6 z-Transforms

This chapter is a very brief introduction to the wonderful world of transforms. Transforms come in many varieties. There are z-transforms, moment-generating functions, characteristic functions, Fourier transforms, Laplace transforms, and more. All are very similar in their function. In this chapter, we will study *z-transforms*, a variant particularly well suited to common *discrete* random variables. In Chapter 11, we will study Laplace transforms, a variant ideally suited to common continuous random variables.

Transforms are an extremely powerful analysis technique. In this chapter we will cover two of the most common uses of transforms:

1. Computing higher moments of random variables (see Sections 6.1 – 6.6).
2. Solving recurrence relations, particularly recurrences that will come up later when we study Markov chains. This will be discussed in Section 6.7 and then again when we get to infinite-state Markov chains in Chapter 26.

6.1 Motivating Examples

Suppose that you want to know the third moment of a Binomial(n, p) distribution. Let $X \sim \text{Binomial}(n, p)$. Then,

$$\mathbf{E}\left[X^3\right] = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} i^3.$$

This is a daunting expression.

As another example, you might want to know the fifth moment of a Poisson(λ) distribution. Let $Y \sim \text{Poisson}(\lambda)$. Then,

$$\mathbf{E}\left[X^5\right] = \sum_{i=0}^{\infty} \frac{e^{-\lambda}\lambda^i}{i!} \cdot i^5.$$

Again, it's not easy to see how to derive this. One of the most important uses of transforms is that they allow us to obtain *all* moments of a random variable, as we now explain.

6.2 The Transform as an Onion

One can think of the transform of a random variable as an onion, shown in Figure 6.1. This onion contains inside it all the moments of the random variable. Getting the moments out of the onion is not an easy task, however, and may involve some tears as the onion is peeled, where the "peeling process" involves differentiating the transform. The first moment is stored in the outermost layer of the onion and thus does not require too much peeling to reach. The second moment is stored a little deeper, the third moment even deeper (more tears), etc. Although getting the moments is painful, it is entirely straightforward how to do it – just keep peeling the layers.

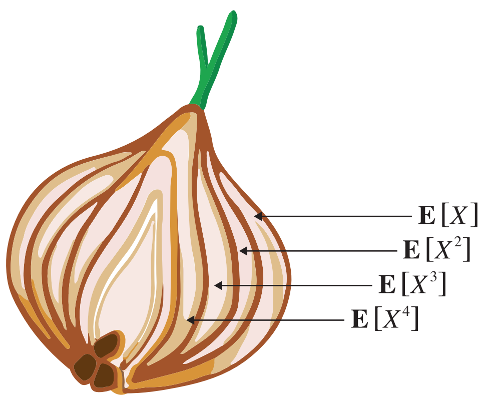

Figure 6.1 *The z-transform onion.*

> **Definition 6.1** *The* **z-transform**, $G_p(z)$, *of a discrete function,* $p(i)$, $i = 0, 1, 2, \ldots$ *is defined as*
>
> $$G_p(z) = \sum_{i=0}^{\infty} p(i) z^i.$$

Observe that the z-transform is a polynomial in z. Here, z should be thought of as a placeholder that keeps $p(i)$ separated from $p(i+1)$, by multiplying $p(i)$ by z^i and multiplying $p(i+1)$ by z^{i+1}.

When we speak of the z-transform of a discrete random variable (r.v.) X, we are referring to the z-transform of the probability mass function (p.m.f.), $p_X(\cdot)$, associated with X.

Definition 6.2 *Let X be a non-negative discrete r.v. with p.m.f. $p_X(i)$, where $i = 0, 1, 2, \ldots$. Then the* **z-transform of r.v. X** *is written as $\widehat{X}(z)$, where*

$$\widehat{X}(z) = G_{p_X}(z) = \mathbf{E}\left[z^X\right] = \sum_{i=0}^{\infty} p_X(i) z^i.$$

Throughout, we assume that z is a constant and we will assume that $|z| \leq 1$.

Note that the z-transform can be defined for any random variable. However, convergence is *guaranteed* when the r.v. is *non-negative* and $|z| \leq 1$, as we'll see in Theorem 6.6. That does not mean that convergence doesn't happen in other settings as well.

Question: What is $\widehat{X}(1)$?

Theorem 6.3 *For all discrete random variables, X,*

$$\widehat{X}(1) = 1.$$

Proof:

$$\widehat{X}(1) = \mathbf{E}\left[z^X\right]\Big|_{z=1} = \sum_{i=-\infty}^{\infty} p_X(i) \cdot 1^i = 1. \qquad \blacksquare$$

6.3 Creating the Transform: Onion Building

The z-transform is defined so as to be really easy to compute for all the commonly used discrete random variables. Below are some examples.

