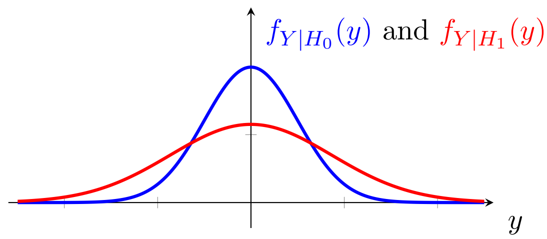
Problem 1 (Detection)

Surprisingly, you have been asked to evaluate the following continuous binary hypothesis testing scenario during your summer internship. The hypotheses are equally likely, $\mathbb{P}[H_0] = \mathbb{P}[H_1]$. Under H_0, Y is Gaussian with mean 0 and variance 1. Under H_1, Y is Gaussian with mean 0 and variance 3. The resulting likelihoods are written out and sketched below.

$$f_{Y|H_0}(y) = \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{y^2}{2}\Big) \qquad\qquad f_{Y|H_1}(y) = \frac{1}{\sqrt{6\pi}} \exp\Big(-\frac{y^2}{6}\Big)$$



(a) Compute $\mathbb{P}[Y^2 < 1|H_0]$.
(You may leave your answer in terms of the standard normal CDF $\Phi(z)$.)

Solution:

$$\begin{aligned}\mathbb{P}[Y^2 < 1|H_0] &= \mathbb{P}[-1 < Y < 1|H_0] = F_{Y|H_0}(1) - F_{Y|H_0}(-1) \\ &= \Phi\Big(\frac{1-0}{1}\Big) - \Phi\Big(\frac{-1-0}{1}\Big) = \Phi(1) - \Phi(-1)\end{aligned}$$

(b) Compute $\mathbb{P}[Y^2]$.
(You may leave your answer in terms of the standard normal CDF $\Phi(z)$.)

Solution:

$$\begin{aligned}\mathbb{P}[Y^2 < 1|H_1] &= \mathbb{P}[-1 < Y < 1|H_1] = F_{Y|H_1}(1) - F_{Y|H_1}(-1) \\ &= \Phi\Big(\frac{1-0}{\sqrt{3}}\Big) - \Phi\Big(\frac{-1-0}{\sqrt{3}}\Big) = \Phi\Big(\frac{1}{\sqrt{3}}\Big) - \Phi\Big(-\frac{1}{\sqrt{3}}\Big) \\ \mathbb{P}[Y^2 < 1] &= \mathbb{P}[Y^2 < 1|H_0]\,\mathbb{P}[H_0] + \mathbb{P}[Y^2 < 1|H_1]\,\mathbb{P}[H_1] \\ &= \frac{1}{2}\Big(\mathbb{P}[Y^2 < 1|H_0]\,\mathbb{P}[H_0] + \mathbb{P}[Y^2 < 1|H_1]\Big) \\ &= \frac{1}{2}\Big(\Phi(1) - \Phi(-1) + \Phi\Big(\frac{1}{\sqrt{3}}\Big) - \Phi\Big(-\frac{1}{\sqrt{3}}\Big)\Big)\end{aligned}$$

(c) Evaluate the log-likelihood ratio $\ln(L(y))$.

Solution:

$$\ln(L(y)) = \ln\left(\frac{f_{Y|H_1}(y)}{f_{Y|H_0}(y)}\right) = \ln\left(\frac{1}{\sqrt{3}}\exp\left(y^2\Big(\frac{1}{2}-\frac{1}{6}\Big)\right)\right)$$
$$= \ln\left(\frac{1}{\sqrt{3}}\right) + \frac{y^2}{3} = \frac{1}{3}y^2 - \frac{1}{2}\ln(3)$$

(d) Evaluate the ML rule $D^{\mathrm{ML}}(y)$.

