Stat S-110 Homework 1, Summer 2015

Due: Wednesday 7/1 at noon (Eastern Time), submitted as a PDF file via the course website or as a hard copy in class. Show your work and explain your steps.

Collaboration Policy: You are welcome to discuss the problems with others, but *you must write up your solutions yourself and in your own words.* Copying someone else's solution, or just making trivial changes for the sake of not copying verbatim, is not acceptable. For example, in problems where you have to make up a "story" or example, two students should not have the exact same answer, or almost the same answer except one has an example with dogs chasing cats and the other has an example with cats chasing mice, with the same structure and the same numbers.

The following problems are from Chapters 1-2 of *Introduction to Probability* by Joe Blitzstein and Jessica Hwang. We write BH 1.6, for example, to denote Exercise 6 of Chapter 1 of the book.

1. (BH 1.2) (a) How many 7-digit phone numbers are possible, assuming that the first digit can't be a 0 or a 1?

(b) Re-solve (a), except now assume also that the phone number is not allowed to start with 911 (since this is reserved for emergency use, and it would not be desirable for the system to wait to see whether more digits were going to be dialed after someone has dialed 911).

2. (BH 1.6) There are 20 people at a chess club on a certain day. They each find opponents and start playing. How many possibilities are there for how they are matched up, assuming that in each game it *does* matter who has the white pieces (in a chess game, one player has the white pieces and the other player has the black pieces)?

3. (BH 1.7) Two chess players, A and B, are going to play 7 games. Each game has three possible outcomes: a win for A (which is a loss for B), a draw (tie), and a loss for A (which is a win for B). A win is worth 1 point, a draw is worth 0.5 points, and a loss is worth 0 points.

(a) How many possible outcomes for the individual games are there, such that overall player A ends up with 3 wins, 2 draws, and 2 losses?

(b) How many possible outcomes for the individual games are there, such that A ends up with 4 points and B ends up with 3 points?

(c) Now assume that they are playing a best-of-7 match, where the match will end when either player has 4 points or when 7 games have been played, whichever is first.

For example, if after 6 games the score is 4 to 2 in favor of A, then A wins the match and they don't play a 7th game. How many possible outcomes for the individual games are there, such that the match lasts for 7 games and A wins by a score of 4 to 3?

4. (BH 1.17) Give a story proof that

$$\sum_{k=1}^{n} k \binom{n}{k}^2 = n \binom{2n-1}{n-1},$$

for all positive integers n.

Hint: Consider choosing a committee of size n from two groups of size n each, where only one of the two groups has people eligible to become president.

5. (BH 1.21) Three people get into an empty elevator at the first floor of a building that has 10 floors. Each presses the button for their desired floor (unless one of the others has already pressed that button). Assume that they are equally likely to want to go to floors 2 through 10 (independently of each other). What is the probability that the buttons for 3 consecutive floors are pressed?

6. (BH 1.49) For a group of 7 people, find the probability that all 4 seasons (winter, spring, summer, fall) occur at least once each among their birthdays, assuming that all seasons are equally likely.

7. (BH 2.4) Fred is answering a multiple-choice problem on an exam, and has to choose one of n options (exactly one of which is correct). Let K be the event that he knows the answer, and R be the event that he gets the problem right (either through knowledge or through luck). Suppose that if he knows the right answer he will definitely get the problem right, but if he does not know then he will guess completely randomly. Let $P(K) = p$.

(a) Find $P(K|R)$ (in terms of p and n).

(b) Show that $P(K|R) \geq p$, and explain why this makes sense intuitively. When (if ever) does $P(K|R)$ equal p?

8. (BH 2.6)

A hat contains 100 coins, where 99 are fair but one is double-headed (always landing Heads). A coin is chosen uniformly at random. The chosen coin is flipped 7 times, and it lands Heads all 7 times. Given this information, what is the probability that the chosen coin is double-headed? (Of course, another approach here would be to *look at both sides of the coin*—but this is a metaphorical coin.)

9. (BH 2.7) A hat contains 100 coins, where *at least* 99 are fair, but there may be one that is double-headed (always landing Heads); if there is no such coin, then all 100 are fair. Let D be the event that there is such a coin, and suppose that $P(D) = 1/2$. A coin is chosen uniformly at random. The chosen coin is flipped 7 times, and it lands Heads all 7 times.

(a) Given this information, what is the probability that one of the coins is double-headed?

(b) Given this information, what is the probability that the chosen coin is double-headed?

10. (BH 2.10) Fred is working on a major project. In planning the project, two milestones are set up, with dates by which they should be accomplished. This serves as a way to track Fred's progress. Let A_1 be the event that Fred completes the first milestone on time, A_2 be the event that he completes the second milestone on time, and A_3 be the event that he completes the project on time.

Suppose that $P(A_{j+1}|A_j) = 0.8$ but $P(A_{j+1}|A_j^c) = 0.3$ for $j = 1, 2$, since if Fred falls behind on his schedule it will be hard for him to get caught up. Also, assume that the second milestone supersedes the first, in the sense that once we know whether he is on time in completing the second milestone, it no longer matters what happened with the first milestone. We can express this by saying that A_1 and A_3 are conditionally independent given A_2 and they're also conditionally independent given A_2^c.

(a) Find the probability that Fred will finish the project on time, given that he completes the first milestone on time. Also find the probability that Fred will finish the project on time, given that he is late for the first milestone.

(b) Suppose that $P(A_1) = 0.75$. Find the probability that Fred will finish the project on time.

11. (BH 2.11) An *exit poll* in an election is a survey taken of voters just after they have voted. One major use of exit polls has been so that news organizations can try to figure out as soon as possible who won the election, before the votes are officially counted. This has been notoriously inaccurate in various elections, sometimes because of *selection bias*: the sample of people who are invited to and agree to participate in the survey may not be similar enough to the overall population of voters.

Consider an election with two candidates, Candidate A and Candidate B. Every voter is invited to participate in an exit poll, where they are asked whom they voted for; some accept and some refuse. For a randomly selected voter, let A be the event that they voted for A, and W be the event that they are willing to participate in the exit poll. Suppose that $P(W|A) = 0.7$ but $P(W|A^c) = 0.3$. In the exit poll, 60% of the respondents say they voted for A (assume that they are all honest), suggesting a comfortable victory for A. Find $P(A)$, the true proportion of people who voted for A.

12. (BH 2.13) Company A has just developed a diagnostic test for a certain disease. The disease afflicts 1% of the population. As defined in Example 2.3.9, the *sensitivity* of the test is the probability of someone testing positive, given that they have the disease, and the *specificity* of the test is the probability that of someone testing negative, given that they don't have the disease. Assume that, as in Example 2.3.9, the sensitivity and specificity are both 0.95.

Company B, which is a rival of Company A, offers a competing test for the disease. Company B claims that their test is faster and less expensive to perform than Company A's test, is less painful (Company A's test requires an incision), and yet has a higher overall success rate, where overall success rate is defined as the probability that a random person gets diagnosed correctly.

(a) It turns out that Company B's test can be described and performed very simply: no matter who the patient is, diagnose that they do not have the disease. Check whether Company B's claim about overall success rates is true.

(b) Explain why Company A's test may still be useful.

(c) Company A wants to develop a new test such that the overall success rate is higher than that of Company B's test. If the sensitivity and specificity are equal, how high does the sensitivity have to be to achieve their goal? If (amazingly) they can get the sensitivity equal to 1, how high does the specificity have to be to achieve their goal? If (amazingly) they can get the specificity equal to 1, how high does the sensitivity have to be to achieve their goal?