1. Suppose Y is a discrete random variable with probability function $p(y) = ky(1/4)^y$, $y = 0, 1, 2, 3, \ldots$. Find

   (a) k and (b) $E(Y)$ and $V(Y)$.

   Solution. Let $p = 1/4$.

   (a) We have that

$$\begin{aligned}
1 &= \sum_{y=0}^{\infty} kyp^y \\
&= \sum_{y=0}^{\infty} kp\left(\frac{d}{dp}p^y\right) \\
&= kp\frac{d}{dp}\sum_{y=0}^{\infty} p^y \\
&= kp\frac{d}{dp}\left(\frac{1}{1-p}\right) \\
&= \frac{kp}{(1-p)^2}.
\end{aligned}$$

   It follows that $k = \dfrac{(1-p)^2}{p} = \dfrac{9}{4}$.

   (b) We have that

$$\begin{aligned}
E(Y) &= \sum_{y=0}^{\infty} ky^2p^y \\
&= \sum_{y=0}^{\infty} ky^2p^y - \sum_{y=0}^{\infty} kyp^y + \sum_{y=0}^{\infty} kyp^y \\
&= \sum_{y=0}^{\infty} k(y^2-y)p^y + \sum_{y=0}^{\infty} kyp^y \\
&= \sum_{y=0}^{\infty} kp^2\left(\frac{d^2}{dp^2}p^y\right) + \sum_{y=0}^{\infty} kyp^y \\
&= kp^2\frac{d^2}{dp^2}\sum_{y=0}^{\infty} p^y + 1 \\
&= kp^2\frac{d^2}{dp^2}\left(\frac{1}{1-p}\right) + 1 \\
&= \frac{2kp^2}{(1-p)^3} + 1 \\
&= \frac{5}{3},
\end{aligned}$$

and

$$\begin{aligned}
E(Y^2) &= \sum_{y=0}^{\infty} ky^3p^y \\
&= \sum_{y=0}^{\infty} ky^3p^y - 3\sum_{y=0}^{\infty} ky^2p^y + 2\sum_{y=0}^{\infty} kyp^y + 3\sum_{y=0}^{\infty} ky^2p^y - 2\sum_{y=0}^{\infty} kyp^y \\
&= \sum_{y=0}^{\infty} k(y^3 - 3y^2 + 2y)p^y + 3E(Y) - 2 \\
&= \sum_{y=0}^{\infty} ky(y-1)(y-2)p^y + 3 \\
&= \sum_{y=0}^{\infty} kp^3\left(\frac{d^3}{dp^3}p^y\right) + 3 \\
&= kp^3\frac{d^3}{dp^3}\sum_{y=0}^{\infty} p^y + 3 \\
&= \frac{6kp^3}{(1-p)^4} + 3.
\end{aligned}$$

Since $V(Y) = E(Y^2) - E(Y)^2$, it follows that

$$\begin{aligned}
V(Y) &= E(Y^2) - E(Y)^2 \\
&= \frac{6kp^3}{(1-p)^4} + 3 - \frac{25}{9} \\
&= \frac{8}{9}.
\end{aligned}$$

2. Verify the identity $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$ and use it to show that $E[Y^k] = npE[(X+1)^{k-1}]$ where Y is a binomial random variable with parameters n and p and X is a binomial random variable with parameters $n-1$ and p.

Proof. We have that

$$\begin{aligned}
\binom{n}{k} &= \frac{n!}{k!(n-k)!} \\
&= \frac{n(n-1)!}{k(k-1)!(n-k)!} \\
&= \frac{n}{k}\frac{(n-1)!}{(k-1)!(n-k)!} \\
&= \frac{n}{k}\frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} \\
&= \frac{n}{k}\binom{n-1}{k-1}.
\end{aligned}$$

Now

$$
\begin{aligned}
E[Y^k] &= \sum_{y=0}^{n} y^k p(y) && \text{[Definition]} \\
&= \sum_{y=0}^{n} y^k \binom{n}{y} p^y (1-p)^{n-y} \\
&= \sum_{y=1}^{n} y^k \binom{n}{y} p^y (1-p)^{n-y} \\
&= \sum_{y=1}^{n} y^{k-1} n \binom{n-1}{y-1} p^y (1-p)^{n-y} \\
&= np \sum_{y=1}^{n} y^{k-1} \binom{n-1}{y-1} p^{y-1} (1-p)^{n-y} \\
&= np \sum_{x=0}^{n-1} (x+1)^{k-1} \binom{n-1}{x} p^x (1-p)^{(n-1)-x} && \text{[Let } y = x+1\text{]} \\
&= np \sum_{x=0}^{n-1} (x+1)^{k-1} p(x) \\
&= npE[(X+1)^{k-1}],
\end{aligned}
$$

which is what we wanted to show. □

3. Using the recursion relation found in problem 2 for the binomial random variable with parameters n and p, find $E[Y^2]$ and then $V(Y)$.

Solution. If we set $k = 2$ in the formula in problem 2, we get

$$
\begin{aligned}
E[Y^2] &= npE[X+1] \\
&= np(E[X] + E[1]) \\
&= np((n-1)p + 1) \\
&= (np)^2 - np^2 + np.
\end{aligned}
$$

Thus

$$
\begin{aligned}
V(Y) &= E[Y^2] - E[Y]^2 \\
&= (np)^2 - np^2 + np - (np)^2 \\
&= np - np^2 \\
&= np(1-p).
\end{aligned}
$$

4. Using the identity from problem 2, show that

(a) if Y is a negative binomial random variable with parameters r and p, then

$$E[Y^k] = \frac{r}{p} E[(X-1)^{k-1}],$$

where X is a negative binomial random variable with parameters $r+1$ and p.

(b) Use the relation in (a) to find $E[Y]$ and $V(Y)$.

Solution.

(a) From problem 2, we know that

$$\frac{1}{r}\binom{y-1}{r-1} = \frac{1}{y}\binom{y}{r};$$

thus

$$\begin{aligned}
E[Y^k] &= \sum_{y=r}^{\infty} y^k p(y) \\
&= \sum_{y=r}^{\infty} y^k \binom{y-1}{r-1} p^r (1-p)^{y-r} \\
&= \frac{r}{p}\frac{p}{r}\sum_{y=r}^{\infty} y^k \binom{y-1}{r-1} p^r (1-p)^{y-r} \\
&= \frac{r}{p}\sum_{y=r}^{\infty} y^k \frac{1}{r}\binom{y-1}{r-1} p^{r+1} (1-p)^{y-r} \\
&= \frac{r}{p}\sum_{y=r}^{\infty} y^k \frac{1}{y}\binom{y}{r} p^{r+1} (1-p)^{y-r} \\
&= \frac{r}{p}\sum_{y=r}^{\infty} y^{k-1} \binom{y}{r} p^{r+1} (1-p)^{y-r} \\
&= \frac{r}{p}\sum_{x=r+1}^{\infty} (x-1)^{k-1} \binom{x-1}{r} p^{r+1} (1-p)^{(x-1)-r} \\
&= \frac{r}{p}\sum_{x=r+1}^{\infty} (x-1)^{k-1} p(x) \\
&= \frac{r}{p} E[(X-1)^{k-1}].
\end{aligned}$$

(b) If we set $k = 1$ in (a), we immediately get that $E[Y] = \frac{r}{p}$. Now

$$\begin{aligned}
V(Y) &= E[Y^2] - E[Y]^2 \\
&= \frac{r}{p}E[X-1] - \frac{r^2}{p^2} \\
&= \frac{r}{p}(E[X] - E[1]) - \frac{r^2}{p^2} \\
&= \frac{r}{p}\left(\frac{r+1}{p} - 1\right) - \frac{r^2}{p^2} \\
&= \frac{r^2+r}{p^2} - \frac{r}{p} - \frac{r^2}{p^2} \\
&= \frac{r-pr}{p^2} \\
&= \frac{r(1-p)}{p^2}.
\end{aligned}$$

5. Using the identity from problem 2, show that if Y is a hypergeometric random variable with parameters N, r, amd n, then

$$E[Y^k] = \frac{nr}{N}E[(X+1)^{k-1}],$$

where X is a hypergeometric random variable with parameters $N-1$, $r-1$, and $n-1$.

Proof. We have that

$$\begin{aligned}
E[Y^k] &= \sum_{y=0}^{n} y^k p(y) \\
&= \sum_{y=0}^{n} y^k \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} \\
&= \sum_{y=1}^{n} y^k \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} \\
&= \sum_{y=1}^{n} y^k \frac{\frac{r}{y}\binom{r-1}{y-1}\binom{N-r}{n-y}}{\frac{N}{n}\binom{N-1}{n-1}} \\
&= \frac{nr}{N}\sum_{y=1}^{n} y^{k-1} \frac{\binom{r-1}{y-1}\binom{N-r}{n-y}}{\binom{N-1}{n-1}} \\
&= \frac{nr}{N}\sum_{x=0}^{n-1} (x+1)^{k-1} \frac{\binom{r-1}{x}\binom{(N-1)-(r-1)}{(n-1)-x}}{\binom{N-1}{n-1}} && [\text{Let } y = x+1] \\
&= \frac{nr}{N}\sum_{x=0}^{n-1} (x+1)^{k-1} p(x) \\
&= \frac{nr}{N} E[(X+1)^{k-1}],
\end{aligned}$$

which is what we wanted to prove. □

6. If Y is a hypergeometric random variable with parameters N, r, and n, use the recursion relation found in problem 5 to find $E[Y]$ and $V(Y)$.

Solution. Plugging in $k = 1$ in the formula from problem 5 immediately shows us that

$E[Y] = \frac{nr}{N}$. Using this same formula, we have that

$$\begin{aligned}
V(Y) &= E[Y^2] - E[Y]^2 \\
&= \frac{nr}{N} E[X+1] - \left(\frac{nr}{N}\right)^2 \\
&= \frac{nr}{N} (E[X] + E[1]) - \left(\frac{nr}{N}\right)^2 \\
&= \frac{nr}{N} \left(\frac{(n-1)(r-1)}{N-1} + 1\right) - \left(\frac{nr}{N}\right)^2 \\
&= \frac{nr}{N} \left(\frac{(n-1)(r-1) + N - 1}{N-1}\right) - \left(\frac{nr}{N}\right)^2 \\
&= \frac{nr}{N} \left(\frac{(n-1)(r-1) + N - 1}{N-1} - \frac{nr}{N}\right) \\
&= \frac{nr}{N} \left(\frac{N(n-1)(r-1) + N^2 - N - nr(N-1)}{N(N-1)}\right) \\
&= \frac{nr}{N} \left(\frac{Nnr - Nn - Nr + N + N^2 - N - Nnr + nr}{N(N-1)}\right) \\
&= \frac{nr}{N} \left(\frac{-Nn - Nr + N^2 + nr}{N(N-1)}\right) \\
&= \frac{nr}{N} \left(\frac{N^2 - Nr - Nn + nr}{N(N-1)}\right) \\
&= \frac{nr}{N} \left(\frac{N(N-r) - n(N-r)}{N(N-1)}\right) \\
&= \frac{nr}{N} \left(\frac{N-r}{N}\right) \left(\frac{N-n}{N-1}\right).
\end{aligned}$$

7. The number of chocolate chips in 1 cup of chocolate chip ice cream has a Poisson distribution with a mean of 10 chips per cup.

   (a) What is the probability that a cup of chocolate chip ice cream has 9 chocolate chips.

   (b) What is the probability that a half cup of chocolate chip ice cream has at least 3 chocolate chips?

   Solution.

   (a) We have that $\lambda = 10$, so that

   $$P(Y = 9) = p(9) = \frac{10^9}{9!} e^{-10} \approx 0.12511.$$

(b) Now $\lambda = 5$, so that

$$\begin{aligned}
P(Y \geq 3) &= 1 - P(Y < 3) \\
&= 1 - p(0) - p(1) - p(2) \\
&= 1 - e^{-5} - 5e^{-5} - \frac{5^2}{2}e^{-5} \\
&\approx 0.875348.
\end{aligned}$$

8. Suppose the distribution function of Y is given by

$$F(y) = \begin{cases} 0 & \text{if } y < 0 \\ y^3 & \text{if } 0 \leq y \leq 1 \\ 1 & \text{if } y > 1. \end{cases}$$

Find (a) $P\left(Y > \frac{3}{4}\right)$ (b) $E[Y]$ and (c) $V(Y)$.

Solution.