Solution:

$$D^{\mathrm{ML}}(y) = \begin{cases} 1 & \ln(L(y)) \geq 0 \\ 0 & \ln(L(y)) < 0 \end{cases} = \begin{cases} 1 & y^2 \geq \frac{3}{2}\ln(3) \\ 0 & y^2 < \frac{3}{2}\ln(3) \end{cases}$$

(e) Calculate the probability of error for the ML rule.
(You may leave your answer in terms of the standard normal CDF $\Phi(z)$.)

Solution:

Below, we take advantage of the fact that the PDF of a zero-mean Gaussian is symmetric about the origin. For conciseness, define $\beta = \frac{3}{2}\ln(3)$.

$$\begin{aligned}
P_{\mathrm{FA}} &= \mathbb{P}[Y^2 \geq \beta | H_0] = 2\mathbb{P}[Y < -\beta | H_0] = 2F_{Y|H_0}(-\beta) \\
&= 2\Phi\Big(\frac{-\beta - 0}{1}\Big) = 2\Phi(-\beta) \\
P_{\mathrm{MD}} &= \mathbb{P}[Y^2 < \beta | H_1] = \mathbb{P}[-\beta < Y < \beta | H_1] = F_{Y|H_1}(\beta) - F_{Y|H_1}(-\beta) \\
&= \Phi\Big(\frac{\beta - 0}{\sqrt{3}}\Big) - \Phi\Big(\frac{-\beta - 0}{\sqrt{3}}\Big) = \Phi\Big(\frac{\beta}{\sqrt{3}}\Big) - \Phi\Big(-\frac{\beta}{\sqrt{3}}\Big) \\
P_{\mathrm{error}} &= P_{\mathrm{FA}}\,\mathbb{P}[H_0] + P_{\mathrm{MD}}\,\mathbb{P}[H_1] \\
&= \frac{1}{2}\left(2\Phi(-\beta) + \Phi\Big(\frac{\beta}{\sqrt{3}}\Big) - \Phi\Big(-\frac{\beta}{\sqrt{3}}\Big)\right) \\
&= \frac{1}{2}\left(2\Phi(-\frac{3}{2}\ln(3)) + \Phi\Big(\frac{\sqrt{3}}{2}\ln(3)\Big) - \Phi\Big(-\frac{\sqrt{3}}{2}\ln(3)\Big)\right)
\end{aligned}$$

Problem 2 (Estimation)

On the second day of your increasingly strange summer internship, you are asked to construct estimators for jointly Gaussian random variables.

(a) Let X and Y_1 be jointly Gaussian random variables with the following parameters: $\mathbb{E}[X] = -2\,,\ \mathbb{E}[Y_1] = 0\,,\ \mathrm{Var}[X] = 3\,,\ \mathrm{Var}[Y_1] = 4\,,\ \mathrm{Cov}[X, Y_1] = 2\,.$
Determine the MMSE estimator of X given $Y_1 = y_1\,.$

Solution:

Since X and Y_1 are jointly Gaussian, the MMSE estimator is the LLSE estimator:

$$\begin{aligned}\hat{x}_{\mathrm{MMSE}}(y_1) = \mathbb{E}[X] + \frac{\mathrm{Cov}[X, Y_1]}{\mathrm{Var}[Y_1]}(y_1 - \mathbb{E}[Y_1]) &= -2 + \frac{2}{4}(y_1 - 0) \\ &= \frac{y_1}{2} - 2\end{aligned}$$

(b) Calculate the mean-squared error of the MMSE estimator.

