Mathematics Primer for Introduction to Mathematical Statistics 8th Edition

Joseph W. McKean

Contents

Chapter 1

Introduction

The purpose of this appendix is to review some of the calculus concepts which are used in the text. It is not meant to be complete and readers who have difficulty with some of these concepts are advised to consult a calculus text for more explanation. We do reference Buck's (1965) advanced calculus book but our discussion concerns fundamental results which can be found in any calculus book.

For the most part, we state and discuss the concepts. In the section on sequences, we spend a little time on proofs since the reader may not be familiar with a few of the concepts such as limit supremums.

Chapter 2

Sequences

Let $\{a_n\}$ be a sequence of real numbers. Recall from calculus that $a_n \to a$ ($\lim_{n\to\infty} a_n = a$) if and only if

$$\text{for every } \epsilon > 0, \text{ there exists an } N_0 \text{ such that } n \geq N_0 \implies |a_n - a| < \epsilon. \quad (2.0.1)$$

Let A be a set of real numbers which is bounded from above; that is, there exists an $M \in R$ such that $x \leq M$ for all $x \in A$. Recall that a is the **supremum** of A if a is the least of all upper bounds of A. From calculus, we know that the supremum of a set bounded from above exists. Furthermore, we know that a is the supremum of A if and only if, for all $\epsilon > 0$, there exists an $x \in A$ such that $a - \epsilon < x \leq a$. Similarly, we can define the **infimum** of A. The first theorem on limits is the Sandwich Theorem.

Theorem 2.0.1 (Sandwich Theorem)**.** *Suppose for sequences $\{a_n\}$, $\{b_n\}$, and $\{c_n\}$ that $c_n \leq a_n \leq b_n$, for all n, and that $\lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n = a$. Then $\lim_{n\to\infty} a_n = a$.*

Proof: Let $\epsilon > 0$ be given. Because both $\{b_n\}$ and $\{c_n\}$ converge, we can choose N_0 so large that $|c_n - a| < \epsilon$ and $|b_n - a| < \epsilon$, for $n \geq N_0$. Because $c_n \leq a_n \leq b_n$, it is easy to see that

$$|a_n - a| \leq \max\{|c_n - a|, |b_n - a|\},$$

for all n. Hence, if $n \geq N_0$, then $|a_n - a| < \epsilon$. ■

The second fact concerns subsequences. Recall that $\{a_{n_k}\}$ is a **subsequence** of $\{a_n\}$ if the sequence $n_1 \leq n_2 \leq \cdots$ is an infinite subset of the positive integers. Note that $n_k \geq k$.

Theorem 2.0.2. *The sequence $\{a_n\}$ converges to a if and only if every subsequence $\{a_{n_k}\}$ converges to a.*

Proof: Suppose the sequence $\{a_n\}$ converges to a. Let $\{a_{n_k}\}$ be any subsequence. Let $\epsilon > 0$ be given. Then there exists an N_0 such that $|a_n - a| < \epsilon$, for $n \geq N_0$. For the subsequence, take k' to be the first index of the subsequence beyond N_0.

Because for all k, $n_k \geq k$, we have that $n_k \geq n_{k'} \geq k' \geq N_0$, which implies that $|a_{n_k} - a| < \epsilon$. Thus, $\{a_{n_k}\}$ converges to a. The converse is immediate because a sequence is also a subsequence of itself. ■

The third theorem concerns **monotonic sequences**.

Theorem 2.0.3. *Let $\{a_n\}$ be a nondecreasing sequence of real numbers; i.e., for all n, $a_n \leq a_{n+1}$. Suppose $\{a_n\}$ is bounded from above; i.e., for some $M \in R$, $a_n \leq M$ for all n. Then the limit of a_n exists.*

Proof: Let a be the supremum of $\{a_n\}$. Let $\epsilon > 0$ be given. Then there exists an N_0 such that $a - \epsilon < a_{N_0} \leq a$. Because the sequence is nondecreasing, this implies that $a - \epsilon < a_n \leq a$, for all $n \geq N_0$. Hence, by definition, $a_n \to a$. ■

2.0.1 Limits Supremum and Infimum

Let $\{a_n\}$ be a sequence of real numbers and define the two subsequences

$$\begin{aligned} b_n &= \sup\{a_n, a_{n+1}, \ldots\}, \quad n = 1, 2, 3 \ldots &(2.0.2)\\ c_n &= \inf\{a_n, a_{n+1}, \ldots\}, \quad n = 1, 2, 3 \ldots. &(2.0.3) \end{aligned}$$

It is obvious that $\{b_n\}$ is a nonincreasing sequence. Hence, if $\{a_n\}$ is bounded from below, then the limit of b_n exists. In this case, we call the limit of $\{b_n\}$ the **limit supremum** (limsup) of the sequence $\{a_n\}$ and write it as

$$\overline{\lim_{n\to\infty}}\, a_n = \lim_{n\to\infty} b_n. \tag{2.0.4}$$

Note that if $\{a_n\}$ is not bounded from below, then $\overline{\lim}_{n\to\infty} a_n = -\infty$. Also, if $\{a_n\}$ is not bounded from above, we define $\overline{\lim}_{n\to\infty} a_n = \infty$. Hence, the $\overline{\lim}$ of any sequence always exists. Also, from the definition of the subsequence $\{b_n\}$, we have

$$a_n \leq b_n, \quad n = 1, 2, 3, \ldots. \tag{2.0.5}$$

On the other hand, $\{c_n\}$ is a nondecreasing sequence. Hence, if $\{a_n\}$ is bounded from above, then the limit of c_n exists. We call the limit of $\{c_n\}$ the **limit infimum** (liminf) of the sequence $\{a_n\}$ and write it as

$$\underline{\lim_{n\to\infty}}\, a_n = \lim_{n\to\infty} c_n. \tag{2.0.6}$$

Note that if $\{a_n\}$ is not bounded from above, then $\underline{\lim}_{n\to\infty} a_n = \infty$. Also, if $\{a_n\}$ is not bounded from below, $\underline{\lim}_{n\to\infty} a_n = -\infty$. As with $\overline{\lim}$, the $\underline{\lim}$ of any sequence always exists. Also, from the definition of the subsequences $\{c_n\}$ and $\{b_n\}$, we have

$$c_n \leq a_n \leq b_n, \quad n = 1, 2, 3, \ldots. \tag{2.0.7}$$

Also, because $c_n \leq b_n$ for all n, we have

$$\underline{\lim_{n\to\infty}}\, a_n \leq \overline{\lim_{n\to\infty}}\, a_n. \quad ■ \tag{2.0.8}$$

Example 2.0.1. Here are two examples. More are given in the exercises.