Example 6.4 *Derive the z-transform of $X \sim$ Binomial(n, p):*

$$\begin{aligned}
\widehat{X}(z) = \mathbf{E}\left[z^X\right] &= \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} z^i \\
&= \sum_{i=0}^{n} \binom{n}{i} (zp)^i (1-p)^{n-i} \\
&= (zp + (1-p))^n .
\end{aligned}$$

Example 6.5 *Derive the z-transform of $X \sim$ Geometric(p):*

$$\begin{aligned}
\widehat{X}(z) = \mathbf{E}\left[z^X\right] &= \sum_{i=1}^{\infty} p(1-p)^{i-1} z^i \\
&= zp \sum_{i=1}^{\infty} (z(1-p))^{i-1} \\
&= zp \sum_{i=0}^{\infty} (z(1-p))^{i} \\
&= \frac{zp}{1 - z(1-p)}.
\end{aligned}$$

Question: Can you see where we used the fact that $|z| \le 1$ above?

Answer: We needed $|z(1-p)| < 1$ to get $\sum_{i=1}^{\infty} (z(1-p))^{i-1}$ to converge.

In both the above cases, notice how much easier it is to create the transform than to compute higher moments.

One might wonder if the series defined by $\widehat{X}(z)$ might in some cases diverge. This is not the case.

Theorem 6.6 (Convergence of z-transform) *$\widehat{X}(z)$ is bounded for any non-negative discrete r.v. X, assuming $|z| \le 1$.*

Proof: We are given that

$$-1 \le z \le 1.$$

Because $i \ge 0$, this implies that

$$-1 \le z^i \le 1.$$

Multiplying all terms by $p_X(i)$, we have

$$-p_X(i) \le z^i p_X(i) \le p_X(i).$$

Now summing over all i, we have

$$-\sum_i p_X(i) \le \sum_i z^i p_X(i) \le \sum_i p_X(i),$$

which evaluates to

$$-1 \le \widehat{X}(z) \le 1.$$

So $\widehat{X}(z)$ is bounded between -1 and 1. ■

6.4 Getting Moments: Onion Peeling

Once we have created the onion corresponding to a r.v., we can "peel its layers" to extract the moments of the random variable.

Theorem 6.7 (Onion Peeling Theorem) *Let X be a discrete, integer-valued, non-negative r.v. with p.m.f. $p_X(i)$, $i = 0, 1, 2, \ldots$. Then we can get the moments of X by differentiating $\widehat{X}(z)$ as follows:*

$$\begin{aligned}
\widehat{X}'(z)\Big|_{z=1} &= \mathbf{E}[X] \\
\widehat{X}''(z)\Big|_{z=1} &= \mathbf{E}[X(X-1)] \\
\widehat{X}'''(z)\Big|_{z=1} &= \mathbf{E}[X(X-1)(X-2)] \\
\widehat{X}''''(z)\Big|_{z=1} &= \mathbf{E}[X(X-1)(X-2)(X-3)] \\
&\vdots
\end{aligned}$$

Note: If the above moments are not defined at $z = 1$, one can instead consider the limit as $z \to 1$, where evaluating the limit may require using L'Hospital's rule.

Proof: Below we provide a sketch of the proof argument. This can be obtained formally via induction and can also be expressed more compactly. However, we choose to write it out this way so that you can visualize exactly how the moments "pop" out of the transform when it's differentiated:

$$\widehat{X}(z) = p_X(0)z^0 + p_X(1)z^1 + p_X(2)z^2 + p_X(3)z^3 + p_X(4)z^4 + p_X(5)z^5 + \cdots$$

$$\begin{aligned}
\widehat{X}'(z) &= p_X(1) + 2p_X(2)z^1 + 3p_X(3)z^2 + 4p_X(4)z^3 + 5p_X(5)z^4 + \cdots \\
\widehat{X}'(z)\Big|_{z=1} &= 1p_X(1) + 2p_X(2) + 3p_X(3) + 4p_X(4) + 5p_X(5) + \cdots \\
&= \mathbf{E}[X] \checkmark
\end{aligned}$$

$$\begin{aligned}
\widehat{X}''(z) &= 2p_X(2) + 3\cdot 2p_X(3)z + 4\cdot 3p_X(4)z^2 + 5\cdot 4p_X(5)z^3 + \cdots \\
\widehat{X}''(z)\Big|_{z=1} &= 2\cdot 1p_X(2) + 3\cdot 2p_X(3) + 4\cdot 3p_X(4) + 5\cdot 4p_X(5) + \cdots \\
&= \mathbf{E}[X(X-1)] \checkmark
\end{aligned}$$