Solution:

$$\begin{aligned}\mathrm{MSE}_{\mathrm{MMSE}} = \mathrm{Var}[X] - \frac{(\mathrm{Cov}[X, Y_1])^2}{\mathrm{Var}[Y_1]} &= 3 - \frac{2^2}{4} \\ &= 2\end{aligned}$$

(c) You are now given an additional observation Y_2 satisfying $\mathbb{E}[Y_2] = 0\,,\ \mathrm{Var}[Y_2] = 1\,,$ $\mathrm{Cov}[X, Y_2] = 1\,,\ \mathrm{Cov}[Y_1, Y_2] = 0\,.$ You may assume that $X\,,\ Y_1\,,$ and Y_2 are jointly Gaussian. Let $\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$. Write down the cross-covariance matrix $\mathbf{\Sigma}_{X\underline{Y}}$ and the covariance matrix $\mathbf{\Sigma}_{\underline{Y}}\,.$

Solution:

$$\begin{aligned}\mathbf{\Sigma}_{X\underline{Y}} &= \begin{bmatrix} \mathrm{Cov}[X, Y_1] & \mathrm{Cov}[X, Y_2] \end{bmatrix} = \begin{bmatrix} 2 & 1 \end{bmatrix} \\ \mathbf{\Sigma}_{\underline{Y}} &= \begin{bmatrix} \mathrm{Var}[Y_1] & \mathrm{Cov}[Y_1, Y_2] \\ \mathrm{Cov}[Y_2, Y_1] & \mathrm{Var}[Y_2] \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}\end{aligned}$$

(d) Determine the vector LLSE estimator $\hat{x}_{\mathrm{LLSE}}(\underline{y})\,.$

Solution:

Since X and Y_1 are jointly Gaussian, the MMSE estimator is the LLSE estimator:

$$\begin{aligned}\hat{x}_{\mathrm{LLSE}}(\underline{y}) &= \mathbb{E}[X] + \mathbf{\Sigma}_{X\underline{Y}}\mathbf{\Sigma}_{\underline{Y}}^{-1}\underline{y} \\ &= -2 + \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1/4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= \frac{y_1}{2} + y_2 - 2\end{aligned}$$

Problem 3 (Sums of Random Variables)

Let $X_1, \dots, X_{10}$ be i.i.d. continuous Uniform$(2,5)$ random variables. After your lunch break, you are tasked to compute ssome quantities related to $Y = \sum_{i=1}^{10} 2X_i$.

(a) Calculate $\mathbb{P}[X_1 \geq 4,\ X_2 \geq 3,\ X_3 \leq 4]$.

Solution:

The PDF of X_i has constant height $1/3$ between 2 and 5. Therefore, we have that

$$\begin{aligned}\mathbb{P}[X_1 \geq 4,\ X_2 \geq 3,\ X_3 \leq 4] &= \mathbb{P}[X_1 \geq 4]\,\mathbb{P}[X_2 \geq 3]\,\mathbb{P}[X_3 \leq 4] \\ &= \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{27}\end{aligned}$$

(b) Compute the mean and variance of Y.

Solution:

$$\begin{aligned}\mathbb{E}[X_i] &= \frac{2+5}{2} = \frac{7}{2} \\ \mathrm{Var}[X_i] &= \frac{(5-2)^2}{12} = \frac{9}{12} = \frac{3}{4} \\ \mathbb{E}[2X_i] &= 2\mathbb{E}[X_i] = 7 \\ \mathrm{Var}[2X_i] &= 4\mathrm{Var}[X_i] = 3 \\ \mathbb{E}[Y] &= 10 \cdot \mathbb{E}[2X_i] = 70 \\ \mathrm{Var}[Y] &= 10 \cdot \mathrm{Var}[2X_i] = 30\end{aligned}$$

(c) Approximate $\mathbb{P}\big[|Y - \mathbb{E}[Y]| \geq 9\big]$ using the Central Limit Theorem.

