CS 70 Discrete Mathematics and Probability Theory
Fall 2015 Jean Walrand HW 12

Due Wednesday November 18 at 10PM

Please note that understanding all questions on this homework is necessary to prepare for your midterm 3. Since midterm 3 is on November 16, solution will not be provided before the midterm. Please reach out to your TAs for help or ask questions on Piazza.

1. **Geometric Distribution**

   Two faulty machines, M_1 and M_2, are repeatedly run synchronously in parallel (i.e., both machines execute one run, then both execute a second run, and so on). On each run, M_1 fails with probability p_1 and M_2 fails with probability p_2, all failure events being independent. Let the random variables X_1, X_2 denote the number of runs until the first failure of M_1, M_2 respectively; thus X_1, X_2 have geometric distributions with parameters p_1, p_2 respectively. Let X denote the number of runs until the first failure of *either* machine. Show that X also has a geometric distribution, with parameter $p_1 + p_2 - p_1 p_2$.

   Answer: We have that $X_1 \sim \text{Geom}(p_1)$ and $X_2 \sim \text{Geom}(p_2)$. Also, X_1, X_2 are independent r.v.'s. We also use the following definition of the minimum:

$$\min(x,y) = \begin{cases} x & \text{if } x \leq y; \\ y & \text{if } x > y. \end{cases}$$

   Now, for all $k \in \{1, 2, \ldots\}$, $\min(X_1, X_2) = k$ is equivalent to $(X_1 = k) \cap (X_2 \geq k)$ or $(X_2 = k) \cap (X_1 > k)$. Hence,

$$\begin{aligned}
\Pr[X = k] &= \Pr[\min(X_1, X_2) = k] \\
&= \Pr[(X_1 = k) \cap (X_2 \geq k)] + \Pr[(X_2 = k) \cap (X_1 > k)] \\
&= \Pr[X_1 = k] \cdot \Pr[X_2 \geq k] + \Pr[X_2 = k] \cdot \Pr[X_1 > k] && \text{since } X_1, X_2 \text{ are independent} \\
&= [(1-p_1)^{k-1} p_1](1-p_2)^{k-1} + [(1-p_2)^{k-1} p_2](1-p_1)^k && \text{since } X_1, X_2 \text{ are geometric} \\
&= ((1-p_1)(1-p_2))^{k-1} (p_1 + p_2(1-p_1)) \\
&= (1 - p_1 - p_2 + p_1 p_2)^{k-1} (p_1 + p_2 - p_1 p_2).
\end{aligned}$$

   But this final expression is precisely the probability that a geometric r.v. with parameter $p_1 + p_2 - p_1 p_2$ takes the value k. Hence $X \sim \text{Geom}(p_1 + p_2 - p_1 p_2)$, and $\mathbb{E}(X) = \frac{1}{p_1 + p_2 - p_1 p_2}$.

   An alternative, slightly cleaner approach is to work with the *tail probabilities* of the geometric distribution, rather than with the usual point probabilities as above. In other words, we can work with $\Pr[X \geq k]$ rather than with $\Pr[X = k]$; clearly the values $\Pr[X \geq k]$ specify the values $\Pr[X = k]$ since $\Pr[X = k] = \Pr[X \geq k] - \Pr[X \geq (k+1)]$, so it suffices to calculate them instead. We then get the following argument:

$$\begin{aligned}
\Pr[X \geq k] &= \Pr[\min(X_1, X_2) \geq k] \\
&= \Pr[(X_1 \geq k) \cap (X_2 \geq k)] \\
&= \Pr[X_1 \geq k] \cdot \Pr[X_2 \geq k] && \text{since } X_1, X_2 \text{ are independent} \\
&= (1-p_1)^{k-1}(1-p_2)^{k-1} && \text{since } X_1, X_2 \text{ are geometric} \\
&= ((1-p_1)(1-p_2))^{k-1} \\
&= (1-p_1-p_2+p_1p_2)^{k-1}.
\end{aligned}$$

This is the tail probability of a geometric distribution with parameter $p_1+p_2-p_1p_2$, so we are done.

