Exam 3 Solutions

First Name	Last Name	UID
		U

Honor Code: This exam represents only my own work. I did not give or receive help.

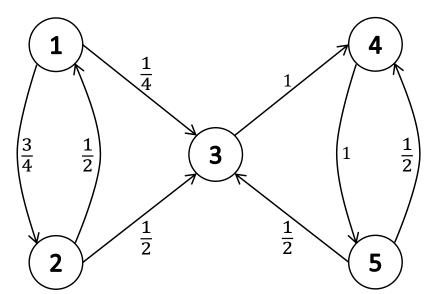
Signature: ____________________________

- You have exactly **2 hours** to complete this exam.
- **No devices are allowed**, including no phones and no calculators. No form of collaboration is allowed.
- You can use the provided formula sheet handouts – no other materials are allowed.
- All work to be graded must be included in this document. Submit no extra sheets.
- $\boxed{\text{Box}}$ your final answers.
- There are 5 problems in total, each having 5 parts, with each part worth 4 points for a total of 100 points.
- **Partial Credit:** There will be partial credit for solution attempts even if not all the mathematical manipulations are completed correctly. To maximize your chances for partial credit, attempt every problem.
- **Explanation:** In order to receive full credit, all work should be supported by a concise explanation that is clear, relevant, specific, logical, and correct. In particular, for each part, you must clearly outline the key steps and provide proper justification for your calculations.

*** Good Luck! ***

Problem 1 (Markov Chains) *20 points*

Consider the discrete-time, homogeneous Markov chain $X_0, X_1, X_2, \ldots,$ with state transition diagram shown in the figure. Assume that X_0 is equally likely to be in all five states.



(a) (4 pts) Calculate $\mathbb{P}[X_1 = 3]$. Explain your work.

Solution: $\boxed{\mathbb{P}[X_1 = 3] = \frac{1}{4}}$

$\mathbb{P}[X_1 = 3] = \sum_{i=1}^{5} \mathbb{P}[X_1 = 3 | X_0 = i] \cdot \mathbb{P}[X_0 = i] = \frac{1}{5}(\frac{1}{4} + \frac{1}{2} + 0 + \frac{1}{2}) = \frac{1}{4}$.

(b) (4 pts) Identify the communicating classes and whether they are transient or recurrent.

Solution: $\boxed{\text{Two communicating classes: } C_1 = \{1, 2\} \text{ transient, } C_2 = \{3, 4, 5\} \text{ recurrent.}}$

(c) (4 pts) Determine the period of each state. Explain your work.

Solution: $\boxed{\text{States 1, 2 have period 2. States 3, 4, 5 have period 1.}}$

All states in a communicating class have the same period. The lengths of all cycles in communicating class $\{1, 2\}$ are *all* the multiples of $2 \Rightarrow$ their greatest common divisor (gcd) is 2. Thus states 1 and 2 have a period of 2. In communicating class $\{3, 4, 5\}$, there is a cycle of length 2 and a cycle of length 3 starting at state 4, namely, $4 \to 5 \to 4$ and $4 \to 5 \to 3 \to 4$, and their gcd is 1. Thus states $3, 4, 5$ have a period of 1.

(d) (4 pts) Compute the steady-state probability distribution of this Markov chain. Explain your work.

Solution: $\boxed{\pi_1 = \pi_2 = 0, \pi_3 = \frac{1}{5}, \pi_4 = \pi_5 = \frac{2}{5}}$

The steady-state probabilities of transient states is zero $\Rightarrow \pi_1 = \pi_2 = 0$.

In the steady state we have

$\pi_3 = \frac{1}{2} \times \pi_5$ and $\pi_4 = 1 \times \pi_3 + \frac{1}{2} \times \pi_5$ which implies that $\pi_4 = \pi_5 = 2\pi_3$.

The normalization condition $\sum_{i=1}^{5} \pi_i = 1$ gives us: $0 + 0 + \pi_3 + 2\pi_3 + 2\pi_3 = 1 \Rightarrow 5\pi_3 = 1 \Rightarrow \pi_3 = \frac{1}{5}$.

