Solutions 3

1. For each statement below, either show that it is true or give a counterexample. Throughout, X, Y, Z are discrete random variables.

   (a) If X and Y are independent and Y and Z are independent, then X and Z are independent.
   False: for a simple example, take X = Z.

   (b) If X and Y are independent, then they are conditionally independent given Z.
   False: this was discussed in class (the fire-popcorn example) in terms of events, for which we can let X, Y, Z be indicators.

   (c) If X and Y are conditionally independent given Z, then they are independent.
   False: this was discussed in class in terms of events (the chess opponent of unknown strength example); a coin with a random bias (as on HW 2) is another simple, useful example to keep in mind.

   (d) If X and Y have the same distribution given Z, i.e., for all a and z, we have P(X = a|Z = z) = P(Y = a|Z = z), then X and Y have the same distribution.
   True: by the law of total probability, conditioning on Z gives
   $$P(X = a) = \sum_z P(X = a|Z = z)P(Z = z).$$
   Since P X and Y have the same conditional distribution given Z, this becomes $\sum_z P(Y = a|Z = z)P(Z = z) = P(Y = a)$.

2. Consider the following conversation from an episode of The Simpsons:

   Lisa: Dad, I think he's an ivory dealer! His boots are ivory, his hat is ivory, and I'm pretty sure that check is ivory.
   Homer: Lisa, a guy who's got lots of ivory is less likely to hurt Stampy than a guy whose ivory supplies are low.

   Here Homer and Lisa are debating the question of whether or not the man (named Blackheart) is likely to hurt Stampy the Elephant if they sell Stampy to him. They clearly disagree about how to use their observations about Blackheart to learn about the probability (conditional on the evidence) that Blackheart will hurt Stampy.

   (a) Define clear notation for the various events of interest here.

   Solution: Let H be the event that the man will hurt Stampy, let L be the event that a man has lots of ivory, and let D be the event that the man is an ivory dealer.

(b) Express Lisa's and Homer's arguments (Lisa's is partly implicit) as conditional probability statements in terms of your notation from (a).

Solution: Lisa observes that L is true. She suggests (reasonably) that this evidence makes D more likely, i.e., P(D|L) > P(D). Implicitly, she suggests that this makes it likely that the man will hurt Stampy, i.e., P(H|L) > P(H|L^c). Homer argues that P(H|L) < P(H|L^c).

(c) Assume it is true that someone who has a lot of a commodity will have less desire to acquire more of the commodity. Explain what is wrong with Homer's reasoning that the evidence about Blackheart makes it less likely that he will harm Stampy.

Solution: Homer does not realize that observing that Blackheart has so much ivory makes it much more likely that Blackheart is an ivory dealer, which in turn makes it more likely that the man will hurt Stampy. (This is an example of Simpson's Paradox.) It may be true that, controlling for whether or not Blackheart is a dealer, having high ivory supplies makes it less likely that he will harm Stampy: P(H|L, D) < P(H|L^c , D) and P(H|L, D^c) < P(H|L^c , D^c). However, this does not imply that P(H|L) < P(H|L^c).

3. A gambler repeatedly plays a game where in each round, he wins a dollar with probability 1/3 and loses a dollar with probability 2/3. His strategy is "quit when he is ahead by \$2," though some suspect he is a gambling addict anyway. Suppose that he starts with a million dollars. Show that the probability that he'll ever be ahead by \$2 is less than 1/4.

Solution: Let A_1 be the event that he is successful on the first play and let W be the event that he is ever ahead by \$2 before being ruined. Then by the law of total probability, we have
P(W) = P(W|A_1)P(A_1) + P(W|A_1^c)P(A_1^c).
Let a_i be the probability that the gambler achieves a profit of \$2 before being ruined, starting with a fortune of \$i. For our setup, P(W) = a_i, P(W|A_1) = a_{i+1} and P(W|A_1^c) = a_{i-1}. Therefore,
a_i = a_{i+1}/3 + 2a_{i-1}/3, with boundary conditions a_0 = 0 and a_{i+2} = 1. We can then solve this difference equation for a_i (directly or using the result of the gambler's ruin problem):
$a_i = \frac{2^i - 1}{2^{2+i} - 1}$. This is always less than 1/4 since $\frac{2^i-1}{2^{2+i}-1} < \frac{1}{4}$ is equivalent to $4(2^i - 1) < 2^{2+i} - 1$, which is equivalent to the true statement $2^{2+i} - 4 < 2^{2+i} - 1$.

