AMS 206 – Homework Assignment #1
due Wednesday, January 23

Required Problems

1. **Do the parts of this question in order, and do each part before reading the next part!** Consider the four events A_1, A_2, A_3, and A_4, which are described as follows:
   A_1 is the event that the area of the state of Pennsylvania is less than 5,000 sq. miles
   A_2 is the event that the area of the state of Pennsylvania is between 5,000 and 50,000 sq. miles
   A_3 is the event that the area of the state of Pennsylvania is between 50,000 and 100,000 sq. miles
   A_4 is the event that the area of the state of Pennsylvania is more than 100,000 sq. miles

   (a) Without using any outside information, assign your subjective probabilities to these four events. Do this before reading the rest of the question!

   (b) You are now given the information that the area of Alaska (the largest of the 50 states) is 586,000 sq. miles and the area of Rhode Island (the smallest) is 1,214 sq. miles. In light of this information, determine your subjective probablities of the four events.

   (c) You are now given the information that when area is considered, Pennsylvania is the 33rd largest of the 50 states (i.e., 32 states are larger). Update your probabilities.

   (d) You are now given the information that the area of New York, the 30th largest state, is 49,576 sq. miles. Update your probabilities.

2. The loop shuttle runs every 10 minutes. Suppose the waiting time for the next shuttle is exponentially distributed. Find the probability that you have to wait between 5 and 9 minutes (do this by evaluating the integral of the pdf).

3. Suppose we observe 7 artichoke plants and see that they produced the following number of artichokes per plant:

   28 16 48 16 14 35 32

   Let y_i be the number of artichokes on plant i, and suppose we want to model the y_i as independent realizations from a Gamma distribution with parameters α and β (i.e., with mean α/β; $f(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}$).

   (a) Suppose that α is known to be 9. Find the maximum likelihood estimator of β.

   (b) Now suppose that α is unknown, but β is known to be 0.3.

      i. Try to find the maximum likelihood estimator of α and see that you cannot find it in closed form.

      ii. Plot the likelihood as a function of α.

      iii. Use the plot to find an approximate maximum likelihood estimator of α. (For this problem, it is sufficient to just eyeball the plot; optionally, you can do the maximization numerically.)

Optional Problem

4. Consider the case where you observe data $Y_1, \ldots, Y_n$ which are modeled as *iid* $N(\mu, \frac{1}{\tau})$. Note that using the precision (instead of the variance) will make the computations easier.

   (a) Write out the likelihood function for μ and τ, $L(\mu, \tau|\mathbf{y}) = f(\mathbf{y}|\mu, \tau)$.

   (b) Find the marginal likelihood for τ, i.e., $L(\tau|\mathbf{y}) = f(\mathbf{y}|\tau) = \int f(\mathbf{y}|\mu, \tau)d\mu$, the likelihood function for τ when μ has been integrated out. You will need to expand the square in the exponential term and then complete the square for μ. After you have integrated out μ, simplify the resulting expression for τ to see that it has the form of a gamma distribution (but without the right normalizing constant).

   (c) Find the maximum likelihood estimate for τ using (i) the original likelihood (with τ and μ) and (ii) the marginal likelihood.

   (d) The two MLEs should be very similar, but not quite the same. Explain how and why they differ in this case.