(a) We have that

$$\begin{aligned}
P\left(Y > \frac{3}{4}\right) &= 1 - P\left(Y \leq \frac{3}{4}\right) \\
&= 1 - F\left(\frac{3}{4}\right) \\
&= 0.578125.
\end{aligned}$$

(b) The probability density function, $f(y)$, is

$$\frac{dF(y)}{dy} = \begin{cases} 0 & \text{if } y < 0 \\ 3y^2 & \text{if } 0 \leq y < 1 \\ 0 & \text{if } y > 1, \end{cases}$$

so that

$$\begin{aligned}
E[Y] &= \int_{-\infty}^{\infty} yf(y)dy \\
&= \int_0^1 yf(y)dy \\
&= \int_0^1 3y^3 dy = 0.75.
\end{aligned}$$

(c) By definition

$$\begin{aligned} V(Y) &= E[Y^2] - E[Y]^2 \\ &= \int_{-\infty}^{\infty} y^2 f(y)dy - 0.75^2 \\ &= \int_0^1 3y^4 dy - 0.75^2 \\ &= 0.0375. \end{aligned}$$

9. Let Y be a continuous random variable with density function

$$f(y) = \frac{k}{1+y^2}, \infty < y < \infty.$$

Find (a) k (b) $E[Y]$ and $V(Y)$, if they exist. (Such a distribution is called a Cauchy distribution.)

Solution.

(a) We must have that

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f(y)dy \\ &= \int_{-\infty}^{\infty} \frac{k}{1+y^2}dy \\ &= \int_{-\infty}^{a} \frac{k}{1+y^2}dy + \int_{a}^{\infty} \frac{k}{1+y^2}dy \\ &= \lim_{t\to-\infty}\int_{t}^{a} \frac{k}{1+y^2}dy + \lim_{s\to\infty}\int_{a}^{s} \frac{k}{1+y^2}dy \\ &= k\lim_{t\to-\infty}(\arctan(a) - \arctan(t)) + k\lim_{s\to\infty}(\arctan(s) - \arctan(a)) \\ &= k\left(\arctan(a) + \frac{\pi}{2}\right) + k\left(\frac{\pi}{2} - \arctan(a)\right) \\ &= k\pi, \end{aligned}$$

so that $k = \dfrac{1}{\pi}$.

(b)

$$\begin{aligned} E[Y] &= \int_{-\infty}^{\infty} yf(y)dy \\ &= \int_{-\infty}^{\infty} \frac{ky}{1+y^2}dy \\ &= \lim_{t\to-\infty}\int_{t}^{a} \frac{ky}{1+y^2}dy + \lim_{s\to\infty}\int_{a}^{s} \frac{ky}{1+y^2}dy \\ &= \frac{k}{2}\lim_{t\to-\infty}(\ln(1+a^2) - \ln(1+t^2)) + \frac{k}{2}\lim_{s\to\infty}(\ln(1+s^2) - \ln(1+a^2)) \\ &= \text{Does Not Exist.} \end{aligned}$$

Since $E[Y]$ does not exist, it follows that $V(Y)$ does not exist.

10. A point is chosen at random on a line segment of length L. Find the probability that the ratio of the shorter segment to the longer segment is less than $1/4$.

    Solution Let Y be a random variable that represents the point on the line. Let x be a point on the line segment. Then we have two cases:

    Case 1. The left segment is no bigger than the right. Thus we want $\frac{x}{L-x} < \frac{1}{4}$, so that $4x < L - x$; that is, $x < \frac{L}{5}$.

    Case 2. The right segment is no bigger than the left. Thus we want $\frac{L-x}{x} < \frac{1}{4}$, so that $4L - 4x < x$; that is, $x > \frac{4L}{5}$.

    If X is a random variable representing a point on the line segment, we want $P\left(X < \frac{L}{5}\right)$ or $P\left(X > \frac{4L}{5}\right)$. Thus

$$\begin{aligned}
P\left(X < \frac{L}{5}\right) + P\left(X > \frac{4L}{5}\right) &= P\left(X < \frac{L}{5}\right) + 1 - P\left(X \leq \frac{4L}{5}\right) \\
&= \frac{L/5}{L} + 1 - \frac{4L/5}{L} \\
&= \frac{2}{5}.
\end{aligned}$$