Supplement 1: Series fun, or some sums

Computing the mean and variance of discrete distributions often involves summing infinite series. That was the most difficult and my least favorite topic in my calculus course. Here are a few useful derivations. They aren't always clever, but they tend to follow an obvious pattern, which means that even non-clever people like me may have a hope of re-deriving them.[1] For a justification of some of the operations on infinite series of functions used, see Apostol [1, Chapter 11].

S1.1 Geometric series

You already know this series. I am including it for the sake of completeness.

Let $0 < p < 1$.

$$\sum_{k=0}^{n} p^k = \frac{1-p^{n+1}}{1-p} \tag{1}$$

$$\sum_{k=1}^{n} p^k = \frac{p-p^{n+1}}{1-p} \tag{2}$$

$$\sum_{k=0}^{\infty} p^k = \frac{1}{1-p} \tag{3}$$

$$\sum_{k=1}^{\infty} p^k = \frac{p}{1-p} \tag{4}$$

Proof: It is enough to prove (1), so let $x = 1 + p + \cdots + p^n$. Then simply expanding $(1-p)x$ yields $(1-p)x = (1-p)(1 + p + \cdots + p^n) = 1 - p^{n+1}$, from which (1) follows. ∎

S1.2 A weighted geometric sum

S1.2.1 Proposition *For* $0 < p < 1$,

$$\sum_{k=1}^{\infty} kp^k = \frac{p}{(1-p)^2}. \tag{5}$$

[1] It may not seem very Caltech-like to put down cleverness. Many of you are very clever, and that is good. But relying on cleverness has a downside. My favorite comments on why one should avoid clever solutions is from the programmer Mark Jason Dominus in his brilliant *Higher Order PERL* [2, p. 229]: "These three tactics are presented in increasing order of 'cleverness.' Such cleverness should be used only when necessary, since it requires a corresponding application of cleverness on the part of the maintenance programmer eight weeks later, and such cleverness may not be available."

The elementary approach: To prove (5), first fix n and let

$$x = p + 2p^2 + 3p^3 + \cdots + np^n = \sum_{k=1}^{n} kp^k.$$

Then $(1-p)x$ expands to

$$\begin{aligned}(1-p)x &= p + 2p^2 + 3p^3 + \cdots + np^n \\ &\quad - \ p^2 - 2p^3 - \cdots - (n-1)p^n - np^{n+1} = \\ &= p + p^2 + p^3 + \cdots + p^n - np^{n+1} \\ &= \frac{p - p^{n+1}}{1-p} - np^{n+1} \qquad \text{by (2).}\end{aligned}$$

Dividing both sides by $1-p$ gives

$$\sum_{k=1}^{n} kp^k = x = \frac{p - p^{n+1}}{(1-p)^2} - n\frac{p^{n+1}}{1-p},$$

and letting $n \to \infty$ gives

$$\sum_{k=1}^{\infty} kp^k = \frac{p}{(1-p)^2},$$

as desired. ∎

Generating function approach to (5): Let $f(p) = 1/(1-p)$. For $0 < p < 1$, by (3):

$$\frac{1}{1-p} = f(p) = 1 + p + p^2 + \cdots$$

Differentiating term-by-term we have

$$\begin{aligned}\frac{1}{(1-p)^2} = f'(p) &= 0 + 1 + 2p + 3p^2 + \cdots \\ &= \frac{1}{p}\Big(p + 2p^2 + 3p^3 + \cdots\Big)\end{aligned}$$

So multiplying both sides by p gives (5). ∎

S1.3 Expected value of a geometric random variable

A geometric random variable is the epoch of the first success in a sequence of independent repetitions of a Bernoulli trial with probability of success p. (It is also a special case of the negative binomial distribution.) The pmf is given by

$$P(X = k) = p(1-p)^{k-1}, \quad k = 1, 2, \ldots$$

Rewriting $p(1-p)^{k-1} = \frac{p}{1-p}(1-p)^k$, we see that (3) implies these probabilities sum to 1. To lighten the notation, let $q = 1 - p$.

I claim the expectation is

$$\boldsymbol{E}\, X = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{1}{p}. \tag{6}$$

For example, the expected length of the St. Petersburg game (toss a coin until the first Tails) has $p = 1/2$, so the expected length is $1/(1/2) = 2$.