4. A sequence of n independent experiments is performed. Each experiment is a success with probability p and a failure with probability q = 1- p. Show that conditional on the number of successes, all possibilities for the list of outcomes of the experiment are equally likely (of course, we only consider lists of outcomes where the number of successes is consistent with the information being conditioned on).

Solution: Let X_j be 1 if the jth experiment is a success and 0 otherwise, and let X = X_1 + ⋯ + X_n be the total number of successes. Then for any k and any a_1,..., a_n 2 {0, 1} with a_1+...+ a_n = k, P($X_1 = a_1, \ldots, X_n = a_n$|X = k) = $\frac{\mathrm{P}(X_1 = a_1,\ldots,X_n = a_n, \mathrm{X} = \mathrm{k})}{P(X=k)}$ = $\frac{\mathrm{P}(X_1 = a_1,\ldots,X_n = a_n)}{P(X=k)}$ = $\frac{p^k q^{n-k}}{C_n^k p^k q^{n-k}}$ = $\frac{1}{C_n^k}$.

This does not depend on a_1,..., a_n. Thus, for n independent Bernoulli trials, given that there are exactly k successes, the C_n^k possible sequences consisting of k successes and n - k failures are equally likely. Interestingly, the conditional probability above also does not depend on p.

5. (a) Consider the following 7-door version of the Monty Hall problem. There are 7 doors, behind one of which there is a car (which you want), and behind the rest of which there are goats (which you don't want). Initially, all possibilities are equally likely for where the car is. You choose a door. Monty Hall then opens 3 goat doors, and offers you the option of switching to any of the remaining 3 doors.
Assume that Monty Hall knows which door has the car, will always open 3 goat doors and offer the option of switching, and that Monty chooses with equal probabilities from all his choices of which goat doors to open. Should you switch? What is your probability of success if you switch to one of the remaining 3 doors?

   Solution: Assume the doors are labeled such that you choose Door 1 (to simplify notation), and suppose first that you follow the "stick to your original choice" strategy. Let S be the event of success in getting the car, and let C_j be the event that the car is behind Door j. Conditioning on which door has the car, we have P(S) = P(S|C_1)P(C_1) + ⋯ + P(S|C_7)P(C_7) = P(C_1) = 1/7 . Let M_{ijk} be the event that Monty opens Doors i, j, k. Then
   P(S) = $\sum_{i,j,k} \mathrm{P}(\mathrm{S}|M_{ijk})\mathrm{P}(M_{ijk})$
   P(S|M_{ijk}) = P(S) = 1/7
   for all i, j, k with $2 \le i < j < k \le 7$. Thus, the conditional probability that the car is behind 1 of the remaining 3 doors is 6/7, which gives 2/7 for each. So you should switch, thus making your probability of success 2/7 rather than 1/7.

   (b) Generalize the above to a Monty Hall problem where there are n 3 doors, of which Monty opens m goat doors, with 1 $\le$m $\le$ n - 2.

   Solution: By the same reasoning, the probability of success for "stick to your original choice" is 1/n, both unconditionally and conditionally. Each of the n – m - 1 remaining doors has conditional probability $\frac{n-1}{(n-m-1)n}$ of having the car. This value is greater than 1/n, so you should switch, thus obtaining probability $\frac{n-1}{(n-m-1)n}$ of success (both conditionally and unconditionally).

6. Calvin and Hobbes play a match consisting of a series of games, where Calvin has probability p of winning each game (independently). They play with a "win by two" rule: the first player to win two games more than his opponent wins the match. Find the probability that Calvin wins the match (in terms of p), in two different ways:

   (a) by conditioning, using the law of total probability.

   Solution: Let C be the event that Calvin wins the match, X ~ Bin(2, p) be how many of the first 2 games he wins, and q = 1- p. Then P(C) = P(C|X = 0)q^2 + P(C|X = 1)(2pq) + P(C|X = 2)p^2 = 2pqP(C) + p^2 , so P(C) = $\frac{p^2}{1-2pq}$. This can also be written as $\frac{p^2}{p^2+q^2}$, since p + q = 1.

   (b) by interpreting the problem as a gambler's ruin problem.

   Solution: The problem can be thought of as a gambler's ruin where each player starts out with \$2. So the probability that Calvin wins the match is $\frac{1-(\frac{q}{p})^2}{1-(\frac{q}{p})^4} = \frac{(p^2-q^2)/p^2}{(p^4-q^4)/p^4} = \frac{(p^2-q^2)/p^2}{(p^2-q^2)(p^2+q^2)/p^4} = \frac{p^2}{p^2+q^2}$, which agrees with the above.