Hints for Quiz 2 Review Problems

A Few Problem Types and Corresponding Strategies

1. **Build-a-sequence problems**: These problems ask you how many ways there are to construct a particular sequence, or the probability of seeing a particular sequence. For "how many" problems, it is best to think about filling each slot in the sequence in order, counting how many options you have for each slot given the slots you have filled so far, and multiply them together. For "probability" problems, divide by the total number of possible sequences.

2. **Independent replications**: These are problems where a simple, identical experiment is repeated several times independently. The event whose probability you need to calculate is a summary of the total set of experiments (e.g., probability that something never happens). The strategy here should be to rewrite the event in terms of a single experiment's outcome, then raise it to a power (applying the multiplication rule).

3. **Compound event**: These are problems where the event whose probability you need to calculate can happen one of several ways that have equal probability and are easy to enumerate. The strategy here is to calculate the probability of each scenario and add them up (applying the addition rule). For example, in a series of fair coin tosses, the probability that all of the outcomes are the same is equal to the probability that all of the outcomes are heads *plus* the probability that all outcomes are tails.

4. **Binomial**: This is a specific example that we discussed a lot in class. We have a binomial problem when we have a set of n independent, identical experiments with binary outcome (e.g., heads or tails, red or not red, up or down), where the probability of one outcome is p for each trial, and the summary of interest X is the *number of of times* a particular outcome occurred. So problems like this include "What is the probability of getting exactly k heads out of n coin tosses". A couple of things to note:

   - The derivation of the binomial distribution involves an application of both the multiplication rule (for calculating the probability of any particular sequence that has k successes) and the addition rule (for adding up the probabilities of each of the $\binom{n}{k}$ sequences that have k successes).

   - Binomial problems may also ask you to calculate the probability of getting at least or at most k successes – you should think of this as a compound event, and calculate the probability as the sum of the probabilities that X takes the value x for each x that satisfies the event. For example, the probability of seeing at least k successes is the sum of the probabilities that X takes each value between k and n.

5. **Max/min (German airplanes)**: These problems ask you to calculate probabilities involving a random variables X defined as the maximum or minimum of a random sample from a set of sequential numbers. These can involve sampling with or without replacement. In these problems, it is easier to directly calculate the probability that X is greater than or less than a particular number, rather than directly calculating the probability that X is exactly equal to a particular number. The particular trick here is to note that:

- If X is the maximum of the sample, then $P(X \leq k) = P(\text{All samples } \leq k)$.
- If X is the minimum of the sample, then $P(X \geq k) = P(\text{All samples } \geq k)$.

If you need to compute the probability that X is exactly equal to a number k, you can use the fact that $P(X = k) = P(X \leq k) - P(X < k) = P(X \geq k) - P(X > k)$, and calculate the probability of an exact value in terms of the probability of being greater than or less than that value.

Note: There was an error in lecture about this sort of problem. See Piazza note @57. The notes on the German airplane problem are correct.

6. **Simple complement (Birthday problem, Gambler's rule)**: In some problems, the events whose probability you need to calculate is a complex compound event where there are many ways for the event to occur. For these problems, the probability of the event *not* occurring may have a probability that's easier to compute. If this is the case, use the fact that $P(A) = 1 - P(A^C)$ (probability of an event occurring is equal to one minus the probability of that event not occurring).

   For example, in the birthday problem, the event is "at least two people share the same birthday", which is very complex (many different people could share the same birthday, or multiple pairs of people could share birthdays, etc); it is easier to calculate the probability of this *not* occurring, because this translates to "all people have distinct birthdays". Likewise the gambler's rule problem, the event that "I win at least once" is very complex (I could win once, twice, three times...); on the other hand, the opposite of this event "I never win" has a probability that is simple to compute.

7. **Sampling without replacement**: These problems ask you the probability of drawing a sample of size n with some special properties when sampling from a population of size N without replacement. It is easiest to attack these problems by counting the number of ways you can construct a sample with the special property, divided by the total number of ways of constructing a sample of size n. So the denominator in these answers should be $\binom{N}{n}$. For constructing the numerator, you should divide up your sample of size n into "slots" or "seats" that need to be filled, and consider the number of ways that you can fill those slots with suitable candidates from the population (these will all look like "something choose something"). It can be helpful here to *not* take shortcuts and to include binomial coefficients even if they are "something choose 1" or "something choose 0".

   For example, if a professor choosing a sample of n students from a classroom of size N, the probability that the sample includes you, but not your friend Pat involves:

   - $\binom{1}{1}$: The number of ways to select you from the set of people in the class who are you (a set of size 1).
   - $\binom{1}{0}$: The number of ways to *not* select your friend from the set of people in the class who are your friend (again, a set of size 1).
   - $\binom{N-2}{n-1}$: The number of ways to fill the remaining $n-1$ slots in your sample with the people in the class who are not you or your friend.

   The final expression is:

   $$P(\text{Select a sample including you but not Pat}) = \frac{\binom{1}{1}\binom{1}{0}\binom{N-2}{n-1}}{\binom{N}{n}} = \frac{\binom{N-2}{n-1}}{\binom{N}{n}}.$$

Note that *if you don't simplify the statement* that the top numbers in the binomial coefficients in the numerator add up to N and the bottom numbers in the binomial coefficients in the numerator add up to n, but once you simplify by dropping factors that are equal to 1, that rule no longer applies. We encourage you to avoid simplifying in these problems until the very end (and if you don't simplify in your final answer, we won't hold it against you).

Review Problems

We'll give hints for solving the review problems here, making reference to the problem types discussed above:

1. A die has one red face and five blue faces. If the die is rolled 12 times, what is the chance that the red face never appears?

