MONT 105Q – Mathematical Journeys
Information on Midterm Exam
March 17, 2016

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 1.

- This will be a full period, individual exam. No sharing of information 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 be similar to things you have seen on the problem sets. Some sample exam questions are given later in this document. The material to be covered is basically all the material from Chapters 1,2,3,5,5^{bis},6,7,8 in *Naked Statistics and the supplementary material we have discussed in class.*
- I will provide copies of the normal curve area table.
- 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 15-20 minutes, so it will be necessary to spend some of your preparation time on deciding what you want to say for each topic.

Possible Essay Topics

1. What was *Russell's Paradox* and what role did that play in the "crisis in foundations" in mathematics in the early 20th century. How do you think that discovery probably affected Russell personally and how might it have influenced what he thought about mathematics? (For instance, might it have changed his mind about the status of theorems proved in Euclidean geometry?) What were the three main schools of work devoted to trying to resolve that crisis and what were the ultimate results of each of them?
2. What is the central thesis of the article *The Dawning of the Age of Stochasticity* by David Mumford? How does Mumford think mathematics should develop in the current century? How is his proposal related to the outcomes of the various attempts to solve the "crisis of foundations" we read about in *Logicomix*? Would Mumford agree or disagree with this statement: "what mathematicians should be doing is closer to the kind of science Darwin practiced on the voyage of the *Beagle* than it is to a Euclidean search for absolute certainty." Explain.

Sample Mathematical Questions

Note: There are somewhat 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. 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 (min, 1st quartile, median, 3rd quartile, max).
2) Draw the corresponding "box and whisker" plot.
3) Construct a frequency histogram for the data using 4 equal bins on the range 20 to 60. Explain how you are treating data that falls at a bin boundary. (Any consistent method is OK.)
4) One measure of "skewness" of a data set is the quantity

$$B = \frac{Q_3 - 2 * \text{median} + Q_1}{Q_3 - Q_1}.$$

   Compute that statistic for this data set. What does it seem to indicate? Does this seem reasonable from the box plot?
5) Compute the SD of the data set. How many of the data values lie within two SD's of the mean?

B. The aces, kings, and queens from a standard deck of cards are removed and placed in a stack of twelve cards by themselves (there are three hearts, three spades, three diamonds and three clubs).

1) If a single card is selected at random from the stack, what is the probability that it is a heart?
2) If a single card is selected at from from the stack, what is the conditional probability that the card is a queen, given that it is a spade?
3) If a single card is selected from the stack, consider the two events: (1) the card is an ace, and (2) the card is a diamond. Are these independent?

3') Same question as 3) but on the stack of 11 cards obtained by removing the queen of diamonds.

4) Three random draws *with replacement* are made from this stack of cards. What is the probability that all three cards are spades?
5) Three random draws *without replacement* are made from the whole stack of 12 cards. What is the probability that all three are clubs?

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 value derived from particular data (a "statistic") or is it a property of the population as a whole (a "parameter")?
3) Would the researcher *know* the population mean in this circumstance?

4) What additional information would the researcher need in order to understand the distribution of shell widths? Describe how that would be estimated and how that would be interpreted.
5) What could we say about the distribution of sample means of samples of size 80 (even if the distribution of the individual shell widths was unknown)?
6) 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*? What kind of bias might this process introduce?

D. Suppose that a large data set of air temperature readings is normally distributed with $\mu = 18.6°$C and $\sigma = .2°$ C. The "z-score" of a temperature reading x is computed by

$$z = \frac{x - \mu}{\sigma}$$

1) What would be the z-score of a reading of $x = 17.9°$?
2) What temperature reading would correspond to a z-score of 1.4?
3) Based on this information, if a temperature reading x is selected at random from the data set, what is the probability that $18.2° \leq x \leq 18.9°$?
4) Based on this information, if a temperature reading x is selected at random from the data set, what is the probability that $x \geq 18.9°$?
5) Based on this information, if a temperature reading x is selected at random from the data set, what is the probability that $x \leq 18.2°$?
6) Based on this information, if $n = 10$ measurements are taken, what is the probability that exactly 5 of them are greater than 18.8?