Common Instructions: You must submit answers via Google Classroom's assignment submission interface before the appropriate deadline. The format of the submission file is `pdf` (for e.g., scan your hand-written answers using mobile app and make a pdf from it). Please write legibly and concisely. Always simplify your answers as much as possible. For coding/simulations you are free to use any programming language.

Week1

1. Exercises 2.1, 2.4, 2.7, 2.10, 2.16, 2.17 in Kevin Murphy's textbook.

* Exercises 27, 37, 38 in Chapter 2 in Sheldon Ross's book[1].

Week1a

1. In section 2 in Murphy's book many common discrete (sec 2.3), and common continuous (sec 2.4) distributions are presented. For each of these cases implement the parameter estimation algorithm using (any variant of) method of moments. In each case, your code must take as input the training data and output the parameters of the distribution. Also, repeat this exercise for Multivariate Gaussian (sec 2.5.2) and the Dirichlet distribution (sec 2.5.4). You are free to use any numerical equation solver or numerical optimization package. Finally, you need to submit the following (in the same pdf):

   (a) For each univariate distribution plot the estimated parameter values vs. the number of training datapoints. Plot all parameters in the same graph (different colors/markers for each parameter). One graph per distribution. The details of your training data will be communicated separately.

   (b) For each multivariate distribution (Gaussian and Dirichlet), compute the (joint) likelihood of the training dataset[2] with the estimated parameter values. Plot likelihood of training data vs. the number of training datapoints.

   (c) Note your observations (if any) from the plots.

   Please do NOT submit anything else than the above list.

Week2

1. Exercises 3.6, 3.8(a), 3.11(a), 3.11(b), 4.1, 4.2, 4.3, 4.4.

[1] We follow the 10^{th} edition of Sheldon Ross's "Introduction to Probability Models".

[2] From iid assumption, it follows that the (joint) likelihood of the training data is product of likelihoods of each training datapoint (datum).

* Consider the random variable corresponding to the MLE[3]. Analogous to the Central Limit Theorem, can something be said about the asymptotic distribution of the MLE? Provide details.

Week3

1. Implement the parameter estimation algorithm for the Gaussian model based regression model we studied in Thursday's (16-08-2018) lecture. Your code should take in as input a regression training dataset (with n_1 input variables and n_2 output variables) and build the regression function that predicts the output for any input. The details of the dataset and the simulation plots you need to submit will be released on the day of the deadline. Meanwhile verify the correctness of your code using some simulated data of your choice.

Week4

1. Exercise 4.2, 7.3, 7.4, 7.5, 7.6.

* Exercise 7.7.

Week5

1. Implement the Gaussian Discriminant (Bayes classifier;eqn.4.33) and Naive bayes classifier (use Gaussian model for each input feature/attribute). Compare the accuracy[4] of these two classifiers with that of Logistic Regression[5] on your dataset.

2. Exercise 4.19, 4.21, 4.22.

* Exercise 4.20, 8.6, 8.7.

---

[3] Because samples are random, the estimate is also random.

[4] Accuracy is percentage of correctly classified test data samples.

[5] Download code from `https://www.csie.ntu.edu.tw/~cjlin/liblinear/` and use "train -s 0 -B 1 -C trainingsetfile modelfile" command and use "predict testfile modelfile outputfile".