Homework 6: Covariance, Correlation

Instructions: Your answers are due at the beginning of class on the due date. You can either turn in a paper copy, or a pdf version through canvas. We highly recommend using latex (`http://www.cs.utah.edu/~jeffp/teaching/latex/`) for producing the assignment answers. If the answers are too hard to read (e.g. **pdf's from your phone's camera must be very clear and careful**) you will lose points (entire questions may be given 0).

Please make sure your name appears at the top of the page.

You may discuss the concepts with your classmates, but write up the answers entirely on your own. **Be sure to show all the work involved in deriving your answers! If you just give a final answer without explanation, you may not receive credit for that question.**

1. A polling station in Salt Lake City has 50 registered Republicans and 50 registered Democrats. After 3 random people show up to vote, let R be the number of Republicans and D be the number of Democrats that have voted.

   (a) Compute and list all the entries of the joint probability table $P(R = a, D = b)$ that are non-zero. *(Hint: it should only have 4 squares in total.)*

   (b) What is $\text{Cov}(R, D)$?

   (c) What is $\rho(R, D)$?

   Hint: These computations may seem daunting, but the final answers should come out to be simple numbers. You can use a calculator of course, but don't just give the final answer – show your work using the mathematical formulas from class/book!

2. In January, let X represent the fraction of time it is snowing at Snowbird, and Y represent the fraction of time it is snowing at the SLC airport. Say they have the following joint density of X and Y

   $$f(x, y) = \begin{cases} \frac{2}{11}(10x + y) & \text{for } x, y \in [0, 1] \times [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

   (a) What is the probability that the it is snowing at Snowbird less than half of the time?

   (b) What is $\text{Cov}(X, Y)$?

   (c) What is $\rho(X, Y)$?

3. **(Extra Credit: 10pts)** Consider a more general situation for Problem 1, where instead of 50 Republicans and 50 Democrats, we have n people from each party, and instead of the first 3 people that show up, we have the first k people that show up. Again, let R be the number of Republicans and D be the number of Democrats that show up out of the first k. What is $\rho(R, D)$ in this case? To get the extra credit, you have to derive $\rho(R, D)$ using equations, not just give the final answer!