Course: M362K Intro to Stochastic Processes
Term: Fall 2014
Instructor: Gordan Zitkovic

Lecture 5
Generating Functions

The path-counting method used in the previous lecture only works for computations related to the first n steps of the random walk, where n is given in advance. We will see later that most of the interesting questions do *not* fall into this category. For example, the distribution of *the time it takes* for the random walk to hit the level $l \neq 0$ is like that. There is no way to give an a-priori bound on the number of steps it will take to get to l. To deal with a wider class of properties of random walks (and other processes), we need to develop some new mathematical tools.

DEFINITION AND FIRST PROPERTIES

As we know, the distribution of an $\mathbb{N}_0$-valued random variable X is completely determined by the sequence $\{p_k\}_{k\in\mathbb{N}_0}$ of numbers in $[0,1]$ given by

$$p_k = \mathbb{P}[X = k],\ k \in \mathbb{N}_0.$$

As a sequence of real numbers, $\{p_k\}_{k\in\mathbb{N}_0}$ can be used to construct a power series:

$$P_X(s) = \sum_{k=0}^{\infty} p_k s^k. \tag{5.1}$$

It follows from the fact that $\sum_k |p_k| \leq 1$ that the radius of convergence[1] of $\{p_k\}_{k\in\mathbb{N}_0}$ is at least equal to 1. Therefore, P_X is well defined for $s \in [-1,1]$, and, perhaps, elsewhere, too.

Definition 5.1. The function P_X given by $P_X(s) = \sum_{k=0}^{\infty} p_k s^k$ is called the **generating function** of the random variable[2] X.

Before we proceed, let us derive expressions for the generating functions of some of the popular $\mathbb{N}_0$-valued random variables.

[1] Remember, that the **radius of convergence** of a power series $\sum_{k=0}^{\infty} a_k x^k$ is the largest number $R \in [0,\infty]$ such that $\sum_{k=0}^{\infty} a_k x^k$ converges absolutely whenever $|x| < R$.

[2] more precisely, of its pmf $\{p_k\}_{k\in\mathbb{N}_0}$.

Example 5.2.

1. **Bernoulli** ($b(p)$). Here $p_0 = q$, $p_1 = p$, and $p_k = 0$, for $k \geq 2$. Therefore,
$$P_X(s) = ps + q.$$

2. **Binomial** ($b(n,p)$). Since $p_k = \binom{n}{k} p^k q^{n-k}$, $k = 0, \dots, n$, we have
$$P_X(s) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} s^k = (ps+q)^n,$$
by the binomial theorem.

3. **Geometric** ($g(p)$). For $k \in \mathbb{N}_0$, $p_k = q^k p$, so that
$$P_X(s) = \sum_{k=0}^{\infty} q^k s^k p = p \sum_{k=0}^{\infty} (qs)^k = \frac{p}{1-qs}.$$

4. **Poisson** ($p(\lambda)$). Given that $p_k = e^{-\lambda} \frac{\lambda^k}{k!}$, $k \in \mathbb{N}_0$, we have
$$P_X(s) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} s^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(s\lambda)^k}{k!} = e^{-\lambda} e^{s\lambda} = e^{\lambda(s-1)}.$$

Some of the most useful analytic properties of P_X are listed in the following proposition

Proposition 5.3. *Let X be an $\mathbb{N}_0$-valued random variable, $\{p_k\}_{k \in \mathbb{N}_0}$ its pmf, and P_X be its generating function. Then*

1. *$P_X(s) = \mathbb{E}[s^X]$, $s \in [-1,1]$,*
2. *$P_X(s)$ is convex and non-decreasing with $0 \leq P_X(s) \leq 1$ for $s \in [0,1]$*
3. *$P_X(s)$ is infinitely differentiable on $(-1,1)$ with*
$$\frac{d^n}{ds^n} P_X(s) = \sum_{k=n}^{\infty} k(k-1)\dots(k-n+1) s^{k-n} p_k, \ n \in \mathbb{N}. \tag{5.2}$$
In particular,
$$p_n = \frac{1}{n!} \frac{d^n}{ds^n} P_X(s) \Big|_{s=0}$$
and so $s \mapsto P_X(s)$ uniquely determines the sequence $\{p_k\}_{k \in \mathbb{N}_0}$.

Proof. Statement 1. follows directly from the formula
$$\mathbb{E}[g(X)] = \sum_{k=0}^{\infty} g(k) p_k,$$
applied to $g(x) = s^x$. As far as (3) is concerned, we only note that the expression (5.2) is exactly what you would get if you differentiated

the expression (5.1) term by term. The rigorous proof of the fact this is allowed is beyond the scope of these notes. With 3. at our disposal, 2. follows by the fact that the first two derivatives of the function P_X are non-negative and that $P_X(1) = 1$. □

Remark 5.4.