2. **Poisson Distribution**

   (a) It is fairly reasonable to model the number of customers entering a shop during a particular hour as a Poisson random variable. Assume that this Poisson random variable X has mean λ. Suppose that whenever a customer enters the shop they leave the shop without buying anything with probability p. Assume that customers act independently, i.e. you can assume that they each simply flip a biased coin to decide whether to buy anything at all. Let us denote the number of customers that buy something as Y and the number of them that do not buy anything as Z (so $X = Y + Z$). What is the probability that $Y = k$ for a given k? How about $\Pr[Z = k]$? Prove that Y and Z are Poisson random variables themselves.
   Hint: you can use the identity $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

   (b) Assume that you were given two independent Poisson random variables X_1, X_2. Assume that the first has mean λ_1 and the second has mean λ_2. Prove that $X_1 + X_2$ is a Poisson random variable with mean $\lambda_1 + \lambda_2$.

   Answer:

   (a) We consider all possible ways that the event $Y = k$ might happen: namely, $k + j$ people enter the shop ($X = k + j$) and then exactly k of them choose to buy something. That is,

$$\begin{aligned}
\Pr[Y=k] &= \sum_{j=0}^{\infty} \Pr[X=k+j] \cdot \Pr[Y=k \mid X=k+j] \\
&= \sum_{j=0}^{\infty} \left(\frac{\lambda^{k+j}}{(k+j)!}e^{-\lambda}\right) \cdot \left(\binom{k+j}{k} p^j (1-p)^k\right) \\
&= \sum_{j=0}^{\infty} \frac{\lambda^{k+j}}{(k+j)!}e^{-\lambda} \cdot \frac{(k+j)!}{k!j!} p^j (1-p)^k \\
&= \frac{(\lambda(1-p))^k e^{-\lambda}}{k!} \cdot \sum_{j=0}^{\infty} \frac{(\lambda p)^j}{j!} \\
&= \frac{(\lambda(1-p))^k e^{-\lambda}}{k!} \cdot e^{\lambda p} \\
&= \frac{(\lambda(1-p))^k e^{-\lambda(1-p)}}{k!}
\end{aligned}$$

   Hence, Y follows the Poisson distribution with parameter $\lambda(1-p)$. The case for Z is completely analogous: $\Pr[Z=k] = \frac{(\lambda p)^k e^{-\lambda p}}{k!}$ and Z follows the Poisson distribution with parameter λp.

(b) To show that X_1+X_2 is a Poisson random variable with mean $\lambda_1+\lambda_2$, we have show that $\Pr[(X_1+X_2)=i] = \frac{(\lambda_1+\lambda_2)^i}{i!}e^{-(\lambda_1+\lambda_2)}$.

$$\begin{aligned}
\Pr[(X_1+X_2)=i] &= \sum_{k=0}^{i} \Pr[X_1=k, X_2=(i-k)] \\
&= \sum_{k=0}^{i} \frac{\lambda_1^k}{k!}e^{-\lambda_1} \cdot \frac{\lambda_2^{i-k}}{(i-k)!}e^{-\lambda_2} \\
&= e^{-\lambda_1}e^{-\lambda_2} \sum_{k=0}^{i} \frac{1}{k!(i-k)!}\lambda_1^k\lambda_2^{i-k} \\
&= \frac{e^{-\lambda_1}e^{-\lambda_2}}{i!} \sum_{k=0}^{i} \frac{i!}{k!(i-k)!}\lambda_1^k\lambda_2^{i-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{i!} \sum_{k=0}^{i} \binom{i}{k}\lambda_1^k\lambda_2^{i-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{i!}(\lambda_1+\lambda_2)^i
\end{aligned}$$

To go from the second-to-last line to the last line, we use the binomial expansion.

3. **Variance**

This problem will give you practice using the "standard method" to compute the variance of a sum of random variables that are not pairwise independent (so you cannot use "linearity" of variance). If you don't even know what is "linearity" of variance, read the lecture note and slides first.

(a) A building has n floors numbered $1,2,\ldots,n$, plus a ground floor G. At the ground floor, m people get on the elevator together, and each gets off at a uniformly random one of the n floors (independently of everybody else). What is the *variance* of the number of floors the elevator *does not* stop at? (In fact, the variance of the number of floors the elevator *does* stop at must be the same (do you see why?) but the former is a little easier to compute.)