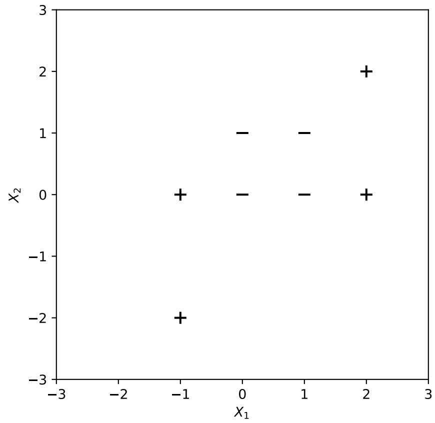
Therefore $\pi_4 = \pi_5 = 2\pi_3 = \frac{2}{5}$.

(e) (4 pts) Suppose you are in state 4 at time 0, i.e., $X_0 = 4$. Let $T > 0$ be the first time you return back to state 4, i.e., $X_T = 4$ and $X_t \neq 4$ for all $0 < t < T$. Compute $\mathbb{E}[T]$. Explain your work.

Solution: $\boxed{\mu_4 = 2.5}$

There are only two paths for returning back: $X_1 = 5, X_2 = 4$ or $X_1 = 5, X_2 = 3, X_3 = 4$. For the first path $T = 2$ and the probability of this path is $1 \times 0.5 = 0.5$. For the second path, $T = 3$ and the probability of this path is $1 \times 0.5 \times 1 = 0.5$. Thus, $\mathbb{E}[T] = 2 \times 0.5 + 3 \times 0.5 = 2.5$.

Problem 2 (Machine Learning) *20 points*



Training Data

x_1	x_2	label
-1	-2	+
-1	0	+
2	0	+
2	2	+
1	1	-
0	1	-
1	0	-
0	0	-

Testing Data

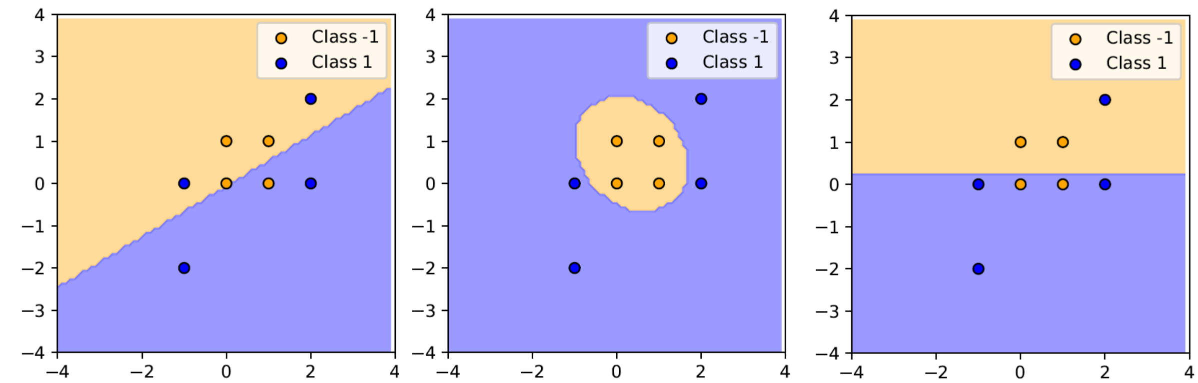
x_1	x_2	label
0	2	-
0	0	+
1	-1	+

You are given the training data on the figure and table and the testing data in the table.

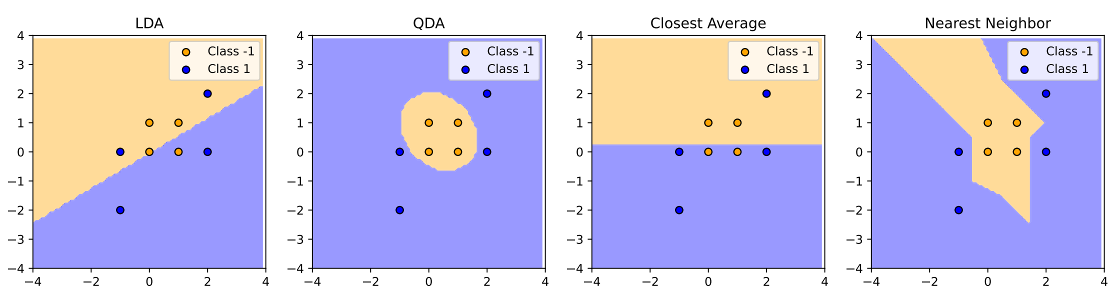
(a) (4 pts) Compute the sample mean vectors $\underline{\mu}_+$ and $\underline{\mu}_-$ and add them to the plot above.