1. If you know about moment-generating functions, you will notice that $P_X(s) = M_X(\log(s))$, for $s \in (0,1)$, where $M_X(\lambda) = \mathbb{E}[\exp(\lambda X)]$ is the moment-generating function of X.

2. Generating functions can be used with sequences $\{a_k\}_{k\in\mathbb{N}_0}$ which are not necessarily pmf's of random variables. The method is useful for any sequence $\{a_k\}_{k\in\mathbb{N}_0}$ such that the power series $\sum_{k=0}^{\infty} a_k s^k$ has a positive (non-zero) radius of convergence.

3. The name *generating function* comes from the last part of the property (3). The knowledge of P_X implies the knowledge of the whole sequence $\{p_k\}_{k\in\mathbb{N}_0}$. Put differently, P_X generates the whole distribution of X.

Remark 5.5. Note that the true radius of convergence varies from distribution to distribution. It is infinite in 1. , 2. and 4., and equal to $1/q > 1$ in 4. in Example 5.2. For the distribution with pmf given by $p_k = \frac{C}{(k+1)^2}$, where $C = (\sum_{k=0}^{\infty} \frac{1}{(k+1)^2})^{-1}$, the radius of convergence is exactly equal to 1. Can you see why?

CONVOLUTION AND MOMENTS

The true power of generating functions comes from the fact that they behave very well under the usual operations in probability.

Definition 5.6. Let $\{p_k\}_{k\in\mathbb{N}_0}$ and $\{q_k\}_{k\in\mathbb{N}_0}$ be two probability-mass functions. The **convolution** $p * q$ of $\{p_k\}_{k\in\mathbb{N}_0}$ and $\{q_k\}_{k\in\mathbb{N}_0}$ is the sequence $\{r_k\}_{k\in\mathbb{N}_0}$, where

$$r_n = \sum_{k=0}^{n} p_k q_{n-k}, n \in \mathbb{N}_0.$$

This abstractly-defined operation will become much clearer once we prove the following proposition:

Proposition 5.7. *Let X, Y be two* independent *$\mathbb{N}_0$-valued random variables with pmfs $\{p_k\}_{k\in\mathbb{N}_0}$ and $\{q_k\}_{k\in\mathbb{N}_0}$. Then the sum $Z = X + Y$ is also $\mathbb{N}_0$-valued and its pmf is the convolution of $\{p_k\}_{k\in\mathbb{N}_0}$ and $\{q_k\}_{k\in\mathbb{N}_0}$ in the sense of Definition 5.6.*

Proof. Clearly, Z is $\mathbb{N}_0$-valued. To obtain an expression for its pmf, we use the law of total probability:

$$\mathbb{P}[Z = n] = \sum_{k=0}^{n} \mathbb{P}[X = k]\mathbb{P}[Z = n | X = k].$$

On the other hand,

$$\begin{aligned}\mathbb{P}[Z = n | X = k] &= \mathbb{P}[X + Y = n | X = k] = \mathbb{P}[Y = n - k | X = k] \\ &= \mathbb{P}[Y = n - k],\end{aligned}$$

where the last equality follows from independence of X and Y. Therefore,

$$\mathbb{P}[Z = n] = \sum_{k=0}^{n} \mathbb{P}[X = k]\mathbb{P}[Y = n - k] = \sum_{k=0}^{n} p_k q_{n-k}. \qquad \square$$

Corollary 5.8. *Let $\{p_k\}_{k\in\mathbb{N}_0}$ and $\{p_k\}_{k\in\mathbb{N}_0}$ be any two pmfs.*

1. *Convolution is commutative, i.e., $p * q = q * p$.*
2. *The convolution $r = p * q$ of two pmfs is a pmf, i.e. $r_k \geq 0$, for all $k \in \mathbb{N}_0$ and $\sum_{k=0}^{\infty} r_k = 1$.*

Corollary 5.9. *Let $\{p_k\}_{k\in\mathbb{N}_0}$ and $\{p_k\}_{k\in\mathbb{N}_0}$ be any two pmfs, and let*

$$P(s) = \sum_{k=0}^{\infty} p_k s^k \text{ and } Q(s) = \sum_{k=0}^{\infty} q_k s^k$$

be their generating functions. Then the generating function $R(s) = \sum_{k=0}^{\infty} r_k s^k$, of the convolution $r = p * q$ is given by

$$R(s) = P(s)Q(s).$$

Equivalently, the generating function P_{X+Y} of the sum of two independent $\mathbb{N}_0$-valued random variables is equal to the product

$$P_{X+Y}(s) = P_X(s)P_Y(s),$$

of the generating functions P_X and P_Y of X and Y.

Example 5.10.

1. The binomial $b(n, p)$ distribution is a sum of n independent Bernoullis $b(p)$. Therefore, if we apply Corrolary 5.9 n times to the generating function $(q + ps)$ of the Bernoulli $b(p)$ distribution we immediately get that the generating function of the binomial is

$$(q + ps) \dots (q + ps) = (q + ps)^n.$$

2. More generally, we can show that the sum of m *independent* random variables with the $b(n, p)$ distribution has a binomial distribution $b(mn, p)$. If you try to sum binomials with different values of the parameter p you will not get a binomial.

3. What is even more interesting, the following statement can be shown: Suppose that the sum Z of two independent $\mathbb{N}_0$-valued random variables X and Y is binomially distributed with parameters n and p. Then both X and Y are binomial with parameters n_X, p and n_y, p where $n_X + n_Y = n$. In other words, the only way to get a binomial as a sum of independent random variables is the trivial one.

Another useful thing about generating functions is that they make the computation of moments easier.

Proposition 5.11. *Let $\{p_k\}_{k\in\mathbb{N}_0}$ be a pmf of an $\mathbb{N}_0$-valued random variable X and let $P(s)$ be its generating function. For $n \in \mathbb{N}$ the following two statements are equivalent*

1. $\mathbb{E}[X^n] < \infty$,
2. $\frac{d^n P(s)}{ds^n}\Big|_{s=1}$ *exists (in the sense that the left limit* $\lim_{s\nearrow 1} \frac{d^n P(s)}{ds^n}$ *exists)*

In either case, we have

$$\mathbb{E}[X(X-1)(X-2)\dots(X-n+1)] = \frac{d^n}{ds^n}P(s)\Big|_{s=1}.$$

The quantities

$$\mathbb{E}[X], \quad \mathbb{E}[X(X-1)], \quad \mathbb{E}[X(X-1)(X-2)], \dots$$

are called **factorial moments** of the random variable X. You can get the classical moments from the factorial moments by solving a system of linear equations. It is very simple for the first few:

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X], \\
\mathbb{E}[X^2] &= \mathbb{E}[X(X-1)] + \mathbb{E}[X], \\
\mathbb{E}[X^3] &= \mathbb{E}[X(X-1)(X-2)]] + 3\mathbb{E}[X(X-1)] + \mathbb{E}[X], \dots
\end{aligned}$$

A useful identity which follows directly from the above results is the following:

$$\mathrm{Var}[X] = P''(1) + P'(1) - (P'(1))^2,$$

and it is valid if the first two derivatives of P at 1 exist.

Example 5.12. Let X be a Poisson random variable with parameter λ. Its generating function is given by

$$P_X(s) = e^{\lambda(s-1)}.$$

Therefore, $\frac{d^n}{ds^n} P_X(1) = \lambda^n$, and so, the sequence $(\mathbb{E}[X], \mathbb{E}[X(X-1)], \mathbb{E}[X(X-1)(X-2)], \dots)$ of factorial moments of X is just $(\lambda, \lambda^2, \lambda^3, \dots)$. It follows that

$$\begin{aligned}
\mathbb{E}[X] &= \lambda, \\
\mathbb{E}[X^2] &= \lambda^2 + \lambda, \ \mathrm{Var}[X] = \lambda \\
\mathbb{E}[X^3] &= \lambda^3 + 3\lambda^2 + \lambda, \dots
\end{aligned}$$

RANDOM SUMS

Our next application of generating functions in the theory of stochastic processes deals with the so-called *random sums*. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a sequence of random variables, and let N be a random time (a **random time** is simply an $\mathbb{N}_0 \cup \{+\infty\}$-value random variable). We can define the random variable

$$Y = \sum_{k=0}^{N} \xi_k$$

for $\omega \in \Omega$ by

$$Y(\omega) = \begin{cases} 0, & N(\omega) = 0, \\ \sum_{k=1}^{N(\omega)} \xi_k(\omega), & N(\omega) \geq 1. \end{cases}$$

More generally, for an arbitrary stochastic process $\{X_k\}_{k\in\mathbb{N}_0}$ and a random time N (with $\mathbb{P}[N = +\infty] = 0$), we define the *random variable* X_N by $X_N(\omega) = X_{N(\omega)}(\omega)$, for $\omega \in \Omega$. When N is a constant ($N = n$), then X_N is simply equal to X_n. In general, think of X_N as a value of the stochasti process X taken at the time which is itself random. If $X_n = \sum_{k=1}^{n} \xi_k$, then $X_N = \sum_{k=1}^{N} \xi_k$.