$$\begin{aligned}
\widehat{X}'''(z) &= 3\cdot 2p_X(3) + 4\cdot 3\cdot 2p_X(4)z + 5\cdot 4\cdot 3p_X(5)z^2 + \cdots \\
\widehat{X}'''(z)\Big|_{z=1} &= 3\cdot 2\cdot 1p_X(3) + 4\cdot 3\cdot 2p_X(4) + 5\cdot 4\cdot 3p_X(5) + \cdots \\
&= \mathbf{E}[X(X-1)(X-2)] \checkmark
\end{aligned}$$

And so on ... ■

Question: What is the insight behind the above proof? How does the transform hold all these moments?

Answer: The insight is that the "z" term separates the layers, allowing us to get each successive moment when differentiating. One can think of the z's as the pasta in the lasagna that keeps everything from running together.

Let's consider an example of applying the Onion Peeling Theorem.

Example 6.8 (Variance of Geometric) *Let $X \sim$ Geometric(p). Compute* **Var**(X).

$$\begin{aligned}
\widehat{X}(z) &= \frac{zp}{1 - z(1-p)} \\
\mathbf{E}[X] &= \frac{d}{dz}\left(\frac{zp}{1-z(1-p)}\right)\bigg|_{z=1} = \frac{p}{(1-z(1-p))^2}\bigg|_{z=1} = \frac{1}{p} \\
\mathbf{E}\left[X^2\right] &= \widehat{X}''(z)\Big|_{z=1} + \mathbf{E}[X] = \frac{2p(1-p)}{(1-z(1-p))^3}\bigg|_{z=1} + \frac{1}{p} = \frac{2-p}{p^2} \\
\mathbf{Var}(X) &= \mathbf{E}\left[X^2\right] - (\mathbf{E}[X])^2 = \frac{1-p}{p^2}.
\end{aligned}$$

Question: As we've seen, the z-transform of X is an onion that contains all moments of X. But does it also contain the distribution of X? Specifically, can you get the p.m.f., $p_X(i)$, from the onion, $\widehat{X}(z)$?

Answer: The answer is yes! In Exercise 6.14 you will derive an algorithm for extracting $p_X(i)$ from $\widehat{X}(z)$ for any non-negative discrete r.v. X. This tells us that there is an injective mapping from the set of discrete non-negative random variables to the set of z-transforms. Put another way, the z-transform uniquely determines the distribution.

6.5 Linearity of Transforms

Since transforms are just expectations, it makes sense that one might have a law similar to Linearity of Expectation. However, since transforms encompass *all* moments, it also makes sense that such a law might require independence of the random variables being added. Theorem 6.9 encapsulates these points.

Theorem 6.9 (Linearity) *Let X and Y be independent discrete random variables. Let $W = X + Y$. Then the z-transform of W is $\widehat{W}(z) = \widehat{X}(z) \cdot \widehat{Y}(z)$.*

Proof:

$$\begin{aligned}
\widehat{W}(z) &= \mathbf{E}\left[z^W\right] \\
&= \mathbf{E}\left[z^{X+Y}\right] \\
&= \mathbf{E}\left[z^X \cdot z^Y\right] \\
&= \mathbf{E}\left[z^X\right] \cdot \mathbf{E}\left[z^Y\right] \\
&= \widehat{X}(z) \cdot \widehat{Y}(z).
\end{aligned}$$

■

Question: Where did we use the fact that $X \perp Y$?

Answer: In splitting up the expectation into a product of expectations.

Example 6.10 (From Bernoulli to Binomial)

Let $X \sim \text{Bernoulli}(p)$. Let $Y \sim \text{Binomial}(n, p)$.

Question: (a) What is $\widehat{X}(z)$? (b) How can we use $\widehat{X}(z)$ to get $\widehat{Y}(z)$?

Answer:

(a) $\widehat{X}(z) = (1-p) \cdot z^0 + p \cdot z^1 = 1 - p + pz$.
(b) $Y = \sum_{i=1}^{n} X_i$. Given that $X_i \sim X$, for all i, and the X_i's are independent,

$$\widehat{Y}(z) = \left(\widehat{X}(z)\right)^n = (1 - p + pz)^n.$$

Example 6.11 (Sum of Binomials)

Let $X \sim \text{Binomial}(n, p)$ and $Y \sim \text{Binomial}(m, p)$, where $X \perp Y$.

Question: What is the distribution of $Z = X + Y$?

