MONT 105N – Analyzing Environmental Data
Information on Midterm Exam
March 26, 2013

General Information and Groundrules

As announced in class and on the course schedule online, the midterm exam for our Montserrat seminar will be given in class on Friday, April 12.

- This will be a full period, individual exam. No sharing of information in any form will be permitted during the exam.
- The material to be covered is basically everything we have studied since the start of the semester (including our discussions of *The Hockey Stick and the Climate Wars*, through the material on confidence intervals from Chapter 13 of the text.
- You may use a calculator for numerical computations during the exam, but you may not use stored text or other electronic devices for accessing online information.
- There will be three or four mathematical problems (each possibly with a few separate parts). These questions will be similar to things you have seen on the problem sets or the group projects. Some sample exam questions are given later in this document.
- I will provide copies of the normal curve area table from page 308 and the t-table from page 316.
- The exam will also include an essay question on a set topic chosen from the topics listed below. This part of the exam should take you about 15 minutes, so it will be necessary to spend some of your preparation time on deciding what you want to say for each topic.
- I will see if the small classroom Swords 330 is available so that you can have a better space for taking this test than our regular classroom.

Possible Essay Topics

1. Exactly what is the "hockey stick" graph? How was it originally generated? When and where did it first appear in the published scientific record? Have other studies confirmed this general pattern or called into question the original conclusions? When and why did the "hockey stick" become such a contentious issue in the climate change debate? What is the IPCC and what do they do? What was the role of the IPCC here?
2. What was the "climategate scandal?" Who was involved? What charges were leveled against Mann and other climate scientists? On further review, were the attacks justified? Why do you think that Mann has persevered in trying to work on these scientific problems? (After all, it would have been very easy just to give up and do something else.) Why did he write *The Hockey Stick and the Climate Wars*? Do you see him as just bullheaded? heroic? neither?
3. Science is often said to be "self-correcting." What does that mean and how does it work? How are results communicated within the scientific community? What "quality control" mechanisms are built into the everyday practice of science? What does Mann

say about role of skepticism and healthy debate in good science in *The Hockey Stick and the Climate Wars*? How can one tell the difference between healthy debate and an attack that is directed at the scientist personally or motivated by political and/or ideological differences? As an example, consider the various criticisms of Mann's work by Ross McKittrick and Stephen McIntyre. Where do they fall?

Sample Mathematical Questions

Note: There are more questions and more parts to these questions than I could reasonably ask you to do in the actual exam (because of the limited time available). This should give you an idea of the range of possible topics and styles of questions I might include in the exam, though.

A. Waterbuck are large mammals in the deer family native to eastern South Africa. In 1976, biologists launched a three-year study of waterbuck ecology and determined the annual survival and fertility rates for females given in the following table:

	A = calves	B = yearlings	C = adults
fertility	0	0.048	0.081
survival	0.94	1.00	0.75

1. Express the information given above as a life-cycle graph.
2. Express the information given above as a system of difference equations (time n in years).
3. If the initial conditions are $A(0) = 100$, $B(0) = 2000$, $C(0) = 3000$, find the population of each group in years $n = 1, 2, 3, 4$.

B. The following data set has $n = 9$: $23, 28, 40, 44, 47, 50, 51, 54, 55$

1. Find the "5-number" summary for this data set.
2. Draw the corresponding box plot.
3. Compute the (Bowley) measure of skewness. Does this seem reasonable from the box plot?
4. Compute the SD of the data set. How many of the points lie within two SD's of the mean? Is Chebyshev's Rule satisfied here? (Say what that rule says, and determine whether or not it is satisfied.)

C. (Short answer) Suppose that a researcher collects 80 individuals of the Atlantic surf clam. These clams can be found at levels down to about a meter in the sand, and larger clams tend to live at deeper levels. The researcher finds an average shell width 10.2 cm. Think of this as a sampling process.

1. What is the population? What is the sample?
2. Is the 10.2 a statistic or a parameter of the population?
3. Would the researcher *know* the population mean in this circumstance?

4. What additional information would the researcher need in order to find a *confidence interval* for the population mean? Describe how that would be determined and how that would be interpreted.
5. If you knew that reasearcher was being lazy about digging and the clams he collected were all taken from sand levels no deeper than 10cm, would that be a *simple random sample*?

D. Suppose that a large data set of air temperature readings is normally distributed with $\overline{x} = 18.6°$C and $SD = .2°$ C.

1. What would be the z-score of a reading of 17.9°?
2. What temperature reading would correspond to a z-score of 1.4?
3. Based on this information, if a temperature reading T is selected at random from the data set, what is the probability that $18.2° \leq T \leq 18.9°$?
4. Based on this information, if a temperature reading T is selected at random from the data set, what is the probability that $T > 19.0°$?

E. Physicians measured the blood lead levels in 373 bridge workers employed by painting contractors in eight states. The lead levels had $\overline{x} = 27.2$ micrograms per liter of blood, with an SD of 16.1 micrograms per liter.

1. Determine a 95% confidence interval for the average lead level in bridge workers.
2. A health objective of a federal regulatory agency was the elimination of blood lead levels of 28 micrograms per liter or higher for these workers. From the evidence given by your confidence interval, does it seem that that objective was being met? Explain, by describing the way we interpret the meaning of a confidence interval of this sort.

F. A study shows that a 95% confidence interval for the average amount X of hazardous waste generated by a single hospital is $210 \leq X \leq 260$ (in units of kg/day). This interval was computed using the formulas we have discussed.

1. What was the sample mean $\overline{x}$ used to generate this confidence interval? What was the margin of error?
2. If the sample size was $n = 100$, what was the SD of the waste amounts in the sample?
3. If the sample size was $n = 16$, what was the SD of the waste amounts in the sample?