

# Part I

---

## Fundamentals and Probability on Events

In this first part of the book we focus on some basic tools that we will need throughout the book.

We start, in Chapter 1, with a review of some mathematical basics: series, limits, integrals, counting, and asymptotic notation. Rather than attempting an exhaustive coverage, we instead focus on a select “toolbox” of techniques and tricks that will come up over and over again in the exercises throughout the book. Thus, while none of this chapter deals with probability, it is worth taking the time to master its contents.

In Chapter 2, we cover the fundamentals of probability. Here we define probability based on an experiment and events. We discuss the axioms of probability, conditioning, independence, the Law of Total Probability, and Bayes’ Law.

# 1 Before We Start ... Some Mathematical Basics

---

This book assumes some mathematical skills. The reader should be comfortable with high school algebra, including logarithms. Basic calculus (integration, differentiation, limits, and series evaluation) is also assumed, including nested (3D) integrals and sums. We also assume that the reader is comfortable with sets and with simple combinatorics and counting (as covered in a discrete math class). Finally, we assume versatility with “big-O” and “little-o” notation. To help the reader, in this chapter we review a few basic concepts that come up repeatedly throughout the book. Taking the time to understand these *now* will make it much easier to work through the book.

## 1.1 Review of Simple Series

There are several series that come up repeatedly in the book, starting in Chapter 3.

**Question:** Try evaluating the following in closed form. (Don’t peek at the answers until you’ve tried these yourself.) We provide the full derivations below.

(a)  $S = 1 + x + x^2 + x^3 + \cdots + x^n$ .

(b)  $S = 1 + x + x^2 + x^3 + \cdots$ , where  $|x| < 1$ .

(c)  $S = 1 + 2x + 3x^2 + 4x^3 + \cdots + nx^{n-1}$ .

(d)  $S = 1 + 2x + 3x^2 + 4x^3 + \cdots$ , where  $|x| < 1$ .

**Example 1.1** Evaluate:  $S = 1 + x + x^2 + x^3 + \cdots + x^n$ .

**Solution:** The trick here is to multiply both sides by the quantity  $(1 - x)$ :

$$\begin{aligned}(1 - x)S &= S - xS \\ &= 1 + x + x^2 + x^3 + \cdots + x^n \\ &\quad - x - x^2 - x^3 - \cdots - x^{n+1} \\ &= 1 - x^{n+1}.\end{aligned}$$

Hence,

$$S = \frac{1 - x^{n+1}}{1 - x}. \quad (1.1)$$

Note that (1.1) assumes that  $x \neq 1$ . If  $x = 1$ , then the answer is clearly  $S = n + 1$ .

**Example 1.2** Evaluate:  $S = 1 + x + x^2 + x^3 + \dots$  where  $|x| < 1$ .

**Solution:** This is the same as series (a) except that we need to take the limit as  $n \rightarrow \infty$ :

$$S = \lim_{n \rightarrow \infty} 1 + x + x^2 + \dots + x^n = \lim_{n \rightarrow \infty} \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x}. \quad (1.2)$$

**Question:** Why did we need  $|x| < 1$ ? What would happen if  $|x| \geq 1$ ?

**Answer:** If  $|x| \geq 1$ , then the infinite sum diverges.

**Example 1.3** Evaluate:  $S = 1 + 2x + 3x^2 + 4x^3 + \dots + nx^{n-1}$ .

**Approach 1:** One approach is to again use the  $(1 - x)$  trick:

$$\begin{aligned} (1 - x)S &= 1 + 2x + 3x^2 + 4x^3 + \dots + nx^{n-1} \\ &\quad -x - 2x^2 - 3x^3 - 4x^4 - \dots - nx^n \\ &= 1 + x + x^2 + x^3 + \dots + x^{n-1} - nx^n \\ &= \frac{1 - x^n}{1 - x} - nx^n \\ &= \frac{1 - (n+1)x^n + nx^{n+1}}{1 - x}. \end{aligned}$$

Hence,

$$S = \frac{1 - (n+1)x^n + nx^{n+1}}{(1 - x)^2}. \quad (1.3)$$

**Approach 2:** An easier approach is to view the sum as the derivative of a known sum:

$$\begin{aligned} S &= \frac{d}{dx} (1 + x + x^2 + x^3 + \dots + x^n) \\ &= \frac{d}{dx} \left( \frac{1 - x^{n+1}}{1 - x} \right) \\ &= \frac{(1 - x) \cdot (-(n+1)x^n) + (1 - x^{n+1})}{(1 - x)^2} \\ &= \frac{1 - (n+1)x^n + nx^{n+1}}{(1 - x)^2}. \end{aligned}$$

The above assumes that  $x \neq 1$ . If  $x = 1$ , then the answer is  $S = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ .

**Example 1.4** Evaluate:  $S = 1 + 2x + 3x^2 + 4x^3 + \cdots$  where  $|x| < 1$ .

**Solution:** We again view  $S$  as a derivative of a sum:

$$\begin{aligned} S &= \frac{d}{dx} (1 + x + x^2 + x^3 + \cdots) \\ &= \frac{d}{dx} \left( \frac{1}{1-x} \right) \\ &= \frac{1}{(1-x)^2}. \end{aligned} \tag{1.4}$$

## 1.2 Review of Double Integrals and Sums

Integrals, nested integrals, and nested sums come up throughout the book, starting in Chapter 7. When evaluating these, it is important to pay attention to the area over which you're integrating and also to remember tricks like integration by parts.

**Question:** Try deriving the following three expressions (again, no peeking at the answers).

- (a)  $\int_0^\infty ye^{-y} dy$ .
- (b)  $\int_0^\infty \int_0^y e^{-y} dx dy$ . Do this both with and without changing the order of integration.
- (c)  $\int_1^e \int_0^{\ln x} 1 dy dx$ . Do this both with and without changing the order of integration.

Below we provide the derivations.

**Example 1.5** Derive:  $\int_0^\infty ye^{-y} dy$ .

**Solution:** We start by reviewing integration by parts:

$$\int_a^b u dv = (uv) \Big|_a^b - \int_a^b v du. \tag{1.5}$$

Applying (1.5), let  $u = y$ , and  $du = dy$ . Let  $dv = e^{-y} dy$ , and  $v = -e^{-y}$ . Then,

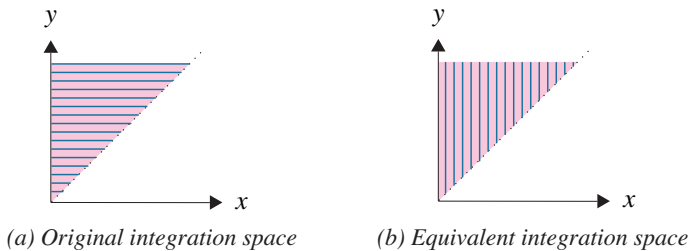
$$\begin{aligned} \int_0^{\infty} ye^{-y} dy &= -ye^{-y} \Big|_{y=0}^{y=\infty} - \int_{y=0}^{\infty} (-e^{-y}) dy \\ &= 0 - (-0) - e^{-y} \Big|_{y=0}^{\infty} \\ &= 0 + 0 - (0 - 1) \\ &= 1. \end{aligned}$$

**Example 1.6** Derive:  $\int_0^{\infty} \int_0^y e^{-y} dx dy$ .

**Solution:** Without changing the order of integration, we have:

$$\begin{aligned} \int_{y=0}^{y=\infty} \int_{x=0}^{x=y} e^{-y} dx dy &= \int_{y=0}^{y=\infty} xe^{-y} \Big|_{x=0}^{x=y} dy \\ &= \int_{y=0}^{y=\infty} ye^{-y} dy \\ &= 1. \end{aligned}$$

To change the order of integration, we first need to understand the space over which we're integrating. The original region of integration is drawn in Figure 1.1(a), where  $y$  ranges from 0 to  $\infty$ , and, for each particular value of  $y$ , we let  $x$  range from 0 to  $y$ .



**Figure 1.1** Region of integration drawn two ways.

We can visualize this instead as shown in Figure 1.1(b), where  $x$  now ranges from 0 to  $\infty$ , and, for each particular value of  $x$ , we let  $y$  range from  $x$  to  $\infty$ :

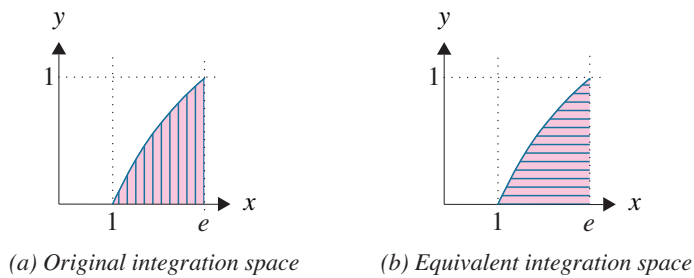
$$\begin{aligned} \int_{x=0}^{x=\infty} \int_{y=x}^{y=\infty} e^{-y} dy dx &= \int_{x=0}^{x=\infty} -e^{-y} \Big|_{y=x}^{y=\infty} dx \\ &= \int_{x=0}^{x=\infty} (0 + e^{-x}) dx \\ &= -e^{-x} \Big|_{x=0}^{x=\infty} \\ &= 1. \end{aligned}$$