1. Suppose $a_n = -n$ for all $n = 1, 2, \ldots$. Then $b_n = \sup\{-n, -n-1, \ldots\} = -n \to -\infty$ and $c_n = \inf\{-n, -n-1, \ldots\} = -\infty \to -\infty$. So, $\underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n = -\infty$.

2. Suppose $\{a_n\}$ is defined by

$$a_n = \begin{cases} 1 + \frac{1}{n} & \text{if } n \text{ is even} \\ 2 + \frac{1}{n} & \text{if } n \text{ is odd.} \end{cases}$$

   Then $\{b_n\}$ is the sequence $\{3, 2+(1/3), 2+(1/3), 2+(1/5), 2+(1/5), \ldots\}$, which converges to 2, while $\{c_n\} \equiv 1$, which converges to 1. Thus, $\underline{\lim}_{n\to\infty} a_n = 1$ and $\overline{\lim}_{n\to\infty} a_n = 2$. ■

It is useful that the $\underline{\lim}_{n\to\infty}$ and $\overline{\lim}_{n\to\infty}$ of every sequence exists. The sandwich effects of expressions (2.0.7) and (2.0.8) lead to the following theorem.

Theorem 2.0.4. *Let $\{a_n\}$ be a sequence of real numbers. Then the limit of $\{a_n\}$ exists if and only if $\underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n$, in which case, $\lim_{n\to\infty} a_n = \underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n$.*

Proof: Suppose first that $\lim_{n\to\infty} a_n = a$. Because the sequences $\{c_n\}$ and $\{b_n\}$ are subsequences of $\{a_n\}$, Theorem 2.0.2 implies that they converge to a also. Conversely, if $\underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n$, then expression (2.0.7) and the Sandwich Theorem, 2.0.1, imply the result. ■

Based on this last theorem, we have two interesting applications which are frequently used in statistics and probability. Let $\{p_n\}$ be a sequence of probabilities and let $b_n = \sup\{p_n, p_{n+1}, \ldots\}$ and $c_n = \inf\{p_n, p_{n+1}, \ldots\}$. For the first application, suppose we can show that $\overline{\lim}_{n\to\infty} p_n = 0$. Then, because $0 \le p_n \le b_n$, the Sandwich Theorem implies that $\lim_{n\to\infty} p_n = 0$. For the second application, suppose we can show that $\underline{\lim}_{n\to\infty} p_n = 1$. Then, because $c_n \le p_n \le 1$, the Sandwich Theorem implies that $\lim_{n\to\infty} p_n = 1$.

We list some other properties in a theorem and ask the reader to provide the proofs in Exercise **??**:

Theorem 2.0.5. *Let $\{a_n\}$ and $\{d_n\}$ be sequences of real numbers. Then*

$$\overline{\lim_{n\to\infty}} (a_n + d_n) \le \overline{\lim_{n\to\infty}} a_n + \overline{\lim_{n\to\infty}} d_n \quad (2.0.9)$$

$$\underline{\lim_{n\to\infty}} a_n = -\overline{\lim_{n\to\infty}} (-a_n). \quad (2.0.10)$$

In showing convergence of a sequence, the following property will prove useful.

Definition 2.0.1. *We say a sequence $\{a_n\}$ is* **Cauchy***, if for any $\epsilon > 0$, there is an N such that $|a_n - a_m| < \epsilon$, if $n, m \ge N$.*

As the following theorem states, the Cauchy property is a necessary and sufficient condition for the convergence of a sequence.

Theorem 2.0.6. *Let $\{a_n\}$ be a sequence. Then $\{a_n\}$ converges if and only if $\{a_n\}$ is Cauchy.*

Proof: Suppose $\{a_n\}$ converges to a. Let $\epsilon > 0$ be given. Then there exists an N such that $|a_n - a| < \frac{\epsilon}{2}$, for $n \geq N$. Hence,

$$|a_n - a_m| = |(a_n - a) - (a_m - a)| \leq |a_n - a| + |a_m - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2} < \epsilon,$$

so the sequence is Cauchy. Next, suppose $\{a_n\}$ is Cauchy. Then these exists an N_1 such that $|a_n - a_{N_1}| \leq 1$, for $n \geq N_1$. Thus, it follows that $|a_n| = |a_n - a_{N_1} + a_{N_1}| \leq 1 + |a_{N_1}|$. Hence, the sequence is bounded, so $\overline{\lim}_{n\to\infty} a_n = a$ for some $a \in R$. Let a_{n_k} be the subsequence in expression (2.0.2) which converges to a. Let $\epsilon > 0$ be given. Then since $\{a_{n_k}\}$ converges to a and the sequence is Cauchy, there exists an N such that for n and n_k greater than N we have

$$|a_n - a| = |(a_n - a_{n_k}) + (a_{n_k} - a)| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

Thus, $a_n \to a$, as $n \to \infty$. ■

We shall make use of this theorem in the next subsection.

