HW14: Joint distributions
(assigned on Thur Nov 17, due on Wed Nov 30)

You may work with in a group of as many as three students on this assignment, handing in one report with all names, provided that you all contribute to the work.

1. As in HW7, suppose that 4 men and 4 women take an exam and are then ranked from highest score (rank 1) to lowest score (rank 8). Suppose that no ties occur and all orderings are equally likely. Let $X =$ rank for highest-scoring woman, and let $Y =$ rank for highest-scoring man.

a) Determine the joint pmf of $(X,Y)$. (As always, show your work.)

b) Determine the covariance between $X$ and $Y$.

c) Is the covariance between $X$ and $Y$ positive, negative, or zero? Explain why this makes sense intuitively, based on the context.

2. Suppose that Earl’s bowling score follows a normal distribution with mean 150 and SD 20 for the first game that he bowls in a day. But for the second game, he generally does better and is more consistent, so his score for the second game follows a normal distribution with mean 165 and SD 15. Also suppose that the correlation between his two scores is 0.6.

a) Determine the probability that his second score exceeds his first score.

b) Determine the probability that the sum of his two scores exceeds 350.

3. Suppose that $X$ has an exponential distribution with mean 1. Also suppose that conditional on $X = x$, $Y$ has an exponential distribution with mean $x$.

a) Write R code to simulate $N$ repetitions of $(X,Y)$ pairs. The code should also produce histograms and calculate means and standard deviations of the individual variables, and the code should produce a scatterplot and calculate the correlation between the variables. Submit your code. [Hint: Remember that the parameter of an exponential distribution is the reciprocal of the mean.]

b) Run the code for $N = 1000$ repetitions. Produce (and submit) a scatterplot of the results. Comment on what the scatterplot reveals about whether/how $X$ and $Y$ are correlated.

c) Run the code for $N = 100,000$ repetitions. Report the means and standard deviations of each variable, and the correlation between the variables.

d) Based on the simulation results, do you think that the (marginal) distribution of $Y$ is an exponential distribution? Explain, based on the graphs and calculations from the simulation.