**Example 1.7** Derive:  $\int_1^e \int_0^{\ln x} 1 dy dx$ .

**Solution:** Without changing the order of integration, we have:

$$\begin{aligned} \int_{x=1}^{x=e} \int_{y=0}^{y=\ln x} 1 dy dx &= \int_{x=1}^{x=e} \ln x dx \\ &\quad \text{(applying integration by parts)} \\ &= (\ln x \cdot x) \Big|_{x=1}^{x=e} - \int_{x=1}^{x=e} x \cdot \frac{1}{x} dx \\ &= e - 0 - (e - 1) \\ &= 1. \end{aligned}$$

To change the order of integration, we first need to understand the space over which we're integrating. This is drawn in Figure 1.2(a).



**Figure 1.2** Region of integration drawn two ways.

We can visualize this instead as shown in Figure 1.2(b), which leads to the nested integrals:

$$\begin{aligned} \int_{y=0}^{y=1} \int_{x=e^y}^{x=e} 1 dx dy &= \int_{y=0}^{y=1} x \Big|_{x=e^y}^{x=e} dy \\ &= \int_{y=0}^{y=1} (e - e^y) dy \\ &= (ey - e^y) \Big|_{y=0}^{y=1} \\ &= (e - e) - (0 - e^0) \\ &= 1. \end{aligned}$$

## 1.3 Fundamental Theorem of Calculus

The Fundamental Theorem of Calculus (FTC) will come up in the book starting in Chapter 7. We state it here and provide some intuition for why it holds.

**Theorem 1.8 (FTC and extension)** *Let  $f(t)$  be a continuous function defined on the interval  $[a, b]$ . Then, for any  $x$ , where  $a < x < b$ ,*

$$\frac{d}{dx} \int_a^x f(t) dt = f(x). \quad (1.6)$$

*Furthermore, for any differentiable function  $g(x)$ ,*

$$\frac{d}{dx} \int_a^{g(x)} f(t) dt = f(g(x)) \cdot g'(x). \quad (1.7)$$

**We start with intuition for (1.6):**

The integral  $\int_a^x f(t) dt$  represents the area under the curve  $f(t)$  between  $t = a$  and  $t = x$ . We are interested in the rate at which this area changes for a small change in  $x$ .

It helps to think of the integral as a “box” parameterized by  $x$ .

$$\text{Box}(x) = \boxed{\int_a^x f(t) dt}.$$

$$\begin{aligned} \frac{d}{dx} \int_a^x f(t) dt &= \frac{d}{dx} \text{Box}(x) = \lim_{\Delta \rightarrow 0} \frac{\text{Box}(x + \Delta) - \text{Box}(x)}{\Delta} \\ &= \lim_{\Delta \rightarrow 0} \frac{\int_a^{x+\Delta} f(t) dt - \int_a^x f(t) dt}{\Delta} \\ &= \lim_{\Delta \rightarrow 0} \frac{\int_x^{x+\Delta} f(t) dt}{\Delta} \\ &\approx \lim_{\Delta \rightarrow 0} \frac{f(x) \cdot \cancel{\Delta}}{\cancel{\Delta}} \quad f(x) \approx f(x + \Delta) \text{ for tiny } \Delta \\ &= f(x). \end{aligned}$$

**The same argument applies to (1.7):**

$$\text{Box}(x) = \boxed{\int_a^{g(x)} f(t) dt}.$$

$$\begin{aligned}
\frac{d}{dx} \int_a^{g(x)} f(t) dt &= \frac{d}{dx} \text{Box}(x) = \lim_{\Delta \rightarrow 0} \frac{\text{Box}(x + \Delta) - \text{Box}(x)}{\Delta} \\
&= \lim_{\Delta \rightarrow 0} \frac{\int_a^{g(x+\Delta)} f(t) dt - \int_a^{g(x)} f(t) dt}{\Delta} \\
&= \lim_{\Delta \rightarrow 0} \frac{\int_{g(x)}^{g(x+\Delta)} f(t) dt}{\Delta} \\
&\approx \lim_{\Delta \rightarrow 0} \frac{f(g(x)) \cdot (g(x + \Delta) - g(x))}{\Delta} \\
&= f(g(x)) \cdot \lim_{\Delta \rightarrow 0} \frac{g(x + \Delta) - g(x)}{\Delta} \\
&= f(g(x)) \cdot g'(x).
\end{aligned}$$

## 1.4 Review of Taylor Series and Other Limits

There are certain limits and limiting series which come up repeatedly in this book, so we discuss these here.

