CS 70 Discrete Mathematics and Probability Theory
Fall 2015 Jean Walrand HW 11

Due Wednesday Nov 11 at 10PM

1. **Independence**

   Let's start with the basics. You have two friends, Olivia and Fitz. We're interested in the things they do. Specifically, we're interested in whether Olivia eats popcorn and whether Fitz retires to Vermont on any given day. For the purposes of this question, we can assume these events are independent.

   (a) Show that Olivia not eating popcorn and Fitz retiring to Vermont are independent. That is, $\bar{O}$ and F are independent.

   Answer:
   Recall that if A and B are independent then $Pr[A \cap B] = Pr[A] \cdot Pr[B]$.
   $Pr[\bar{O} \cap F] = Pr[F] - Pr[O \cap F]$
   Since O and F are independent, $Pr[O \cap F] = Pr[O] \cdot Pr[F]$. Thus:
   $Pr[F] - Pr[O \cap F] = Pr[F] - Pr[O] \cdot Pr[F]$
   $Pr[F] - Pr[O] \cdot Pr[F] = Pr[F] \cdot (1 - Pr[O]) = Pr[F] \cdot Pr[\bar{O}]$
   We conclude $Pr[\bar{O} \cap F] = Pr[F] \cdot Pr[\bar{O}]$. They are independent.

   (b) Show that Olivia eating popcorn and Fitz not retiring to Vermont are independent. That is, O and $\bar{F}$ are independent.

   Answer:
   As above.
   $Pr[O \cap \bar{F}] = Pr[O] - Pr[O \cap F] = Pr[O] - Pr[O] \cdot Pr[F] = Pr[O] \cdot (1 - Pr[F]) = Pr[O] \cdot Pr[\bar{F}]$

   (c) Show that Olivia not eating popcorn and Fitz not retiring to Vermont are independent. That is, $\bar{O}$ and $\bar{F}$ are independent.

   Answer:
   As above.
   $Pr[\bar{O} \cap \bar{F}] = Pr[\bar{O}] - Pr[\bar{O} \cap F] = Pr[\bar{O}] - Pr[\bar{O}] \cdot Pr[F] = Pr[\bar{O}] \cdot (1 - Pr[F]) = Pr[\bar{O}] \cdot Pr[\bar{F}]$

   (d) You've made a new friend Cyrus, who sometimes yells passionately at people. Olivia eating popcorn, Fitz retiring to Vermont, and Cyrus yelling are mutually independent events. Show that Olivia eating popcorn and Fitz retiring to Vermont $\cap$ Cyrus yelling are independent. That is, O and $(F \cap C)$ are independent.

   Answer:
   $Pr[O \cap (F \cap C)] = Pr[O \cap F \cap C]$. From the definition of mutual independence, $Pr[O \cap F \cap C] = Pr[O] \cdot Pr[F] \cdot Pr[C]$. And $Pr[F \cap C] = Pr[F] \cdot Pr[C]$, so $Pr[O] \cdot Pr[F] \cdot Pr[C] = Pr[O] \cdot Pr[F \cap C]$. Thus $Pr[O \cap (F \cap C)] = Pr[O] \cdot Pr[F \cap C]$. These events are independent.

   (e) Show that Olivia eating popcorn and Fitz retiring to Vermont Δ Cyrus yelling are independent. That is, O and $(F \Delta C)$ are independent. Recall that Δ is the symmetric difference of two sets. $A \Delta B = (B - A) \cup (A - B)$.