2.0.2 Infinite Series

Let $\{a_n\}$ be a sequence of real numbers. The corresponding **infinite series** is $\sum_{n=1}^{\infty} a_n$ The sequence of partial sums $\{S_n\}$ is given by

$$S_n = \sum_{i=1}^{n} a_i.$$

If $S_n \to S$, as $n \to \infty$, then we say the series $\sum_{i=1}^{n} a_i$ converges to S and write S as the infinite series, i.e.,

$$S = \sum_{n=1}^{\infty} a_n. \tag{2.0.11}$$

Assume that the series $\sum_{n=1}^{\infty} a_n$ converges to S. Then, using (2.0.11), for any $\epsilon > 0$, there exists an N such that

$$\left| \sum_{n=N+1}^{\infty} a_n \right| = \left| S - \sum_{j=1}^{n} a_j \right| < \epsilon; \tag{2.0.12}$$

i.e., the tail of a convergent series becomes arbitrarily small as n gets large. Further, the sequence of partial sums, $S_n = \sum_{j=1}^{n} a_j$, is Cauchy and for partial sums the Cauchy condition, Definition 2.0.1, simplifies to

$$\left| \sum_{j=n}^{n+m} a_j \right|. \tag{2.0.13}$$

Since the series converges, this can be made arbitrarily small for n sufficiently large. In this sense, for a convergent series, both the tail and middle of the series can be made arbitrarily small for n sufficiently large.

For the geometric series presented next, the partial sums are easily obtained in close form. This is rarely the case.

Example 2.0.2 (Geometric Series)**.** For $a \in R$, consider the sequence of partial sums given by

$$S_n = \sum_{i=0}^{n} a^i.$$

Note that

$$(1-a)S_n = 1 - a^{n+1};$$

hence,

$$S_n = \frac{1-a^{n+1}}{1-a}. \tag{2.0.14}$$

If $|a| < 1$ then $a^{n+1} \to 0$, as $n \to \infty$. Thus for $|a| < 1$, the geometric series converges to $(1-a)^{-1}$, which we write as

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}, \quad \text{for } |a| < 1. \blacksquare \tag{2.0.15}$$

We say that a series $\sum a_n$ **converges absolutely** if $\sum |a_n|$ converges. As the next theorem states, absolute convergence implies convergence.

Theorem 2.0.7. *If $\sum_{n=1}^{\infty} |a_n|$ converges then $\sum_{n=1}^{\infty} a_n$ converges.*

Proof: We show that the sequence of partial sums $\{S_n\}$ of the series $\sum_{n=1}^{\infty} a_n$ satisfies the Cauchy condition given in Definition 2.0.1. For partial sums the Cauchy condition is

$$\left| \sum_{j=n}^{n+m} a_j \right| \leq \sum_{j=n}^{n+m} |a_j|.$$

The sum on the right-side is the Cauchy condition for the the series $\sum_{n=1}^{\infty} |a_n|$. Since this series converges, by Theorem 2.0.6 the right-side can be made arbitrarily small for n and m sufficiently large. Hence the series $\sum_{n=1}^{\infty} a_n$ is Cauchy. Therefore by Theorem 2.0.6, the series $\sum_{n=1}^{\infty} a_n$ converges. $\blacksquare$

Using this theorem, we can establish convergence of a series by showing the convergence of the corresponding series of absolute values. For example, suppose we can establish boundedness, that is, $\sum_{n=1}^{\infty} |a_n| \leq B$, for some $B \in R$. Then because the sequence of partial sums of $\sum_{n=1}^{\infty} |a_n|$ are nondecreasing, by Theorem 2.0.3 it follows that $\sum_{n=1}^{\infty} |a_n|$ converges. As another example, suppose $0 \leq b_n \leq c_n$ and $\sum_{n=1}^{\infty} c_n$ converges. Then using a similar argument, it follows that $\sum_{n=1}^{\infty} b_n$ converges. There are many such tests for convergence which are discussed in most calculus books. We record one other, the ratio test, which we have made use of in the text; see page 162 of Buck (1965) for a proof.

Theorem 2.0.8. *Suppose* $0 \leq a_n$ *for all* n. *Suppose for some* r *that the following limit exists* $\lim_{n\to\infty}(a_{n+1}/a_n) = r$. *Then* $\sum_{n=1}^{\infty} a_n$ *converges or diverges depending on whether* $r < 1$ *or* $r > 1$, *respectively.*

As a final result we have the following theorem on the rearrangement of the terms in an absolutely convergent series. The proof is taken from Buck (1965), page 169.

Theorem 2.0.9. *Suppose* $\sum_{n=1}^{\infty} a_n$ *converges to* S *and* $\sum_{n=1}^{\infty} |a_n|$ *converges. Let* $\sum_{n=1}^{\infty} a_{r_n}$, $r_1, r_2, \ldots$, *be any rearrangement of the terms in the series. Then* $\sum_{n=1}^{\infty} a_{r_n}$, *converges to* S, *also.*

Proof: Let $\epsilon > 0$ be given. Then, since $\sum_{n=1}^{\infty} |a_n|$ converges, choose N sufficiently large so that $\sum_{n=N+1}^{\infty} |a_n| < \epsilon/2$. Notice that the integers $1, \ldots, N$ appear once and only once in $r_1, r_2, \ldots$. Choose n_0 such that $\{1, \ldots, N\} \subset \{r_1, r_2, \ldots, r_{n_0}\}$. Then

$$\left| S - \sum_{n=1}^{n_0} a_{r_n} \right| \leq \left| S - \sum_{n=1}^{N} a_n \right| + \left| \sum_{n=1}^{N} a_n - \sum_{n=1}^{n_0} a_{r_n} \right|.$$

The first term on the right-side is less than $\epsilon/2$. For the second term, after the subtraction, the only terms left are for $r_n > N$, so the second term is less than $\epsilon/2$, also. ∎

Multiple Series

Besides univariate series, multiple series are used in this text. It suffices to briefly consider double series. Consider the sequence of partial sums $\{S_{m,n}\}$

$$S_{m,n} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij},$$

where $a_{ij} \in R$. We say that $S_{m,n}$ converges to S if

$$\lim_{m\to\infty} \lim_{n\to\infty} S_{m,n} = S.$$

Since the sums involved with $S_{m,n}$ are finite, we can always iterate them as

$$S_{m,n} = \sum_{i=1}^{m} \left[\sum_{j=1}^{n} a_{ij} \right] = \sum_{j=1}^{n} \left[\sum_{i=1}^{m} a_{ij} \right].$$

This is useful because we can compute a double sum as iterated single sums. Using the same theorems (Fubini's and Tonelli's Theorems) in analysis, as cited in Section 4.0.3, the double series converges provided that the series converges absolutely. If so, either order of iteration sums to the same limit. Furthermore, we can establish absolute convergence by showing that either iterated infinite double series converges absolutely. Again, either order of iteration converges to the same value. The following example serves as an illustration.

Example 2.0.3. Consider the double infinite series with terms

$$a_{m.n} = \binom{m}{n}\left(\frac{1}{3}\right)^n \left(\frac{3}{2}\right)^m \frac{1}{m!}, \quad n = 0, \ldots, m, m = 0, 1, \ldots.$$

The region of summation is shown in Figure 2.0.1. We cannot overemphasize the importance of sketching a graph of the summation region when computing double infinite series. For our solution, we sum with respect to m first; i,e., as in the figure, n is fixed, but arbitrary, and we sum m from n to ∞. We then sum out n from 0 to ∞. For the solution, we do make use of the result (3.0.15).

$$\sum_{n=0}^{\infty}\left(\frac{1}{3}\right)^n \left[\sum_{m=n}^{\infty}\frac{m!}{(m-n)!n!}\left(\frac{3}{2}\right)^m \frac{1}{m!}\right]$$
$$= \sum_{n=0}^{\infty}\left(\frac{1}{3}\right)^n \frac{1}{n!}\left[\sum_{m=n}^{\infty}\frac{1}{(m-n)!}\left(\frac{3}{2}\right)^{m-n}\left(\frac{3}{2}\right)^n\right] = e^{3/2}\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{2}\right)^n = e^2.$$

We leave it as an exercise to show that the other way of iteration leads to the same solution. ■

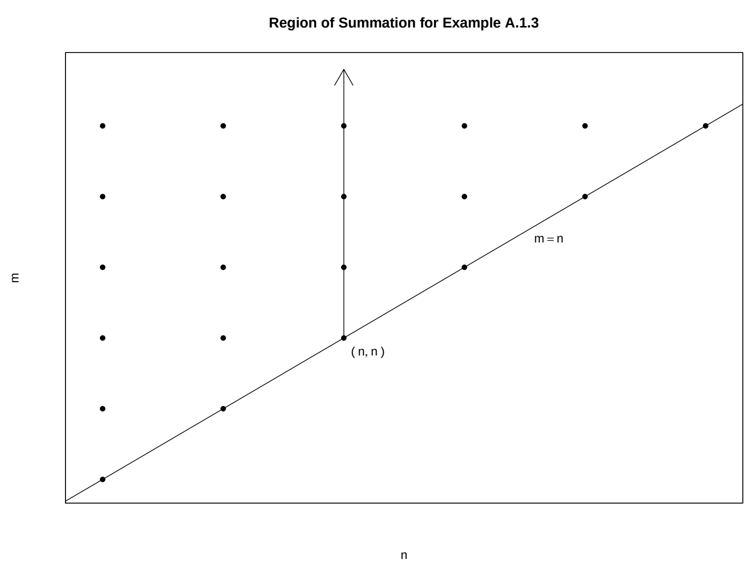

Figure 2.0.1: Region of summation for Example 2.0.3. It depicts summation with respect to m at a fixed but arbitrary n.

Chapter 3

Derivatives

We assume that the reader is familiar with differentiable calculus. The following discussion serves only as a brief review of it.

Let $f(x)$ be a real valued function defined on an interval of real numbers (a, b) which can be $(-\infty, \infty)$. Recall that f is **differentiable** at x with **derivative** $f'(x)$ if the following limit exists:

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = f'(x). \tag{3.0.1}$$

We often express $f'(x)$ as $f'(x) = \frac{d}{dx} f(x)$. It follows immediately from this definition that if $f(x)$ is differentiable at x then it is continuous at x.

Next we list some useful formulas for derivatives. Their derivations can be found in any calculus book. Assume that functions $f(x)$ and $g(x)$ are differentiable at x.

$$\frac{d}{dx} x = 1; \frac{d}{dx} x^r = r x^{r-1} \text{ (for } r < 1 \text{ not differentiable at 0)} \tag{3.0.2}$$

$$\frac{d}{dx} [a f(x) + b g(x)] = a f'(x) + b g'(x) \tag{3.0.3}$$

$$\frac{d}{dx} [f(x) g(x)] = f'(x) g(x) + f(x) g'(x) \tag{3.0.4}$$

$$\frac{d}{dx} [f(x)/g(x)] = [f'(x) g(x) - f(x) g'(x)]/g(x)^2 \tag{3.0.5}$$

$$\frac{d}{dx} [f(x)^a)] = a f'(x)^{a-1} \tag{3.0.6}$$

$$\frac{d}{dx} f[g(x)] = f'[g(x)] g'(x) \tag{3.0.7}$$

$$\frac{d}{dx} e^{f(x)} = f'(x) e^{f(x)} \tag{3.0.8}$$

Assume in addition that $f(x)$ is strictly increasing or decreasing on (a, b). Then the inverse function $f^{-1}(y)$ exists on (a, b) and its derivative is given by

$$\frac{d}{dy} f^{-1}(y) = \frac{1}{f'[f^{-1}(y)]}. \tag{3.0.9}$$

Recall the Mean Value Theorem which is given by:

Theorem 3.0.1 (Mean Value Theorem)**.** *Let $f(x)$ be continuous on the interval $[a, b]$ and differentiable on (a, b). Then there is a point ξ such that $a < \xi < b$ and*

$$f(b) = f(a) + (b - a)f'(\xi). \tag{3.0.10}$$

Note that if $f(x)$ is differentiable in an open neighborhood of x_0 and $f'(x)$ is continuous at x_0 then for some ξ between x and x_0 we can write equation (3.0.10) as

$$\begin{aligned} f(x) &= f(x_0) + (x - x_0)f'(\xi) \\ &= f(x_0) + (x - x_0)f'(x_0) + (x - x_0)[f'(\xi) - f'(x_0)] \\ &= f(x_0) + (x - x_0)f'(x_0) + o(x - x_0), \end{aligned} \tag{3.0.11}$$

where the **little** o notation is defined by

$$a = o(b) \text{ if and only if } \tfrac{a}{b} \to 0 \text{ when } b \to 0. \tag{3.0.12}$$

Suppose $f'(x)$ is differentiable; i.e., $(d/dx)f'(x)$ exists. Then we call this the **second derivative** of $f(x)$ and write it as $(d^2/dx^2)f(x) = f''(x) = f^{(2)}(x)$. Continuing in this way, if it exists, we denote the nth derivative of $f(x)$ by $f^{(n)}(x)$ for $n \geq 1$.