(b) A group of three friends has n books they would all like to read. Each friend (independently of the other two) picks a random permutation of the books and reads them in that order, one book per week (for n consecutive weeks). Let X be the number of weeks in which all three friends are reading the same book. Compute VarX.

Answer:

(a) Let X be the number of floors the elevator does not stop at. As in the previous homework, we can represent X as the sum of the indicator variables $X_1,\ldots,X_n$, where $X_i=1$ if no one gets off on floor i. Thus, we have

$$\mathbb{E}(X_i) = \Pr[X_i=1] = \left(\frac{n-1}{n}\right)^m,$$

and from linearity of expectation,

$$\mathbb{E}(X) = \sum_{i=1}^{n}\mathbb{E}(X_i) = n\left(\frac{n-1}{n}\right)^m.$$

To find the variance, we cannot simply sum the variance of our indicator variables. However, we can still compute $\mathrm{Var X} = \mathbb{E}(\mathrm{X}^2) - \mathbb{E}(\mathrm{X})^2$ directly using linearity of expectation, but now how can we find $\mathbb{E}(X^2)$? Recall that

$$\begin{aligned}
\mathbb{E}(X^2) &= \mathbb{E}((X_1 + \ldots + X_n)^2) \\
&= \mathbb{E}(\sum_{i,j} X_i X_j) \\
&= \sum_{i,j} \mathbb{E}(X_i X_j) \\
&= \sum_{i}^{n} \mathbb{E}(X_i^2) + \sum_{i \neq j} \mathbb{E}(X_i X_j).
\end{aligned}$$

The first term is simple to calculate: $\mathbb{E}(X_i^2) = 1^2 \Pr[X_i = 1] = \left(\frac{n-1}{n}\right)^m$, meaning that

$$\sum_{i=1}^{n} \mathbb{E}(X_i^2) = n \left(\frac{n-1}{n}\right)^m.$$

$X_i X_j = 1$ when both X_i and X_j are 1, which means no one gets off the elevator on floor i and floor j. This happens with probability

$$\Pr[X_i = X_j = 1] = \Pr[X_i = 1 \cap X_j = 1] = \left(\frac{n-2}{n}\right)^m.$$

Thus, we can now compute

$$\sum_{i \neq j} \mathbb{E}(X_i X_j) = n(n-1) \left(\frac{n-2}{n}\right)^m.$$

Finally, we plug in to see that

$$\mathrm{Var X} = \mathbb{E}(\mathrm{X}^2) - \mathbb{E}(\mathrm{X})^2 = \mathrm{n} \left(\frac{\mathrm{n}-1}{\mathrm{n}}\right)^{\mathrm{m}} + \mathrm{n}(\mathrm{n}-1) \left(\frac{\mathrm{n}-2}{\mathrm{n}}\right)^{\mathrm{m}} - \left(\mathrm{n} \left(\frac{\mathrm{n}-1}{\mathrm{n}}\right)^{\mathrm{m}}\right)^2.$$

(b) Let $X_1, \ldots, X_n$ be indicator variables such that $X_i = 1$ if all three friends are reading the same book on week i. Thus, we have

$$\mathbb{E}(X_i) = \Pr[X_i = 1] = \left(\frac{1}{n}\right)^2,$$

and from linearity of expectation,

$$\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X_i) = n \left(\frac{1}{n}\right)^2 = \frac{1}{n}.$$

As before, we know that

$$\mathbb{E}(X^2) = \sum_{i}^{n} \mathbb{E}(X_i^2) + \sum_{i \neq j} \mathbb{E}(X_i X_j).$$

Furthermore, because X_i is an indicator variable, $\mathbb{E}(X_i^2) = 1^2 \Pr[X_i = 1] = \left(\frac{1}{n}\right)^2$, and

$$\sum_{i}^{n} \mathbb{E}(X_i^2) = n \left(\frac{1}{n}\right)^2 = \frac{1}{n}.$$

Again, because X_i and X_j are indicator variables, we are interested in

$$\Pr[X_i = X_j = 1] = \Pr[X_i = 1 \cap X_j = 1] = \frac{1}{(n(n-1))^2},$$

the probability that all three friends pick the same book on week i and week j. Thus,

$$\sum_{i \neq j} \mathbb{E}(X_i X_j) = n(n-1)\left(\frac{1}{(n(n-1))^2}\right) = \frac{1}{n(n-1)}.$$

