1. (continuing problem 2, Assignment 6): Consider the random variable X with the probability mass distribution

$$P[X=1]=0.3, \qquad P[X=4]=0.25, \qquad P[X=7]=0.4, \qquad P[X=10]=0.05.$$

Calculate the variance of X and Y with $Y=3X+2$.

$E[X] = 4.6$

$E[X^2] = 0.3 + 4 + \frac{49.2}{5} + 5 = 28.9$

$Var[X] = E[X^2] - E[X]^2 = 28.9 - 21.16 = 7.74$

$Var[Y] = Var[3X+2] = 3^2 Var[X] = 69.66$

2. Suppose we pick a month at random from a non leap-year calendar and let X be the number of days in that month. Find the mean and the variance of X.

$P[X=28] = 1/12$, $P[X=30] = 4/12$, $P[X=31] = 7/12$

$E[X] = (28 + 120 + 217)/12 = 365/12 = 30.41\overline{6}$

$E[X^2] = (28^2 + 30^2 \cdot 4 + 31^2 \cdot 7)/12$

$= (784 + 3600 + 6727)/12$

$= 11{,}111/12$

$= 925.91\overline{6}$

$Var[X] = E[X^2] - E[X]^2 = \frac{11{,}111}{12} - \frac{133{,}225}{144} = \frac{107}{144} = 0.7430\overline{5}$

3. Let Y be a binomial distributed random variable with n trials of success probability p. Show that

$$\mathbb{V}ar[Y] = np(1-p).$$

$P[X=k] = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, 2, \ldots$, $E[X] = np$

$E[X^2] = \sum_{k=0}^{\infty} k^2 P[X=k] = \sum_{k=0}^{\infty} k^2 \binom{n}{k} p^k (1-p)^{n-k}$

$= \sum_{k=1}^{\infty} k^2 \binom{n}{k} p^k (1-p)^{n-k}$ $\qquad k\binom{n}{k} = n\binom{n-1}{k-1}$

$= \sum_{k=1}^{\infty} kn \binom{n-1}{k-1} p^k (1-p)^{n-k}$

$= np \sum_{k=1}^{\infty} k \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k}$ $\qquad k = j+1$

$= np \sum_{j=0}^{\infty} (j+1) \binom{n-1}{j} p^j (1-p)^{n-j-1}$

$= np\, E[Y+1]$ where $Y \sim Binom(n-1, p)$

$= np(E[Y]+1)$

$= np((n-1)p + 1)$

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

4. Let Z be a geometric distributed random variable with success probability p. Calculate $\mathbb{V}ar[Z]$.

$P[X=k] = (1-p)^k p \quad , k = 0, 1, 2, \ldots$

$E[X] = \frac{1-p}{p}$

$$
\begin{aligned}
E[X^2] &= \sum_{k=0}^{\infty} k^2 p (1-p)^k \\
&= p \sum_{k=1}^{\infty} ((k-1)+1)^2 (1-p)^k \\
&= p \sum_{k=1}^{\infty} (k-1)^2 (1-p)^k + p \sum_{k=1}^{\infty} 2(k-1)(1-p)^k + p \sum_{k=1}^{\infty} (1-p)^k \\
&= (1-p) \sum_{j=0}^{\infty} j^2 p (1-p)^j + 2(1-p) \sum_{j=0}^{\infty} j p (1-p)^j + 1 - p \\
&= (1-p) E[X^2] + 2(1-p) E[X] + (1-p)
\end{aligned}
$$

$E[X^2](1 - 1 + p) = 2(1-p)\frac{1-p}{p} + (1-p)$

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

$Var[X] = \frac{(1-p)(2-p)}{p^2} - \frac{(1-p)^2}{p^2} = \frac{1-p}{p^2}\left(2 - p - (1-p)\right) = \frac{1-p}{p^2}$

5. Assume that X is a random variable taking values on the non-negative integers that satisfies

$$\mathbb{P}[X \geq n+i \,|\, X \geq n] = \mathbb{P}[X \geq i].$$

Show that X is a geometric distributed random variable.

$P[X \geq i] = P[X \geq n+i \,|\, X \geq n] = \frac{P[X \geq n+i]}{P[X \geq n]}$

$P[X \geq i] = \frac{P[X \geq n+i]}{P[X \geq n]}$

In particular if $i = 1$,

$$1 - P[X=0] = P[X \geq 1] = \frac{P[X \geq n+1]}{P[X \geq n]} = 1 - \frac{P[X=n]}{P[X \geq n]}$$

$$P[X=0]P[X \geq n] = P[X=n]$$

Let $p = P[X=0]$.

$\boxed{\text{Identity: } P[X=n] = p\,P[X \geq n] \text{ for any nonnegative integer } n.}$

Claim: $P[X=n] = p(1-p)^n$, $n = 0, 1, 2, \ldots$

Proof. For the base case $n=0$,

$$P[X=0] = pP[X \geq 0] = p = p(1-p)^0$$

using the identity above

Assume $P[X=n] = p(1-p)^n$ for some $n \geq 0$.

$$\begin{aligned}
P[X=n+1] &= pP[X \geq n+1] \quad \text{(Identity)} \\
&= p(P[X \geq n] - P[X=n]) \\
&= p\left(\frac{P[X=n]}{p} - P[X=n]\right) \quad \text{(Identity)} \\
&= p((1-p)^n - p(1-p)^n) \quad \text{(Inductive hypothesis)} \\
&= p(1-p)^n(1-p) \\
&= p(1-p)^{n+1}
\end{aligned}$$

6. The number of errors on a book page follow a Poisson distribution. It has been determined that on 10% of the pages there is at least one error.

   a) Determine the parameter of the Poisson distribution.

   b) What is the expected number of errors on a page?

$$P[X=k] = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

a) Let $P[X=k]$ be the probability that there are k errors on a page. Then $P[X>0] = 0.1$.

$$e^{-\lambda}\frac{\lambda^0}{0!} = P[X=0] = 1 - P[X>0] = 0.9$$

$$-\lambda = \ln 9/10$$

$$\lambda = \ln 10/9 \approx 0.10536$$

b) $E[X] = \sum_{k=0}^{\infty} k e^{-\lambda}\frac{\lambda^k}{k!} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = \lambda e^{-\lambda}e^{\lambda} = \lambda$

$E[X] = \lambda = \ln 10/9$