $E(X^{-1})$  exists.  $\square$

**EXAMPLE 6.9.5.** Let  $X$  be a normal random variable with a mean  $\mu$  and a variance  $\sigma^2$ . Its density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x - \mu)^2\right], \quad -\infty < x < \infty.$$

In this example,  $f(0) > 0$ . Hence,  $E(X^{-1})$  does not exist. Consequently,  $E(1/|\bar{y}|)$  in Example 6.9.3 does not exist if the population  $\mathcal{P}$  is normally distributed, since the density function of  $|\bar{y}|$  is positive at zero. Also, in Example 6.9.4,  $E(1/\hat{\beta}_1)$  does not exist, because  $\hat{\beta}_1$  is normally distributed if the response  $y$  satisfies the assumption of normality.

**EXAMPLE 6.9.6.** Let  $X$  be a continuous random variable with the density function

$$f(x) = \frac{1}{\Gamma(n/2)2^{n/2}} x^{n/2-1} e^{-x/2}, \quad 0 < x < \infty,$$

where  $n$  is a positive integer and  $\Gamma(n/2)$  is the value of the gamma function,  $\int_0^\infty e^{-x} x^{n/2-1} dx$ . This is the density function of a chi-squared random variable with  $n$  degrees of freedom.

Let us consider the limit

$$\lim_{x \rightarrow 0^+} \frac{f(x)}{x^\alpha} = \frac{1}{\Gamma(n/2)2^{n/2}} \lim_{x \rightarrow 0^+} x^{n/2-\alpha-1} e^{-x/2}$$

for  $\alpha > 0$ . This limit exists and is equal to zero if  $n/2 - \alpha - 1 > 0$ , that is, if  $n > 2(1 + \alpha) > 2$ . Thus by Theorem 6.9.1,  $E(X^{-1})$  exists if the number of degrees of freedom exceeds 2.

More recently, Khuri and Casella (2002) presented several extensions and generalizations of the results in Piegorsch and Casella (1985), including a necessary and sufficient condition for the existence of  $E(X^{-1})$ .

### 6.9.2. Transformation of Continuous Random Variables

Let  $Y$  be a continuous random variable with a density function  $f(y)$ . Let  $W = \psi(Y)$ , where  $\psi(y)$  is a function whose derivative exists and is continuous on a set  $A$ . Suppose that  $\psi'(y) \neq 0$  for all  $y \in A$ , that is,  $\psi(y)$  is strictly monotone. We recall from Section 4.5.1 that the density function  $g(w)$  of  $W$  is given by

$$g(w) = f[\psi^{-1}(w)] \left| \frac{d\psi^{-1}(w)}{dw} \right|, \quad w \in B, \quad (6.73)$$

where  $B$  is the image of  $A$  under  $\psi$  and  $y = \psi^{-1}(w)$  is the inverse function of  $w = \psi(y)$ . This result can be easily obtained by applying the change of variables technique in Section 6.4.1. This is done as follows: We have that for any  $w_1$  and  $w_2$  such that  $w_1 \leq w_2$ ,

$$P(w_1 \leq W \leq w_2) = \int_{w_1}^{w_2} g(w) dw. \quad (6.74)$$

If  $\psi'(y) > 0$ , then  $\psi(y)$  is strictly monotone increasing. Hence,

$$P(w_1 \leq W \leq w_2) = P(y_1 \leq Y \leq y_2), \quad (6.75)$$

where  $y_1$  and  $y_2$  are such that  $w_1 = \psi(y_1)$ ,  $w_2 = \psi(y_2)$ . But

$$P(y_1 \leq Y \leq y_2) = \int_{y_1}^{y_2} f(y) dy. \quad (6.76)$$

Let us now apply the change of variables  $y = \psi^{-1}(w)$  to the integral in (6.76).