Finally, we compute

$$\mathrm{Var X} = \mathbb{E}(\mathrm{X}^2) - \mathbb{E}(\mathrm{X})^2 = \frac{1}{\mathrm{n}} + \frac{1}{\mathrm{n(n-1)}} - \left(\frac{1}{\mathrm{n}}\right)^2.$$

4. **Coupon Collection**

   Suppose you take a deck of n cards and repeatedly perform the following step: take the current top card and put it back in the deck at a uniformly random position. (I.e., the probability that the card is placed in any of the n possible positions in the deck — including back on top — is $1/n$.) Consider the card that starts off on the bottom of the deck. What is the expected number of steps until this card rises to the top of the deck? (Hint: Let T be the number of steps until the card rises to the top. We have $T = T_n + T_{n-1} + \cdots + T_2$, where the random variable T_i is the number of steps until the bottom card rises from position i to position $i-1$. Thus, for example, T_n is the number of steps until the bottom card rises off the bottom of the deck, and T_2 is the number of steps until the bottom card rises from second position to top position. What is the distribution of T_i?

   Answer: Since a card at location i moves to location $i-1$ when the current top card is placed in any of the locations $i, i+1, \ldots, n$, it will rise with probability $p = \frac{n-i+1}{n}$. Thus, $T_i \sim \mathrm{Geom}(p)$, and $\mathbb{E}(T_i) = \frac{1}{p} = \frac{n}{n-i+1}$. We now can see how this is exactly the coupon collector's problem, but with one fewer term (namely, without T_1). Finally, we can apply linearity of expectation to compute

$$\mathbb{E}(T) = \sum_{i=2}^{n} \mathbb{E}(T_i) = \sum_{i=2}^{n} \frac{n}{n-i+1} = n \sum_{i=2}^{n} \frac{1}{n-i+1} \approx n \ln(n-1)$$

5. **Markov's Inequality and Chebyshev's Inequality**

   A random variable X has variance $\mathrm{Var X} = 9$ and expectation $\mathbb{E}(X) = 2$. Furthermore, the value of X is never greater than 10. Given this information, provide either a proof or a counterexample for the following statements.

   (a) $\mathbb{E}(X^2) = 13$.

   (b) $\Pr[X = 2] > 0$.

   (c) $\Pr[X \geq 2] = \Pr[X \leq 2]$.

   (d) $\Pr[X \leq 1] \leq 8/9$.

   (e) $\Pr[X \geq 6] \leq 9/16$.

   (f) $\Pr[X \geq 6] \leq 9/32$.

   Answer:

(a) TRUE. Since $9 = \mathrm{Var X} = \mathbb{E}(\mathrm{X}^2) - \mathbb{E}(\mathrm{X})^2 = \mathbb{E}(\mathrm{X}^2) - 2^2$, we have $\mathbb{E}(X^2) = 9+4 = 13$.

(b) FALSE. Construct a random variable X that satisfies the conditions in the question but does not take on the value 2. A simple example would be a random variable that takes on 2 values, where $\Pr[X=a] = \frac{1}{2}, \Pr[X=b] = \frac{1}{2}$, and $a \neq b$. The expectation must be 2, so we have $\frac{1}{2}a + \frac{1}{2}b = 2$. The variance is 9, so $\mathbb{E}(X^2) = 13$ (from part (a)) and $\frac{1}{2}a^2 + \frac{1}{2}b^2 = 13$. Solving for a and b, we get $\Pr[X=-1] = \frac{1}{2}, \Pr[X=5] = \frac{1}{2}$ as a counterexample.