The mean value theorem provides a first-order approximation of $f(x)$ in a neighborhood of x_0. Higher order expansions are called **Taylor series**.

Theorem 3.0.2 (Taylor Series)**.** *Suppose that $f(x)$ has at least $n + 1$ derivatives in an open interval I of x_0 and $f^{(n+1)}$ is continuous in I. Then for all $x \in I$ there is some ξ between x and x_0 such that*

$$f(x) = f(x_0) + \sum_{j=1}^{n} \frac{f^{(j)}(x_0)}{j!}(x - x_0)^j + \frac{f^{(n+1)}(\xi)}{(n + 1)!}(x - x_0)^{n+1}. \tag{3.0.13}$$

For $x_0 = 0$, a Taylor series is called a **Maclaurin series**.

The last term in the Taylor series, (3.0.13), is called the remainder of the series. Functions $f(x)$ which are infinitely differentiable in an open interval I of x_0 and for which the remainder goes to 0 as $n \to \infty$ for all $x \in I$ are said to be **analytic** in I. For these functions, $f(x)$ is the infinite Taylor series; i.e., for all $x \in I$,

$$f(x) = f(x_0) + \sum_{n=1}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \tag{3.0.14}$$

Examples of analytic functions and their intervals of convergence are:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad -\infty < x < \infty, \tag{3.0.15}$$

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad -\infty < x < \infty, \tag{3.0.16}$$

$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \quad -\infty < x < \infty, \tag{3.0.17}$$

$$(1-x)^{-r} = \sum_{n=0}^{\infty} \binom{n+r-1}{r-1} x^n,, \quad |x| < 1, \ \ r > 0. \tag{3.0.18}$$

Partial Derivatives

For functions of several variables, we briefly discuss the concept of **partial differentiation**. These derivatives are taken with respect to a variable while the other variables are held fixed. For the partial with respect to x, instead of the d/dx notation of the univariate case, we use the notation $\partial/\partial x$. For example let $f(x,y) = 2x^2 \exp\{-x^2 - y^2\}$. Then

$$\frac{\partial}{\partial x} f(x,y) = 4x \exp\{-x^2 - y^2\} - 4x^3 \exp\{-x^2 - y^2\}$$
$$\frac{\partial}{\partial y} f(x,y) = -4x^2 y \exp\{-x^2 - y^2\}$$

Partial derivatives of the partials can be performed, also. These are called **mixed** partials. Continuing with the last example, we have:

$$\frac{\partial^2}{\partial x \partial y} f(x,y) = \frac{\partial}{\partial y}\{\frac{\partial}{\partial x} f(x,y)\} = -8xy \exp\{-x^2 - y^2\} + 8x^3 y \exp\{-x^2 - y^2\}$$
$$\frac{\partial^2}{\partial y^2} f(x,y) = -4x^2 \exp\{-x^2 - y^2\} + 8x^2 y^2 \exp\{-x^2 - y^2\}$$

For mixed partials, the order of differentiation does not matter as long as the partials exist and are continuous. For instance, for such functions, $(\partial^2/\partial x \partial y)\, f(x,y) = (\partial^2/\partial y \partial x)\, f(x,y)$.

Chapter 4

Integration

One way of motivating the process of **integration** of functions is to consider **Riemann sums**. Let $F(x)$ be a function defined on the interval $[a, b]$. Partition $[a, b]$ into the n subintervals

$$[a + (i-1)\Delta x, a + i\Delta x], \quad i = 1, \ldots, n; \Delta x = \frac{b-a}{n}.$$

Let ξ_i be the midpoint of $[a+(i-1)\Delta x, a+i\Delta x]$. Then the corresponding **Riemann sum** is

$$R_n = R_n(f, a, b, \Delta x) = \sum_{i=1}^{n} f(\xi_i)\Delta x. \tag{4.0.1}$$

Note that this is an average. In calculus, it is proved that if $f(x)$ is continuous on $[a, b]$ then R_n converges to a limit as $n \to \infty$ and we denote it as

$$\lim_{n\to\infty} R_n = \int_a^b f(x)\, dx. \tag{4.0.2}$$

We call this limit[1] the **definite integral** of $f(x)$ over $[a, b]$. Also, if $f(x) \geq 0$ on $[a, b]$ then R_n is an approximation to the area under the curve $y = f(x)$ between a and b. Hence, the area under the curve $y = f(x)$ between a and b is defined as the integral of f from a to b.

For many of the integrals in this text, the bounds in the interval $[a, b]$ may be infinite. For example, the integral of interest is $\int_a^\infty f(x)\, dx$. Provided the following limit exists, the value of the integral is given by the limit

$$\lim_{h\to\infty} \int_a^h f(x)\, dx = \int_a^\infty f(x)\, dx. \tag{4.0.3}$$

In these cases, we have from the theory of calculus that

$$\int_a^b |f(x)|\, dx < \infty \Rightarrow \int_a^b f(x)\, dx \text{ exists,} \tag{4.0.4}$$

[1]The limit is the same for all subdivisions provided $\Delta x \to 0$.

that is, absolute convergence implies convergence.

As the next theorem states integration is the process of anti-differentiation.