   Independent replications

2. A random number generator draws uniformly at random with replacement from the 10 digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9.

   (a) Suppose you run the generator five times. Find the chance that the same digit is drawn all five times.

   Compound event (union of simple events that specify the digit).

   (b) What is the probability that the largest of these draws is exactly 5?

   Max/min problem. Calculate probability that largerst draw is no larger than 5 first.

   (c) What is the probability that the smallest draw is no smaller than 3?

   Max/min problem.

   (d) What is the probability that at least one number is repeated?

   Simple complement problem. What is the opposite of this event?

3. A roulette wheel has 38 distinct pockets. Each time the wheel is spun, one of the pockets is the winner.

   (a) Suppose you spin the wheel 10 times and keep track of the sequence of winning pockets. How many possible sequences are there?

   Build-a-sequence.

   (b) There are exactly 2 green pockets on the wheel. What is the probability that exactly 2 of the 10 spins lands in a green pocket?

   Binomial problem.

4. A standard deck consists of 13 hearts, 13 diamonds, 13 clubs, and 13 spades, making 52 cards in all.

   (a) Suppose cards are dealt one by one at random without replacement. What is the chance that the 10th card is a heart, given that the 7th card is a spade and the 32nd card is a diamond?

   Subset selection, but easier as a symmetric permutation problem. See Chapter 5.3.

   (b) A poker hand consists of 5 cards picked at random without replacement from the deck. Find the chance that the king of hearts and the king of diamonds are both picked in a poker hand.

   Sampling without replacement problem.

5. A monkey is tapping at a keyboard that has one key for each of the 26 letters of the English alphabet. Assume that each time the monkey is equally likely to pick any one of the 26 letters, regardless of what it has picked at other times.

   (a) Find the chance that the first eight letters the monkey picks can form the word KEYBOARD, by rearrangement if necessary.

   Build-a-sequence.

   (b) A palindrome is a sequence of letters that reads the same forwards and backwards, like RADAR (a five-letter palindrome). For our purposes, it can be a nonsense "word" like WSBSW or AAAAA. How many five-letter palindromes are there? What is the chance that the first five letters that the monkey picks, in sequential order, form a palindrome?

   Build-a-sequence.

6. In what follows, assume that n and N are positive integers. A bet is such that you have chance $1/N$ of winning it each time you play, regardless of the results of all other times. Suppose you play n times.

   (a) Find the exact chance that you win at least one of the n bets, and provide an exponential approximation for the chance. Show your calculations; you do not need to prove math facts about the exponential or logarithmic functions.

   Simple complement. Review exponential approximation in Chapter 4.3 and 4.4.

   (b) In terms of N, what is the smallest value of n so that the approximate chance of winning is at least 2/3? Prove your answer.

   Algebra. See method used in Chapter 4.3.

7. You are conducting a survey to see how faculty, staff, and students in the Statistics department view the increasing popularity of data science. Suppose there are 150 people invovled with the department: 100 students, 30 faculty, and 20 staff. For your survey, you are selection 15 people, without replacement. Professors D'Amour and Adikhari are included among the faculty.

   These are all sampling without replacement problems.

   (a) What is the probability that your sample includes no staff members?

   (b) What is the probability that your sample includes exactly 2 faculty?

   (c) What is the probability that your sample includes at least 10 students?

   (d) What is the probability that your sample includes exactly 2 staff but no faculty?

   (e) What is the probability that your sample includes Prof. Adikhari, but does not include Prof. D'Amour.

   (f) What is the probability that your sample includes exactly one of Prof. Adhikari and Prof. D'Amour?

   (g) What is the probability that your sample includes exactly one of Prof. Adhikari and Prof. D'Amour, and includes exactly 10 students?

8. Suppose cars can only have one of four colors in Berkeley: red, blue, green, or white. The distribution of colors is as follows: red 10%, blue 35%, green 25%, white 30%. You stand on a street corner and watch cars 20 cars drive by. Assume that this observation process approximates sampling cars from Berkeley with replacement. (Hint: When counting cars of a specific color, it can help to think of car colors as, e.g., "red" or "not red".)

   These are binomial problems. Think about how to express it that way.

   (a) What is the probability that you see exactly 5 red cars?

   (b) What is the probability that you see at least 10 blue cars?

9. Halloween is coming soon and the children in your neighborhood will soon come trick-or-treating (that is, they will come to your door to demand candy). Suppose there are 20 children in your neighborhood and 30 houses (one of which is yours). Each child independently chooses 10 houses at random without replacement to visit.

   (a) What is the probability that a specific child will visit your house?
       Simple sampling without replacement problem.

   (b) What is the probability that exactly 10 children visit your house?
       Binomial problem, using the probability from the last part.

10. (Continuing the last question.) You have bunch of candy in your house, ranging from very cheap to very fancy. Suppose that, if you ordered the candy in your house from cheapest to fanciest, the ith candy cost $25i$ cents to buy (so the cheapest candy cost 25 cents, the 10th cheapest cost $2.50, etc). You want to keep the fancy candy for yourself (the kids wouldn't appreciate the fancy stuff anyway), so when a child comes to your door, you give them the cheapest remaining candy – thus, if 10 children come to your door, you will give out the 10 cheapest pieces of candy that you have.

   (a) If k children come to your door, what is the total cost of the candy that you will give out? (Hint: This uses a summation identity from before the probability chapters.)
       Use "common summation" 2 from the Appendix

   (b) What is the probability that you will give away more than $15 worth of candy?
       Function of a random variable. Use the binomial construction from the last question to answer.