Midterm (75 minutes)

No calculators allowed. Just set up your answers, for example, P = 49/52. BE SURE TO EXPLAIN YOUR REASONING. If you want extra time, you can buy time at a price of 1 point a minute; for example, if your test is handed in 10 minutes after the scheduled finish time, 10 points will be subtracted from the test score.

1. A Vanguard study of all mutual funds operating in 2013 found that 62% had beaten the overall stock market during the previous five years. What statistical problem do you see here?

2. Identify three misleading things about this figure that is intended to demonstrate an abnormally large increase after 1980 in the number of households earning more $100,000 a year:
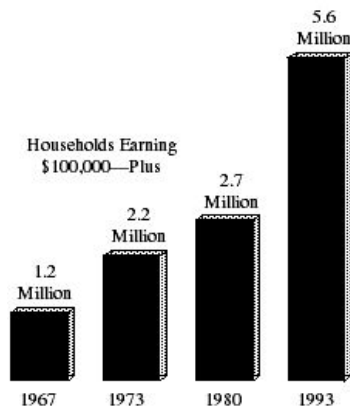
3. The annual return from Stock 1 can be modeled as a random draw from a normal distribution with a mean of 5% and standard deviation of 10%, while the annual return from Stock 2 can be modeled as a random draw from a normal distribution with a mean of 10% and standard deviation of 20%. The annual returns are independent. Which stock has a higher probability of a positive return?

4. Consider a pick-6 Lotto game where a player picks 6 different numbers from 1 to 47 and wins the grand prize if these 6 numbers match (not necessarily in order) the 6 numbers that the contest organizer draws randomly from ping pong balls numbered 1 to 47. What is the probability that the 6 numbers that are drawn will be the same 6 numbers (not necessarily in the same order) that were drawn in the previous contest?

5. A box has three drawers. One drawer contains two gold coins; another drawer contains two lead coins; and the third drawer contains one gold coin and one lead coin. You pick a drawer at random and a coin from that drawer at random. It is a gold coin, which is set aside and you do not get to keep. Now, your options are to either (a) take the other coin in the drawer that you chose; or (b) randomly pick a coin from one of the two drawers you did not choose. If you choose Option (a), what is the probability that you will get a gold coin?

6. The Black Swan problem involves
   a. outliers
   b. law of averages
   c. law of large lumbers
   d. self-selection bias

7. In 1999 Sally Clark, an English solicitor, was accused of murdering her two infant sons; she said that both children had died of sudden infant death syndrome (SIDS). A pediatric professor testified for the prosecution that the probability that a child in an affluent family would suffer SIDS is 1/8500 and, therefore, the probability that two children in the family would suffer SIDS is (1/8500)(1/8500) = 1/72,250,000, or about 1 in 72 million. As an expert witness for the defense,
   a. how might you challenge the 1-in-72 million calculation?

   b. how would you challenge the prosector's conclusion that, if the probability of two SIDS deaths is 1 in 72 million, then the probability that the Clark is innocent is 1 in 72 million?

8. Two roommates, Ocean and Cove, are taking a statistics class together. Each week, the 16 students in the class are randomly separated into eight 2-person teams.
   a. What is the probability that Ocean and Cove will be on the same team the first week of the term?

   b. If there are 12 weeks in the term, what is the probability that Ocean and Cove will be on the same team at least once?

9. A hot-hand study looked at all 383 shots taken by Jeremy Lin when he played basketball for the New York Knicks during the 2011-2012 season. The researcher separated these shots into 38 groups of 10: the first 10 shots he took that season, the next 10 shots he took that season, and so on (with the final 3 shots discarded). The researcher then looked at these 38 sets of 10 shots and calculated how frequently Lin made 0 of 10 shots, 1 of 10 shots, 2 of 10 shots, and so on, in order to see whether Lin made 0, 1, 2, 8, 9, or 10 shots more often than predicted by the binomial distribution. Which assumptions made by the binomial model are likely to be violated in this study?

10. In the television show, *The Sopranos*, one psychiatrist tells another that treating sociopaths with talk therapy was counterproductive. In fact, a study of the criminal personality concluded that

   > *talk therapy, while not only being useless with sociopaths, actually serves to validate them….They sharpen their skills as con men on their therapy. He showed data. The one-year reconviction rates were higher than for those receiving no treatment at all.*

   Why, as a statistician, would you question this conclusion?