Solution: $\underline{\mu}_+ = \left[\frac{-1-1+2+2}{4}, \frac{-2+0+0+2}{4}\right]^\top = [0.5, 0]^\top$ $\underline{\mu}_- = \left[\frac{1+0+1+0}{4}, \frac{1+1+0+0}{4}\right]^\top = [0.5, 0.5]^\top$

(b,c,d) (12 pts) In each figure shown below, determine if the decision boundary is for the Closest-Average, LDA, QDA, or Nearest-Neighbor classifier or none of them. Explain your choices. In the figures below, Class -1 and Class 1 points correspond to labels $-$ and $+$, respectively.



Solution:



Explanation: The Closest-Average classifier is the perpendicular bisector of the line joining $\underline{\mu}_+ = [0.5, 0]^\top$ and $\underline{\mu}_- = [0.5, 0.5]^\top$ which is the horizontal line with equation $x_2 = 0.25$. The QDA classifier can have a quadratic (elliptical) boundary. The LDA classifier has a linear boundary whose orientation is determined by the class-spreads and the class-means. The Nearest-Neighbor (NN) classifier has a jagged, piece-wise linear boundary.

(e) (4 pts) Determine the number of *training errors* for the decision boundary shown in each figure above. Determine the number of *test errors* for the *Nearest-Neighbor* classifier. In each figure, the decision is $+$ in the darker shaded region and is $-$ in the lighter shaded one.

Solution: The number of training errors is 0 for the QDA classifier and 3 for both the LDA and Closest-Average classifiers.

The testing error of the NN classifier is 2: The nearest neighbor of $(0,2)^\top$ is $(0,1)^\top$, of $(0,0)^\top$ is $(0,0)^\top$, and of $(1,-1)^\top$ is $(1,0)^\top$. Hence, only $(0,2)^\top$ is correctly classified.

Grading note: It is not sufficient to label one of the graphs above as Nearest-Neighbor (NN) and then use that to calculate the test errors. Each of the test points is exactly distance 1 from one of the training points, resulting in 2 test errors as calculated. A correct solution to problem 4(e) would then allow for correcting any mislabeling of one of the figures from 4(bcd) as NN.

Problem 3 (Statistics) *20 points*

You observe cars as they pass by on Comm Ave and note their arrival times. Let X_i be the time (in minutes) between the arrival of the i^{th} and $(i+1)^{\text{st}}$ cars, $i = 1, 2, 3, \ldots$. According to historical traffic data, $X_1, X_2, X_3, \ldots$, are i.i.d. with $X_i \sim \text{Exponential}(\lambda)$. You do not know the value of λ. Let Y_i be the time between the arrival of the i^{th} and $(i+100)^{\text{th}}$ cars, $i = 1, 2, 3, \ldots$.

Possibly useful:

$$\Phi(-1.64) = F_{T_2}(-2.92) = F_{T_{99}}(-1.66) = 0.05$$
$$\Phi(-1.96) = F_{T_2}(-4.30) = F_{T_{99}}(-1.98) = 0.025$$
$$\Phi(-2.57) = F_{T_2}(-9.93) = F_{T_{99}}(-2.63) = 0.01$$

(a) (4 pts) Are Y_1, Y_2, and Y_3 i.i.d.? Are Y_1, Y_{101}, and Y_{201} i.i.d.? Clearly explain your reasoning in both cases.

Solution: $\boxed{Y_1, Y_2, \text{ and } Y_3 \text{ are NOT i.i.d. but } Y_1,\ Y_{101}, \text{ and } Y_{201} \text{ are i.i.d.}}$

$Y_1 = X_1 + \ldots + X_{100}, Y_2 = X_2 + \ldots + X_{101}$, and $Y_3 = X_3 + \ldots + X_{102}$, all depend on $X_3, X_4, \ldots, X_{100}$ and thus are not independent. The Y_i's are identically distributed since the X_i's are i.i.d.

Y_1, Y_{101}, Y_{201} are identically distributed because each of them is the sum of 100 i.i.d. exponential random variables. They are also independent because they are the sum of 3 non-overlapping sets of 100 i.i.d. random variables: $Y_1 = X_1 + \ldots + X_{100}$, $Y_{101} = X_{101} + \ldots + X_{200}$, and $Y_{201} = X_{201} + \ldots + X_{300}$.

(b) (4 pts) Compute $\mathbb{E}[Y_1]$ in terms of λ. Simplify its form as much as you can. Explain your work.

