Homework 7

Vineet Rai

Due date: Thursday, April 4

1. (2 points) **Exercise 5.23** If X is a random variable having a geometric distribution show that

$$P(X = x + n | X > n) = P(X = x).$$

If A is the event that $(X > n)$ and B is the event that $(X = x + n)$ then we can apply the formula for conditional probability:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(X > n \quad \cap \quad X = x + n)}{P(X > n)}$$

Since X is a random variable with a geometric distribution, it is supported for values of $x = 1, 2, 3\dots$ and cannot have a value of 0. Therefore, B is a subset of sample space A and there are no events of B which lie outside of the space of A. In other words, if $X = x + n$ and $x = 1, 2, 3\cdots$ then X must be greater than n. This means the numerator of the fraction reduces to $P(B)$.

$$= \frac{P(X = x + n)}{P(X > n)}$$

We can decompose the denominator by subtracting the probability that $X = n$. We now have three probability terms in the fraction.

$$= \frac{P(X = x + n)}{P(X \geq n) - P(X = n)}$$

The following formulas are defined for the Geometric Distribution:

$$P(X = x) = \theta(1-\theta)^{x-1} \text{ for } x = 1, 2, 3\dots$$
$$P(X \geq x) = (1-\theta)^{x-1}$$

We can now substitute each of the three terms with the appropriate formula from the Geometric Distribution, as follows:

$$P(X = x + n) = \theta(1-\theta)^{x+n-1}$$
$$P(X \geq n) = (1-\theta)^{n-1}$$
$$P(X = n) = \theta(1-\theta)^{n-1}$$

$$\frac{P(X = x + n)}{P(X \geq n) - P(X = n)} = \frac{\theta(1-\theta)^{x+n-1}}{(1-\theta)^{n-1} - \theta(1-\theta)^{n-1}}$$

Finally, we can simplify the expression.

$$= \frac{\theta(1-\theta)^{x+n-1}}{(1-\theta)(1-\theta)^{n-1}} = \frac{\theta(1-\theta)^{x+n-1}}{(1-\theta)^{n}} = \theta(1-\theta)^{x-1} = P(X = x)$$

2. (2 points) **Exercise 5.82 (a)** Use a computer program to calculate the ***exact*** probability of obtaining one or more defectives in a sample of 100 taken from a lot of 1000 manufactured products assumed to contain six defectives.

```
> phyper(0, 6, 1000-6, 100, lower.tail = FALSE)
[1] 0.4694474
>
> ## Another way to calculate
> s <- 0:99
> 1 - prod((994-s)/(1000-s))
[1] 0.4694474
```

3. (2 points) Let X be a random variable that has the following probability density function:

$$f(x) = \frac{1}{x\log(3)}, \text{ for } 1 < x < 3.$$

Find $E(X)$ and $Var(X)$.

$$\begin{aligned}
E(X) = \mu &= \int_{-\infty}^{\infty} x \cdot f(x)\ dx \\
&= \int_{1}^{3} \frac{x}{x\log(3)}\ dx \\
&= \left[\frac{x}{\log(3)}\right]_1^3 \\
E(X) &= \frac{2}{\log(3)} \approx 1.82
\end{aligned}$$

$$\begin{aligned}
V(X) = E(X^2) - \mu^2 &= \int_{-\infty}^{\infty} x^2 \cdot f(x)\ dx\ - \mu^2 \\
&= \int_{1}^{3} \frac{x^2}{x\log(3)}\ dx\ - \left(\frac{2}{\log(3)}\right)^2 \\
&= \left[\frac{x^2}{2\log(3)}\right]_1^3 - \frac{4}{\log^2(3)} \\
&= \frac{4}{\log(3)} - \frac{4}{\log^2(3)} \\
V(X) &= \frac{4\log(3) - 4}{\log^2(3)} \approx 0.327
\end{aligned}$$

4. (2 points) Use the uniform random number generator in ‘R‘ to generate 1000 random numbers from the probability density function in #3. Save the random vector in ‘x‘. Print summary(x) and var(x).

$$
\begin{aligned}
\text{CDF} = F(x) &= \int f(x)\,dx = \int \frac{1}{x\log(3)}\,dx \\
&= \frac{\log(x)}{\log(3)} + C \\
&\frac{\log(1)}{\log(3)} = 0 \quad \text{and} \quad \frac{\log(3)}{\log(3)} = 1 \quad \text{so} \quad C = 0, \quad \text{therefore} \\
F(x) &= \frac{\log(x)}{\log(3)}
\end{aligned}
$$

$$
\begin{aligned}
&\text{Inverse Transform Method:} \quad F(F^{-1}(u)) = u \\
u &= \frac{\log(F^{-1}(u))}{\log(3)} \\
u\log(3) &= \log(F^{-1}(u)) \\
F^{-1}(u) &= e^{u\log(3)} \\
F^{-1}(u) &= e^{\log(3^u)} \\
F^{-1}(u) &= 3^u
\end{aligned}
$$

```
> set.seed(20190404)
> u <- runif(1e3)
> y <- function(t) 3^t
> x <- y(u)
> summary(x)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.002   1.302   1.716   1.805   2.279   2.999
> var(x)
[1] 0.3238541
```

5. (2 points) **Exercise 6.64** If Z is a random variable having the standard normal distribution, find the respective values z_1 and z_2 such that

a. $P(z_1 \leq Z \leq 0) = 0.4306$.

```
> -qnorm(0.5 + 0.4306)
[1] -1.480275
```

b. $P(Z \geq z_2) = 0.7704$.

```
> qnorm(0.7704, lower.tail = FALSE)
[1] -0.7401648
```