Solution:

$$\begin{aligned}\mathbb{P}\big[|Y - \mathbb{E}[Y]| \geq 9\big] &= \mathbb{P}[Y - \mathbb{E}[Y] \geq 9] + \mathbb{P}[Y - \mathbb{E}[Y] \leq -9] \\ &= \mathbb{P}[Y \geq \mathbb{E}[Y] + 9] + \mathbb{P}[Y \leq \mathbb{E}[Y] - 9] \\ &= 1 - \mathbb{P}[Y < \mathbb{E}[Y] + 9] + \mathbb{P}[Y \leq \mathbb{E}[Y] - 9] \\ &\approx 1 - \Phi\left(\frac{\mathbb{E}[Y] + 9 - \mathbb{E}[Y]}{\sqrt{30}}\right) + \Phi\left(\frac{\mathbb{E}[Y] - 9 - \mathbb{E}[Y]}{\sqrt{30}}\right) \\ &= 1 - \Phi\left(\frac{9}{\sqrt{30}}\right) + \Phi\left(-\frac{9}{\sqrt{30}}\right) \\ &= 2\Phi\left(-\frac{9}{\sqrt{30}}\right)\end{aligned}$$

where the last step uses $\Phi(-z) = 1 - \Phi(z)$.

Problem 4 (Statistics) *15 points*

Your friend claims to have a trick coin that comes up as heads more often than tails. You flip the coin 100 times where X_i is 0 if the i^{th} flip is tails and 1 if it is heads. You may assume that the X_i are independent of one another. You propose the following null hypothesis: the coin is actually fair, meaning that heads and tails are equally likely and each X_i is Bernoulli$(1/2)$.

(a) Under the null hypothesis, what kind of random variable is $Y = \sum_{i=1}^{100} X_i$? (Don't forget the parameters.)

Solution:

Binomial$(100, 1/2)$

(b) Say you observe a total of 39 heads. Write down a sum corresponding to the *exact* probability that $Y = \sum_{i=1}^{100} X_i$ is less than or equal to 39 under the null hypothesis. (You do not need to calculate a numerical value for this sum.)

Solution:

$$\sum_{i=0}^{39} \binom{100}{i} \left(\frac{1}{2}\right)^{100}$$

(c) For 100 coin flips, it is reasonable to approximate the sum $Y = \sum_{i=1}^{100} X_i$ as Gaussian. Should we reject the the null hypothesis with significance level 0.05? Show your work for full credit. You may find one of the following values useful.

- $\Phi(-1.3) = 0.1$
- $\Phi(-1.6) = 0.05$
- $\Phi(-2.0) = 0.025$

Solution:

Since the variance is known $\sigma^2 = 1/4$ and we assume that under the null the data is Gaussian, we can use a one-sample Z-test to determine if the mean is $\mu = 1/2$. We first compute the Z-statistic:

$$Z = \sqrt{n}\frac{M_n - \mu}{\sigma} = \sqrt{100}\frac{\frac{1}{100}\sum_{i=1}^{100} X_i - \frac{1}{2}}{\sqrt{1/4}} = \frac{\frac{39}{10} - 5}{1/2} = \frac{39 - 50}{5} = -\frac{11}{5} = -2.2$$

Now, we evaluate the p-value:

$$\text{p} - \text{value} = 2\Phi(-|Z|) = 2\Phi(-2.2) < 2\Phi(-2.0) = 2 \cdot 0.025 = 0.05$$

Thus, we reject the null hypothesis and declare that the coin is not fair.

Problem 5 (Machine Learning)

In this problem, you will work through the process of constructing and evaluating an LDA binary classifier by hand. (All of the values below were carefully chosen to make the calculations easy since MATLAB and other computing software is not allowed.) You have been given the following 1-dimensional training and test datasets:

$$\mathbf{X}_{\text{train}} = \begin{bmatrix} +3 \\ +1 \\ 0 \\ -2 \end{bmatrix} \quad \underline{Y}_{\text{train}} = \begin{bmatrix} +1 \\ +1 \\ -1 \\ -1 \end{bmatrix} \quad \mathbf{X}_{\text{test}} = \begin{bmatrix} 0 \\ -3 \end{bmatrix} \quad \underline{Y}_{\text{test}} = \begin{bmatrix} +1 \\ -1 \end{bmatrix}$$

(a) Compute the sample means $\hat{\mu}_+$ and $\hat{\mu}_-$ as well as the sample covariance matrix $\hat{\mathbf{\Sigma}}$, which in this 1-dimensional setting is just a sample variance (and could be denoted by $\hat{\sigma}^2$ instead if you wish). Show your work for full credit.