**Question:** What is the famous limit in (1.8) called, and how should we interpret it?

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n. \quad (1.8)$$

**Answer:** Expression (1.8) is the definition of Euler's number,  $e$ , which is an irrational, transcendental number having value approximately 2.7183.

It helps to think about (1.8) in terms of money. Suppose you have  $m$  dollars. You are promised a 100% interest rate yearly. If the interest is compounded annually, you will have  $2m$  dollars after one year. If the interest is compounded every 6 months, you will have  $\left(1 + \frac{1}{2}\right)^2 m = \frac{9}{4}m$  dollars after one year. If the interest is compounded every 4 months, you will have  $\left(1 + \frac{1}{3}\right)^3 m = \frac{64}{27}m$  dollars after one year. Notice how this keeps going up. If the interest is compounded continuously, you will have

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \cdot m = e \cdot m$$

dollars after one year. Big difference!

**Question:** What, then, is this limit (assume  $x$  is a positive constant):

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n ?$$

**Answer:**

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x. \quad (1.9)$$

To see this, let  $a = \frac{n}{x}$ . As  $n \rightarrow \infty$ , we also have  $a \rightarrow \infty$ :

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \lim_{a \rightarrow \infty} \left(1 + \frac{1}{a}\right)^{ax} = \lim_{a \rightarrow \infty} \left(\left(1 + \frac{1}{a}\right)^a\right)^x = e^x. \quad (1.10)$$

**Question:** Let  $0 < x < 1$ . Let's do some comparisons:

- (a) What is bigger,  $1 + x$  or  $e^x$ ?
- (b) What is bigger,  $1 - x$  or  $e^{-x}$ ?

**Hint:** It helps to think about the Taylor series expansion of  $e^x$  around  $x = 0$ .

**Answer:** For  $0 < x < 1$ , it turns out that  $e^x > 1 + x$  and  $e^{-x} > 1 - x$ . To see this, we start with a brief reminder of the Taylor series expansion around 0, also known as a Maclaurin series. Consider any function  $f(x)$  which is infinitely differentiable at  $x = 0$ . Let us define

$$p(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \dots$$

Observe that the multiplier  $\frac{x^n}{n!}$  gets very small for large  $n$ . It is easy to see that  $p(x)$  is a polynomial that approximates  $f(x)$  very well around  $x = 0$ . In particular, you can see via differentiation that the following are true:

$$\begin{aligned} p(0) &= f(0) \\ p'(0) &= f'(0) \\ p''(0) &= f''(0) \\ p'''(0) &= f'''(0) \\ &\text{etc.} \end{aligned}$$

In fact, Taylor's theorem [72, p.678] says roughly that if  $x$  is within the radius of convergence of  $p(\cdot)$ , then  $p(x)$  approaches  $f(x)$  as we write out more and more terms of  $p(x)$ . Expressing  $p(x)$  with an infinite number of terms allows us to say that  $f(x) = p(x)$ .

Returning to our question, we can see that the function  $f(x) = e^x$  is infinitely differentiable around 0, and thus, for any  $x$ , we can express:

$$e^x = f(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots . \quad (1.11)$$

Thus clearly for any  $x > 0$ , we have that

$$e^x > 1 + x, \quad (1.12)$$

where  $1 + x$  is a very good approximation for  $e^x$  when  $x$  is very small.

Likewise, we can express  $f(x) = e^{-x}$  as

$$e^{-x} = 1 - \frac{x}{1!} + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} - \cdots . \quad (1.13)$$

Now, when  $0 < x < 1$ , we see that

$$e^{-x} > 1 - x, \quad (1.14)$$

because  $\frac{x^2}{2!} > \frac{x^3}{3!} > \frac{x^4}{4!} > \cdots$ . Again,  $1 - x$  is a very good approximation for  $e^{-x}$  when  $x$  is very small.

We end with a discussion of the harmonic series.

**Definition 1.9** *The  $n$ th harmonic number is denoted by  $H_n$ , where*

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n}. \quad (1.15)$$

**Example 1.10 (Approximating  $H_n$ )**

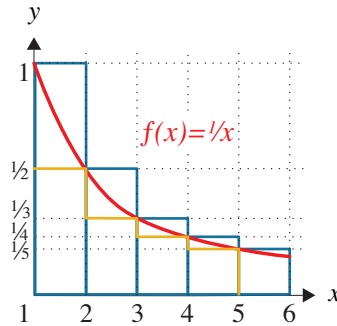
**Question:** How can we find upper and lower bounds on  $H_n$ ?

