Homework 4:
Variational Inference

Instructions:

1. Questions 1, 3-5 in this assignment require coding in python. Submit a jupyter notebook with both your code and your answers to the questions.
2. For question 2 you may either include your answers in your jupyter notebook or in a separate file.
3. You may discuss this assignment with other students in the class, but you must submit your own answers to the questions below.
4. Include an honor pledge with your submission.
5. Submit on-line.
6. This homework is worth 100 points and the point totals for each question are shown in parentheses.

Assignment:

1. (20) Use the EM implementation for Gaussian mixture models (GMM) in sklearn to segment the image Haiti_Image.tif into four classes. Then answer the following questions.
   - Compare the classification results obtained with the GMM to those obtained from K-means (KM) segmentation of the same image. Again use four classes for KM. Show the results of the two approaches to segmentation as images with the separate classes shown in different colors. Comment on the similarities and differences in the two segmentations.
   - Write the equation for the posterior prediction of the probability of the class label for each data points in the image.
   - Show four separate images, one for each class, that color code the probabilities of a class label for each data point in the image. Comment on the insights provided by this probabilistic view of the segmentation. Very briefly describe example of how this probabilistic perspective might be used by governmental or non-profit agencies concerned with improving the environment or the living conditions in Haiti.
2. (20) Explain how the Evidential Lower Bound (ELBO) has components that trade-off between bias and variance.
3. (20) Provide displays that illustrate variational approximation for the following two examples.

- Suppose we have a two dimensional Gaussian with an arbitrary mean, μ, and with a variance-covariance matrix with $\sigma_{11} = 1.0, \sigma_{22} = \sqrt{0.5}$, and the correlation, $\rho = 0.75$ Show the variational approximation that minimizes the Kullback-Liebler divergence with this Gaussian using the mean field assumption. Explain the shape of this approximation.
- Suppose we have the following results from tests of infections in the sore throats of our patients: $\{0, 1, 0, 1, 1, 0, 0, 0, 0, 0\}$, where 1 indicates an infection and 0 indicates no infection. We assume a uniform prior for the probability of infection. Show the plot of the true posterior and use the ADVI optimization to obtain the variational approximation of the posterior. Overlay the plot of this approximation on the true posterior.
- (+3 extra credit) Show the Laplace approximation of the posterior and briefly comment on the visible differences with the variational approximation.

4. (20) With the CHD data set (CHDdata.csv) from the previous two homework assignments develop a variational inference estimate using ADVI optimization for the posterior distributions of the parameters in a main effects logistic regression model. Use all predictor variables in the data set and standardize all of the numeric, continuous predictors using the mean and standard deviation. For each of the parameters in the regression use a Gaussian prior with a mean vector of $\mathbf{0}$ and the identity matrix as the variance-covariance matrix. Show the ELBO plot from the optimization and discuss how it looks. Use plots to discuss the marginals for each of the regression model parameters obtained from variational inference and compare them to the ones you obtained using sampling.

5. (20) With the data set bangladesh.csv develop a partial-pooled model for all districts to predict contraception usage. Do not use any other predictor variable besides district. Use variational inference with ADVI optimization to develop estimates of the proportions of women using contraception in each district. Plot each of these estimates by district. Compare your results with those you obtained using sampling.