Probability and Statistics 1 - Surgery Hours class (Andres Villegas)
Exercise Sheet 5 Solutions: Discrete Random Variables 2

1. Let X be the number of consecutive days of rain, $X \sim Poisson(0.6)$. Let Y be the amount paid by the insurance company. We have

$$E(Y) = 1000P(X = 1) + 2000P(X \geq 2)$$

But $P(X = 1) = e^{-0.6}(0.6) = 0.3293$ and

$$P(X \geq 2) = 1 - P(X = 1) - P(X = 0) = 1 - e^{-0.6}(0.6) - e^{-0.6} = 0.1219$$

Hence,

$$\begin{aligned} E(Y) &= 1000P(X = 1) + 2000P(X \geq 2) \\ &= 1000(0.3293) + 2000(0.1219) \\ &= 573.09 \end{aligned}$$

The variance of Y is given by $Var(Y) = E(Y^2) - E(Y)^2$. We have

$$\begin{aligned} E(Y^2) &= 1000^2P(X = 1) + 2000^2P(X \geq 2) \\ &= 1000^2(0.3293) + 2000^2(0.1219) \\ &= 816892.51 \end{aligned}$$

Then $Var(Y) = E(Y^2) - E(Y)^2 = 816892.51 - 573.09^2 = 488460.65$ and the standard deviation of the amount paid by the insurance company is $\sqrt{488460.65} = 698.90$

2. Let N be the number of storms that shut down business in one year, $N \sim Poisson(1.5)$. The expected value of the amount paid to the company is

$$\begin{aligned} E(amount\ paid) &= 10000P(N = 2) + 20000P(N = 3) + \cdots \\ &= \sum_{i=2}^{\infty} 10000(i - 1)P(N = i) \\ &= 10000 \sum_{i=2}^{\infty} (i - 1)P(N = i) \\ &= 10000 \left(\sum_{i=2}^{\infty} iP(N = i) - \sum_{i=2}^{\infty} P(N = i) \right) \\ &= 10000 \left(\sum_{i=1}^{\infty} iP(N = i) - P(N = 1) - \sum_{i=0}^{\infty} P(N = i) + P(N = 0) + P(N = 1) \right) \\ &= 10000\big(E(N) - 1 + P(N = 0)\big) \\ &= 10000(1.5 - 1 + e^{-1.5}) \\ &= 7231.30 \end{aligned}$$

3. Let X be the number of patients tested until the first patient with tuberculosis is found. X has a geometric distribution with parameter $\theta = 0.15$.

a) $P(X=5)=(1-\theta)^4\theta=(1-0.15)^4(0.15)=0.0783$

b) $P(X=10)=(1-\theta)^9\theta=(1-0.15)^9(0.15)=0.0347$

c) $E(X)=\frac{1}{\theta}=\frac{1}{0.15}=6.6667$

d) This can be computed from first principles as follows

$$P\left(15^{th}\ patient\ tested\ be\ the\ 3^{rd}\ with\ tuberculosis\right)$$
$$=P\left(2\ out\ of\ the\ first\ 14\ patients\ have\ tuberculosis, 15^{th}\ patient\ has\ tuberculosis\right)$$
$$=P(2\ out\ of\ the\ first\ 14\ patients\ have\ tuberculosis)P\left(15^{th}\ patient\ has\ tuberculosis\right)$$
$$=\binom{14}{2}(0.15)^2(0.85)^{12}(0.15)$$
$$=0.0437$$

Alternative, this could be computed using the negative binomial distribution.

e) The total number of patients tested required to find the sixth patient with tuberculosis, Y, is the sum of 6 geometric random variable each one representing the number patients tested to find one patient with tuberculosis. That is, $Y=X_1+\cdots+X_6$, with X_i, $i=1,\ldots,6$ having a geometric distribution with parameter $\theta=0.15$. Therefore

$$E(Y)=E(X_1+\cdots+X_6)$$
$$=E(X_1)+\cdots E(X_6)$$
$$=6\frac{1}{\theta}=6\frac{1}{0.15}=40$$

But this includes the 6 patients with tuberculosis; hence, the expected number of patients without tuberculosis tested before the sixth patient with tuberculosis is tested is $E(Y)-6=40-6=34$.

4.

a)

$$P\left(5^{th}\ sale\ on\ the\ 16^{th}\ call\right)=P\left(4\ sales\ in\ the\ first\ 15\ calls, 16^{th}\ call\ is\ a\ sale\right)$$
$$=\ P(4\ sales\ in\ the\ first\ 15\ calls)P\left(16^{th}\ call\ is\ a\ sale\right)$$
$$=\binom{15}{4}(0.2)^4(0.8)^{11}(0.2)$$
$$=0.0375$$

b) In order to reach £2,000 in sales the telemarketer needs to make $\frac{2000}{250}$=8 sales. Let Y be the total number of calls that need to be made to make 8 sales. We have that $Y=X_1+\cdots+X_8$ where each X_i is a geometric random variable with parameter $\theta=0.2$, representing the number of calls needed to make one sale. Hence, the expected number of calls is

$$E(Y)=E(X_1+\cdots+X_8)=E(X_1)+\cdots E(X_8)=8\left(\frac{1}{0.20}\right)=40$$

of which 32 are not successful and 8 are sales.

5. Let X be the number of policies with claims, $X\sim Binomial(25,p)$. The Poisson approximation is $X\approx Poisson(25p)$

a) i) Using $Poisson(2.5)$ $P(X\leq\ 4)=0.89118$ from tables (or evaluation), ii) Using $Poisson(5)$ $P(X\leq\ 4)=0.44049$ again from tables (or evaluation)

b) In i) the error is $0.8912-0.9020=-0.0108$, in ii) the error is $0.4405-0.4207=0.0198$

The approximation is valid for “small p”, and, as p is smaller in $i)$, this gives a better approximation as noted with the smaller error.
The sample size 25 is not “large” and so we would not expect the Poisson approximation to be very good anyway. However the key to the approximation is the small p and here the given approximations are quite good.