Homework 06 Solution

STAT 509 Statistics for Engineers
Summer 2017 Section 001
Instructor: Tahmidul Islam

Question 01

One day Shiwen is very hungry and Jeff gives Shiwen a big box of chocolates. When Shiwen starts to eat, Jeff says 5 percent of chocolates are made with Carolina Reaper. Shiwen is too hungry to stop eating.

a Let X denote the number of chocolates Shiwen eats until the first one made with Carolina Reaper. Find $P(X > 3)$. Interpret what this probability means in words.

b In part (a), calculate $P(X > 5|X > 2)$. Interpret what this probability means in words. (Hint: use conditional probability formula).

c Compare $P(X > 3)$ and $P(X > 5|X > 2)$, which one is larger? Can you explain the reason?

d Suppose on that day Shiwen eats 30 pieces of chocolates. Find the probability that no more than two of those chocolates are mode with Carolina Reaper.

Solution:

(a)

$$\begin{aligned} X \sim& Geom(p = 0.05) \\ P(X > 3) =& 1 - P(X \leq 3) \\ =& 1 - (P(X = 1) + P(X = 2) + P(X = 3)) \\ =& 1 - (0.05 + 0.95 \times 0.05 + 0.95^2 \times 0.05) \\ =& 0.857375 \end{aligned}$$

```
#R code
1 - pgeom(2,p = .05)
```

Interpretation: There is about 86% chance that Shiwen has to eat more than 3 chocolates until he eats the first one made with Carolina Reaper.

Using R:

```
1 - pgeom(2,p = .05)
```

(b)

$$
\begin{aligned}
P(X > 5|X > 2) =& \frac{P(X > 5 \cap X > 2)}{P(X > 2)} \\
=& \frac{P(X > 5)}{P(X > 2)} \\
P(X > 5) =& 1 - P(X \leq 5) \\
=& 1 - \sum_{x=1}^{5} P(X = x) \\
=& 1 - \sum_{x=1}^{5} (1-p)^{x-1} p \\
=& 1 - p \sum_{x=1}^{5} (1-p)^{x-1} \\
=& 1 - 0.05(1 + 0.95 + 0.95^2 + 0.95^3 + 0.95^4) \\
=& 0.7737809 \\
P(X > 2) =& 1 - P(X \leq 2) \\
=& 1 - \sum_{x=1}^{2} P(X = x) \\
=& 1 - \sum_{x=1}^{2} (1-p)^{x-1} p \\
=& 1 - p \sum_{x=1}^{2} (1-p)^{x-1} \\
=& 1 - 0.05(1 + 0.95) \\
=& 0.9025 \\
P(X > 5|X > 2) =& \frac{0.7737809}{0.9025} \\
=& 0.857375
\end{aligned}
$$

```
#R code
(1 - pgeom(4,.05)) / ((1 - pgeom(1,.05)))
```

Interpretation: If Shiwen has already eaten more than 2 chocolates, there is about 86% chance that he will eat the first one made with Carolina Reaper after eating in total 5 chocolates (or after 3 more chocolates).

(c)

We have found that $P(X > 3) = P(X > 5|X > 2)$. That means, the probability of having the first Carolina Reaper after 3 tries is same as having the Reaper after 5 tries if we already know there was no Reaper in the first two tries. Geometric distribution has this special property called the "memoryless" property. Its like, the process forgot what happens in the first two tries and probability remains same as if the process restarted after 2nd try. (This is an interesting case, you can read from book or online to know more about it).

(d)

Let Y = Number of Carolina Reaper in those 30 chocolates. $Y \sim Bin(n = 30, p = 0.05); Y \in \{0, 1, 2, \ldots, 30\}$.

$$\begin{aligned}P(Y \leq 2) =& P(Y = 0) + P(Y = 1) + P(Y = 2)\\ =& \binom{30}{0} p^0 (1-p)^{30-0} + \binom{30}{1} p^1 (1-p)^{30-1} + \binom{30}{2} p^2 (1-p)^{30-2}\\ =& 0.8121788\end{aligned}$$

```
#R code
pbinom(2,30,.05)
```

Question 02

A recent geological study in western Texas indicates that exploratory oil wells strike with probability 0.20. (i.e., oil is found).

a Treating each well as a "trial", suppose that drilling wells in this region obeys the three Bernoulli trial assumptions. State what this would imply (i.e., describe the three assumptions in the background of the problem).

b What is the probability that the 1st successful well is found on the 2th well drilled?

c What is the probability that the 2th successful well is found on the 4th well drilled?

d What is the probability that it will take more than 2 wells to find the 1nd successful well?

Solution:

(a)

1. Drilling a well is a "trial". Each drilling can result in either "finding oil" or "not finding oil". We denote "finding oil" as success.
2. Result of drilling one well does not impact result of drilling another well (that is, trials are independent though may not be realistic in this situation).
3. Probability of oil strike is 20% for all the wells.