By Theorem 6.4.9 we have

$$\int_{y_1}^{y_2} f(y) dy = \int_{w_1}^{w_2} f[\psi^{-1}(w)] \frac{d\psi^{-1}(w)}{dw} dw. \quad (6.77)$$

From (6.75) and (6.76) we obtain

$$P(w_1 \leq W \leq w_2) = \int_{w_1}^{w_2} f[\psi^{-1}(w)] \frac{d\psi^{-1}(w)}{dw} dw. \quad (6.78)$$

On the other hand, if  $\psi'(y) < 0$ , then  $P(w_1 \leq W \leq w_2) = P(y_2 \leq Y \leq y_1)$ . Consequently,

$$\begin{aligned} P(w_1 \leq W \leq w_2) &= \int_{y_2}^{y_1} f(y) dy \\ &= \int_{w_2}^{w_1} f[\psi^{-1}(w)] \frac{d\psi^{-1}(w)}{dw} dw \\ &= - \int_{w_1}^{w_2} f[\psi^{-1}(w)] \frac{d\psi^{-1}(w)}{dw} dw. \end{aligned} \quad (6.79)$$

By combining (6.78) and (6.79) we obtain

$$P(w_1 \leq W \leq w_2) = \int_{w_1}^{w_2} f[\psi^{-1}(w)] \left| \frac{d\psi^{-1}(w)}{dw} \right| dw. \quad (6.80)$$

Formula (6.73) now follows from comparing (6.74) and (6.80).

#### *The Case Where $w = \psi(y)$ Has No Unique Inverse*

Formula (6.73) requires that  $w = \psi(y)$  has a unique inverse, the existence of which is guaranteed by the nonvanishing of the derivative  $\psi'(y)$ . Let us now consider the following extension: The function  $\psi(y)$  is continuously differentiable, but its derivative can vanish at a finite number of points inside its domain. We assume that the domain of  $\psi(y)$  can be partitioned into a finite number, say  $n$ , of disjoint subdomains, denoted by  $I_1, I_2, \dots, I_n$ , on each of which  $\psi(y)$  is strictly monotone (decreasing or increasing). Hence, on each  $I_i$  ( $i = 1, 2, \dots, n$ ),  $\psi(y)$  has a unique inverse. Let  $\psi_i$  denote the restriction of the function  $\psi$  to  $I_i$ , that is,  $\psi_i(y)$  has a unique inverse,  $y = \psi_i^{-1}(w)$ ,  $i = 1, 2, \dots, n$ . Since  $I_1, I_2, \dots, I_n$  are disjoint, for any  $w_1$  and  $w_2$  such that  $w_1 \leq w_2$  we have

$$P(w_1 \leq W \leq w_2) = \sum_{i=1}^n P[Y \in \psi_i^{-1}(w_1, w_2)], \quad (6.81)$$

where  $\psi_i^{-1}(w_1, w_2)$  is the inverse image of  $[w_1, w_2]$ , which is a subset of  $I_i$ ,  $i = 1, 2, \dots, n$ . Now, on the  $i$ th subdomain we have

$$\begin{aligned} P[Y \in \psi_i^{-1}(w_1, w_2)] &= \int_{\psi_i^{-1}(w_1, w_2)} f(y) dy \\ &= \int_{T_i} f[\psi_i^{-1}(w)] \left| \frac{d\psi_i^{-1}(w)}{dw} \right| dw, \quad i = 1, 2, \dots, n, \end{aligned} \quad (6.82)$$

where  $T_i$  is the image of  $\psi_i^{-1}(w_1, w_2)$  under  $\psi_i$ . Formula (6.82) follows from applying formula (6.80) to the function  $\psi_i(y)$ ,  $i = 1, 2, \dots, n$ . Note that  $T_i = \psi_i[\psi_i^{-1}(w_1, w_2)] = [w_1, w_2] \cap \psi_i(I_i)$ ,  $i = 1, 2, \dots, n$  (why?). Thus  $T_i$  is a subset of both  $[w_1, w_2]$  and  $\psi_i(I_i)$ . We can therefore write the integral in (6.82) as

$$\int_{T_i} f[\psi_i^{-1}(w)] \left| \frac{d\psi_i^{-1}(w)}{dw} \right| dw = \int_{w_1}^{w_2} \delta_i(w) f[\psi_i^{-1}(w)] \left| \frac{d\psi_i^{-1}(w)}{dw} \right| dw, \quad (6.83)$$

where  $\delta_i(w) = 1$  if  $w \in \psi_i(I_i)$  and  $\delta_i(w) = 0$  otherwise,  $i = 1, 2, \dots, n$ . Using (6.82) and (6.83) in formula (6.81), we obtain

$$\begin{aligned} P(w_1 \leq W \leq w_2) &= \sum_{i=1}^n \int_{w_1}^{w_2} \delta_i(w) f[\psi_i^{-1}(w)] \left| \frac{d\psi_i^{-1}(w)}{dw} \right| dw \\ &= \int_{w_1}^{w_2} \sum_{i=1}^n \delta_i(w) f[\psi_i^{-1}(w)] \left| \frac{d\psi_i^{-1}(w)}{dw} \right| dw, \end{aligned}$$

from which we deduce that the density function of  $W$  is given by

$$g(w) = \sum_{i=1}^n \delta_i(w) f[\psi_i^{-1}(w)] \left| \frac{d\psi_i^{-1}(w)}{dw} \right|. \quad (6.84)$$