Theorem 4.0.1 (Fundamental Theorem of Calculus)**.** *Let the function $F(x)$ be differentiable on the interval $[a, b]$ with derivative $f(x)$. If $f(x)$ is continuous on $[a, b]$ then*

$$F(b) - F(a) = \int_a^b f(x)\, dx. \quad \blacksquare \tag{4.0.5}$$

For example, since x^n is the derivative of $x^{n+1}/(n+1)$, we have

$$\int_a^b x^n\, dx = \left.\frac{x^{n+1}}{n+1}\right|_a^b = \frac{1}{n+1}[b^{n+1} - a^{n+1}].$$

The following is a list of properties of integration, (proofs of which can be found in any calculus book). Assume that all the integrals exist.

$$\int_a^b [cf(x) + dg(x)]\, dx = c\int_a^b f(x)\, dx + d\int_a^b g(x)\, dx \tag{4.0.6}$$

$$a < c < b \Rightarrow \int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx \tag{4.0.7}$$

$$\int_a^b f(x)\, dx = -\int_b^a f(x)\, dx \tag{4.0.8}$$

$$f(x) \ge g(x) \text{ on } [a, b] \Rightarrow \int_a^b f(x)\, dx \ge \int_a^b g(x)\, dx \tag{4.0.9}$$

We often use the technique of **change-of-variable** to facilitate the computation of an integral. Suppose that $g(x)$ is differentiable on $[a, b]$ and that $F(x)$ is differentiable on the range of $g(x)$. By the chain rule for differentiation, we have

$$D_x\, F[g(x)] = F'[g(x)]g'(x).$$

So,

$$\int_a^b F'[g(x)]g'(x)\, dx = F[g(x)]\Big|_a^b = F[g(b)] - F[g(a)] = \int_{g(a)}^{g(b)} F'(u)\, du. \tag{4.0.10}$$

This formulation is cumbersome, so we generally write the transformation as $u = g(x)$ and use the notation $du = g'(x)\, dx$. For example, suppose we want to compute $\int_2^3 x \exp\{-x^2\}\, dx$. Let $u = x^2$ then $du = 2x\, dx$ and we obtain

$$\int_2^3 x\exp\{-x^2\}\, dx = \frac{1}{2}\int_2^3 2x\exp\{-x^2\}\, dx = \frac{1}{2}\int_4^9 e^{-u}\, du = \frac{1}{2}(e^{-4} - e^{-9}).$$

For monotne transformations, the change-of-variable techniques produces a useful formula. Consider $\int_a^b f(x)\, dx$. Let $y = g(x)$ denote a continuous and differentiable transformation where $g(x)$ is a strictly monotone function on $[a, b]$. Suppose that $g(x)$ is strictly decreasing. Then the transformation maps $[a,b]$ into $[g(b),g(a)]$. Note that $x = g^{-1}(y)$ and by (3.0.9),

$$\frac{dx}{dy} = \frac{1}{g'[g^{-1}(y)]}.$$

Because $g(x)$ is decreasing, dx/dy is negative and, hence, $|dx/dy| = -dx/dy$. So, by the change-of-variable technique and expression (4.0.8) we have

$$\int_a^b f(x)\,dx = \int_{g(a)}^{g(b)} f[g^{-1}(y)]\frac{dx}{dy}\,dy = \int_{g(b)}^{g(a)} f[g^{-1}(y)]\left|\frac{dx}{dy}\right|\,dy. \tag{4.0.11}$$

This same formula holds if the transformation is strictly increasing.

Another technique for computation of integrals is **integration by parts**. Let $u(x)$ and $v(x)$ be differentiable functions on $[a,b]$ whose second derivatives exist, also. By the product rule for differention (3.0.4),

$$\frac{d}{dx}[u(x)v(x)] = u'(x)v(x) + u(x)v'(x).$$

Integrating both sides and solving for the second term on the right, we have

$$\begin{aligned} \int_a^b u(x)v'(x)\,dx &= \int_a^b \frac{d}{dx}[u(x)v(x)]\,dx - \int_a^b u'(x)v(x)\,dx \\ &= u(x)v(x)\Big|_a^b - \int_a^b u'(x)v(x)\,dx. \end{aligned} \tag{4.0.12}$$

For instance, suppose we want to compute $\int_a^b x\exp\{-x\}\,dx$. Take $u(x) = x$ and $v'(x) = \exp\{-x\}$ then $u'(x) = 1$ and $v(x) = -\exp\{-x\}$ and we obtain

$$\int_a^b xe^{-x}\,dx = -xe^{-x}\,|_a^b + \int_a^b e^{-x}\,dx = e^{-a}(a+1) - e^{-b}(b+1).$$

4.0.3 Multiple Integration

The concept of integration can be extended to n-dimensions. We shall briefly review its generalizationto two dimensions. Let $f(x,y)$ be a continuous function of two variables which is defined on a bounded rectangle A. In this setting, the concept of a Reimann sum easily generaizes and we only briefly discuss it. Partition A into mn subrectangles and form the sum over all the subrectangles (i,j),

$$R_{m,n} = R_{m,n}(f, A, \Delta x, \Delta y) = \sum_{i=1}^{m}\sum_{j=1}^{n} f(x_i, y_j)\Delta x_i \Delta y_j, \tag{4.0.13}$$

where (x_i, y_j) is a point in the interior of the (i,j)th rectangle. If the largest diagonal of the subrectangles converges to 0 as m and n converge to ∞, then $R_{m,n}$ has a limit which we call the **double integral** of f over A and write it as

$$\iint\limits_A f(x,y)\,dxdy.$$

The limit is the same for all Riemann sums provided the largest diagonal of the subrectangles converges to 0. If $f(x,y) \geq 0$ on A then a Riemann sum approximates the volume under the surface $z = f(x,y)$ over A and we define this volume to be the $\iint_A f(x,y)\,dxdy$. If A is not a rectangle then we can embed A in a rectangle, define $f(x,y)$ to be 0 when $(x,y) \notin A$, and proceed as above.