Solution:

$$\hat{\mu}_+ = \frac{1}{2}(+3+1) = +2 \qquad \hat{\mu}_- = \frac{1}{2}(0-2) = -1$$

$$\hat{\mathbf{\Sigma}}_+ = (3-2)^2 + (1-2)^2 = 2 \qquad \hat{\mathbf{\Sigma}}_- = (0-(-1))^2 + (-2-(-1))^2 = 2$$

$$\hat{\mathbf{\Sigma}} = \frac{1}{4-2}((2-1)\hat{\mathbf{\Sigma}}_+ + (2-1)\hat{\mathbf{\Sigma}}_-) = 2$$

(b) Work out the LDA classifier. Try to simplify the expression as much as you can. Show your work for full credit.

Solution:

$$D_{\text{LDA}}(x) = \begin{cases} +1 & 2(\hat{\mu}_+ - \hat{\mu}_-)\hat{\mathbf{\Sigma}}^{-1}x \geq \hat{\mu}_+\hat{\mathbf{\Sigma}}^{-1}\hat{\mu}_+ - \hat{\mu}_-\hat{\mathbf{\Sigma}}^{-1}\hat{\mu}_- \\ -1 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} +1 & 2(+2-(-1))\tfrac{1}{2}x \geq 2 \cdot \tfrac{1}{2} \cdot 2 - (-1) \cdot \tfrac{1}{2} \cdot (-1) \\ -1 & \text{otherwise.} \end{cases} = \begin{cases} +1 & x \geq +\tfrac{1}{2} \\ -1 & \text{otherwise.} \end{cases}$$

(c) Apply your classifier to the datasets in order to calculate the training and test error rates. Show your work for full credit.

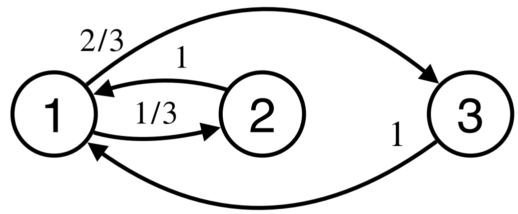
Solution:

$$\underline{Y}_{\text{train,guess}} = \begin{bmatrix} +1 \\ +1 \\ -1 \\ -1 \end{bmatrix} \qquad \underline{Y}_{\text{test,guess}} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

Training Error Rate is 0% and Test Error Rate is 50%.

Problem 6 (Markov Chains)

Your last task at your summer internship is to evaluate the following Markov chain



Throughout the problem, you may assume that $\underline{p}_1 = \begin{bmatrix} P_{X_1}(1) \\ P_{X_1}(2) \\ P_{X_1}(3) \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 1/3 \end{bmatrix}$.

(a) Determine $\underline{p}_2$.

Solution:

$$\mathbf{P} = \begin{bmatrix} 0 & 1/3 & 2/3 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\underline{p}_2 = \mathbf{P}^T \underline{p}_1 = \begin{bmatrix} 0 & 1 & 1 \\ 1/3 & 0 & 0 \\ 2/3 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

(b) Calculate the probability that $X_1 = 2$ given that $X_2 = 1$.

Solution:

This is actually a Bayes' rule question.

$$\begin{aligned} \mathbb{P}[X_1 = 2 | X_2 = 1] &= \frac{\mathbb{P}[X_2 = 1 | X_1 = 2]\, \mathbb{P}[X_1 = 2]}{\mathbb{P}[X_2 = 1]} \\ &= \frac{1 \cdot \frac{2}{3}}{1 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3}} = \frac{2}{3} \end{aligned}$$

(c) Does a unique limiting state probability vector $\underline{\pi}$ exist? If so, argue why and solve for it. If not, argue why.

Solution:

The Markov chain has period 2, and thus does not have a unique limiting state probability vector.