EXAMPLE 6.9.7. Let  $Y$  have the standard normal distribution with the density function

$$f(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right), \quad -\infty < y < \infty.$$

Define the random variable  $W$  as  $W = Y^2$ . In this case, the function  $w = y^2$

has two inverse functions on  $(-\infty, \infty)$ , namely,

$$y = \begin{cases} -\sqrt{w}, & y \leq 0, \\ \sqrt{w}, & y > 0. \end{cases}$$

Thus  $I_1 = (-\infty, 0]$ ,  $I_2 = (0, \infty)$ , and  $\psi_1(I_1) = [0, \infty)$ ,  $\psi_2(I_2) = (0, \infty)$ . Hence,  $\delta_1(w) = 1$ ,  $\delta_2(w) = 1$  if  $w \in (0, \infty)$ . By applying formula (6.84) we then get

$$\begin{aligned} g(w) &= \frac{1}{\sqrt{2\pi}} e^{-w/2} \left| \frac{-1}{2\sqrt{w}} \right| + \frac{1}{\sqrt{2\pi}} e^{-w/2} \left| \frac{1}{2\sqrt{w}} \right| \\ &= \frac{1}{\sqrt{2\pi}} \frac{e^{-w/2}}{\sqrt{w}}, \quad w > 0. \end{aligned}$$

This represents the density function of a chi-squared random variable with one degree of freedom.

### 6.9.3. The Riemann–Stieltjes Integral Representation of the Expected Value

Let  $X$  be a random variable with a cumulative distribution function  $F(x)$ . Suppose that  $h(x)$  is Riemann–Stieltjes integrable with respect to  $F(x)$  on  $(-\infty, \infty)$ . Then the expected value of  $h(X)$  is defined as

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) dF(x). \quad (6.85)$$

Formula (6.85) provides a unified representation of expected values for both discrete and continuous random variables.

If  $X$  is a continuous random variable with a density function  $f(x)$ , that is,  $F'(x) = f(x)$ , then

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) f(x) dx.$$

If, however,  $X$  has a discrete distribution with a probability mass function  $p(x)$  and takes the values  $c_1, c_2, \dots, c_n$ , then

$$E[h(X)] = \sum_{i=1}^n h(c_i) p(c_i). \quad (6.86)$$

Formula (6.86) follows from applying Theorem 6.8.4. Here,  $F(x)$  is a step

function with jump discontinuities at  $c_1, c_2, \dots, c_n$  such that

$$F(x) = \begin{cases} 0, & -\infty < x < c_1 \\ p(c_1), & c_1 \leq x < c_2, \\ p(c_1) + p(c_2), & c_2 \leq x < c_3, \\ \vdots \\ \sum_{i=1}^{n-1} p(c_i), & c_{n-1} \leq x < c_n, \\ \sum_{i=1}^n p(c_i) = 1, & c_n \leq x < \infty. \end{cases}$$

Thus, by formula (6.56) we obtain

$$\int_{-\infty}^{\infty} h(x) dF(x) = \sum_{i=1}^n p(c_i) h(c_i). \quad (6.87)$$

For example, suppose that  $X$  has the discrete uniform distribution with  $p(x) = 1/n$  for  $x = c_1, c_2, \dots, c_n$ . Its cumulative distribution function  $F(c)$  can be expressed as

$$F(x) = P[X \leq x] = \frac{1}{n} \sum_{i=1}^n I(x - c_i),$$

where  $I(x - c_i)$  is equal to zero if  $x < c_i$  and is equal to one if  $x \geq c_i$  ( $i = 1, 2, \dots, n$ ). The expected value of  $h(X)$  is

$$E[h(X)] = \frac{1}{n} \sum_{i=1}^n h(c_i).$$

EXAMPLE 6.9.8. The moment generating function  $\phi(t)$  of a random variable  $X$  with a cumulative distribution function  $F(x)$  is defined as the expected value of  $h_t(X) = e^{tX}$ , that is,