Answer:
$Pr[O\cap(F\Delta C)] = Pr[O\cap((C-F)\cup(F-C))] = Pr[O\cap(C-F)] + Pr[O\cap(F-C))] = Pr[O\cap C\cap\bar{F}] + Pr[O\cap F\cap\bar{C}]$.
Let's attack $Pr[O\cap C\cap\bar{F}]$ first. $Pr[O\cap C\cap\bar{F}] = Pr[O\cap C] - Pr[O\cap C\cap F] = Pr[O]\cdot Pr[C] - Pr[O]\cdot Pr[C]\cdot Pr[F] = Pr[O]\cdot Pr[C]\cdot(1-Pr[F]) = Pr[O]\cdot Pr[C]\cdot Pr[\bar{F}]$.
Likewise $Pr[O\cap F\cap\bar{C}] = Pr[O]\cdot Pr[F]\cdot Pr[\bar{C}]$.
Thus $Pr[O\cap(F\Delta C)] = Pr[O]\cdot Pr[C]\cdot Pr[\bar{F}] + Pr[O]\cdot Pr[F]\cdot Pr[\bar{C}] = Pr[O]\cdot(Pr[C]\cdot Pr[\bar{F}] + Pr[F]\cdot Pr[\bar{C}]) = Pr[O]\cdot Pr[(C-F)\cup(F-C)] = Pr[O]\cdot Pr[F\Delta C]$. These events are independent.

(f) More generally, let $\{A_j, j\in J\}$ be mutually independent. Show that $\{\cap_{j\in J_k}A_j, k\in K\}$ are mutually independent if the subsets J_k of J are pairwise disjoint.

Answer:
For each k, $Pr[\cap_{j\in J_k}A_j] = \prod_{j\in J_k}Pr[A_j]$ because $\{A_j, j\in J_k\}$ are mutually independent. Then because the subsets J_k of J are pairwise disjoint and $\{A_j, j\in J\}$ are mutually independent $Pr[\cap_{k\in K}(\cap_{j\in J_k}A_j)] = Pr[\prod_{k\in K}(\prod_{j\in J_k}Pr[A_j])]$. Thus $\{\cap_{j\in J_k}A_j, k\in K\}$ are mutually independent.

2. Fundamentals

True or false? For the following statements, provide either a proof or a simple counterexample. Let X,Y,Z be arbitrary random variables.

(a) If (X,Y) are independent and (Y,Z) are independent, then (X,Z) are independent.

Answer: FALSE. Let X,Y be i.i.d Bernoulli(1/2) random variables, and let $Z=X$. Then (X,Y) and (Y,Z) are independent by construction, but (X,Z) are not independent because $Z=X$.

(b) If (X,Y) are dependent and (Y,Z) are dependent, then (X,Z) are dependent.

Answer: FALSE. Let X,Z be i.i.d Bernoulli(1/2) random variables, and let $Y=XZ$. Then (X,Z) are independent by construction, but (X,Y) are not independent, since $\Pr[X=0\wedge Y=1] = 0 \neq \Pr[X=0]\Pr[Y=1] = (1/2)(1/4)$.

(c) Assume X is discrete. If $\mathrm{Var}(X)=0$, then X is a constant.

Answer: TRUE. Let $\mu = \mathbf{E}[X]$. By definition,

$$0 = \mathrm{Var}(X) = \mathbf{E}[(X-\mu)^2] = \sum_{\omega\in\Omega}\Pr[\omega](X(\omega)-\mu)^2$$

The RHS is the sum of non-negative numbers, so if the sum is 0, each term must be 0.
So $\Pr[\omega] > 0 \implies (X(\omega)-\mu)^2 = 0 \implies X(\omega) = \mu$. Therefore X is constant (equal to $\mu = \mathbf{E}[X]$).

(d) $\mathbf{E}[X]^4 \leq \mathbf{E}[X^4]$

Answer: TRUE. First, for an arbitrary random variable Y, we have:

$$0 \leq \mathbf{E}[(Y-\mathbf{E}[Y])^2] = \mathbf{E}[Y^2] - \mathbf{E}[Y]^2$$

So $\mathbf{E}[Y]^2 \leq \mathbf{E}[Y^2]$. Now applying this twice, once for $Y=X$ and once for $Y=X^2$:

$$\mathbf{E}[X]^4 = (\mathbf{E}[X]^2)^2 \leq (\mathbf{E}[X^2])^2 \leq \mathbf{E}[(X^2)^2] = \mathbf{E}[X^4]$$

3. Linearity of expectation

Solve each of the following problems using linearity of expectation. Explain your methods clearly.

