MONT 105N – Analyzing Environmental Data
Information for Midterm Exam

General Information

To leave Friday, March 22 for follow-up discussion of our cluster event the night before (and some review in case that does not take the whole period), I have decided to move the midterm exam in MONT 105N one class day later, to *Monday, March 25*.

- This will be a full period, individual exam. No sharing of information with others in any form will be permitted during the exam.
- You may use a calculator during the exam, but *no other electronic devices*.
- There will be three or four mathematical problems (each possibly with a few separate parts). These questions will draw on the material we have studied from *Chapters 9 and 10* of our text. Some sample exam questions are given later in this document. *Note:*
  - The mathematical questions will be similar to the sample questions.
  - I'm not going to ask you to do large computations because we have done most of that in spreadsheets. The exam questions will focus on the concepts behind the various descriptive statistics we have studied and the probability ideas we have learned.
  - There will be a question where you need to use the standard normal probability table we discussed in class on March 13. A copy of the table will be provided with the exam.
- The exam will also include an essay question on a set topic given below, worth 25 out of the total 100 points. As you will see, part of this consists of factual questions. You will need to look up some information before coming to the exam to write this part of the essay. Then, the rest of the essay topic asks you to think about the consequences of what you found. There's no one right or wrong answer for that part; the point there is to get you to think about the issues and try to formulate a position of your own. This essay question should take you about 15 minutes or so, so it will be necessary to spend some of your preparation time on deciding what you want to say.

Essay Question Prompt

One very common everyday situation where we are given information phrased in probabilistic terms is when weather forecasts for a given geographical area include an estimated *probability of precipitation* for some time period in the future. But what does one of those probabilities mean, really? What is the exact formula used to calculate a probability of precipitation? Explain by indicating exactly what it means, for instance, if a weather forecast says there is an 80% chance of rain in Worcester tomorrow. Can it mean more than one thing? (Hint: Yes it can! Why?) What does the National Oceanographic and Atmospheric Administration (NOAA) say about this and what is the justification they give

for doing it this way? Do you think just the number gives an adequate indication to an average person? Would it help to educate people better about how these probabilities are actually computed? Is there a better way to indicate how likely certain weather conditions might be?

Some practice/review questions

1) A random sample of 1000 college first-year students in the US is selected and their ages are determined.
   a) What would you expect was the sample mean age, to the nearest whole number?
   b) Give the formula that would be used to compute the sample SD, calling the ages $x_1, \ldots, x_{1000}$.
   c) Would you expect the sample SD to be about 0.1 years, 1 year, or 10 years? Explain.

2) Find the 5-number summary, and draw the "box and whisker plot" for the data set

   5 3 7 8 8 10 1 2 4

   (Use the inclusive method for the quartiles.)

3) Eight birds – two robins, two woodpeckers, two thrushes, and two blackbirds are placed in a covered cage. Three random draws of a single bird are made from the cage *with replacement*.
   a) Are the three draws *independent* in this case? Why or why not?
   b) What are the chances that all three birds are robins?
   c) What are the chances that none of the birds is a robin?
   d) What are the chances that there is a something other than a robin among the three birds?
   e) (More challenging) What would change in your answers to a,b,c,d if the draws were done *without replacement*?

4) A coin is tossed repeatedly.
   a) What is the chance of getting 7 heads and 3 tails in ten tosses if the heads and tails are equally likely?
   b) *Given that the first 3 tosses come up heads*, what is the chance that the first tails comes up on the 7th toss if heads and tails are equally likely? (Note: This is a conditional probability.)
   c) How do your answers to a) and b) change if you know that on each toss, the probability of a head is .6 and the probability of a tail is .4 (that is the coin is weighted)?

5) Assume that everyone starts school at age 5 and stays in school at least until age 16. True or False and explain:
   a) The median years in school would be greater than 11.

b) The average years in school would be greater than the median.

6) One ticket is drawn at random from each of the two boxes: (i) 1 2 3 4 5 and (ii) 1 2 3 4 5 6 . The tickets in the first box are blue and those in the second box are green, so you can tell which box the ticket came from.
   a) What is the probability that one of the numbers is 2 and the other is 5?
   b) What is the probability that the sum of the numbers is 7?
   c) What is the probability that one numbers is (strictly) bigger than twice the other? (Hint: List all the cases where that is true and count!)

7) Body lengths of a population of immature lobsters are *uniformly distributed* on the interval 15 to 22 cm.
   a) If a single lobster is selected at random, what is the probability that its body length is between 19 and 20 cm?
   b) If 6 lobsters are selected with replacement from the population, what is the probability that 4 of them have body lengths between 19 and 20?

8) In a large class, the average score on a test was 75 out of 100 and the SD was 10. The scores followed a normal distribution.
   a) A single test paper is selected randomly out of the pile. What is the probability that the score was between 75 and 85?
   b) Same question as a) but the score is between 60 and 80?
   c) What was the 80th percentile score (the score such that 80% had scores less than or equal to that, and only 20% had scores greater than or equal to that)? Round to the nearest whole number.

9) George is playing roulette (where, in the US version, there are 18 red numbers, 18 greens, and 2 additional slots, for a total of 38 possible ways for the ball to land). Assume the roulette wheel is not biased or "loaded" so that each of the 38 outcomes is equally likely. George starts with \$100. For his first 50 spins of the wheel, he bets \$1 on 17, his favorite number (he can win if either 17 red or 17 green comes up). Unfortunately, he loses each time, losing \$50 in the process. Ever the optimist, George decides his luck must change on the next spin of the wheel because 17 is "due" (it hasn't come up 50 spins in a row, after all). So he bets his remaining \$50 on 17 one more time. Donald's significant other Laura, who always brings George back to earth, says he is a fool and that he's likely to lose again. Is George right, is Laura right, or are neither of them right? Explain your answer.