Example 4.0.4. Suppose $A = \{(x,y) : x^2 + y^2 \leq 1\}$. If $f(x,y) = 10$ on A then

$$\iint\limits_A f(x,y)\,dxdy = \iint\limits_A 10\,dxdy = 10\pi.$$

That is, the volume of a regular cylinder of radius 1 and height 10 is 10π. ■

Many of the properties for univariate integration carry over to multiple integration. For example properties (4.0.6) and (4.0.9) carry over directly, while property (4.0.7) is replaced by

$$\text{If } A \cap B = \phi \text{ then } \iint\limits_{A\cup B} f(x,y)\,dxdy = \iint\limits_A f(x,y)\,dxdy + \iint\limits_B f(x,y)\,dxdy. \quad (4.0.14)$$

Unlike the univariate case, there is not a Fundamental Theorem of Calculus, Theorm 4.0.1, for double (multiple) integration. Reconsider the Reimann sum for double integration, (4.0.13). Note that it can be written iteratively, in two ways, as two univariate sums, i.e.,

$$R_{m,n} = \sum_{i=1}^{m}\left[\sum_{j=1}^{n} f(x_i,y_j)\Delta y_j\right]\Delta x_i = \sum_{j=1}^{n}\left[\sum_{i=1}^{m} f(x_i,y_j)\Delta x_i\right]\Delta y_j. \quad (4.0.15)$$

Theory then shows that double integration can be computed by iterated integrals. Furthermore under conditions stated below, either order (integrate with respect to x first or y first) results in the same value for the double integral. For example, suppose $f(x,y)$ is continuous on a region A which can be defined as $A = \{(x,y) : a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}$ or $A = \{(x,y) : c \leq y \leq d, \omega_1(y) \leq x \leq \omega_2(y)\}$ for continuous functions $\varphi_i(x)$, $\omega_i(y)$, $i = 1,2$. Then

$$\iint\limits_A f(x,y)\,dxdy = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\right] dx = \int_c^d \left[\int_{\omega_1(y)}^{\omega_2(y)} f(x,y)\,dx\right] dy. \quad (4.0.16)$$

Univariate integration techniques can then be used to compute these iterated integrals.

Reconsider Example 4.0.4. The region of integration can be written as $A = \{(x,y) : -1 \leq x \leq 1, -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}\}$ which is shown in Figure 4.0.1. We fix x and integrate with respect to y first. As the arrow in Figure 4.0.1 illustrates, we integrate y from $-\sqrt{1-x^2}$ to $\sqrt{1-x^2}$ and then integrate out x from -1 to 1.. Using symmetry of the integrand; the change-in-variable $x = \sin\theta$, $dx = \cos\theta\,dx$;

and the trigonmetry identity $\cos^2\theta = (1+\cos 2\theta)/2$, we obtain

$$\begin{aligned}
\iint_A 10\,dxdy &= \int_{-1}^{1}\left[\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 10\,dy\right]dx = 10\int_{-1}^{1} 2\sqrt{1-x^2}\,dx \\
&= 40\int_0^1 \sqrt{1-x^2}\,dx = 40\int_0^{\pi/2}\cos^2\theta\,d\theta \\
&= 40\int_0^{\pi/2}\frac{1+\cos 2\theta}{2}\,d\theta = 20\left[\frac{\pi}{2}+\frac{1}{2}\sin 2\theta\Big|_0^{\pi/2}\right] = 10\pi.
\end{aligned}$$

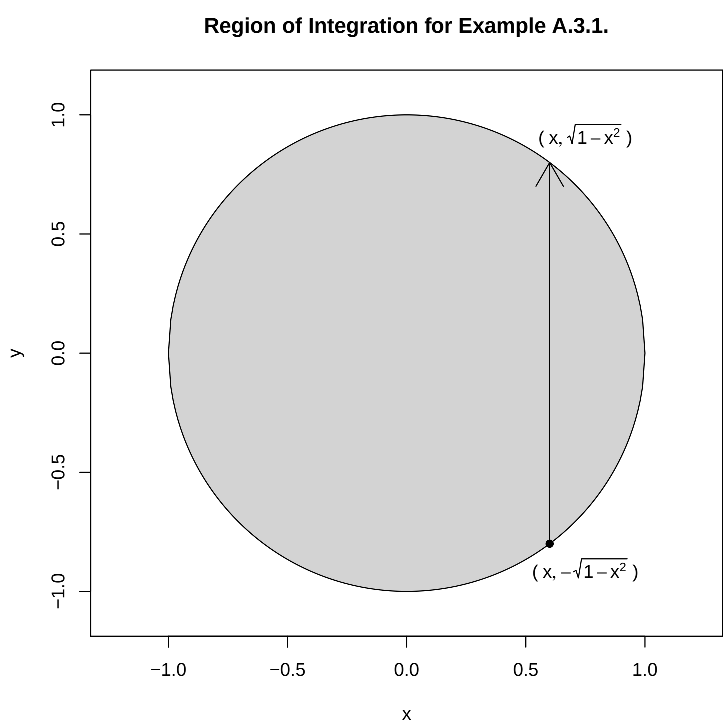

Figure 4.0.1: Region of integration for Example 4.0.4. It depicts integration with respect to y at a fixed but arbitrary x.

We now state two real analysis conditions that permit the the computation of multiple integrals by way of iterated univariate integrals. This is a brief discussion and the interested reader can find detailed discussions in most books on real analysis; eg., Royden and Fitzpatrick (2010). First a measurable function $f(x, y)$ is said to be integrable on a region A, if $\iint_A |f(x, y)|\,dxdy < \infty$. Fubini's Theorem states: if $f(x, y)$ is integrable then expression (4.0.16) is true. The function $f(x, y)$, however, must be integrable, which can be determined by Tonelli's Theorem. Tonelli's

Theorem states that if $f(x,y)$ is a non-negative measurable function on A then the double integral of $f(x,y)$ can be computed iteratively. Hence, Tonelli's Theorem is used to establish that the integral of $|f(x,y)|$ is finite, then Fubini's Theorem is used to compute the double integral of $f(x,y)$ via iterated integrals. Note that all functions used in this text are measurable; see Royden and Fitzpatrick (2010) for discussions on measurable functions.