(a) In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability $1/3$ (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability $1/5$, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

Answer: Let A_i be the indicator you win the i^{th} time you play game A and B_i be the same for game B. The expected value of A_i and B_i are,

$$\mathbf{E}[A_i] = 1 \cdot 1/3 + 0 \cdot 2/3 = 1/3,$$
$$\mathbf{E}[B_i] = 1 \cdot 1/5 + 0 \cdot 4/5 = 1/5.$$

Let T_A be the random variable for the number of tickets you win in game A, and T_B be the number of tickets you win in game B.

$$\begin{aligned}\mathbf{E}[T_A + T_B] &= 3\mathbf{E}[A_1] + \cdots + 3\mathbf{E}[A_{10}] + 4\mathbf{E}[B_1] + \cdots + 4\mathbf{E}[B_{20}] \\ &= 10\left(3 \cdot \frac{1}{3}\right) + 20\left(4 \cdot \frac{1}{5}\right) = 26\end{aligned}$$

□

(b) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "book" appears?

Answer: There are $1,000,000 - 4 + 1 = 999,997$ places where "book" can appear, each with a (non-independent) probability of $\frac{1}{26^4}$ of happening. If A is the random variable that tells how many times "book" appears, and A_i is the indicator variable that is 1 if "book" appears starting at the i^{th} letter, then

$$\begin{aligned}\mathbf{E}[A] &= \mathbf{E}[A_1 + \cdots + A_{999,997}] \\ &= \mathbf{E}[A_1] + \cdots + E[A_{999,997}] \\ &= \frac{999,997}{26^4} \approx 2.19\end{aligned}$$

times. □

(c) 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 expected number of floors the elevator stops at (not counting the ground floor)?

Answer: Let A_i be the indicator that the elevator stopped at floor i.

$$\Pr[A_i = 1] = 1 - \Pr[\text{no one gets off at } i] = 1 - \left(\frac{n-1}{n}\right)^m.$$

If A is the number of floors the elevator stops at, then

$$\begin{aligned}\mathbf{E}[A] &= \mathbf{E}[A_1 + \cdots + A_n] \\ &= \mathbf{E}[A_1] + \cdots + E[A_n] = n \cdot \left(1 - \left(\frac{n-1}{n}\right)^m\right)\end{aligned}$$

□

4. **Company Selection**

Company A produces a particular device consisting of 10 components. Company A can either buy all the components from Company S or Company T, and then uses them to produce the devices without testing every individual component. After that, each device will be tested before leaving the factory. The device works only if every component works properly. Each working device can be sold for x dollars, but each non-working device must be thrown away. Products from Company S have a failure probability of $q = 0.01$ while Company T has a failure probability of $q/2$. However, every component from Company S costs \$10 while it costs \$30 from Company T. Should Company A build the device with components from Company S or Company T in order to maximize its expected profit per device? (Hint: Your answer will depend on x.)

Answer:

Let W denote the event that a device works. Let R be the random variable denoting the profit.

$$\mathbf{E}[R] = P(W)\mathbf{E}[R|W] + P(W^C)\mathbf{E}[R|W^C].$$

Let's first consider the case when we use products from Company S. In this case, a device works with probability $P[W] = (1-q)^{10}$. The profit made on a working device is $x - 100$ dollars while a nonworking device has a profit of -100 dollars. That is, $\mathbf{E}[R|W] = x - 100$ and $\mathbf{E}[R|W^C] = -100$. Using R_S to denote the profit using components from Company S, the expected profit is:

$$\mathbf{E}[R_S] = (1-q)^{10}(x-100) + (1-(1-q)^{10})(-100) = (1-q)^{10}x - 100 = (0.99)^{10}x - 100.$$

If we use products from Company T. The device works with probability $P[W] = (1-q/2)^{10}$. The profit per working device is $\mathbf{E}[R|W] = x - 300$ dollars while the profit for a nonworking device is $\mathbf{E}[R|W^C] = -300$ dollars. The expected profit is:

$$\mathbf{E}[R_T] = (1-q/2)^{10}(x-300) + (1-(1-q/2)^{10})(-300) = (1-q/2)^{10}x - 300 = (0.995)^{10}x - 300.$$

To determine which Company should we use, we solve $\mathbf{E}[R_T] \geq \mathbf{E}[R_S]$, yielding $x \geq 200/[(0.995)^{10} - (0.99)^{10}] = 4280.1$. So for $x < \$4280.1$ using products from Company S results in greater profit, while for $x > \$4280.1$ more profit will be generated by using products from Company T.