**Answer:** Figure 1.3 shows the function  $f(x) = \frac{1}{x}$  in red. We know how to exactly compute the area under the red curve. Now observe that the area under the red curve is upper-bounded by the sum of the areas in the blue rectangles, which form a harmonic sum. Likewise, the area under the red curve is lower-bounded by the sum of the rectangles with the yellow border, which form a related harmonic sum. Specifically, summing the area in the blue rectangles, we have that:

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} > \int_1^{n+1} \frac{1}{x} dx = \ln(n+1).$$

Likewise, summing the area in the yellow rectangles, we have that:

$$\ln n = \int_1^n \frac{1}{x} dx > \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}.$$



**Figure 1.3** The area under the red curve  $f(x) = \frac{1}{x}$  is upper-bounded by  $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}$  (in blue), and it is lower-bounded by  $\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}$  (in yellow).

Adding 1 to both sides,

$$1 + \ln(n) > 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = H_n.$$

Hence, for  $n > 1$ :

$$\ln(n+1) < H_n < 1 + \ln(n). \quad (1.16)$$

From (1.16),

$$H_n \approx \ln(n), \quad \text{for high } n. \quad (1.17)$$

Furthermore,

$$\lim_{n \rightarrow \infty} H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots = \infty. \quad (1.18)$$

## 1.5 A Little Combinatorics

Counting is the basis of discrete probability and will be assumed right from the beginning. It's important to differentiate between combinations and permutations.

### Example 1.11 (Combinations versus permutations)

Suppose Baskin-Robins has  $n$  flavors of ice cream. You are building a cone with  $k < n$  scoops. How many different cones can you make if each flavor can only be used once? There are two cases:

- (a) The ordering of the flavors matters.  
 (b) The ordering of the flavors does not matter.

**Question:** Which of (a) and (b) is counting permutations? Which is counting combinations?

**Answer:**

- (a) Each ordering is called a **permutation**. There are  $n$  choices for the bottom flavor, then  $n - 1$  for the next flavor, then  $n - 2$  for the next one, and so on, with  $n - (k - 1)$  for the  $k$ th and final flavor. Hence there are

$$n \cdot (n - 1) \cdot (n - 2) \cdots (n - (k - 1)) = \frac{n!}{(n - k)!}$$

possible permutations.

- (b) Each choice of  $k$  flavors is called a **combination**. We can think of the number of combinations as equal to the number of permutations divided by  $k!$ , since the ordering of flavors doesn't matter:

$$ABC = ACB = CAB = CBA = BAC = BCA.$$

Thus the number of combinations is:

$$\frac{n!}{(n - k)! \cdot k!}.$$

We write

$$\binom{n}{k} = \frac{n!}{(n - k)! \cdot k!}$$

and call this “ $n$  choose  $k$ ” to denote all the ways of choosing  $k$  flavors out of  $n$ .

Note that the number of combinations is smaller than the number of permutations by a factor of  $k!$ .

There are certain sums of combinations which come up a lot in this book. Try these yourself without peeking at the answers:

**Question:** Evaluate  $S_1 = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}$ .

**Answer:** Imagine you have  $n$  elements. Then  $S_1$  represents the total number of possible subsets of the  $n$  elements (all subsets of size 0 plus all subsets of size 1 plus all subsets of size 2, etc.). But the total number of subsets of  $n$  elements is  $2^n$ , because there are two choices for each of the  $n$  elements: each element can either be “in the subset” or “out of the subset.” Hence  $S_1 = 2^n$ .

**Question:** Evaluate  $S_2 = \binom{n}{0}y^n + \binom{n}{1}xy^{n-1} + \binom{n}{2}x^2y^{n-2} + \cdots + \binom{n}{n}x^n$ .

**Answer:** Consider the *binomial expansion* of  $(x+y)^n$ . In this expansion, the term  $x^k y^{n-k}$  appears exactly  $\binom{n}{k}$  times, because you're picking  $k$  of the  $n$  parentheses from which to draw an  $x$ . Thus,  $S_2 = (x+y)^n$ .

**Question:** Evaluate  $S_3 = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n$ .

**Answer:**  $S_3$  is the same as  $S_2$ , where we substitute  $y = 1$ . Hence  $S_3 = (x+1)^n$ .

It is often useful to be able to approximate  $\binom{n}{k}$ . Theorem 1.12 provides easy-to-use upper and lower bounds.

**Theorem 1.12 (Simple bounds on  $\binom{n}{k}$ )**

$$\left(\frac{n}{k}\right)^k < \binom{n}{k} < \left(\frac{ne}{k}\right)^k. \quad (1.19)$$

**Proof:**

*First we prove the upper bound:*

$$\begin{aligned} \binom{n}{k} &= \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} \\ &= \frac{\frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{n-k+1}{n} \cdot n^k}{k!} \\ &= \frac{1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \cdot n^k}{k!} \\ &< \frac{n^k}{k!}. \end{aligned} \quad (1.20)$$

Now the Taylor series expansion of  $e^k$  for positive integer  $k$  tells us that  $e^k > \frac{k^k}{k!}$ .

Thus:

$$\binom{n}{k} < \frac{n^k}{k!} = \frac{n^k}{k^k} \cdot \frac{k^k}{k!} < \frac{n^k}{k^k} \cdot e^k = \left(\frac{ne}{k}\right)^k.$$

*Next we prove the lower bound:*

Starting from (1.20),

$$\begin{aligned}
 \binom{n}{k} &= \frac{1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \cdot n^k}{k \cdot (k-1) \cdot (k-2) \cdots (k - (k-1))} \\
 &= \frac{1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \cdot n^k}{\frac{k}{k} \cdot \frac{k-1}{k} \cdot \frac{k-2}{k} \cdots \frac{k-(k-1)}{k} \cdot k^k} \\
 &= \frac{1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)}{1 \cdot \left(1 - \frac{1}{k}\right) \cdot \left(1 - \frac{2}{k}\right) \cdots \left(1 - \frac{k-1}{k}\right)} \cdot \left(\frac{n}{k}\right)^k \\
 &> \left(\frac{n}{k}\right)^k.
 \end{aligned}$$

The last line follows since  $k < n$ . Thus  $1 - \frac{i}{n} > 1 - \frac{i}{k}$  for  $0 < i < k$ . ■

Another useful combinatorial identity is called **Vandermonde's identity**.

**Theorem 1.13 (Vandermonde's identity)**

$$\binom{m+n}{r} = \sum_{k=0}^r \binom{m}{k} \cdot \binom{n}{r-k}.$$

**Question:** What is the logic behind Vandermonde's identity?

**Answer:** Suppose you are trying to pick  $r$  leaders from a set of  $m$  women and  $n$  men. There are  $\binom{m+n}{r}$  ways of picking the  $r$  leaders, which is the left-hand side of Vandermonde's identity. However, we can also view the picking of the  $r$  leaders as first picking  $k$  leaders from the set of  $m$  women, where  $0 \leq k \leq r$ , and then picking  $r - k$  leaders from the set of  $n$  men. We now need to sum over all possible values of  $k$ . But this latter view represents the right-hand side of Vandermonde's identity.

**Question:** What does Vandermonde's identity say about  $\binom{2n}{n}$ ?

**Answer:** By Vandermonde's identity,

$$\binom{2n}{n} = \sum_{k=0}^n \binom{n}{k}^2. \tag{1.21}$$

We end with a final useful result when working with factorials, known as the **Stirling bounds**, whose proof can be found in [76].

**Theorem 1.14 (Stirling)** For all positive integers  $n$ ,

$$\sqrt{2\pi n} \left(\frac{n}{e}\right)^n \leq n! \leq e\sqrt{n} \left(\frac{n}{e}\right)^n. \quad (1.22)$$

The power of Stirling's result lies in the fact that the upper and lower bounds given in (1.22) differ by a multiplicative constant of less than 1.1.

## 1.6 Review of Asymptotic Notation

Asymptotic notation primarily comes up starting in Chapter 18. Asymptotic notation is a convenient way to summarize the rate at which a function  $f(n)$  grows with  $n$ , without getting into the specifics. For example, we'd like to be able to say that  $f(n) = 3n$  grows linearly with  $n$ , while  $f(n) = \sqrt{5}n^2$  grows quadratically with  $n$ . The asymptotic notation below will help us do this.

In words:

- $O(g(n))$  is the set of functions that grow no faster than  $g(n)$ .
- $o(g(n))$  is the set of functions that grow strictly slower than  $g(n)$ .
- $\Theta(g(n))$  is the set of functions that grow at the same rate as  $g(n)$ .
- $\Omega(g(n))$  is the set of functions that grow no slower than  $g(n)$ .
- $\omega(g(n))$  is the set of functions that grow strictly faster than  $g(n)$ .

Throughout our discussion we will assume that  $f(n)$  and  $g(n)$  are functions that map *positive* integers  $n$  to *positive* real numbers. If  $f(n)$  is negative (e.g.,  $f(n) = -3n$ ) we classify it based on its absolute value.

We typically write  $f(n) = O(g(n))$  to denote that  $f(n)$  is in the set of functions  $O(g(n))$ . We follow this same convention for the other sets.

### 1.6.1 Big-O and Little-o

**Definition 1.15** We say that  $f(n) = O(g(n))$ , pronounced as  $f(n)$  is “**big-O**” of  $g(n)$ , if

$$\limsup_{n \rightarrow \infty} \frac{f(n)}{g(n)} < \infty.$$

It is sometimes easier to use the following alternative definition of big-O:

*Alternative definition:*  $f(n) = O(g(n))$  if there exists a positive constants  $c > 0$  and  $n_0 > 0$  such that  $0 \leq f(n) \leq c \cdot g(n)$ , for all  $n \geq n_0$ .

### Example 1.16 (Big-O)

- $n = O(3n)$
- $3n = O(n)$
- $-3n^2 \neq O(n)$
- $n \lg n + 18n = O(n^2)$
- $\frac{3 \lg \lg n}{\lg n} = O(1)$

**Definition 1.17** We say that  $f(n) = o(g(n))$ , pronounced as  $f(n)$  is “**little-o**” of  $g(n)$ , if

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0.$$

**Corollary 1.18**  $f(n) = o(1)$  if and only if  $f(n) \rightarrow 0$  as  $n \rightarrow \infty$ .

### Example 1.19 (Little-o)

- $3n \neq o(n)$
- $n \lg n + 18n = o(n^2)$
- $n^2 = o(n^3)$
- $\lg n = o(n^{0.01})$
- $\frac{3 \lg \lg n}{\lg n} = o(1)$

**Question:** If  $f(n) = o(g(n))$ , does that imply that  $f(n) = O(g(n))$ ? How about the converse?

**Answer:**  $f(n) = o(g(n))$  implies that  $f(n) = O(g(n))$ . The converse is not true.

When comparing complicated functions, it is often helpful to first think about simpler functions. Consider the following example.

### Example 1.20 (More complicated functions)

Let  $f(n) = n$ . Let  $g(n) = \frac{\log \log n}{\log \log \log n}$ .

**Question:** Is it the case that  $f(n) = o(g(n))$ ? Or is  $g(n) = o(f(n))$ ?

**Hint:** Rather than take limits, consider an easier question: how do  $n$  and  $\log \log n$  compare?

**Answer:** It is easy to see that

$$\log \log n = o(n).$$

Thus it is also true that if we make the left-hand side smaller, the relationship still holds. By dividing the left-hand side by  $\log \log \log n$ , we are dividing it by something that is  $\geq 1$ , and hence we are only making it smaller. Thus we also have

$$\frac{\log \log n}{\log \log \log n} = o(n).$$

## 1.6.2 Big-Omega and Little-omega

**Definition 1.21** We say that  $f(n) = \Omega(g(n))$ , pronounced as  $f(n)$  is “**big-Omega**” of  $g(n)$ , if

$$\liminf_{n \rightarrow \infty} \frac{f(n)}{g(n)} > 0.$$

It is sometimes easier to use the following alternative definition of big-Omega: *Alternative definition:*  $f(n) = \Omega(g(n))$  if there exists a positive constants  $c > 0$  and  $n_0 > 0$  such that  $f(n) \geq c \cdot g(n) \geq 0$ , for all  $n \geq n_0$ .

### Example 1.22 (Big-Omega)

- $3n = \Omega(n)$
- $n \lg n + 18n = \Omega(n \lg n)$
- $n^2 = \Omega(n \lg n)$
- $\frac{3 \lg \lg n}{\lg n} \neq \Omega(1)$

**Definition 1.23** We say that  $f(n) = \omega(g(n))$ , pronounced as  $f(n)$  is “**little-omega**” of  $g(n)$ , if

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \infty.$$

### Example 1.24 (Little-omega)

- $3n \neq \omega(n)$

- $n \lg n + 18n = \omega(n)$
- $n^2 = \omega(n \lg n)$
- $\frac{3 \lg \lg n}{\lg n} \neq \omega(1)$

**Question:** If  $f(n) = \omega(g(n))$ , does that imply that  $f(n) = \Omega(g(n))$ ? How about the converse?

**Answer:**  $f(n) = \omega(g(n))$  implies that  $f(n) = \Omega(g(n))$ , but the converse is not true.

### 1.6.3 Big-Theta

**Definition 1.25** We say that  $f(n) = \Theta(g(n))$ , pronounced as  $f(n)$  is “**big-Theta**” of  $g(n)$ , if both

$$f(n) = O(g(n)) \text{ and } f(n) = \Omega(g(n)).$$

#### Example 1.26 (Big-Theta)

- $3n = \Theta(n)$
- $n \lg n + 18n \neq \Theta(n)$

**Question:** Can we have  $f(n) = o(g(n))$  and  $f(n) = \omega(g(n))$ ?

**Answer:** No. This would require  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)}$  to be both 0 and  $\infty$ .