Answer:

$$\begin{aligned}
\widehat{Z}(z) &= \widehat{X}(z) \cdot \widehat{Y}(z) \\
&= (zp + (1-p))^n \cdot (zp + (1-p))^m \\
&= (zp + (1-p))^{m+n} .
\end{aligned}$$

Observe that $(zp + (1-p))^{m+n}$ is the z-transform of a Binomial r.v. with parameters $m + n$ and p. Thus, the distribution of Z must be $\text{Binomial}(m + n, p)$, which should make sense.

6.6 Conditioning

Theorem 6.12 *Let X, A, and B be discrete random variables where*

$$X = \begin{cases} A & \text{w/prob } p \\ B & \text{w/prob 1 - } p \end{cases}.$$

Then,

$$\widehat{X}(z) = p \cdot \widehat{A}(z) + (1 - p) \cdot \widehat{B}(z).$$

Theorem 6.12 should be interpreted as first tossing a p-coin (coin with probability p of heads). If that coin comes up heads, then set $X = A$. Otherwise set $X = B$.

Proof:

$$\begin{aligned} \widehat{X}(z) &= \mathbf{E}\left[z^X\right] \\ &= \mathbf{E}\left[z^X \middle| X = A\right] \cdot p + \mathbf{E}\left[z^X \middle| X = B\right] \cdot (1 - p) \\ &= \mathbf{E}\left[z^A\right] \cdot p + \mathbf{E}\left[z^B\right] \cdot (1 - p) \\ &= p\widehat{A}(z) + (1 - p)\widehat{B}(z). \end{aligned}$$

■

Question: In the examples in the previous section, we considered the sum of a constant number (n) of random variables. How can we use conditioning to derive the z-transform of the sum of a r.v. number (N) of random variables?

Answer: Exercise 6.10 walks you through the proof of Theorem 6.13, which generalizes Theorem 5.14 to all higher moments.

Theorem 6.13 (Summing a random number of i.i.d. random variables) *Let $X_1, X_2, X_3, \ldots$ be i.i.d. discrete random variables, where $X_i \sim X$. Let N be a positive, integer-valued, discrete r.v., where $N \perp X_i$ for all i. Let*

$$S = \sum_{i=1}^{N} X_i.$$

Then,

$$\widehat{S}(z) = \widehat{N}\left(\widehat{X}(z)\right),$$

that is, we substitute in $\widehat{X}(z)$ as the z-parameter in $\widehat{N}(z)$.

6.7 Using z-Transforms to Solve Recurrence Relations

Recurrence relations are prevalent throughout computer science, biology, signal processing, and economics, just to name a few fields. One of the most common types of recurrence relations is a linear homogeneous recurrence, of the form:

$$f_{i+n} = a_1 f_{i+n-1} + a_2 f_{i+n-2} + \cdots + a_n f_i.$$

A popular example of such a recurrence relation is the following:

$$f_{i+2} = f_{i+1} + f_i, \tag{6.1}$$

where $f_0 = 0$ and $f_1 = 1$.

Question: Do you recognize the relation?

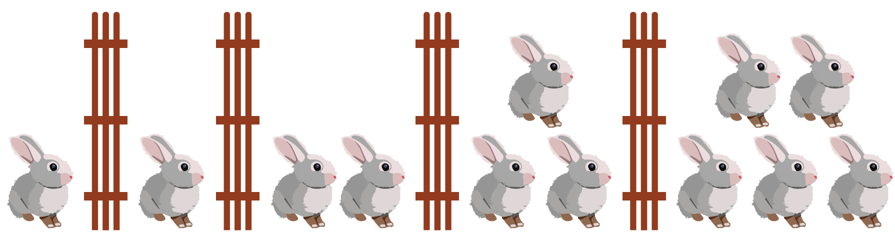

Figure 6.2 *Fibonacci sequence.*

Answer: Equation (6.1) is the Fibonacci sequence. It was used to model the growth in the population of rabbits, where f_i denotes the number of rabbits in month i.

Solving a recurrence relation means finding a closed-form expression for f_n. While (6.1) seems very simple to solve by just "unraveling the recurrence," it turns out to be impossible to do this. It also is hard to imagine how one might "guess" the form of the solution. Fortunately, z-transforms provide an excellent technique for solving these recurrence relations. In this section, we see how to derive a closed-form expression for f_n using z-transforms. This method may seem overly complex. However it's the easiest technique known for handling recurrences. We start by defining the z-transform of a sequence.

