This exam is intended to take about 3 hours; given that you have other final exams to take, I encourage you not to spend much more time on it than that. You may refer to your notes and to the course textbook as you solve these problems. *You may not discuss the problems, through any medium, with any other person, until after 11am on Wednesday, March 18, by which time you must upload your solutions to Gradescope.*

You must submit two separate items to Gradescope:

1. the academic integrity pledge which is on this page and is Problem 1 on the exam, worth 10 points;
2. the rest of the exam, namely Problems 2–6, worth 90 points.

Please complete and upload your academic integrity pledge to Gradescope as soon as possible, to ensure that you know how Gradescope works. You may either write your exam on paper and upload pictures of each page, or you may write your exam using a tablet and upload a .pdf file of your work. All work must be clear and legible, pages must be in the correct order, and you must assign pages to questions as directed by Gradescope. You may be deducted points if you fail to follow any of these directions.

Please show your work and simplify your answers as much as possible. Explain your answers, concisely—coherent explanations are the best evidence you can provide that you understand the material.

1. [10 points] Please write out the following academic integrity pledge and sign it:

I am fair to my classmates and instructors by not using any unauthorized aids.
I respect myself and my university by upholding educational and evaluative goals.
I am honest in my representation of myself and of my work.
I accept responsibility for ensuring my actions are in accord with academic integrity.
I show that I am trustworthy even when no one is watching.

2. [18 points] Robert Newman plans to send a message, X, to John Hancock. The message will be "one" with probability 1/3 and "two" with probability 2/3. Because Robert is using primitive technology (lanterns) there is some probability that the message, Y, John receives will be wrong. Specifically,

$$P(Y = \text{"two"} \mid X = \text{"one"}) = 0.1$$
$$P(Y = \text{"one"} \mid X = \text{"two"}) = 0.5$$

Suppose Y = "two". What probability does John assign to the message Robert sent having been "two"? That is, what is $P(X = \text{"two"} \mid Y = \text{"two"})$?

Apply Bayes' rule:

$$\begin{aligned}
&P(X = \text{"two"} \mid Y = \text{"two"}) \\
&\quad = \frac{P(Y = \text{"two"} \mid X = \text{"two"})P(X = \text{"two"})}{P(Y = \text{"two"} \mid X = \text{"one"})P(X = \text{"one"}) + P(Y = \text{"two"} \mid X = \text{"two"})P(X = \text{"two"})} \\
&\quad = \frac{(1/2)(2/3)}{(1/10)(1/3) + (1/2)(2/3)} = \frac{10}{11}
\end{aligned}$$

3. [15 points] Two points are chosen independently and uniformly at random on a circle (not inside the circle). What is the probability that the line segment connecting them has length less than the radius of the circle?

   For convenience, center the circle at the origin of the standard (x, y)-coordinate system. The probability doesn't change if we rotate the circle after the points are chosen, so rotate it so that one point lies on the positive x-axis. Then the angular coordinate of the other point can lie anywhere in the interval $(-\pi/3, \pi/3)$. The ratio of the length of this arc to the circumference of the circle is the probability: $(2\pi/3)/(2\pi) = 1/3$.

4. [18 points] Let $N(I)$ be the number of points of a Poisson process with intensity λ that are in an interval $I \subset \mathbb{R}$. Let T_1 be the smallest point of the Poisson process in the interval $[0, 2]$. Given that $N([0, 2]) = 1$, what is the probability density function for T_1? [Hint: It is not the p.d.f. of an exponential random variable.]

   Find the cdf using the definition of conditional probability:

$$P\big(T_1 < t \mid N([0,2]) = 1\big) = \frac{P\big(N([0,t)) = 1 \text{ and } N([t,2]) = 0\big)}{P\big(N([0,2]) = 1\big)}$$
$$= \frac{e^{-\lambda t}(\lambda t)^1/1! \cdot e^{-\lambda(2-t)}(\lambda(2-t))^0/0!}{e^{-\lambda \cdot 2}(\lambda \cdot 2)^1/1!} = \frac{t}{2},$$

   for $t \in [0, 2]$. Taking the derivative with respect to t gives the pdf:

$$f_{T_1 \mid N([0,2])=1}(t) = \begin{cases} 1/2 & \text{if } t \in [0,2]; \\ 0 & \text{otherwise.} \end{cases}$$

5. Let $X_i \sim \mathrm{Ber}(p)$ be independent random variables for $i \in \{1,2,3\}$. Let $Y_1 = X_1 - X_2$, $Y_2 = X_2 - X_3$, and $Y_3 = X_3 - X_1$.

   a. [6 points] What is the probability mass function for Y_1?

$$\begin{aligned} P(Y_1 = -1) &= P(X_1 = 0, X_2 = 1) = (1-p)p \\ P(Y_1 = 0) &= P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 1) = (1-p)^2 + p^2 \\ P(Y_1 = 1) &= P(X_1 = 1, X_2 = 0) = p(1-p) \end{aligned}$$

   b. [6 points] What is $\mathsf{Corr}[Y_1, Y_2]$?

   $\mathsf{Var}[Y_1] = \mathsf{Var}[X_1 - X_2] = \mathsf{Var}[X_1] + \mathsf{Var}[X_2] = 2p(1-p) = \mathsf{Var}[Y_2]$. Also

$$\begin{aligned} \mathsf{Cov}[Y_1, Y_2] &= \mathsf{Cov}[X_1 - X_2, X_2 - X_3] \\ &= \mathsf{Cov}[X_1, X_2] - \mathsf{Cov}[X_1, X_3] - \mathsf{Cov}[X_2, X_2] + \mathsf{Cov}[X_2, X_3] \\ &= -\mathsf{Var}[X_2] = -p(1-p), \end{aligned}$$

   so $\mathsf{Corr}[Y_1, Y_2] = -p(1-p)/(2p(1-p)) = -1/2$.

   c. [6 points] What is the joint probability mass function of Y_1 and Y_2?