We end with a harder example:

**Example 1.27** Prove that  $3^{\log_2(n^2)} = o(n^4)$ .

**Solution:**

$$\lim_{n \rightarrow \infty} \frac{3^{\log_2(n^2)}}{n^4} = \lim_{n \rightarrow \infty} \frac{3^{2 \log_2 n}}{2^{\log_2(n^4)}} = \lim_{n \rightarrow \infty} \frac{9^{\log_2 n}}{2^{4 \log_2 n}} = \lim_{n \rightarrow \infty} \frac{9^{\log_2 n}}{16^{\log_2 n}} = \lim_{n \rightarrow \infty} \left( \frac{9}{16} \right)^{\log_2 n} = 0.$$

## 1.7 Exercises

### 1.1 Choosing ice-cream flavors

There are  $n$  distinct flavors of ice-cream. Jin and Jong each pick two distinct

flavors. Count the number of possibilities where Jin and Jong share exactly one flavor. In this problem the order of the flavors does not matter. An example of a possibility that “counts” is:

Jin gets {Chocolate, Vanilla} and Jong gets {Chocolate, Strawberry}.

### 1.2 Evaluating a single integral

Evaluate the following integral:

$$\int_0^{\infty} 5xe^{-5x} dx.$$

### 1.3 Evaluating a double integral

Evaluate this double integral:

$$\int_0^{10} \int_0^y x dx dy.$$

Do this in two different ways. For each way, draw the region of integration:

- Without changing the order of integration.
- By changing the order of integration.

### 1.4 Evaluating a double integral

Evaluate this double integral:

$$\int_1^{10} \int_1^{x^2} dy dx.$$

Do this in two different ways. For each way, draw the region of integration:

- Without changing the order of integration.
- By changing the order of integration.

### 1.5 Practice with Taylor series

Let  $f(x) = \ln(1+x)$ . Follow the procedure in the chapter to express the Taylor series expansion of  $f(\epsilon)$ , around 0, where  $0 < \epsilon \leq 1$ .

### 1.6 Counting balls and bins

You are throwing  $n$  identical balls into  $n$  distinct (numbered) bins, where each ball is thrown into one bin.

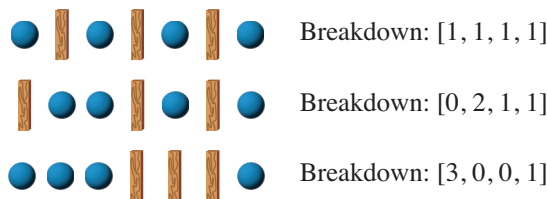


Figure 1.4 Some examples of Exercise 1.6 where  $n = 4$ .

- (a) How many ways are there to distribute the  $n$  balls among the  $n$  bins?  
[Hint: Figure 1.4 illustrates that each way of distributing balls can be viewed as an arrangement of  $n - 1$  “sticks” and  $n$  balls.]
- (b) How many ways are there to distribute the  $n$  balls among the  $n$  bins such that bin 1 has  $\geq k$  balls?

### 1.7 Permutations

Consider all permutations of the numbers  $\{1, 2, 3, \dots, n\}$ .

- (a) How many permutations are there?
- (b) What fraction of all permutations have a 1 in the first position?
- (c) What fraction of all permutations have a 1 in the first position and a 2 in the second position?
- (d) What fraction of all permutations have a 1 somewhere before a 2?

### 1.8 Practice with asymptotic notation

- (a) Let  $a(n) = n$ . Let  $b(n) = n \log n$ . Which of the following are true?
- $a(n) = O(b(n))$
  - $a(n) = o(b(n))$
  - $a(n) = \Theta(b(n))$
  - $a(n) = \omega(b(n))$
  - $a(n) = \Omega(b(n))$
- (b) Repeat the problem where now  $a(n) = \frac{3 \log 3}{\ln \ln n}$  and  $b(n) = 1$ .
- (c) Repeat the problem where now  $a(n) = 2^n$  and  $b(n) = n^{\ln n}$ .

### 1.9 Harder practice with asymptotic notation

Use the definitions for asymptotic notation to prove the following statements:

- (a) Show that  $5^{\log_2(n^3)} = \omega(n^6)$ .
- (b) Show that  $(\ln n)^{\frac{3 \ln n}{\ln \ln n}} = \Omega(n^3)$ .
- (c) Given that  $k = \frac{9 \ln n}{\ln \ln n}$ , show that for large  $n$  we have  $k \ln k \geq 8 \ln n$ .