Example 5.13. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be the increments of a symmetric simple random walk (coin-tosses), and let N have the following distribution

$$N \sim \begin{pmatrix} 0 & 1 & 2 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

which is *independent* of $\{\xi_n\}_{n\in\mathbb{N}}$ (it is very important to specify the dependence structure between N and $\{\xi_n\}_{n\in\mathbb{N}}$ in this setting!). Let us compute the distribution of $Y = \sum_{k=0}^{N} \xi_k$ in this case. This is where we, typically, use the formula of total probability:

$$\begin{aligned}
\mathbb{P}[Y = m] &= \mathbb{P}[Y = m | N = 0]\, \mathbb{P}[N = 0] \\
&\quad + \mathbb{P}[Y = m | N = 1]\, \mathbb{P}[N = 1] \\
&\quad + \mathbb{P}[Y = m | N = 2]\, \mathbb{P}[N = 2] \\
&= \mathbb{P}[\sum_{k=0}^{N} \xi_k = m | N = 0]\, \mathbb{P}[N = 0] \\
&\quad + \mathbb{P}[\sum_{k=0}^{N} \xi_k = m | N = 1]\, \mathbb{P}[N = 1] \\
&\quad + \mathbb{P}[\sum_{k=0}^{N} \xi_k = m | N = 2] \mathbb{P}[N = 2] \\
&= \frac{1}{3} \left(\mathbb{P}[0 = m] + \mathbb{P}[\xi_1 = m] + \mathbb{P}[\xi_1 + \xi_2 = m] \right).
\end{aligned}$$

When $m = 1$ (for example), we get

$$\mathbb{P}[Y = 1] = \frac{0 + \frac{1}{2} + 0}{3} = 1/6.$$

Perform the computation for some other values of m for yourself.

What happens when N and $\{\xi_n\}_{n \in \mathbb{N}}$ are dependent? This will usually be the case in practice, as the value of the time N when we stop adding increments will typically depend on the behaviour of the sum itself.

Example 5.14. Let $\{\xi_n\}_{n \in \mathbb{N}}$ be as above - we can think of a situation where a gambler is repeatedly playing the same game in which a coin is tossed and the gambler wins a dollar if the outcome is *heads* and looses a dollar otherwise. A "smart" gambler enters the game and decides on the following tactic: *Let's see how the first game goes. If I lose, I'll play another 2 games and hopefully cover my losses, and If I win, I'll quit then and there.* The described strategy amounts to the choice of the random time N as follows:

$$N = \begin{cases} 1, & \xi_1 = 1, \\ 3, & \xi_1 = -1. \end{cases}$$

Then

$$Y = \begin{cases} 1, & \xi_1 = -1, \\ -1 + \xi_2 + \xi_3, & \xi_1 = 1. \end{cases}$$

Therefore,

$$\begin{aligned}
\mathbb{P}[Y = 1] &= \mathbb{P}[Y = 1 | \xi_1 = 1] \mathbb{P}[\xi_1 = 1] + \mathbb{P}[Y = 1 | \xi_1 = -1] \mathbb{P}[\xi_1 = -1] \\
&= 1 \cdot \mathbb{P}[\xi_1 = 1] + \mathbb{P}[\xi_2 + \xi_3 = 2] \mathbb{P}[\xi_1 = -1] \\
&= \tfrac{1}{2}(1 + \tfrac{1}{4}) = \tfrac{5}{8}.
\end{aligned}$$

Similarly, we get $\mathbb{P}[Y=-1]=\frac{1}{4}$ and $\mathbb{P}[Y=-3]=\frac{1}{8}$. The expectation $\mathbb{E}[Y]$ is equal [3] to $1\cdot\frac{5}{8}+(-1)\cdot\frac{1}{4}+(-3)\cdot\frac{1}{8}=0$.

[3] This is not an accident. One of the first powerful results of the beautiful *martingale theory* states that no matter how smart a strategy you employ, you cannot beat a fair gamble.

We will return to the general (non-independent) case in the next lecture. Let us use generating functions to give a full description of the distribution of $Y=\sum_{k=0}^{N}\xi_k$ in this case.

