Problem 1

For each of the following parts, indicate whether the statement is always true or it can be false by clearly writing "True" or "False." Briefly explain the reasoning behind your answer for partial credit (in case your choice is wrong). Diagrams are welcome.

(a) Let Y be a discrete random variable, with probability mass function generated according to two possible hypotheses H_0 or H_1, as $P_{Y|H_1}(Y|H_1)$ or $P_{Y|H_0}(Y|H_0)$. Assume that the prior probability of H_1, $P(H_1) = 2/3$. Then, the probability of false alarm for the maximum likelihood detector (ML) is higher than the probability of false alarm for the maximum a posteriori detector (MAP).

(b) Let X and Y be jointly continuous random variables. If $\widehat{x}(Y)$ is an estimator of X, then $\mathbb{E}[(\widehat{x}(Y) - X)^2] \geq \mathbb{E}[(\widehat{x}_{\text{MMSE}}(Y) - X)^2]$.

(c) If $\rho_{X,Y} = 0$, then $\mathbb{E}[(\widehat{x}_{\text{MMSE}}(Y) - X)^2] = \text{Var}[X]$.

(d) Let $X_1, \dots, X_n$ be i.i.d. random variables with mean $\mathbb{E}[X_i] = \mu$. Then, the value of sample mean $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ always gets closer to the true mean as we include additional samples: for all $k < n$, $|M_n - \mu| < |M_k - \mu|$.

(e) Let $X_1, \dots, X_{n_1}$ and $Y_1, \dots, Y_{n_2}$ be two independent datasets of different sizes, n_1 and n_2, collected under the same conditions with sample means $M_{n_1}^{(1)}$ and $M_{n_2}^{(2)}$, respectively. Under the null model, we assume the data is i.i.d. Gaussian with mean μ and variance 1. We calculate the Z-statistic for each dataset, $Z^{(1)}$ and $Z^{(2)}$, as well as the p-values. If the p-value for the first dataset is smaller, then its sample mean must be closer to the true mean, $|M_{n_1}^{(1)} - \mu| < |M_{n_2}^{(2)} - \mu|$.

(f) $\text{Var}(X)\text{Var}(Y) \geq \big(\text{Cov}(X, Y)\big)^2$.

(g) Consider a binary classification dataset $(\underline{X}_1, Y_1), \dots, (\underline{X}_n, Y_n)$ and two classifiers $D_{\text{A}}(\underline{x})$ and $D_{\text{B}}(\underline{x})$. If classifier A has a lower error rate than classifier B on the first half of the dataset $(\underline{X}_1, Y_1), \dots, (\underline{X}_{n/2}, Y_{n/2})$, then it also will have a lower error rate on the second half of the dataset $(\underline{X}_{n/2+1}, Y_{n/2+1}), \dots, (\underline{X}_n, Y_n)$.

(h) In a Markov chain with transition probability matrix $\mathbf{P}$, if $P_{jj} > 0$, then the period of state j is 1.

(i) Suppose we have a Markov chain with two communicating classes. Then, there must be multiple limiting probability distributions for the state probability vector $\underline{p}_t$ as $t \to \infty$.

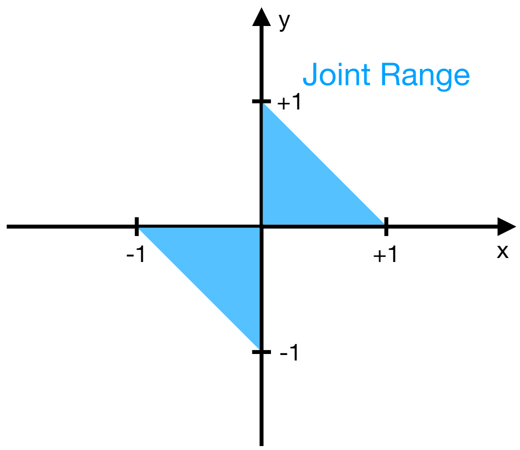
(j) Determine the MAP rule.

(k) Determine the probability of error under the ML rule.

Problem 2

On your second day at your new data scientist job, you consider the following estimation problem. The joint PDF of jointly continuous random variables X and Y is

$$f_{X,Y}(x,y) = \begin{cases} 1 & x \geq 0,\ y \geq 0,\ x+y \leq 1 \quad \text{OR} \quad x \leq 0,\ y \leq 0,\ x+y \geq -1 \\ 0 & \text{otherwise.} \end{cases}$$



(a) Determine $\mathbb{P}[X \geq Y]$. Your answer can be an integral, but you can also take advantage of the simple structure of the joint PDF.

(b) What is the MMSE estimator of X given $Y = y$? Your answer can be an integral, but you can also take advantage of the simple structure of the joint PDF.

(c) Determine $\mathbb{E}[X]$. Your answer can be an integral, but you can also take advantage of the simple structure of the joint PDF.

(d) What is the LLSE estimator of X given $Y = y$? Your answer can be an integral.

Problem 3

During your coffee break, you are trying to bound the probability that the total number of successes exceeds a threshold. Let $X_1, \ldots, X_{100}$ be i.i.d. Bernoulli $(1/2)$ random variables.

(a) What is the mean of $Y = X_1 + X_2$?

(b) What is the PMF of $Y = X_1 + X_2$?

(c) Use the Central Limit Theorem approximation to estimate the probability $\mathbb{P}\left[\sum_{i=1}^{100} X_i \geq 60\right]$. You can leave your answer in terms of the standard normal CDF $\Phi(z)$.

Problem 4

You measure the sulfate concentration in the local water reservoir over 9 consecutive days and obtain values $X_1, \ldots, X_9$, which are assumed to be i.i.d. Gaussian. (The units are mg/L and omitted below.) The sample mean is $M_9 = 5.1$ and the sample variance is $V_9 = 0.36$.

Let W have a t-distribution with 8 degrees-of-freedom. You may find one or more of the following values useful

- $F_W(-1.4) = \Phi(-1.3) = Q(1.3) = 0.1\,, \quad F_W(1.4) = \Phi(1.3) = 0.9$
- $F_W(-1.9) = \Phi(-1.6) = Q(1.6) = 0.05\,, \quad F_W(1.9) = \Phi(1.6) = 0.95$
- $F_W(-2.3) = \Phi(-2.0) = Q(2.0) = 0.025\,, \quad F_W(2.3) = \Phi(2.0) = 0.975$

(a) Construct a confidence interval for the mean with confidence level 0.95.

(b) Is your sample significantly different from the baseline concentration $\mu = 5.4$ at a significance level of 0.1? Justify your approach and support your answer numerically.

(c) Say you also go out on the 10^{th} day and collect measurement $X_{10} = 5$. What is the new sample mean M_{10}?

Problem 5

You have managed to model an interesting three-state system via a Markov chain with the following state transition matrix and initial probability state vector:

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 1 \\ 2/3 & 1/3 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad \qquad \underline{p}_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

(a) Draw the Markov chain, labeling the states as 1, 2, and 3, as well as labeling the arcs with the appropriate transition probabilities.

(b) What is the period of state 1?

(c) Determine $\mathbb{P}[X_0 = 2, X_1 = 2, X_2 = 3]$.