# Class Exercises. Set 1. Probability and Statistics

1. In Python, simulate the sum of two random variables, X and Y, both uniformly distributed between 0 and 1.

2. Assume that X is distributed according to the Gaussian density with mean $\mu_X = 0$ and variance $\sigma_X^2 = 1$; Y is distributed according to the Gaussian density with mean $\mu_Y = 1$ and variance $\sigma_Y^2 = 1$.
   a) What is the probability density function $f_X$ at X = 0?
   b) What is the probability density function $f_Y$ at Y = 0?
   c) Assume the distribution P(Z=z) = 0.5P(X = z)+0.5P(Y = z) known as mixture (i.e., 1/2 of the time points are generated by the X process and 1/2 of the time by the Y process). If Z = 0, what is the probability that the X process generated this data point?

3. In Python, simulate the mixture of two random variables, X and Y, X uniformly distributed between 0 and 1 and Y uniformly distributed between 4 and 6. The mixing coefficients are 0.4 and 0.6 respectively. Play with different values for the mixing coefficients.

4. To estimate the probability of getting heads, $p$, on the throw of a coin, we repeat the experiment $n$ times, independently, and we count the number of successes (heads), $Y$. If $X$ is a random variable with

$$\begin{cases} X = 1, \text{ if head} \\ X = 0, \text{ if tails} \end{cases}$$

   then $Y = \sum_{i=1}^{n} X_i$.
   Consider the following two estimators of $p$: $\hat{P}_1 = Y/n$ e $\hat{P}_2 = (Y+1)/(n+2)$.

   (a) Check if $\hat{P}_1$ and $\hat{P}_2$ are unbiased estimators of $p$.

   (b) Compute the mean-square error (MSE) of the estimators $\hat{P}_1$ and $\hat{P}_2$. Confirm that the MSE of $\hat{P}_2$ is less than the MSE of $\hat{P}_1$ when the true (but unknown in practice) $p$ if equal to 0.5.

5. Load the height/weight data using data = np.genfromtxt('heightWeightData.txt', delimiter=','). The first column is the class label (1=male, 2=female), the second column is height, the third weight.
   a) Write a Python script to fit a Gaussian model to each class using all the data for training. What's the training error?
   b) Repeat a) imposing the same covariance matrix for both classes.
   c) Repeat a) imposing the diagonal covariance matrices.