Proposition 5.15. *Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a sequence of independent $\mathbb{N}_0$-valued random variables, all of which share the same distribution with pmf $\{p_k\}_{k\in\mathbb{N}_0}$ and generating function $P_\xi(s)$. Let N be a random time independent of $\{\xi_n\}_{n\in\mathbb{N}}$. Then the generating function P_Y of the random sum $Y=\sum_{k=0}^{N}\xi_k$ is given by*

$$P_Y(s)=P_N(P_\xi(s)).$$

Proof. We use the idea from Example 5.13 and condition on possible values of N. We also use the following fact (Tonelli's theorem) without proof:

$$\text{If } a_{ij}\geq 0, \text{ for all } i,j, \text{ then } \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}a_{ij}=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}a_{ij}. \tag{5.3}$$

$$\begin{aligned}
P_Y(s)&=\sum_{k=0}^{\infty}s^k\mathbb{P}[Y=k]\\
&=\sum_{k=0}^{\infty}s^k\Big(\sum_{i=0}^{\infty}\mathbb{P}[Y=k|N=i]\mathbb{P}[N=i]\Big)\\
&=\sum_{k=0}^{\infty}s^k\Big(\sum_{i=0}^{\infty}\mathbb{P}[\sum_{j=0}^{i}\xi_j=k]\mathbb{P}[N=i]\Big) &&\text{(by independence)}\\
&=\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}s^k\mathbb{P}[\sum_{j=0}^{i}\xi_j=k]\mathbb{P}[N=i] &&\text{(by Tonelli)}\\
&=\sum_{i=0}^{\infty}\mathbb{P}[N=i]\sum_{k=0}^{\infty}s^k\mathbb{P}[\sum_{j=0}^{i}\xi_j=k] &&\text{(by (5.3))}
\end{aligned}$$

By (iteration of) Corollary 5.9, we know that the generating function of the random variable $\sum_{j=0}^{i}\xi_j$ - which is exactly what the second sum above represents - is $(P_\xi(s))^i$. Therefore, the chain of equalities above can be continued as

$$\begin{aligned}
&=\sum_{i=0}^{\infty}\mathbb{P}[N=i](P_\xi(s))^i\\
&=P_N(P_\xi(s)).
\end{aligned}$$

$\square$

Corollary 5.16 (Wald's Identity I)**.** *Let $\{\xi_n\}_{n\in\mathbb{N}}$ and N be as in Proposition 5.15. Suppose, also, that $\mathbb{E}[N] < \infty$ and $\mathbb{E}[\xi_1] < \infty$. Then*

$$\mathbb{E}[\sum_{k=0}^{N} \xi_k] = \mathbb{E}[N]\,\mathbb{E}[\xi_1].$$

Proof. We just apply the composition rule for derivatives to the equality $P_Y = P_N \circ P_\xi$ to get

$$P_Y'(s) = P_N'(P_\xi(s))P_\xi'(s).$$

After we let $s \nearrow 1$, we get

$$\mathbb{E}[Y] = P_Y'(1) = P_N'(P_\xi(1))P_\xi'(1) = P_N'(1)P_\xi'(1) = \mathbb{E}[N]\,\mathbb{E}[\xi_1]. \qquad \square$$

Example 5.17. Every time Springfield Wildcats play in the Superbowl, their chance of winning is $p \in (0,1)$. The number of years between two Superbowls they get to play in has the Poisson distribution $p(\lambda)$, $\lambda > 0$. What is the expected number of years Y between the consequtive Superbowl wins?

Let $\{\xi_n\}_{n\in\mathbb{N}}$ be the sequence of independent $p(\lambda)$-random variables modeling the number of years between consecutive Superbowl appearances by the Wildcat. Moreover, let N be a geometric $g(p)$ random variable with success probability p. Then

$$Y = \sum_{k=0}^{N} \xi_k.$$

Indeed, every time the Wildcats lose the Superbowl, another $\xi_\cdot$ years have to pass before they get another chance and the whole thing stops when they finally win. To compute the expectation of Y we use Corollary 5.16

$$\mathbb{E}[Y] = \mathbb{E}[N]\mathbb{E}[\xi_k] = \frac{1-p}{p}\lambda.$$

PROBLEMS

Problem 5.1. If $P(s)$ is the generating function of the random variable X, then the generating function of $2X+1$ is

(a) $P(2s+1)$

(b) $2P(s)+1$

(c) $P(s^2+1)$

(d) $sP(s^2)$

(e) none of the above

Solution: The correct answer is (d): using the representation $P_Y(s) = \mathbb{E}[s^Y]$, we get $P_{2X+1} = \mathbb{E}[s^{2X+1}] = \mathbb{E}[s(s^{2X})] = s\mathbb{E}[(s^2)^X] = sP_X(s^2)$. Alternatively, we can use the fact that

$$\mathbb{P}[2X+1=k] = \begin{cases} 0, & k \in \{0,2,4,\dots\}, \\ p_{\frac{k-1}{2}}, & k \in \{1,3,5,\dots\}, \end{cases}$$

where $p_n = \mathbb{P}[X=n]$, $n \in \mathbb{N}_0$, so that

$$P_{2X+1}(s) = 0 + p_0 s + 0s^2 + p_1 s^2 + 0s^3 + p_2 s^4 + \dots = sP_X(s^2).$$

Problem 5.2. Let X be an $\mathbb{N}_0$-valued random variable, and let $P(s)$ be its generating function. Then the generating function of $4X$ is given by

1. $(P(s))^4$,
2. $P(P(P(P(s))))$,
3. $P(s^4)$,
4. $4P(s)$,
5. none of the above.