Proof: To prove (6), rewrite (5) by replacing p with q to get

$$\sum_{k=1}^{\infty} kq^k = \frac{q}{(1-q)^2}.$$

Multiplying both sides by $(1-q)/q$ gives

$$\sum_{k=1}^{\infty} k(1-q)q^{k-1} = \frac{1}{1-q}.$$

Now let $p = 1 - q$ to get (6). ∎

S1.4 An inverse expectation

I claim that for a geometric X as above,

$$\boldsymbol{E}\,\frac{1}{X} = \sum_{k=1}^{\infty} \frac{1}{k} p(1-p)^{k-1} = \frac{p}{1-p} \ln\left(\frac{1}{p}\right). \tag{7}$$

Proof provided the wise TA Victor Kasatkin: Let

$$f(q) = \sum_{k=1}^{\infty} \frac{q^k}{k}.$$

It is analytic for $|q| < 1$. So for $|q| < 1$ we may compute the derivative term-by-term:

$$f'(q) = \sum_{k=1}^{\infty} q^{k-1} = \sum_{j=0}^{\infty} q^j = \frac{1}{1-q}.$$

Now, $f(0) = 0$, and thus

$$f(q) = \int_0^q f'(t)\,dt = \int_0^q \frac{1}{1-t}\,dt = -\ln(1-q).$$

In other words,

$$\sum_{k=1}^{\infty} \frac{q^k}{k} = f(q) = -\ln(1-q) = \ln\left(\frac{1}{1-q}\right).$$

Now multiply both sides by p/q and replace q by $1-p$ to get (7) ∎

S1.5 Variance of the geometric distribution

If X is a geometric random variable, we can compute its variance (and higher moments). Recall that

$$\boldsymbol{Var}\,X = \boldsymbol{E}(X^2) - (\boldsymbol{E}\,X)^2.$$

So let us first compute

$$x = \sum_{k=1}^{\infty} k^2 q^k.$$

Because of the constant of normalization, $\boldsymbol{E}(X^2) = \frac{1-q}{q}x$. So write

$$\begin{aligned}
(1-q)x &= \sum_{k=1}^{\infty} k^2 q^k - q\sum_{k=1}^{\infty} k^2 q^k \\
&= \sum_{k=1}^{\infty} k^2 q^k - \sum_{k=1}^{\infty} k^2 q^{k+1} \\
&= \sum_{k=1}^{\infty} k^2 q^k - \sum_{k=0}^{\infty} k^2 q^{k+1} \\
&= \sum_{k=1}^{\infty} k^2 q^k - \sum_{k=1}^{\infty} (k-1)^2 q^k \\
&= \sum_{k=1}^{\infty} \big(k^2 - (k-1)^2\big) q^k \\
&= \sum_{k=1}^{\infty} (2k-1) q^k \\
&= 2\frac{q}{(1-q)^2} - \frac{q}{1-q} = \frac{q(1+q)}{(1-q)^2},
\end{aligned}$$

where the last line follows from (5) and (4). The variance can now be computed as $(1-q)x/q - (1/p)^2$, or

$$\boldsymbol{Var}\, X = \frac{p}{(1-p)^2} \tag{8}$$

S1.6 Sums related to higher geometric moments

The calculation of the variance suggested a recursive way of computing the following series:

$$S(n) = \sum_{k=1}^{\infty} k^n q^k.$$

I don't have a lot of use for this beyond $n = 2$, but I thought I'd write it down before I forgot it.

Start by writing

$$\begin{aligned}
(1-q)S(n+1) &= \sum_{k=1}^{\infty} k^n q^k - \sum_{k=1}^{\infty} k^n q^{k+1} \\
&= \sum_{k=1}^{\infty} k^n q^k - \sum_{k=0}^{\infty} k^n q^{k+1} \\
&= \sum_{k=1}^{\infty} k^n q^k - \sum_{k=1}^{\infty} (k-1)^n q^k \\
&= \sum_{k=1}^{\infty} \big(k^n - (k-1)^n\big) q^k \\
&= \sum_{k=1}^{\infty} \left(-\sum_{j=1}^{n-1} \binom{n}{j} k^j (-1)^{n-j} \right) q^k,
\end{aligned}$$

where the last line is just the Binomial Theorem. Now rearrange the terms to get

$$(1-q)S(n+1) = -\sum_{j=1}^{n-1}\binom{n}{j}\sum_{k=1}^{\infty}k^j(-1)^{n-j}q^k$$
$$-\sum_{j=1}^{n-1}\binom{n}{j}(-1)^{n-j}S(j),$$

or

$$S(n+1) = \frac{-1}{1-q}\sum_{j=1}^{n-1}\binom{n}{j}(-1)^{n-j}S(j).$$

We already know that

$$S(0) = \frac{q}{1-q},$$

so with enough patience (or Mathematica) we find $S(n)$ for any nonnegative integer n.