Definition 6.14 *Given a sequence of values:* $\{f_0, f_1, f_2, \ldots\}$. *Define*

$$F(z) = \sum_{i=0}^{\infty} f_i z^i.$$

$F(z)$ is the **z-transform of the sequence**. *Note that z just functions as a placeholder, for the purpose of separating out the f_i's. Note that the f_i's here are* not *probabilities, and there is no r.v. associated with this z-transform.*

We illustrate the method on a recurrence relation of this form:

$$f_{i+2} = b f_{i+1} + a f_i, \tag{6.2}$$

where we assume f_0 and f_1 are given and a and b are constants. However, the method can be applied more generally. Our goal is to derive a closed-form expression for f_n.

Step 1: Derive $F(z)$ as a ratio of polynomials.

The goal in Step 1 is to derive $F(z)$. It will be useful to represent $F(z)$ as a ratio of two polynomials in z. From (6.2), we have:

$$\begin{aligned}
f_{i+2} &= b f_{i+1} + a f_i \\
f_{i+2} z^{i+2} &= b f_{i+1} z^{i+2} + a f_i z^{i+2} \\
\sum_{i=0}^{\infty} f_{i+2} z^{i+2} &= b \sum_{i=0}^{\infty} f_{i+1} z^{i+2} + a \sum_{i=0}^{\infty} f_i z^{i+2} \\
F(z) - f_1 z - f_0 &= bz \sum_{i=0}^{\infty} f_{i+1} z^{i+1} + az^2 \sum_{i=0}^{\infty} f_i z^i \\
F(z) - f_1 z - f_0 &= bz (F(z) - f_0) + az^2 F(z) \\
\left(1 - bz - az^2\right) F(z) &= f_1 z + f_0 - bz f_0 \\
F(z) &= \frac{f_0 + z(f_1 - b f_0)}{1 - bz - az^2}.
\end{aligned} \tag{6.3}$$

Step 2: Rewrite $F(z)$ via partial fractions.

The goal in Step 2 is to apply partial fractions to $F(z)$. Specifically, we want to write

$$F(z) = \frac{N(z)}{D(z)} = \frac{A}{h(z)} + \frac{B}{g(z)},$$

where $D(z) = h(z) \cdot g(z)$ and h, g are (hopefully) linear in z.

Lemma 6.15 *If $D(z) = az^2 + bz + 1$, then*

$$D(z) = \left(1 - \frac{z}{r_0}\right)\left(1 - \frac{z}{r_1}\right),$$

where r_0 and r_1 are the (real) roots of $D(z)$.

Proof: To see that the two ways of writing $D(z)$ are equivalent, we note that the two quadratic expressions have the same two roots (r_0 and r_1) and furthermore have the same constant term, 1. ∎

In our case, see (6.3), $D(z) = -az^2 - bz + 1$, so

$$(r_0, r_1) = \left(\frac{-b - \sqrt{b^2 + 4a}}{2a}, \frac{-b + \sqrt{b^2 + 4a}}{2a}\right) \tag{6.4}$$

$$\begin{aligned} D(z) &= h(z) \cdot g(z) \\ h(z) &= 1 - \frac{z}{r_0} \\ g(z) &= 1 - \frac{z}{r_1}. \end{aligned}$$

We now use $N(z) = f_0 + z(f_1 - f_0 b)$ from (6.3) to solve for A and B:

$$F(z) = \frac{A}{1 - \frac{z}{r_0}} + \frac{B}{1 - \frac{z}{r_1}} \tag{6.5}$$

$$= \frac{A\left(1 - \frac{z}{r_1}\right) + B\left(1 - \frac{z}{r_0}\right)}{\left(1 - \frac{z}{r_0}\right)\left(1 - \frac{z}{r_1}\right)}$$

$$= \frac{(A+B) + z\left(-\frac{A}{r_1} - \frac{B}{r_0}\right)}{\left(1 - \frac{z}{r_0}\right)\left(1 - \frac{z}{r_1}\right)} = \frac{N(z)}{D(z)} = \frac{f_0 + z(f_1 - f_0 b)}{D(z)}. \tag{6.6}$$

Matching the z-coefficients in the numerators of (6.6), we have

$$\begin{aligned} A + B &= f_0 \\ -\frac{A}{r_1} - \frac{B}{r_0} &= f_1 - f_0 b, \end{aligned}$$

which solves to

$$B = \frac{r_0 f_0 + (f_1 - f_0 b)\, r_0 r_1}{r_0 - r_1} \tag{6.7}$$

$$A = f_0 - B. \tag{6.8}$$

Step 3: Rewrite $F(z)$ via series expansion.

Returning to (6.5), we assume that z is chosen such that $0 < z < r_0$ and $0 < z < r_1$. This allows us to write:

$$\frac{A}{1 - \frac{z}{r_0}} = A \sum_{i=0}^{\infty} \left(\frac{z}{r_0} \right)^i \quad \text{and} \quad \frac{B}{1 - \frac{z}{r_1}} = B \sum_{i=0}^{\infty} \left(\frac{z}{r_1} \right)^i.$$

Thus, the geometric series expansion of $F(z)$ can be rewritten as follows:

$$F(z) = \sum_{i=0}^{\infty} f_i z^i = A \sum_{i=0}^{\infty} \left(\frac{z}{r_0} \right)^i + B \sum_{i=0}^{\infty} \left(\frac{z}{r_1} \right)^i. \tag{6.9}$$

Step 4: Match terms to obtain f_n.

Finally, we match the z-coefficients in (6.9) to obtain the f_n's:

$$f_n = \frac{A}{r_0^n} + \frac{B}{r_1^n}, \tag{6.10}$$

where A and B are obtained from (6.8) and (6.7) and r_0 and r_1 are obtained from (6.4).

To get a final form, recall that we are given that $f_0 = 0$ and $f_1 = 1$. Furthermore, $a = 1$ and $b = 1$. Then, from (6.4), we have that

$$\begin{aligned} r_0 &= -\phi \\ r_1 &= \phi^{-1} \end{aligned}$$

where

$$\phi = \frac{1 + \sqrt{5}}{2}.$$

Finally, from (6.8) and (6.7), we get:

$$A = -\frac{1}{\sqrt{5}} \qquad \text{and} \qquad B = \frac{1}{\sqrt{5}}.$$

Substituting these into (6.10), we get

$$f_n = \frac{1}{\sqrt{5}} \left(\phi^n - (-\phi)^{-n} \right).$$

6.8 Exercises

6.1 **Moments of Poisson**
Use z-transforms to derive $\mathbf{E}[X(X-1)(X-2)\cdots(X-k+1)]$, for $k = 1, 2, 3, \ldots$, where $X \sim \text{Poisson}(\lambda)$.

6.2 **Sum of Poissons**
Let $X_1 \sim \text{Poisson}(\lambda_1)$. Let $X_2 \sim \text{Poisson}(\lambda_2)$. Suppose $X_1 \perp X_2$. Let $Y = X_1 + X_2$. How is Y distributed? Prove it using z-transforms. Note that the parameter for the Poisson denotes its mean.

6.3 **Moments of Binomial**
Use z-transforms to derive $\mathbf{E}[X(X-1)(X-2)\cdots(X-k+1)]$, for $k = 1, 2, 3, \ldots$, where $X \sim \text{Binomial}(n, p)$.

6.4 **Sums of Binomials**
Suppose that $X \sim \text{Binomial}(n, p)$ and $Y \sim \text{Binomial}(n, q)$ and $X \perp Y$. Let $Z = X + Y$. Can we say that $Z \sim \text{Binomial}(n, p+q)$? If so, prove it via z-transforms. If not, explain why not.

6.5 **z-Transform of linear combination**
Suppose X and Y are independent random variables. What is the z-transform of $aX + bY$, where a and b are arbitrary integers. Express your answer as a function of $\widehat{X}(\cdot)$ and $\widehat{Y}(\cdot)$.

6.6 **Scaling up random variables via transforms**
Let X_1, X_2, and X_3 be i.i.d. random variables, all with distribution X. Let $S = X_1 + X_2 + X_3$ and let $Y = 3X$. Suppose we are told that the z-transform of X is some function $g_X(z)$. What can you say about the z-transform of S? What can you say about the z-transform of Y? Express both of these in terms of $g_X(\cdot)$.

6.7 **Matching random variables and their z-transforms**
Assume that X, Y, Z, and T are independent random variables, where:
- $X, Y, Z \sim \text{Binomial}(3, 0.5)$
- $T \sim \text{Bernoulli}(0.5)$

Match each expression on the left to its z-transform on the right.

1._______ $3T$	a. $2^{-9} \cdot (z+1)^9$
2._______ $X + 3$	b. $2^{-3} \cdot (z+1)^3 \cdot z^3$
3._______ $X + Y + Z$	c. $2^{-4} \cdot (z+1)^3 + 0.5$
4._______ $X \cdot T$	d. $0.5 \cdot (z^3 + 1)$

6.8 **Difference transform**
Let $W = X - Y$, where $X \perp Y$. Which of the following represents $\widehat{W}(z)$:

(a) $\widehat{X}(z) \cdot \widehat{Y}(z)$
(b) $\frac{\widehat{X}(z)}{\widehat{Y}(z)}$
(c) $\widehat{X}(z) \cdot \widehat{Y}\left(\frac{1}{z}\right)$
(d) $\widehat{X}(z) - \widehat{Y}(z)$
(e) None of the above.
Justify your answer.

6.9 **Trip time**
A professor walks all the way down Carnegie St. (Figure 6.3). When she reaches the end, with probability 0.5, she turns down Mellon St. and walks all the way to the end. Otherwise, with probability 0.5 she turns down University St. and walks all the way to the end. We are interested in the professor's total trip time, T.
- Let C be a r.v. denoting the time to walk down Carnegie St.
- Let M be a r.v. denoting the time to walk down Mellon St.
- Let U be a r.v. denoting the time to walk down University St.

Assume that you are given $\mathbf{E}[C]$, $\mathbf{E}\left[C^2\right]$, $\mathbf{Var}(C)$, $\widehat{C}(z)$. You are also given these expressions for M and U. Assume that C, M, and U are independent.
(a) Express $\mathbf{E}[T]$ in terms of the given quantities (and constants).
(b) Express $\mathbf{Var}(T)$ in terms of the given quantities (and constants).
(c) Express $\widehat{T}(z)$ in terms of the given quantities (and constants).
You do *not* have to simplify your answers.

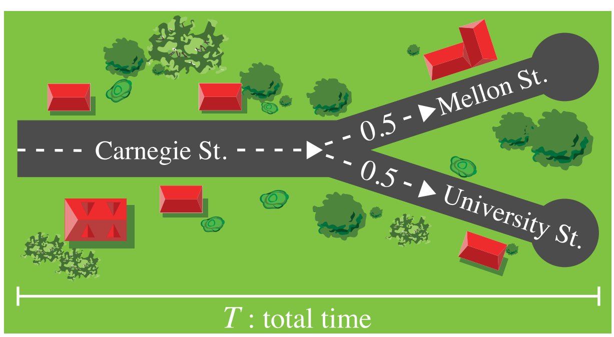

Figure 6.3 *Figure for Exercise 6.9.*

6.10 **Sum of a random number of random variables**
Suppose that $X_1, X_2, \ldots$ are i.i.d. discrete random variables, all distributed as X. Suppose that N is a positive integer-valued discrete r.v., where $N \perp X_i$ for all i. Let

$$S = \sum_{i=1}^{N} X_i.$$

(a) Prove that $\widehat{S}(z) = \widehat{N}\left(\widehat{X}(z)\right)$. [Hint: Condition on N.]
(b) Suppose that each day that the sun shines, I earn 10 dollars with probability $p = \frac{1}{3}$ and 1 dollar with probability $p = \frac{2}{3}$. The sun shines every day with probability $q = \frac{4}{5}$. Today is sunny. Let S denote the total money I earn starting today until it turns cloudy.
  (i) Write an expression for $\widehat{S}(z)$ using part (a).
  (ii) Differentiate your z-transform to get $\mathbf{E}[S]$ and $\mathbf{Var}(S)$.

6.11 **Geometric number of Geometrics**
Suppose that $X_1, X_2, \ldots$ are i.i.d. discrete random variables, all with distribution Geometric(q). Suppose that $N \sim$ Geometric(p), where $N \perp X_i$ for all i. Let

$$Y = \sum_{i=1}^{N} X_i.$$

Derive the z-transform $\widehat{Y}(z)$. What does the transform say about the distribution of Y? Provide some intuition for the result.

6.12 **Mouse in maze with transforms**
A mouse is trapped in a maze. Initially it has to choose one of two directions. If it goes to the right, then it will wander around in the maze for three minutes and will then return to its initial position. If it goes to the left, then with probability $\frac{1}{3}$ it will depart the maze after two minutes of traveling, and with probability $\frac{2}{3}$ it will return to its initial position after five minutes of traveling. Assume that the mouse is at all times equally likely to go to the left or the right. Let T denote the number of minutes that it will be trapped in the maze. In Exercise 5.26 we computed $\mathbf{E}[T]$ and $\mathbf{Var}(T)$. This time compute $\widehat{T}(z)$, and then differentiate it to get $\mathbf{E}[T]$.

6.13 **The wandering frog**
[Proposed by Tianxin Xu] There are three lily pads, A, B, and C. A frog sits on lily pad A. At each time step, the frog has an equal probability of jumping from the lily pad that it is currently on to either of the other pads.
(a) What is the expected number of hops before the frog returns to pad A?
(b) What is the z-transform of the number of hops before the frog returns to A?
(c) What is the probability that the frog is on lily pad A after n hops? Check your answer by thinking about the case where $n \to \infty$.

6.14 **Getting distribution from the transform**
The transform of a r.v. captures all moments of the r.v., but does it also capture the distribution? The answer is yes! You are given the z-transform,

$\widehat{X}(z)$, of a non-negative, discrete, integer-valued r.v., X. Provide an algorithm for extracting the p.m.f. of X from $\widehat{X}(z)$.

6.15 **Using z-transforms to solve recurrences**
This problem will walk you through the process of solving a recurrence relation:

$$a_{n+1} = 2a_n + 3, \qquad a_0 = 1.$$

(a) Define $A(z) = \sum_{n=0}^{\infty} a_n z^n$ to be the z-transform of the sequence of a_n's. Multiply every term of the recurrence relation by z^{n+1} and sum over all n to obtain an expression for $A(z)$ in terms of $A(z)$'s. You should get:

$$A(z) = \frac{1+2z}{(1-2z)(1-z)}. \tag{6.11}$$

(b) Apply partial fractions to determine the constants v and w that allow you to break up (6.11) into simpler terms:

$$A(z) = \frac{1+2z}{(1-2z)(1-z)} = \frac{v}{1-2z} + \frac{w}{1-z}.$$

(c) Recall from Section 1.1 how we can express $\frac{1}{1-z}$ and $\frac{1}{1-2z}$ as power series in z. Use these, and the correct values of v and w, to express $A(z)$ as a power series in z.
(d) Determine a_n by looking at the coefficient of z^n in your power series.

6.16 **Polygon triangulation**
In this problem, we are interested in the number of triangulations of an n-sided polygon. Figure 6.4 shows all possible triangulations for $n = 3, 4, 5$. Let

$$a_n = \text{number of triangulations of an } n+1\text{-sided polygon},$$

where $n \geq 2$. Our goal is to derive a clean expression for a_n. We will use z-transforms, where

$$A(z) = \sum_{n=0}^{\infty} a_n z^n.$$

Follow these steps:
(a) Argue that (6.12) holds for $n \geq 2$:

$$a_n = \sum_{k=0}^{n} a_k a_{n-k}, \qquad n \geq 2, \tag{6.12}$$

where we will set $a_0 = 0$, $a_1 = 1$, and $a_2 = 1$. This is a counting argument. Looking at Figure 6.5, first start by assuming that the pink triangle $(1, k+1, n+1)$ is included in your triangulation. Count the number of ways to triangulate, given that constraint. Now consider all possibilities for the $k+1$ endpoint of the included pink triangle.

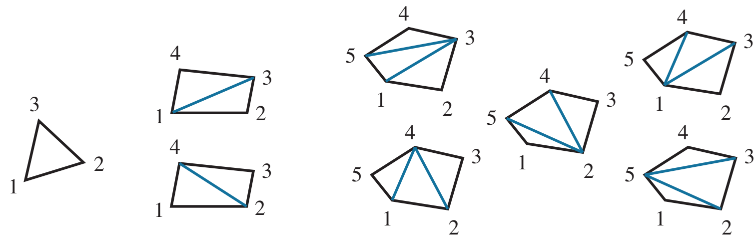

Figure 6.4 *Figure for Exercise 6.16, showing all possible triangulations of n-sided polygons.*

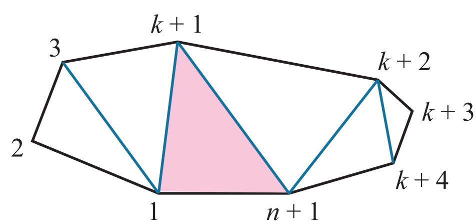

Figure 6.5 *For Exercise 6.16: counting triangulations in an $n + 1$-sided polygon.*

(b) Using (6.12), argue that

$$A(z) = z + (A(z))^2 . \tag{6.13}$$

The first few steps of the derivation are given below:

$$\begin{aligned} A(z) &= z + \sum_{n=2}^{\infty} a_n z^n \\ &= z + \sum_{n=2}^{\infty} \left(\sum_{k=0}^{n} a_k a_{n-k} \right) z^n \\ &= z + \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k a_{n-k} \right) z^n \end{aligned}$$

Explain why each of the above steps is true and then finish the derivation to get (6.13).

(c) Solve (6.13) to get $A(z)$. You will need to use $A(0) = 0$.

(d) All that remains is to express $A(z)$ as a power series of z. To do this, we are providing you with the Taylor series expansion of $\sqrt{1-4z}$ in (6.14):

$$\sqrt{1-4z} = 1 - 2 \sum_{n=1}^{\infty} \frac{1}{n} \binom{2n-2}{n-1} z^n. \tag{6.14}$$

(e) Obtain a_n.