Solution: 3.

Problem 5.3. Let X and Y be two N_0-valued random variables, let $P_X(s)$ and $P_Y(s)$ be their generating functions and let $Z = X - Y$, $V = X + Y$, $W = XY$. Then

1. $P_X(s) = P_Z(s)P_Y(s)$,
2. $P_X(s)P_Y(s) = P_Z(s)$,
3. $P_W(s) = P_X(P_Y(s))$,
4. $P_Z(s)P_V(s) = P_X(s)P_Y(s)$,
5. none of the above.

Solution: 5.

Problem 5.4. Let X be an $\mathbb{N}_0$-valued random variable and $P(s)$ its generating function. If $Q(s) = P(s)/(1-s)$, then

1. $Q(s)$ is a generating function of a random variable,
2. $Q(s)$ is a generating function of a non-decreasing sequence of non-negative numbers,

3. $Q(s)$ is a concave function on $(0,1)$,

4. $Q(0) = 1$,

5. none of the above.

Solution: 2.

Problem 5.5. The generating function of the $\mathbb{N}_0$-valued random variable X is given by

$$P_X(s) = \frac{s}{1+\sqrt{1-s^2}}.$$

1. Compute $p_0 = \mathbb{P}[X=0]$.

2. Compute $p_1 = \mathbb{P}[X=1]$.

3. Does $\mathbb{E}[X]$ exist? If so, find its value; if not, explain why not.

Solution:

1. $p_0 = P_X(0)$, so $p_0 = 0$,

2. $p_1 = P_X'(0)$, and

$$P_X'(s) = \frac{1}{\sqrt{1-s^2}(1+\sqrt{1-s^2})},$$

so $p_1 = \frac{1}{2}$.

3. If $\mathbb{E}[X]$ existed, it would be equal to $P_X'(1)$. However,

$$\lim_{s\nearrow 1} P_X'(s) = +\infty,$$

so $\mathbb{E}[X]$ (and, equivalently, $P_X'(1)$) does not exist.

Problem 5.6. Let $P(s)$ be the generating function of the sequence $(p_0, p_1, \dots)$ and $Q(s)$ the generating function of the sequence $(q_0, q_1, \dots)$. If the sequence $\{r_n\}_{n\in\mathbb{N}_0}$ is defined by

$$r_n = \begin{cases} 0, & n \leq 1 \\ \sum_{k=1}^{n-1} p_k q_{n-1-k}, & n > 1, \end{cases}$$

then its generating function is given by (*Note:* Careful! $\{r_n\}_{n\in\mathbb{N}_0}$ is not exactly the convolution of $\{p_n\}_{n\in\mathbb{N}_0}$ and $\{q_n\}_{n\in\mathbb{N}_0}$.)

(a) $P(s)Q(s) - p_0 q_0$

(b) $(P(s) - p_0)(Q(s) - q_0)$

(c) $\frac{1}{s}(P(s) - p_0)Q(s)$

(d) $\frac{1}{s}P(s)(Q(s) - q_0)$

(e) $s(P(s) - p_0)Q(s)$

Solution: The elements of $\{r_n\}_{n\in\mathbb{N}_0}$ are

$r_0 = 0$, $r_1 = 0$, $r_2 = p_1q_0$, $r_3 = p_1q_1 + p_2q_0$, $r_4 = p_1q_2 + p_2q_1 + p_3q_0, \dots$

If it started from $n = 2$ it would be exactly the convolution of the sequences

$$(p_1, p_2, \dots) \text{ and } (q_0, q_1, \dots)$$

The generating function of the first one is

$$\tfrac{1}{s}(P(s) - p_0),$$

while the generating function of the second one is $Q(s)$. Therefore the generating function of the sequence

$$(r_2, r_3, \dots)$$

is $\frac{1}{s}(P(s) - p_0)Q(s)$. To "put two zeros" in front of everything, we need to multiply it by s^2, so

$$R(s) = s(P(s) - p_0)Q(s),$$

and the correct answer is (e).