According to Mathematica, the function

$$S(n,q) = \sum_{k=1}^{\infty}k^n q^k$$

is known as the `PolyLog[-n, q]` function, which can be expressed in terms of an integral over the interval $[0,1]$.

S1.7 The Taylor series for the exponential

Apostol [1, p. 436] proves that the Taylor series for the exponential function yields the following identity.

For each real number x,

$$e^x = \sum_{k=0}^{\infty}\frac{x^k}{k!}. \tag{9}$$

Consider the function $g(x) = e^x$. Its n^{th} derivative is given by $g^{(n)}(x) = e^x$, so $g^{(n)}(0) = 1$ for every n, and the infinite Taylor's series expansion of g around zero is

$$g(x) = g(0) + \sum_{k=1}^{\infty}\frac{1}{k!}g^{(k)}(0)(x-0)^k = 1 + \sum_{n=1}^{\infty}\frac{x^k}{k!}.$$

So

$$e^x = \sum_{k=0}^{\infty}\frac{x^k}{k!}.$$

S1.8 Series for the logarithm

When a function has representation asa power series on an interval, then its indefinite integral and derivative may be found by differentiating term by term. See Theorems 11.8 and 11.9 in Apostol [1, p. 432].

Equation (3) tells us that the function f defined by the geometric series

$$f(p) = 1 + p + p^2 + \cdots + p^n + \cdots = \frac{1}{1-p}$$

for $|p| < 1$. Replacing p by $-p$ gives

$$1 - p + p^2 - p^3 + \cdots + (-1)^n p^n + \cdots = \frac{1}{1+p}. \tag{10}$$

Since $\int 1/(1+p)\,dp = \ln(1+p)$, integrating (10) term-by-term gives

$$p - \frac{p^2}{2} + \frac{p^3}{3} - \frac{p^4}{4} + \cdots \frac{(-1)^n x^{n+1}}{n+1} + \cdots = \ln(1+p) \tag{11}$$

for $|p| < 1$,

S1.9 A Fun Fibonacci Sum

The **Fibonacci sequence** is defined by the difference equation or recurrence relation

$$F_n = F_{n-1} + F_{n-2} \qquad (n > 1),$$

with initial conditions

$$F_0 = 0, \qquad F_1 = 1.$$

It can be used to define a probability distribution because

$$\sum_{n=1}^{\infty} \frac{F_n}{2^{n+1}} = 1. \tag{12}$$

To see this, observe that

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{F_n}{2^{n+1}} &= \frac{F_1}{4} + \frac{F_2}{8} + \sum_{n=3}^{\infty} \frac{F_n}{2^{n+1}} && \text{regroup} \\
&= \frac{1}{4} + \frac{1}{8} + \sum_{n=3}^{\infty} \frac{F_{n-1}}{2^{n+1}} + \sum_{n=3}^{\infty} \frac{F_{n-2}}{2^{n+1}} && \text{recursion relation} \\
&= \frac{1}{4} + \frac{1}{8} + \sum_{n=2}^{\infty} \frac{F_n}{2^{n+2}} + \sum_{n=1}^{\infty} \frac{F_n}{2^{n+3}} && \text{shift indices} \\
&= \frac{1}{4} + \frac{1}{8} + \left(\sum_{n=1}^{\infty} \frac{F_n}{2^{n+2}} - \frac{F_1}{8} \right) + \sum_{n=1}^{\infty} \frac{F_n}{2^{n+3}} && \text{regroup} \\
&= \frac{1}{4} + \frac{1}{8} + \frac{1}{2} \sum_{n=1}^{\infty} \frac{F_n}{2^{n+1}} - \frac{F_1}{8} + \frac{1}{4} \sum_{n=1}^{\infty} \frac{F_n}{2^{n+1}} && \text{factor} \\
&= \frac{1}{4} + \frac{3}{4} \sum_{n=1}^{\infty} \frac{F_n}{2^{n+1}}, && \text{simplify}
\end{aligned}$$

from which it follows that $\displaystyle\sum_{n=1}^{\infty} \frac{F_n}{2^{n+1}} = 1$.

Bibliography

[1] T. M. Apostol. 1967. *Calculus, Volume I: One-variable calculus with an introduction to linear algebra*, 2d. ed. New York: John Wiley & Sons.

[2] M. J. Dominus. 2005. *Higher order PERL: Transforming programs with programs.* Amsterdam: Morgan Kaufmann.