(b)

Let X = Number of wells needed to be drilled to have first success. $X \sim Geom(p = 0.20)$.

$$P(X = 2) = (1 - .2) = 0.16.$$

```
#R code
dgeom(1, p= .2)
```

(c)

Let Y = Number of wells needed to be drilled to have 2nd success. $Y \sim NegBin(p = 0.20, r = 2)$.

$$\begin{aligned}P(Y = 4) =& \binom{3}{1} 0.20^2 (1 - 0.80)^2\\ =& 0.0768.\end{aligned}$$

```
#R code
dnbinom(2, 2, 0.2)
```

(d)

$$P(X > 2) = 1 - P(X = 1) = 1 - 0.20 = 0.80.$$

Question 03

Jeff gives Shiwen a box of 100 pieces of chocolate, in which 5 are made with Carolina Reaper.

a Shiwen randomly eats 10 pieces of chocolate, what is the probability that none of them is made with Carolina Reaper?

b Shiwen randomly eats 10 pieces of chocolate, what is the probability that at most 2 pieces of chocolates are made with Carolina Reaper?

Solution:

(a)

Let X = Number of Carolina Reaper in those 10 pieces Shiwen ate. $X \sim HyperGeom(N = 100, n = 10, r = 5)$.

$$P(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

$$P(X = 0) = \frac{\binom{5}{0}\binom{95}{10}}{\binom{100}{10}} = 0.5837524.$$

```
#R code
dhyper(0,5,95,10)
```

(b)

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.9933621$$

```
#R code
phyper(2,5,95,10)
```

Question 04

Let Y be the number of calls received per day by the USC Campus Police. Suppose that Y has a Poisson distribution with $\lambda = 6.5$.

a What is the probability that on a given day there are exactly 5 calls? at least 5 calls? at most 5 calls?

b What is the mean of Y? Interpret it using the context of the problem.

c Suppose that the daily cost (in dollars) to respond to Y calls is given by

$$g(Y) = 150 + 100Y + 5Y^2.$$

Find the expected daily cost. (Hint: $var(Y) = E(Y^2) - [E(Y)]^2$).

d Suppose that in a given week, there are 30 calls received. Twelve of the calls involved illegal consumption of alcohol and 18 did not. If administration picks 5 cases (calls) to review at random, what is the probability that at least 4 of these cases will involve illegal consumption of alcohol?

Solution:

$Y \sim Pois(\lambda = 6.5)$.

(a)

$$\begin{aligned} P(Y = 5) =& \frac{e^{-\lambda}\lambda^y}{y!} \\ =& \frac{e^{-6.5}6.5^5}{5!} \\ =& 0.1453689. \\ P(Y \geq 5) =& 1 - P(Y < 5) = 1 - P(Y \leq 4) = 1 - \sum_{y=0}^{4} \frac{e^{-\lambda}\lambda^y}{y!} \\ =& 1 - e^{-\lambda} \sum_{y=0}^{4} \frac{\lambda^y}{y!} \\ =& 1 - e^{-6.5}[\frac{6.5^0}{0!} + \frac{6.5^1}{1!} + \frac{6.5^2}{2!} + \frac{6.5^3}{3!} + \frac{6.5^4}{4!}] \\ =& 0.7763282. \\ P(Y \leq 5) =& P(Y \leq 4) + P(Y = 5) \\ =& (1 - 0.7763282) + 0.1453689 = 0.3690407. \end{aligned}$$

```
#R code
dpois(5,6.5)
1 - ppois(4,6.5)
ppois(5,6.5)
```

(b) $E(Y) = \lambda = 6.5$. The long run average number of calls received by USC police is 6.5 calls per day. If we observe number of calls each day for a large number of days, we will see on average they receive 6.5 calls a day.

(c) Expected daily cost is $E(g(Y))$.

$$\begin{aligned} E(g(Y)) =& E(150 + 100Y + 5Y^2) \\ =& 150 + 100E(Y) + 5E(Y^2) \\ E(Y) =& 6.5. \\ Var(Y) =& \lambda \\ E(Y^2) - (E(Y))^2 =& 6.5 \\ E(Y^2) =& 6.5 + 6.5^2 \\ =& 48.75. \\ E(g(Y)) =& 150 + 100 \times 6.5 + 5 \times 48.75 \\ =& 1043.75. \end{aligned}$$

(d) Let X = Number of calls involving illegal alcohol consumption. $X \sim HyperGeom(N = 30, n = 5, r = 12)$.

$$\begin{aligned} P(X \geq 4) &= P(X = 4) + P(X = 5) \\ &= \frac{\binom{12}{4}\binom{18}{1}}{\binom{30}{5}} + \frac{\binom{12}{5}\binom{18}{0}}{\binom{30}{5}} \\ &= 0.06808134. \end{aligned}$$

```
#R code
dhyper(4, 12, 18, 5) + dhyper(5, 12, 18, 5)
```