   The strings $X_1X_2X_3$ that give each pair of values (Y_1, Y_2) are

$$\begin{array}{c} {}_{Y_1}\backslash^{Y_2} \\ -1 \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{ccc} -1 & 0 & 1 \end{array} \\ \begin{pmatrix} & \scriptstyle 011 & \scriptstyle 010 \\ \scriptstyle 001 & \begin{smallmatrix} 000 \\ 111 \end{smallmatrix} & \scriptstyle 110 \\ \scriptstyle 101 & \scriptstyle 100 & \end{pmatrix} \end{array}, \text{ so the pmf is } \begin{array}{c} {}_{Y_1}\backslash^{Y_2} \\ -1 \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{ccc} -1 & 0 & 1 \end{array} \\ \begin{pmatrix} 0 & p^2(1-p) & p(1-p)^2 \\ p(1-p)^2 & (1-p)^3 + p^3 & p^2(1-p) \\ p^2(1-p) & p(1-p)^2 & \end{pmatrix} \end{array}.$$

   d. Extra Credit [5 points] The *correlation matrix* for Y_1, Y_2, and Y_3 is the 3×3 matrix C, with $C_{ij} = \mathsf{Corr}[Y_i, Y_j]$. Find the eigenvalues of C. Can you make a conjecture about the eigenvalues of any correlation matrix?

   For all $i \in \{1,2,3\}$, $C_{ii} = 1$, and by symmetry, for all $i \neq j \in \{1,2,3\}$, $C_{ij} = -1/2$. To compute the eigenvalues, solve

$$\begin{aligned} 0 = \det \begin{pmatrix} 1-\lambda & -1/2 & -1/2 \\ -1/2 & 1-\lambda & -1/2 \\ -1/2 & -1/2 & 1-\lambda \end{pmatrix} & \\ = (1-\lambda)^3 - \frac{2}{8} - \frac{3}{4}(1-\lambda) & \\ = -\frac{\lambda}{4}(4\lambda^2 - 12\lambda + 9) & \\ = -\frac{\lambda}{4}(2\lambda - 3)^2, & \end{aligned}$$

   so the eigenvalues are 0 and 3/2 (twice).

We proved in class that for any two random variables the correlation $\rho \in [-1, 1]$. In this case the correlation matrix is

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \text{ which has eigenvalues } \lambda \in \{1+\rho, 1-\rho\}.$$

Notice that these are necessarily non-negative. With these two examples, and some imagination, one might conjecture that the eigenvalues of a correlation matrix are always non-negative. (Notice that they must be real because the matrix is real and symmetric.) In fact, this conjecture is true; it is usually stated as the correlation matrix is *positive semi-definite*, which means that for any real vector v of the dimension of the correlation matrix, $v^\intercal C v \geq 0$.

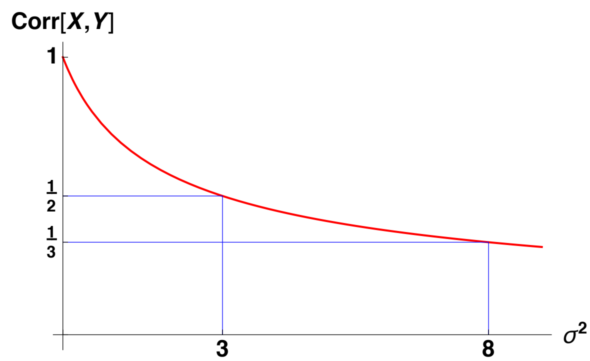
6. Let X and Z be independent random variables, with $\mathsf{Var}[X] = 1$ and $Z \sim \mathcal{N}(0, \sigma^2)$. Let $Y = X + Z$.

   a. [14 points] Find $\mathsf{Corr}[X, Y]$ and draw a plot of it as a function of σ^2.

   $\mathsf{Cov}[X, Y] = \mathsf{Cov}[X, X + Z] = \mathsf{Var}[X] = 1$, and $\mathsf{Var}[Y] = \mathsf{Var}[X + Z] = 1 + \sigma^2$ so

   $$\mathsf{Corr}[X, Y] = \frac{1}{\sqrt{1+\sigma^2}},$$

   and here is a plot:



   b. [7 points] $X \sim \mathrm{Unif}[0, \sqrt{12}]$ has $\mathsf{Var}[X] = 1$. Let $X_1, \ldots, X_N$ be i.i.d. with distribution $\mathrm{Unif}[0, \sqrt{12}]$, and let $Z_1, \ldots, Z_N$ be i.i.d. with distribution $\mathcal{N}(0, 1/4)$. Assume all the random variables $\{X_1, \ldots, X_N, Z_1, \ldots, Z_N\}$ are independent. Let $Y_i = X_i + Z_i$, so that we get a set of N pairs (X_i, Y_i). Draw a plot of what they probably look like in the (x, y) plane.

   The X coordinates of the pairs are chosen uniformly at random between 0 and $\sqrt{12}$, and the Y coordinates will be scattered around the line $Y = X$, with a mean deviation of 0 and a standard deviation of $1/2$. Thus they will look something like: