CS 70 Discrete Mathematics and Probability Theory
Summer 2016 Dinh, Psomas, and Ye Discussion 5B Sol

1. Will I Get My Package?

A deceitful delivery dude is out transporting n packages to n customers. Not only does he hand a random package to each customer, but he also opens a package before delivering with probability $\frac{1}{2}$ (independently of the choice of the package). Let X be the number of customers who receive their own packages unopened.

1. Compute the expectation $\mathbb{E}(X)$.

   Define $X_i = \begin{cases} 1 & \text{if the } i\text{-th customer gets his/her own package unopened;} \\ 0 & \text{otherwise.} \end{cases}$

   $\mathbb{E}(X) = \mathbb{E}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbb{E}(X_i)$.

   $\mathbb{E}(X_i) = \Pr[X_i = 1] = \frac{1}{2n}$ since the i-th customer will get his/her package with probability $\frac{1}{n}$ and it is unopened with probability $\frac{1}{2}$ and packages are opened independently.

   Hence $\mathbb{E}(X) = n \cdot \frac{1}{2n} = \frac{1}{2}$. By linearity: $\mathbb{E}(X) = \mathbb{E}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbb{E}(X_i) = n \cdot \frac{1}{2n} = \frac{1}{2}$.

2. Calculate the probability that two particular customers i, j receive their own packages unopened.

   $\Pr[\text{both customers receive own packages}] = \frac{(n-2)!}{n!}$.

   $\Pr[i = \text{unopened}, j = \text{unopened}] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

   Since the two events are independent, $\Pr[X_i = 1, X_j = 1] = \frac{1}{4n(n-1)}$.

3. Compute $\text{Var}(X)$.

   $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

   $\mathbb{E}(X^2) = \mathbb{E}((X_1 + X_2 + \ldots + X_n)^2) = \mathbb{E}(\sum_{i,j} X_i X_j) = \sum_{i,j} \mathbb{E}(X_i X_j)$, where the last equality follows from using linearity of expectation.

   If $i = j$ then $\mathbb{E}(X_i X_j) = \mathbb{E}(X_i^2) = \frac{1}{2n}$; if $i \neq j$, then $\mathbb{E}(X_i X_j) = \Pr[X_i X_j = 1] \cdot 1 + \Pr[X_i X_j = 0] \cdot 0 = \frac{1}{4n(n-1)}$.

   Hence, $\mathbb{E}(X^2) = \sum_{i,j} \mathbb{E}(X_i X_j) = \sum_i \mathbb{E}(X_i^2) + \sum_{i \neq j} \mathbb{E}(X_i X_j) = n \cdot \frac{1}{2n} + n(n-1) \cdot \frac{1}{4n(n-1)} = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$.

   Thus, $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$.

2. (Telebears) Lydia has just started her Telebears appointment. She needs to register for a marine science class and CS70. There are no waitlists, and she can attempt to enroll once per day in either class or both. The Telebears system is strange and picky, so the probability of enrolling in the marine science class is p_1 and the probability of enrolling in CS70 is p_2. The probabilities are independent. Let M be the number of attempts it takes to enroll in the marine science class, and C be the number of attempts it takes to enroll in CS70.

1. What distribution do M and C follow? Are M and C independent?

   $M \sim \text{Geom}(p)$, $C \sim \text{Geom}(p)$ Yes they are independent.

2. For an integer $k \geq 1$, what is $\Pr[C \geq k]$?

   Question is asking for the probability that it takes at least k tries to enroll in CS70. $(1 - p_2)^{k-1}$.

3. What is the expected number of classes she will be enrolled in if she must enroll within 14 days (inclusive)?

   Let X be a random variable that is the number of classes she enrolls in within 14 days, and Y and Z be the indicator r.v.'s of whether she enrolls in the marine science class and CS 70, respectively. Then, $X = Y + Z$. So, $\mathbb{E}(X) = 1 \cdot \Pr[M \leq 14] + 1 \cdot \Pr[C \leq 14] = 1 - (1-p_1)^{14} + 1 - (1-p_2)^{14}$.

4. For an integer $k \geq 1$, what is the probability that she is enrolled in both classes before attempt k?

   Use independence. Let X be the number of attempts before she is enrolled in both. $\Pr[X < k] = \Pr[M < k]\Pr[C < k] = (1-(1-p_1)^{k-1})(1-(1-p_2)^{k-1})$.

3. Toujours les poissons Use the Poisson distribution to answer these questions.

1. Suppose that on average, 20 people ride your roller coaster per day. What is the probability that exactly 7 people ride it tomorrow?

   $X \sim \text{Poiss}(20)$. $\Pr[X = 7] = \frac{20^7}{7!}e^{-20} \approx 5.23 \cdot 10^{-4}$.

2. Suppose that on average, you go to Six Flags twice a year. What is the probability that you will go at most once in 2015?

   $X \sim \text{Poiss}(2)$. $\Pr[X \leq 1] = \frac{2^0}{0!}e^{-2} + \frac{2^1}{1!}e^{-2} \approx 0.41$.

3. Suppose that on average, there are 5.7 accidents per day on California roller coasters. (I hope this is not true.) What is the probability there will be *at least* 3 accidents throughout the *next two days* on California roller coasters?

   Let Y be the number of accidents that occur in the next two days. We can approximate Y as a Poisson distribution $Y \sim Poiss(\lambda = 11.4)$, where λ is the average number of accidents over two days. Now, we compute

$$
\begin{aligned}
\Pr[Y \geq 3] &= 1 - \Pr[Y < 3] \\
&= 1 - \Pr[Y = 0 \cup Y = 1 \cup Y = 2] \\
&= 1 - (\Pr[Y = 0] + \Pr[Y = 1] + \Pr[Y = 2]) \\
&= 1 - \left(\frac{11.4^0}{0!}e^{-11.4} + \frac{11.4^1}{1!}e^{-11.4} + \frac{11.4^2}{2!}e^{-11.4}\right) \\
&\approx 0.999.
\end{aligned}
$$

   We can show what we did above formally with the following claim: the sum of two independent Poisson random variables is Poisson. We won't prove this, but from the above, you should intuitively know why this is true. Now, we can model accidents on day i as a Poisson distribution $X_i \sim Poiss(\lambda = 5.7)$. Now, Let X_1 be the number of accidents that happen on the next day, and X_2 be the number of accidents that happen on the day after next. We are interested in $Y = X_1 + X_2$. Thus, we know $Y \sim Poiss(\lambda = 5.7 + 5.7 = 11.4)$.

4. Comparing Geometric Distributions

Suppose $X \sim Geom(p)$ and $Y \sim Geom(q)$. What is $Pr[X \geq Y]$ (This might be useful: $\sum_{i=1}^{\infty} a \cdot r^{i-1} = \frac{a}{1-r}$)?

$$
\begin{aligned}
Pr[X \geq Y] &= 1 - Pr[X < Y] \\
&= 1 - \sum_{k=1}^{\infty} Pr[X = k] Pr[Y > k] \\
&= 1 - \sum_{k=1}^{\infty} (1-p)^{k-1} p (1-q)^{k} \\
&= 1 - p(1-q) \sum_{k=1}^{\infty} (1-p)^{k-1} (1-q)^{k-1} \\
&= 1 - p(1-q) \cdot \frac{1}{1-(1-p)(1-q)} \\
&= \frac{p}{1-(1-p)(1-q)}
\end{aligned}
$$