$$\phi(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} dF(x),$$

where  $t$  is a scalar. If  $X$  is a discrete random variable with a probability mass function  $p(x)$  and takes the values  $c_1, c_2, \dots, c_n, \dots$ , then by letting  $n$  go to infinity in (6.87) we obtain (see also Section 5.6.2)

$$\phi(t) = \sum_{i=1}^{\infty} p(c_i) e^{tc_i}.$$

The moment generating function of a continuous random variable with a density function  $f(x)$  is

$$\phi(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx. \quad (6.88)$$

The convergence of the integral in (6.88) depends on the choice of the scalar  $t$ . For example, for the gamma distribution  $G(\alpha, \beta)$  with the density function

$$f(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha) \beta^\alpha}, \quad \alpha > 0, \quad \beta > 0, \quad 0 < x < \infty,$$

$\phi(t)$  is of the form

$$\begin{aligned} \phi(t) &= \int_0^\infty \frac{e^{tx} x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha) \beta^\alpha} dx \\ &= \int_0^\infty \frac{x^{\alpha-1} \exp[-x(1-\beta t)/\beta]}{\Gamma(\alpha) \beta^\alpha} dx. \end{aligned}$$

If we set  $y = x(1-\beta t)/\beta$ , we obtain

$$\phi(t) = \int_0^\infty \frac{\beta}{(1-\beta t) \Gamma(\alpha) \beta^\alpha} \left( \frac{\beta y}{1-\beta t} \right)^{\alpha-1} e^{-y} dy.$$

Thus,

$$\begin{aligned} \phi(t) &= \frac{1}{(1-\beta t)^\alpha} \int_0^\infty \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)} dy \\ &= (1-\beta t)^{-\alpha}, \end{aligned}$$

since  $\int_0^\infty e^{-y} y^{\alpha-1} dy = \Gamma(\alpha)$  by the definition of the gamma function. We note that  $\phi(t)$  exists for all values of  $\alpha$  provided that  $1-\beta t > 0$ , that is,  $t < 1/\beta$ .

#### 6.9.4. Chebyshev's Inequality

In Section 5.6.1 there was a mention of Chebyshev's inequality. Using the Riemann–Stieltjes integral representation of the expected value, it is now possible to provide a proof for this important inequality.

**Theorem 6.9.2.** Let  $X$  be a random variable (discrete or continuous) with a mean  $\mu$  and a variance  $\sigma^2$ . Then, for any positive constant  $r$ ,

$$P(|X - \mu| \geq r\sigma) \leq \frac{1}{r^2}.$$

*Proof.* By definition,  $\sigma^2$  is the expected value of  $h(X) = (X - \mu)^2$ . Thus,

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 dF(x), \quad (6.89)$$

where  $F(x)$  is the cumulative distribution function of  $X$ . Let us now partition  $(-\infty, \infty)$  into three disjoint intervals:  $(-\infty, \mu - r\sigma]$ ,  $(\mu - r\sigma, \mu + r\sigma)$ ,  $[\mu + r\sigma, \infty)$ . The integral in (6.89) can therefore be written as

$$\begin{aligned} \sigma^2 &= \int_{-\infty}^{\mu - r\sigma} (x - \mu)^2 dF(x) + \int_{\mu - r\sigma}^{\mu + r\sigma} (x - \mu)^2 dF(x) \\ &\quad + \int_{\mu + r\sigma}^{\infty} (x - \mu)^2 dF(x) \\ &\geq \int_{-\infty}^{\mu - r\sigma} (x - \mu)^2 dF(x) + \int_{\mu + r\sigma}^{\infty} (x - \mu)^2 dF(x). \end{aligned} \quad (6.90)$$

We note that in the first integral in (6.90),  $x \leq \mu - r\sigma$ , so that  $x - \mu \leq -r\sigma$ . Hence,  $(x - \mu)^2 \geq r^2\sigma^2$ . Also, in the second integral,  $x - \mu \geq r\sigma$ . Hence,  $(x - \mu)^2 \geq r^2\sigma^2$ . Consequently,

$$\begin{aligned} \int_{-\infty}^{\mu - r\sigma} (x - \mu)^2 dF(x) &\geq r^2\sigma^2 \int_{-\infty}^{\mu - r\sigma} dF(x) = r^2\sigma^2 P(X \leq \mu - r\sigma), \\ \int_{\mu + r\sigma}^{\infty} (x - \mu)^2 dF(x) &\geq r^2\sigma^2 \int_{\mu + r\sigma}^{\infty} dF(x) = r^2\sigma^2 P(X \geq \mu + r\sigma). \end{aligned}$$

From inequality (6.90) we then have

$$\begin{aligned} \sigma^2 &\geq r^2\sigma^2 [P(X - \mu \leq -r\sigma) + P(X - \mu \geq r\sigma)] \\ &= r^2\sigma^2 P(|X - \mu| \geq r\sigma), \end{aligned}$$

which implies that

$$P(|X - \mu| \geq r\sigma) \leq \frac{1}{r^2}. \quad \square$$

## FURTHER READING AND ANNOTATED BIBLIOGRAPHY

DeCani, J. S., and R. A. Stine (1986). "A note on deriving the information matrix for a logistic distribution." *Amer. Statist.*, **40**, 220–222. (This article uses calculus techniques, such as integration and l'Hospital's rule, in determining the mean and variance of the logistic distribution as was seen in Example 6.9.2.)

Fulks, W. (1978). *Advanced Calculus*, 3rd ed. Wiley, New York. (Chap. 5 discusses the Riemann integral; Chap. 16 provides a study of improper integrals.)



- Graybill, F. A. (1976). *Theory and Application of the Linear Model*. Duxbury Press, North Scituate, Massachusetts. (Section 8.5 discusses the calibration problem for a simple linear regression model as was seen in Example 6.9.4.)
- Hardy, G. H., J. E. Littlewood, and G. Pólya (1952). *Inequalities*, 2nd ed. Cambridge University Press, Cambridge, England. (This is a classic and often referenced book on inequalities. Chap. 6 is relevant to the present chapter.)
- Hartig, D. (1991). “L’Hôpital’s rule via integration.” *Amer. Math. Monthly*, **98**, 156–157.
- Khuri, A. I., and G. Casella (2002). “The existence of the first negative moment revisited.” *Amer. Statist.*, **56**, 44–47. (This article demonstrates the utility of the comparison test given in Theorem 6.5.3 in showing the existence of the first negative moment of a continuous random variable.)
- Lindgren, B. W. (1976). *Statistical Theory*, 3rd ed. Macmillan, New York. (Section 2.2.2 gives the Riemann–Stieltjes integral representation of the expected value of a random variable as was seen in Section 6.9.3.)
- Montgomery, D. C., and E. A. Peck (1982). *Introduction to Linear Regression Analysis*. Wiley, New York. (The calibration problem for a simple linear regression model is discussed in Section 9.7.)
- Moran, P. A. P. (1968). *An Introduction to Probability Theory*. Clarendon Press, Oxford, England. (Section 5.9 defines moments of a random variable using the Riemann–Stieltjes integral representation of the expected value; Section 5.10 discusses a number of inequalities pertaining to these moments.)
- Piegorsch, W. W., and G. Casella (1985). “The existence of the first negative moment.” *Amer. Statist.*, **39**, 60–62. (This article gives a sufficient condition for the existence of the first negative moment of a continuous random variable as was seen in Section 6.9.1.)
- Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley, Reading, Massachusetts. (Chap. 9 is concerned with transformations of continuous random variables as was seen in Section 6.9.2. See, in particular, Theorems 2 and 3 in this chapter.)
- Taylor, A. E., and W. R. Mann (1972). *Advanced Calculus*, 2nd ed. Wiley, New York. (Chap. 18 discusses the Riemann integral as well as the Riemann–Stieltjes integral; improper integrals are studied in Chap. 22.)
- Wilks, S. S. (1962). *Mathematical Statistics*. Wiley, New York. (Chap. 3 uses the Riemann–Stieltjes integral to define expected values and moments of random variables; functions of random variables are discussed in Section 2.8.)