Problem 5.7. Let N be a random time, independent of $\{\xi_n\}_{n\in\mathbb{N}_0}$, where $\{\xi_n\}_{n\in\mathbb{N}_0}$ is a sequence of mutually independent Bernoulli ($\{0,1\}$-valued) random variables with parameter $p_B \in (0,1)$. Suppose that N has a geometric distribution $g(p_g)$ with parameter $p_g \in (0,1)$. Compute the distribution of the random sum

$$Y = \sum_{k=1}^{N} \xi_k,$$

i.e., find $\mathbb{P}[Y = i]$, for $i \in \mathbb{N}_0$. (*Note:* You can think of Y as a binomial random variable with "random n".)

Solution: Independence between N and $\{\xi_n\}_{n\in\mathbb{N}}$ allows us a to use a result from class which states that the generating function $P_Y(s)$ of Y is given by

$$P_Y(s) = P_N(P_B(s)),$$

where $P_N(s) = \frac{p_g}{1-q_g s}$ is the generating function of N (geometric distribution) and $P_B(s) = q_B + p_B s$ is the generating function of each ξ_k (Bernoulli distribution). Therefore,

$$P_Y(s) = \frac{p_g}{1 - q_g(q_b + p_B s)} = \frac{\frac{p_g}{1-q_g q_B}}{1 - \frac{q_g p_B}{1-q_g q_B} s} = \frac{p_Y}{1 - q_Y s},$$

where $p_Y = \frac{p_g}{1-q_g q_B}$ and $q_Y = 1 - p_Y$. P_Y can be recognized as the generating function of a geometric random variable with parameter p_Y, so $\mathbb{P}[Y = i] = p_y(q_y)^i$.

Problem 5.8. A fair coin and a fair 6-sided die are thrown repeatedly until the the first time 6 appears on the die. Let X be the number of *heads* obtained (we are including the *heads* that may have occurred together with the first 6) in the count. The generating function of X is

(a) $\frac{1}{2}\frac{s}{6-5s} + \frac{1}{2}$

(b) $\frac{s}{6-5s}$

(c) $\frac{1+s}{7-5s}$

(d) $\frac{1+2s}{7-4s}$

(e) None of the above

Solution: X can be seen as a random sum of $\frac{1}{2}$-Bernoulli random variables with the number of summands given by G, the number of tries it takes to get the first heads. G is clearly geometrically distributed with parameter $\frac{1}{6}$. The generating function of G is $P_G(s) = \frac{s}{6-5s}$, so the generating function of the whole random sum X is

$$P_X(s) = P_G((\tfrac{1}{2} + \tfrac{1}{2}s)) = \frac{1+s}{7-5s}.$$

Problem 5.9. Six fair gold coins are tossed, and the total number of *tails* is recorded; let's call this number N. Then, a set of three fair silver coins is tossed N times. Let X be the total number of times at least two *heads* are observed (among the N tosses of the set of silver coins).

(*Note:* A typical outcome of such a procedure would be the following: out of the six gold coins 4 were *tails* and 2 were *heads*. Therefore $N = 4$ and the 4 tosses of the set of three silver coins may look something like $\{HHT, THT, TTT, HTH\}$, so that $X = 2$ in this state of the world.)

Find the generating function *and* the pmf of X. You don't have to evaluate binomial coefficients.

Solution: Let H_k, $k \in \mathbb{N}$, be the number of heads on the k^{th} toss of the set of three silver coins. The distribution of H_k is binomial so $\mathbb{P}[H_k \geq 2] = \mathbb{P}[H_k = 2] + \mathbb{P}[H_k = 3] = 3/8 + 1/8 = 1/2$. Let ξ_k be the indicator

$$\xi_k = \mathbf{1}_{\{H_k \geq 2\}} = \begin{cases} 1, & H_k \geq 2, \\ 0, & H_k < 2. \end{cases}$$

The generating function $P_\xi(s)$ of each ξ_k is a generating function of a Bernoulli random variable with $p = \frac{1}{2}$, i.e.

$$P_\xi(s) = \tfrac{1}{2}(1+s).$$

The random variable N is has a binomial distribution with parameters $n = 6$ and $p = \frac{1}{2}$. Therefore,

$$P_N(s) = (\tfrac{1}{2} + \tfrac{1}{2}s)^6.$$

By a theorem from the notes, the generating function of the value of the random sum X is given by

$$P_X(s) = P_N(P_\xi(s)) = \left(\tfrac{1}{2} + \tfrac{1}{2}(\tfrac{1}{2} + \tfrac{1}{2}s)\right)^6 = (\tfrac{3}{4} + \tfrac{1}{4}s)^6.$$

Therefore, X is a binomial random variable with parameters $n = 6$ and $p = 1/4$, i.e.

$$p_k = \mathbb{P}[X = k] = \binom{6}{k}\frac{3^{6-k}}{4^6},\ k = 0, \dots, 6.$$

Problem 5.10. Tony Soprano collects his cut from the local garbage management companies. During a typical day he can visit a geometrically distributed number of companies with parameter $p = 0.1$. According to many years' worth of statistics gathered by his consigliere Silvio Dante, the amount he collects from the i^{th} company is random with the following distribution

$$X_i \sim \begin{pmatrix} \$1000 & \$2000 & \$3000 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

The amounts collected from different companies are independent of each other, and of the number of companies visited.

