Week 7

Sam Reeves

1. Let $X_1, X_2, ..., X_n$ be n mutually independent random variables, each of which is uniformly distributed on the integers from 1 to k. Let Y denote the minimum of X_is. Find the distribution of Y.

This pertains to the first section of Chapter 7 in our probability text: Sums of Discrete Random Variables. Because the variables are distributed on the set of integers, we know this is not a continuous distribution.

I think the following statement provides a hint:

> We shall find it convenient to assume here that these distribution functions are defined for all integers, by defining them to be 0 where they are not otherwise defined.

So for $k \in \{1, 2, ..., k\}$, $P(X = i) = \frac{1}{k}$.

$P(X > 0) = 1$, $P(X > 1) = 1 - \frac{1}{k}, \ldots, P(X > k-1) = \frac{1}{k}, P(X > k) = 0$

The text gives:

$$Z = X + Y$$

$$X = k, Z = z$$
$$Y = z - k$$

$$P(Z = z) = \sum_{k=-\infty}^{\infty} P(X = k) \cdot P(Y = z - k)$$

and follows with:

$$m_3(j) = \sum_k m_1(k) \cdot m_2(j - k)$$

suggesting that $m_3(x)$ is the distribution function for the random variable $Z = X + Y$. It is explained that convolutions are commutative and associative... So, we can manipulate the order of these terms.

If $S_n = X_1 + X_2 + ... + X_n$, then $S_n = S_n - 1 + X_n$.

So, the distribution of Y is:

$$P(X_i = Y) = \sum_{i=1}^{k} \frac{(k)^n - (k-i)^n}{k^n}$$

That is, the sum of the ways we can get the minimum Y over the number of possible values of X_i.

2. Your organization owns a copier. This machine has a manufacturer's expected lifetime of 10 years. This means that we expect one failure every ten years. For each of the following: include the probability that the machine will fail after 8 years, the expected value, and the standard deviation.

```
p <- 1/10
n <- 8
```

a. Model as geometric. (Hint: the probability is equivalent to not failing during the first 8 years...)

In this week's lecture, we learned that the pgeom function is very precise in that it calculates $p(n) = pq^n$. These are 8 years of not failing followed by a year of failure. We can specify lower.tail = TRUE to indicate $P(X \leq)$ as opposed to $P(X > x)$, as given in the documentation for pgeom.

For our purposes, we want $P(X > n) = (1-p)^{(n+1)}$, demonstrated:

```
(1-p) ^ (n+1)
```

```
## [1] 0.3874205
```

```
pgeom(n, p, lower.tail = FALSE)
```

```
## [1] 0.3874205
```

Borrowing from my work in week 5:

$$
\begin{aligned}
E(x) = \sum_{x \in \Omega} xm(x) = \frac{1}{p} \\
= 10
\end{aligned}
$$

Standard deviation is $\sigma = \sqrt{\frac{1-p}{p^2}}$

```
sqrt( (1-p) / (p^2) )
```

```
## [1] 9.486833
```

b. Model as an exponential.

The probability that the machine will fail after 8 years is:

$$
\begin{aligned}
P(X) = 1 - (1 - e^{-pn}) \\
= e^{-pn}
\end{aligned}
$$

```
exp(1)^(-p * n)
```

```
## [1] 0.449329
```

```r
pexp(8, p, lower.tail = FALSE)
```

```
## [1] 0.449329
```

The expected failure rate is given by the manufacturer to be 10 years. We can still call this $\frac{1}{p}$.

For $p = \lambda$:

$$
\begin{aligned}
\mu_X &= \frac{1}{\lambda} \\
\sigma_X^2 &= \frac{1}{\lambda^2} \\
\mu_X = \sigma_X \ &= \frac{1}{\lambda} \\
\sigma_X &= 10
\end{aligned}
$$

The standard deviation is equal to the mean.

c. Model as a binomial. (Hint: 0 successes in 8 years...)

Borrowing the concepts from Michael Ippolito's recitation in this week's lecture:

$$
\begin{aligned}
b(n, p, k) &= \binom{n}{k} p^k (1-p)^{n-k} \\
&= \binom{8}{0} \frac{1}{10}^0 \frac{9}{10}^8 \\
&= \binom{8}{0} \frac{9}{10}^8
\end{aligned}
$$

According to the documentation for the binomial distribution functions in R, we can use the dbinom function to return the probability density of the function described above...

```r
k <- 0

choose(8,0) * (1-p)^8
```

```
## [1] 0.4304672
```

```r
dbinom(k, n, p)
```

```
## [1] 0.4304672
```

In this case, the expected value is $E(X) = np = 0.8$

The standard deviation for a binomial distribution is:

$$\sigma = \sqrt{np(1-p)}$$

```
sqrt(n * p * (1-p))
```

```
## [1] 0.8485281
```

d. Model as a Poisson.

According to the documentation for the Poisson distribution functions in R:

$$P(X) = \lambda^x e \frac{-\lambda}{x!}$$

In our case, $\lambda = p$. So, we can use the ppois function with $x = 0$ to return the probability that there is no failure in a year, then multiply that times itself 8 times to find the probability that there are no failures for the first 8 years.

$$P(0) = \lambda^0 e \frac{-\lambda}{x!}$$
$$0! = 1$$
$$\therefore$$
$$P(0, 1, ...7) = (e \times -\lambda)^8$$

```
exp(-p)^8
```

```
## [1] 0.449329
```

```
ppois(0, p)^8
```

```
## [1] 0.449329
```

The expected value is $E(X) = \lambda n$:

```
n*p
```

```
## [1] 0.8
```

In the Poisson distribution, the variance is equal to the expected value for one trial, so the standard deviation is equal to the squareroot of the expected value:

$$E(X) = \sigma^2$$
$$\sigma = \sqrt{\lambda n}$$

```
sqrt(n*p)
```

```
## [1] 0.8944272
```