## EXERCISES

### In Mathematics

- 6.1.** Let  $f(x)$  be a bounded function defined on the interval  $[a, b]$ . Let  $P$  be a partition of  $[a, b]$ . Show that  $f(x)$  is Riemann integrable on  $[a, b]$  if and only if

$$\inf_P US_P(f) = \sup_P LS_P(f) = \int_a^b f(x) dx.$$

- 6.2.** Construct a function that has a countable number of discontinuities in  $[0, 1]$  and is Riemann integrable on  $[0, 1]$ .
- 6.3.** Show that if  $f(x)$  is continuous on  $[a, b]$  except for a finite number of discontinuities of the first kind (see Definition 3.4.2), then  $f(x)$  is Riemann integrable on  $[a, b]$ .
- 6.4.** Show that the function

$$f(x) = \begin{cases} x \cos(\pi/2x), & 0 < x \leq 1, \\ 0, & x = 0, \end{cases}$$

is not of bounded variation on  $[0, 1]$ .

- 6.5.** Let  $f(x)$  and  $g(x)$  have continuous derivatives with  $g'(x) > 0$ . Suppose that  $\lim_{x \rightarrow \infty} f(x) = \infty$ ,  $\lim_{x \rightarrow \infty} g(x) = \infty$ , and  $\lim_{x \rightarrow \infty} f'(x)/g'(x) = L$ , where  $L$  is finite.
- (a) Show that for a given  $\epsilon > 0$  there exists a constant  $M > 0$  such that for  $x > M$ ,

$$|f'(x) - Lg'(x)| < \epsilon g'(x).$$

Hence, if  $\lambda_1$  and  $\lambda_2$  are such that  $M < \lambda_1 < \lambda_2$ , then

$$\left| \int_{\lambda_1}^{\lambda_2} [f'(x) - Lg'(x)] dx \right| < \int_{\lambda_1}^{\lambda_2} \epsilon g'(x) dx.$$

- (b) Deduce from (a) that

$$\left| \frac{f(\lambda_2)}{g(\lambda_2)} - L \right| < \epsilon + \frac{|f(\lambda_1)|}{g(\lambda_2)} + |L| \frac{g(\lambda_1)}{g(\lambda_2)}.$$

- (c) Make use of (b) to show that for a sufficiently large  $\lambda_2$ ,

$$\left| \frac{f(\lambda_2)}{g(\lambda_2)} - L \right| < 3\epsilon,$$

and hence  $\lim_{x \rightarrow \infty} f(x)/g(x) = L$ .

[Note: This problem verifies l'Hospital's rule for the  $\infty/\infty$  indeterminate form by using integration properties without relying on Cauchy's mean value theorem as in Section 4.2 (see Hartig, 1991).]

- 6.6.** Show that if  $f(x)$  is continuous on  $[a, b]$ , and if  $g(x)$  is a nonnegative Riemann integrable function on  $[a, b]$  such that  $\int_a^b g(x) dx > 0$ , then

there exists a constant  $c$ ,  $a \leq c \leq b$ , such that

$$\frac{\int_a^b f(x)g(x) dx}{\int_a^b g(x) dx} = f(c).$$

**6.7.** Prove Theorem 6.5.2.

**6.8.** Prove Theorem 6.5.3.

**6.9.** Suppose that  $f(x)$  is a positive monotone decreasing function defined on  $[a, \infty)$  such that  $f(x) \rightarrow 0$  as  $x \rightarrow \infty$ . Show that if  $f(x)$  is Riemann–Stieltjes integrable with respect to  $g(x)$  on  $[a, b]$  for every  $b \geq a$ , where  $g(x)$  is bounded on  $[a, \infty)$ , then the integral  $\int_a^\infty f(x) dg(x)$  is convergent.

**6.10.** Show that  $\lim_{n \rightarrow \infty} \int_0^{n\pi} |(\sin x)/x| dx = \infty$ , where  $n$  is a positive integer. [Hint: Show first that

$$\int_0^{n\pi} \left| \frac{\sin x}{x} \right| dx = \int_0^\pi \sin x \left[ \frac{1}{x} + \frac{1}{x+\pi} + \cdots + \frac{1}{x+(n-1)\pi} \right] dx.]$$

**6.11.** Apply Maclaurin's integral test to determine convergence or divergence of the following series:

(a) 
$$\sum_{n=1}^{\infty} \frac{\log n}{n\sqrt{n}},$$

(b) 
$$\sum_{n=1}^{\infty} \frac{n+4}{2n^3+1},$$

(c) 
$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n+1}-1}.$$

**6.12.** Consider the sequence  $\{f_n(x)\}_{n=1}^\infty$ , where  $f_n(x) = nx/(1+nx^2)$ ,  $x \geq 0$ . Find the limit of  $\int_1^2 f_n(x) dx$  as  $n \rightarrow \infty$ .