5. **Round the Clock**

You have decided to try an unusual meal scheme. You will have three meals a day, with no exceptions. You will have each meal at the same time everyday. Most importantly, you have decided to pick those times randomly. Today you make the schedule, and then you'll follow the schedule from now on. You will select your three meals' times randomly from the set of 24 hours $\{12{:}00\text{am}, 1{:}00\text{am}, \ldots, 10{:}00\text{pm}, 11{:}00\text{pm}\}$ and from now on you will have that meal at that exact time every day. Note that you might even have two meals at the same time (i.e. you sample with replacement)! But if you schedule two or more meals at the same time, you decide on a completely random ordering over the meals scheduled for that time, and then you'll have your meals according to that ordering every day. Furthermore, for simplicity assume that it takes no time for you to finish a meal.

(a) For each one of your meals, what is the expected amount of time you have to wait after finishing it, until you get to have your next meal?

Answer: Let X_1, X_2, X_3 be the random variables, where X_i shows how long we have to wait until the next meal after we finish meal i. By linearity of expectation we have $\mathbf{E}(X_1 + X_2 + X_3) =$

$\mathbf{E}(X_1)+\mathbf{E}(X_2)+\mathbf{E}(X_3)$. But note that $X_1+X_2+X_3$ is always 24, because if at every point in the day there is exactly one meal that was finished last, and so that point in time could be attributed to that meal. This way we can see that entire 24 hours of a day are perfectly partitioned by the intervals we are waiting after meals 1, 2, and 3. So $X_1+X_2+X_3=24$. But then everything in the problem statement is symmetric about the meals. We don't even label them! So it must be true that X_1, X_2, and X_3 have the same expected value. So we get

$$24=\mathbf{E}(X_1+X_2+X_3)=\mathbf{E}(X_1)+\mathbf{E}(X_2)+\mathbf{E}(X_3)=3\times\mathbf{E}(X_1).$$

This means that $\mathbf{E}(X_1)=8$, and similarly $\mathbf{E}(X_2)=\mathbf{E}(X_3)=8$. So for each meal, we have to wait an expected 8 hours after finishing it for the next meal.

(b) Now suppose that your friend calls you at an exact random hour during the day (24 different possibilities). S/he asks you to go to his/her place and have your next meal with him/her. You do not want to ruin your schedule, so you tell your friend that you'll be at his/her place in X hours to have your next meal, where $X \geq 0$ is the number of hours until your next meal is scheduled. **What is the expected value of X?** (If a phone call falls on the exact hour of a meal, break ties randomly as in part (a)).

Answer: We use the following observation: the phone call is not that different from a meal! Its time is random, just like the meals, and even when its scheduled time is tied with some of the meals, the ordering of the meals and the phone call becomes perfectly uniformly random. Therefore if we think of a phone call as a fourth meal, then by following a similar reasoning as in the previous part we get that the expected time we have to wait after each meal until the next one is $\frac{24}{4}=6$. In particular, after having the fourth meal, the phone call, the expected amount of time we need to wait for the next meal is 6 hours.

Side note: Naively one might think that the answer should be half the answer to the previous part, because the in-between-meals interval in which the phone call happens has expected length 8 and so it might seem OK to say that on average the phone call happens in the middle of the interval, so the time from the phone call until the next meal is 4. But this reasoning is wrong, because when everything is random, there are in-between-meals intervals that are larger than 8 hours and there are also intervals smaller than 8 hours. If the phone call happens in intervals larger than 8 hours, then the average value for X would be larger than 4. One might think that the times the phone call happens in large intervals cancel out with the times the phone call happens in smaller intervals, but this is not true. A phone call is placed at a random time, and therefore it is more likely to fall into a larger interval, than in a small interval.