Stat 110 Homework 3, Fall 2013

Due: Friday 9/27 at 1:10 pm, in the Stat 110 dropbox outside of SC 105. No late homework will be accepted. Please write your name, and staple your homework. Show your work and make sure to give clear, careful, convincing explanations.

Reading: Chapter 2; Sections 3.1-3.3.

1. You are the contestant on the Monty Hall show. Monty is trying out a new version of his game, with rules as follows. You get to choose one of three doors. One door has a car behind it, another has a computer, and the other door has a goat (with all permutations equally likely). Monty, who knows which prize is behind each door, will open a door (but not the one you chose) and then let you choose whether to switch from your current choice to the other unopened door.

Assume that you prefer the car to the computer, the computer to the goat, and (by transitivity) the car to the goat.

(a) Suppose for this part only that Monty always opens the door that reveals your less preferred prize out of the two alternatives, e.g., if he is faced with the choice between revealing the goat or the computer, he will reveal the goat. Monty opens a door, revealing a goat (this is again for this part only). Given this information, should you switch? If you do switch, what is your probability of success in getting the car?

(b) Now suppose that Monty reveals your less preferred prize with probability p, and your more preferred prize with probability $q = 1 - p$. Monty opens a door, revealing a computer. Given this information, should you switch (your answer can depend on p)? If you do switch, what is your probability of success in getting the car (in terms of p)?

2. Simpson's paradox is the phenomenon that says it is possible to have events A, B, C such that

$$P(A|B, C) < P(A|B^c, C)$$
$$P(A|B, C^c) < P(A|B^c, C^c)$$

but

$$P(A|B) > P(A|B^c).$$

(a) Can Simpson's paradox occur if A and B are independent? If so, give a concrete example (with both numbers and an interpretation); if not, prove that it is impossible.

(b) Can Simpson's paradox occur if A and C are independent? If so, give a concrete example (with both numbers and an interpretation); if not, prove that it is impossible.

(c) Can Simpson's paradox occur if B and C are independent? If so, give a concrete example (with both numbers and an interpretation); if not, prove that it is impossible.

3. There are 100 equally spaced points around a circle. At 99 of the points, there are sheep, and at 1 point, there is a wolf. At each time step, the wolf randomly moves either clockwise or counterclockwise by 1 point. If there is a sheep at that point, he eats it. The sheep don't move. What is the probability that the sheep who is initially opposite the wolf is the last one remaining?

Hint: one good approach is to use the result of the gambler's ruin problem.

4. An immortal drunk man wanders around randomly on the integers. He starts at the origin, and at each step he moves 1 unit to the right or 1 unit to the left, with probabilities p and $q = 1 - p$ respectively, independently of all his previous steps. Let S_n be his position after n steps.

(a) Find the PMF of S_n, for fixed n (make sure to specify the support of the distribution).

Hint: What are the distributions of L_n and R_n, the numbers of steps to the left and right among the first n steps (respectively)? How are S_n, L_n, and R_n related?

(b) Let p_k be the probability that the drunk ever reaches the value k, for all $k \geq 0$. Write down a difference equation for p_k (you do not need to solve it for this part).

(c) Find p_k, fully simplified; be sure to consider all 3 cases: $p < 1/2, p = 1/2$, and $p > 1/2$. Feel free to assume that if $A_1, A_2, \ldots$ are events with $A_j \subseteq A_{j+1}$ for all j, then $P(A_n) \to P(\cup_{j=1}^{\infty} A_j)$ as $n \to \infty$ (because it is true; this is known as *continuity of probability*).

Hint: one good approach is to use the result of the gambler's ruin problem to study events of the form "the drunk reaches k before ever reaching $-a$", where a is a positive integer.

5. Two coins are in a hat. One of the coins is fair (lands Heads with probability 1/2), but the other coin lands Heads with probability 2/3. One of the coins is randomly pulled from the hat, without knowing which of the two it is. Call the coin that was chosen Coin C.

(a) Coin C will be tossed 10 times. Are the events "first toss of Coin C is Heads" and "second toss of Coin C is Heads" independent? Give a clear intuitive explanation.

(b) What is the PMF of the number of Heads in 10 tosses of Coin C?

(c) Now suppose that it is observed that Coin C landed Heads exactly 6 times in the 10 tosses. Given this information, what is the probability that Coin C is fair?

6. There are n people eligible to vote in a certain election. Voting requires registration. Decisions are made independently. Each of the n people will register with probability p_1. Given that a person registers, he or she will vote with probability p_2. Given that a person votes, he or she will vote for Kodos (who is one of the candidates) with probability p_3. What is the distribution of the number of votes for Kodos (give the PMF, fully simplified)?

Hint: what is truer than truth?

7. You are the Challenger in a match against the World Champion of a certain sport. The match will consist of a series of n games, where n is an even number. Each game results in victory for you with probability p, and victory for your opponent with probability $q = 1-p$, independently. A tied *game* is not possible in this sport, but a tied *match* is possible. To become World Champion, you need to win a majority of the games; otherwise, the World Champion retains the title. You are negotiating the choice of n (which is an *even* number).

For each of the following values of p, discuss whether you want the match to be as short as possible, as long as possible, or something in between. If something in between, find your preferred value of n numerically. For each part, explain intuitively what is going on, but also provide some numerical calculations of relevant probabilities of winning the match. The recommended environment for these calculations is R (see the last section of each chapter in the book), but you can also use Wolfram Alpha or other possibilities. In any case, state what computing environment you are using, and provide your code (e.g., if you used `dbinom` or `pbinom` in R then say so, and how you used it).

(a) $p = 0.51$.

(b) $p = 0.5$.

(c) $p = 0.49$.