**6.13.** Consider the improper integral  $\int_0^1 x^{m-1}(1-x)^{n-1} dx$ .

- (a) Show that the integral converges if  $m > 0$ ,  $n > 0$ . In this case, the function  $B(m, n)$  defined as

$$B(m, n) = \int_0^1 x^{m-1} (1-x)^{n-1} dx$$

is called the beta function.

- (b) Show that

$$B(m, n) = 2 \int_0^{\pi/2} \sin^{2m-1} \theta \cos^{2n-1} \theta d\theta$$

- (c) Show that

$$B(m, n) = \int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}} dx = \int_0^\infty \frac{x^{n-1}}{(1+x)^{m+n}} dx$$

- (d) Show that

$$B(m, n) = \int_0^1 \frac{x^{m-1} + x^{n-1}}{(1+x)^{m+n}} dx.$$

- 6.14.** Determine whether each of the following integrals is convergent or divergent:

(a) 
$$\int_0^\infty \frac{dx}{\sqrt{1+x^3}},$$

(b) 
$$\int_0^\infty \frac{dx}{(1+x^3)^{1/3}},$$

(c) 
$$\int_0^1 \frac{dx}{(1-x^3)^{1/3}},$$

(d) 
$$\int_0^\infty \frac{dx}{\sqrt{x}(1+2x)}.$$

- 6.15.** Let  $f_1(x)$  and  $f_2(x)$  be bounded on  $[a, b]$ , and  $g(x)$  be monotone increasing on  $[a, b]$ . If  $f_1(x)$  and  $f_2(x)$  are Riemann–Stieltjes integrable with respect to  $g(x)$  on  $[a, b]$ , then show that  $f_1(x)f_2(x)$  is also Riemann–Stieltjes integrable with respect to  $g(x)$  on  $[a, b]$ .

**6.16.** Let  $f(x)$  be a function whose first  $n$  derivatives are continuous on  $[a, b]$ , and let

$$h_n(x) = f(b) - f(x) - (b-x)f'(x) - \cdots - \frac{(b-x)^{n-1}}{(n-1)!} f^{(n-1)}(x).$$

Show that

$$h_n(a) = \frac{1}{(n-1)!} \int_a^b (b-x)^{n-1} f^{(n)}(x) dx$$

and hence

$$\begin{aligned} f(b) &= f(a) + (b-a)f'(a) + \cdots + \frac{(b-a)^{n-1}}{(n-1)!} f^{(n-1)}(a) \\ &\quad + \frac{1}{(n-1)!} \int_a^b (b-x)^{n-1} f^{(n)}(x) dx. \end{aligned}$$

This represents Taylor's expansion of  $f(x)$  around  $x=a$  (see Section 4.3) with a remainder  $R_n$  given by

$$R_n = \frac{1}{(n-1)!} \int_a^b (b-x)^{n-1} f^{(n)}(x) dx.$$

[Note: This form of Taylor's theorem has the advantage of providing an exact formula for  $R_n$ , which does not involve an undetermined number  $\theta_n$  as was seen in Section 4.3.]

**6.17.** Suppose that  $f(x)$  is monotone and its derivative  $f'(x)$  is Riemann integrable on  $[a, b]$ . Let  $g(x)$  be continuous on  $[a, b]$ . Show that there exists a number  $c$ ,  $a \leq c \leq b$ , such that

$$\int_a^b f(x)g(x) dx = f(a) \int_a^c g(x) dx + f(b) \int_c^b g(x) dx.$$

**6.18.** Deduce from Exercise 6.17 that for any  $b > a > 0$ ,

$$\left| \int_a^b \frac{\sin x}{x} dx \right| \leq \frac{4}{a}.$$

**6.19.** Prove Lemma 6.7.1.

- 6.20.** Show that if  $f_1(x), f_2(x), \dots, f_n(x)$  are such that  $|f_i(x)|^p$  is Riemann integrable on  $[a, b]$  for  $i = 1, 2, \dots, n$ , where  $1 \leq p < \infty$ , then

$$\left[ \int_a^b \left| \sum_{i=1}^n f_i(x) \right|^p dx \right]^{1/p} \leq \sum_{i=1}^n \left[ \int_a^b |f_i(x)|^p dx \right]^{1/p}.$$

- 6.21.** Prove Theorem 6.8.1.

- 6.22.** Prove Theorem 6.8.2.