1. Find the (generating function of) the distribution of the amount of money S that Tony will collect on a given day.

2. Compute $\mathbb{E}[S]$ and $\mathbb{P}[S > 0]$.

Solution:

1. The number N of companies visited has the generating function

$$P_N(s) = \frac{0.1}{1 - 0.9s},$$

and the generting function of the amount X_i collected from each one

$$P_X(s) = 0.2s^{1000} + 0.4s^{2000} + 0.4s^{3000}.$$

The total amount collected is given by the random sum

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

so the generating function of S is given by

$$P_S(s) = P_N(P_X(s)) = \frac{0.1}{1 - 0.9(0.2s^{1000} + 0.4s^{2000} + 0.4s^{3000})}.$$

2. The expectation of S is given by

$$\mathbb{E}[S] = P_S'(1) = P_N'(P_X(1))P_X'(1) = \mathbb{E}[N]\,\mathbb{E}[X_1] = 9 \times \$2200 = \$19800.$$

To compute $\mathbb{P}[S > 0]$, we note that $S = 0$ if and only if no collections have been made, i.e., if the value of N is equal to 0. Since N is geometric with parameter $p = 0.1$, so

$$\mathbb{P}[S > 0] = 1 - \mathbb{P}[S = 0] = 1 - \mathbb{P}[N = 0] = 1 - p = 0.9.$$

Problem 5.11 (*Optional, but fun*)**.** The **Fibonacci sequence** $\{F_n\}_{n\in\mathbb{N}_0}$ is defined recursively by

$$F_0 = F_1 = 1, F_{n+2} = F_{n+1} + F_n, n \in \mathbb{N}_0.$$

1. Find the generating function $P(s) = \sum_{k=0}^{\infty} F_k s^k$ of the sequence $\{F_n\}_{n\in\mathbb{N}_0}$.
2. Use P to derive an explicit expression for F_n, $n \in \mathbb{N}_0$. (*Hint:* use partial fractions)

Note: the purpose of this problem is to show that one can use generating functions to do other things, as well. Indeed $\{F_n\}_{n\in\mathbb{N}_0}$ is not a probability distribution, but the generating function techniques still apply.

Solution:

1. If we subtract $sP(s)$ and $s^2P(s)$ from $P(s)$, we get (note the shift in the index k in the third line)

$$\begin{aligned}
(1 - s - s^2)P(s) &= \sum_{k=0}^{\infty} F_k s^k - s\sum_{k=0}^{\infty} F_k s^k - s^2\sum_{k=0}^{\infty} F_k s^k \\
&= \sum_{k=0}^{\infty} F_k s^k - \sum_{k=0}^{\infty} F_k s^{k+1} - \sum_{k=0}^{\infty} F_k s^{k+2} \\
&= \sum_{k=0}^{\infty} F_k s^k - \sum_{k=1}^{\infty} F_{k-1} s^k - \sum_{k=2}^{\infty} F_{k-2} s^k \\
&= F_0 + F_1 s - F_0 + \sum_{k=2}^{\infty} (F_k - F_{k-1} - F_{k-2}) s^k \\
&= s
\end{aligned}$$

Therefore,

$$P(s) = \frac{s}{1 - s - s^2}.$$

2. The equation $1 - s - s^2 = 0$ has two real roots $s_{1,2}$, and we can express them both in terms of the *Golden section*

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

as $s_1 = \phi$ and $s_2 = 1 - \phi$. Therefore, a partial-fraction expansion gives us

$$P(s) = \frac{s}{1 - s - s^2} = \frac{1}{\sqrt{5}} \left(\frac{1}{1 - \phi s} - \frac{1}{1 - (1 - \phi)s} \right).$$

Both fractions can be recognized as sums of geometric series, so

$$P(s) = \sum_{k=0}^{\infty} \left(\frac{\phi^k - (1 - \phi)^k}{\sqrt{5}} \right) s^k.$$

Generating functions uniquely determine the sequence of their coefficients, so

$$F_n = \frac{\phi^n - (1 - \phi)^n}{\sqrt{5}}, \text{ for } n \in \mathbb{N}_0.$$