(c) FALSE. Construct a random variable X that satisfies the conditions in the question but does not have an equal chance of being less than or greater than 2. A simple example would be a random variable that takes on 2 values, where $\Pr[X=a] = p, \Pr[X=b] = 1-p$. Here, we use the same approach as part (b) except with a generic p, since we want $p \neq \frac{1}{2}$. The expectation must be 2, so we have $pa + (1-p)b = 2$. The variance is 9, so $\mathbb{E}(X^2) = 13$ and $pa^2 + (1-p)b^2 = 13$. Solving for a and b, we find the relation $b = 2 \pm \frac{3}{\sqrt{x}}$, where $x = \frac{1-p}{p}$. Then, we can find an example by plugging in values for x so that $a, b \leq 10$ and $p \neq \frac{1}{2}$. One such counterexample is $\Pr[X=-7] = \frac{1}{10}, \Pr[X=3] = \frac{9}{10}$.

(d) TRUE. Let $Y = 10 - X$. Since X is never exceeds 10, Y is a non-negative random variable. By Markov's inequality,

$$\Pr[10 - X \geq a] = \Pr[Y \geq a] \leq \frac{\mathbb{E}(Y)}{a} = \frac{\mathbb{E}(10-X)}{a} = \frac{8}{a}.$$

Setting $a = 9$, we get $\Pr[X \leq 1] = \Pr[10 - X \geq 9] \leq \frac{8}{9}$.

(e) TRUE. Chebyshev's inequality says $\Pr[|X - \mathbf{E}[X]| \geq a] \leq \frac{\mathrm{Var X}}{a^2}$. If we set $a = 4$, we have

$$\Pr[|X-2| \geq 4] \leq \frac{9}{16}.$$

Now we simply observe that $\Pr[X \geq 6] \leq \Pr[|X-2| \geq 4]$, because the event $X \geq 6$ is a subset of the event $|X-2| \geq 4$.

(f) FALSE. We use the same approach as in part (c), except we find a counterexample that fits the inequality $\Pr[X \geq 6] \leq 9/32$. One example is $\Pr[X=0] = \frac{9}{13}, \Pr[X = \frac{13}{2}] = \frac{4}{13}$.

6. **Umbrella Store**

Bob has a store that sells umbrellas. The number of umbrellas that Bob sells on a rainy day is a random variable Y with mean 25 and standard deviation $\sqrt{105}$. But if it is a clear day, Bob doesn't sell any umbrellas at all. The weather forecast for tomorrow says it will rain with probability $\frac{1}{5}$. Let Z be the number of umbrellas that Bob sells tomorrow.

(a) Let X be an indicator random variable that it will rain tomorrow. Write Z in terms of X and Y.

(b) What is the mean and standard deviation of Z?

(c) Use Chebyshev's inequality to bound the probability that Bob sells at least 25 umbrellas tomorrow.

Answer:

(a) $Z = XY$.

(b) We have

$$\mathbb{E}(Z) = \Pr[X = 1] \cdot \mathbb{E}(Y) + \Pr[X = 0] \cdot 0 = \frac{1}{5}\mathbb{E}(Y) = \frac{1}{5} \cdot 25 = 5$$

and

$$\mathbb{E}(Z^2) = \Pr[X = 1] \cdot \mathbb{E}(Y^2) + \Pr[X = 0] \cdot 0 = \frac{1}{5}\mathbb{E}(Y^2) = \frac{1}{5} \cdot 730 = 146$$

since $\mathbb{E}(Y^2) = \mathrm{VarY} + \mathbb{E}(\mathrm{Y})^2 = 105 + 625 = 730$. So

$$\mathrm{VarZ} = \mathbb{E}(\mathrm{Z}^2) - \mathbb{E}(\mathrm{Z})^2 = 146 - 25 = 121.$$

Therefore, the mean of Z is 5 and the standard deviation is $\sqrt{121} = 11$.

(c) Since $\mathbb{E}(Z) = 5$ and $\mathrm{VarZ} = 121$,

$$\Pr[Z \geq 25] \;=\; \Pr[Z - \mathbb{E}(Z) \geq 20] \;\leq\; \Pr[|Z - \mathbb{E}(Z)| \geq 20] \;\leq\; \frac{\mathrm{VarZ}}{400} \;=\; \frac{121}{400}.$$