Expression (4.0.16) gives us a choice on which variable is integrated first. For Example 4.0.4, because of the symmetry of x and y in the problem, it does not matter. In certain cases, however, a little thought before selecting the order can lead to an easier solution. Our next example is one such case. It helps considerably in performing double integrations to always **sketch** the region of integration first.

Example 4.0.5. The region $A = \{(x,y) : 0 < x < 1, x+1 < y < -x+4\}$ is sketched in Figure 4.0.2. Suppose we want to compute $\iint_A x^2 y\, dxdy$. From the figure, if we integrate with respect to y then for any fixed x, $0 < x < 1$, we need to integrate from $y = x+1$ to $y = -x+4$ and then integrate out x from 0 to 1. On the other hand, suppose we decide to intergate with respect to x first. Then, as shown in the figure, the integration over the three regions B_1, B_2, and B_3 must be computed. Since the regions are disjoint, $\iint_A x^2 y\, dxdy$ is the sum of these three integrals. We choose to integrate with respect to y first.

$$\begin{aligned}\iint\limits_A x^2 y\, dxdy &= \int_0^1 x^2 \left[\int_{x+1}^{-x+4} y\, dy\right] dx = \frac{1}{2}\int_0^1 x^2 \left[(-x+4)^2 - (x+1)^2\right] dx \\ &= \frac{1}{2}\int_0^1 (-10x^3 + 15x^2)\, dx = \frac{5}{4}.\end{aligned}$$

If we intergate with respect to x first then by the figure the double integration is:

$$\iint\limits_A x^2 y dxdy = \int_1^2 y \left[\int_0^{y-1} x^2 dx\right] dy + \int_2^3 y \left[\int_0^1 x^2\, dx\right] dy + \int_3^4 y \left[\int_0^{4-y} x^2 dx\right] dy. \tag{4.0.17}$$

The reader is asked to show that ?????????EXER??? these integrations result in the above answer 5/4. ■

There is a multiple variable analog to the change-in-variable technique for univariate integration for a one-to-one transformation; see discussion around expressions (4.0.10) and (4.0.11). Suppose we are intergrating the function $f(x,y)$ over the region D. Let $(u,v) = T(x,y)$ be a one-to-one transformation from D to $T(D)$. Define the Jacobian of the transformation to be the determinant

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}. \tag{4.0.18}$$

Then the multivariate analog of expression (4.0.11) is

$$\iint\limits_D f(x,y)\, dxdy = \iint\limits_{T(D)} f\left[T^{-1}(u,v)\right] \left|\frac{\partial(x,y)}{\partial(u,v)}\right| dudv. \tag{4.0.19}$$

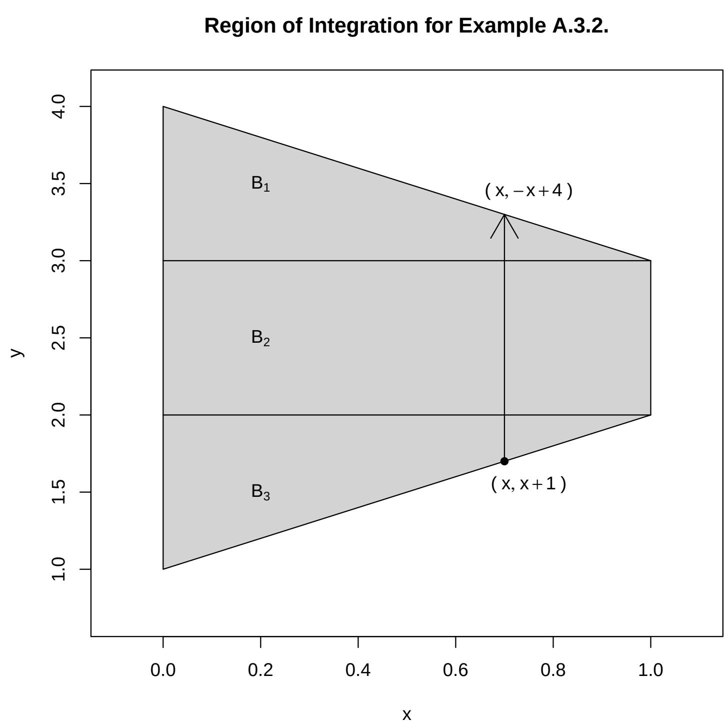

Figure 4.0.2: Region of integration for Example 4.0.5. It depicts integration with respect to y at a fixed but arbitrary x. Note that if integration with respect to x is performed first then separate integrations over the regions B_1, B_2, and B_3 must be performed.

This is proved in most calculus books; see , for example, page 305 of Buck (1965). We illustrate this result by the following example.

Example 4.0.6. Consider integrating the function $f(x, y) = \exp\{-(x^2 + y^2)/2\}$ over the entire plane R^2. Consider the one-to-one transformation $x = u \cos v$ and $y = u \sin v$ for $u > 0$ and $0 < v < 2\pi$. Then the Jacobian is

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \cos v & -u \sin v \\ \sin v & u \cos v \end{vmatrix} = u \cos^2 v + u \sin^2 v = u.$$

Hence, by (4.0.19),

$$\begin{aligned} \iint\limits_{R^2} \exp\{-(x^2 + y^2)/2\}\, dxdy &= \int_0^{2\pi} \int_0^{\infty} \exp\{-(u^2 \cos^2 v + u^2 \sin^2 v)/2\} |u|\, dudv \\ &= \int_0^{2\pi} \left[\int_0^{\infty} u e^{-u^2/2}\, du \right] dv = \int_0^{2\pi} 1\, dv = 2\pi, \end{aligned}$$

where before the last equality the univariate transformation $z = u^2/2$ can be used to show that the inner integral is 1. ■