### In Statistics

- 6.23.** Show that the integral in formula (6.66) is divergent for  $k \geq 1$ .
- 6.24.** Consider the random variable  $X$  that has the logistic distribution described in Example 6.9.2. Show that  $\text{Var}(X) = \pi^2/3$ .
- 6.25.** Let  $\{f_n(x)\}_{n=1}^\infty$  be a family density functions defined by

$$f_n(x) = \frac{|\log^n x|^{-1}}{\int_0^\lambda |\log^n t|^{-1} dt}, \quad 0 < x < \lambda,$$

where  $\lambda \in (0, 1)$ .

- (a) Show that condition (6.72) of Theorem 6.9.1 is not satisfied by any  $f_n(x)$ ,  $n \geq 1$ .
- (b) Show that when  $n = 1$ ,  $E(X^{-1})$  does not exist, where  $X$  is a random variable with the density function  $f_1(x)$ .
- (c) Show that for  $n > 1$ ,  $E(X_n^{-1})$  exists, where  $X_n$  is a random variable with the density function  $f_n(x)$ .
- 6.26.** Let  $X$  be a random variable with a continuous density function  $f(x)$  on  $(0, \infty)$ . Suppose that  $f(x)$  is bounded near zero. Then  $E(X^{-\alpha})$  exists, where  $\alpha \in (0, 1)$ .
- 6.27.** Let  $X$  be a random variable with a continuous density function  $f(x)$  on  $(0, \infty)$ . If  $\lim_{x \rightarrow 0^+} (f(x)/x^\alpha)$  is equal to a positive constant  $k$  for some  $\alpha > 0$ , then  $E[X^{-(1+\alpha)}]$  does not exist.

- 6.28.** The random variable  $Y$  has the  $t$ -distributions with  $n$  degrees of freedom. Its density function is given by

$$f(y) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{y^2}{n}\right)^{-(n+1)/2}, \quad -\infty < y < \infty,$$

where  $\Gamma(m)$  is the gamma function  $\int_0^\infty e^{-x} x^{m-1} dx$ ,  $m > 0$ . Find the density function of  $W = |Y|$ .

- 6.29.** Let  $X$  be a random variable with a mean  $\mu$  and a variance  $\sigma^2$ .  
 (a) Show that Chebyshev's inequality can be expressed as

$$P(|X - \mu| \geq r) \leq \frac{\sigma^2}{r^2},$$

where  $r$  is any positive constant.

- (b) Let  $\{X_n\}_{n=1}^\infty$  be a sequence of independent and identically distributed random variables. If the common mean and variance of the  $X_i$ 's are  $\mu$  and  $\sigma^2$ , respectively, then show that

$$P(|\bar{X}_n - \mu| \geq r) \leq \frac{\sigma^2}{nr^2},$$

where  $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$  and  $r$  is any positive constant.

- (c) Deduce from (b) that  $\bar{X}_n$  converges in probability to  $\mu$  as  $n \rightarrow \infty$ , that is, for every  $\epsilon > 0$ ,

$$P(|\bar{X}_n - \mu| \geq \epsilon) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

- 6.30.** Let  $X$  be a random variable with a cumulative distribution function  $F(x)$ . Let  $\mu'_k$  be its  $k$ th noncentral moment,

$$\mu'_k = E(X^k) = \int_{-\infty}^{\infty} x^k dF(x).$$

Let  $\nu_k$  be the  $k$ th absolute moment of  $X$ ,

$$\nu_k = E(|X|^k) = \int_{-\infty}^{\infty} |x|^k dF(x).$$

Suppose that  $\nu_k$  exists for  $k = 1, 2, \dots, n$ .

- (a) Show that  $\nu_k^2 \leq \nu_{k-1} \nu_{k+1}$ ,  $k = 1, 2, \dots, n-1$ . [*Hint:* For any  $u$  and  $v$ ,

$$\begin{aligned} 0 &\leq \int_{-\infty}^{\infty} \left[ u|x|^{(k-1)/2} + v|x|^{(k+1)/2} \right]^2 dF(x) \\ &= u^2 \nu_{k-1} + 2uv\nu_k + v^2 \nu_{k+1}. \end{aligned}$$

- (b) Deduce from (a) that

$$\nu_1 \leq \nu_2^{1/2} \leq \nu_3^{1/3} \leq \dots \leq \nu_n^{1/n}.$$