Solution: $\boxed{\mathbb{E}[Y_1] = 100/\lambda}$ $\mathbb{E}[Y_1] = \mathbb{E}[X_1] + \ldots + \mathbb{E}[X_{100}] = 100\mathbb{E}[X_1] = 100/\lambda$.

(c) (4 pts) Compute $\mathbb{P}[Y_1 \geq 200/\lambda]$. Assume that the Central Limit Theorem applies. You may leave your answer in the form of a CDF. Explain your work.

Solution: $\boxed{\mathbb{P}[Y_1 \geq 200/\lambda] \overset{\text{CLT}}{=} 1 - \Phi(10)}$

$\mathsf{Var}[Y_1] = \mathsf{Var}[X_1] + \ldots + \mathsf{Var}[X_{100}] = 100\,\mathsf{Var}[X_1] = 100/\lambda^2$.

By the Central Limit Theorem, $Y_1 \sim \text{Gaussian}(100/\lambda, 100/\lambda^2)$.

Thus,

$\mathbb{P}[Y_1 \geq 200/\lambda] = 1 - \mathbb{P}[Y_1 < 200/\lambda] \overset{\text{CLT}}{=} 1 - \Phi[(200/\lambda - 100/\lambda)/\sqrt{(100/\lambda^2)}] = 1 - \Phi(10)$

(d) (4 pts) From the first 101 cars that you observe, you find that $y_1 = 200$. Can you reject the hypothesis that $\lambda = 1$, at a *two-tailed* significance level of $\alpha = 0.05$? Explain your work.

Solution: $\boxed{\text{We can reject the hypothesis.}}$

If the hypothesis is true, then $\mathbb{P}[Y_1 \geq y_1] = \mathbb{P}[Y_1 \geq 200/1] = \mathbb{P}[Y_1 \geq 200/\lambda] \overset{\text{CLT}}{=} 1 - \Phi(10) = \Phi(-10) < \Phi(-1.96) = 0.025 = \alpha/2$. So yes, we can reject the hypothesis.

(e) (4 pts) From observing the first 301 cars, you find that $y_1 = 200, y_{101} = 240$, and $y_{201} = 190$. Find the 95% confidence interval (CI) for $\mathbb{E}[Y_{1000}]$ centered around $M_3 = (Y_1 + Y_{101} + Y_{201})/3$. Explain your work.

Solution: $\boxed{\text{The 95\% CI for } \mu_Y \text{ is } [210 - 43\sqrt{7/3}, 210 + 43\sqrt{7/3}]}$

Firstly, Y_1, Y_{101}, and Y_{201}, and Y_{1000} are i.i.d. with the same mean $\mu_Y = \mathbb{E}[Y_1] = 100/\lambda$ and M_3 is an unbiased estimate of μ_Y since $\mathbb{E}[M_3] = (\mathbb{E}[Y_1] + \mathbb{E}[Y_{101}] + \mathbb{E}[Y_{201}])/3 = \mathbb{E}[Y_1] = \mu_Y$.

The sample variance is $V_3 = [(200-210)^2 + (240-210)^2 + (190-210)^2]/(3-1) = (100+900+400)/2 = 700$.

Since the number of observations is small ($n = 3 \leq 30$), and since Y_1, Y_{101}, and Y_{201} are i.i.d., the confidence interval should use a t-distribution instead of a z-distribution. The number of degrees of freedom is $3 - 1 = 2$. Thus, the 95% CI is $[M_3 - \epsilon, M_3 + \epsilon]$, with $\epsilon = \sqrt{V_3/n}|F_{T_2}^{-1}(-0.025)| = \sqrt{700/3} \times 4.30 = 43\sqrt{7/3}$.

Altogether, the CI is $[210 - 43\sqrt{7/3}, 210 + 43\sqrt{7/3}]$.

Problem 4 (Estimation) *20 points*

Let Y be a discrete random variable (RV) with PMF $p_Y(1) = p_Y(3) = 0.5$ and let Z be a Gaussian RV with zero mean and variance $\sigma^2 > 0$. The RVs Y and Z are independent. Let $X = 2^Y + Z$ where 2^Y means "two to the power of Y".

(a) (4 pts) Compute the MMSE estimator $\widehat{x}_{MMSE}(y)$ of X based on Y. Simplify its form as much as you can. Explain your work.

Solution: $\boxed{\widehat{x}_{MMSE}(y) = 2^y.}$

$\widehat{x}_{MMSE}(y) = \mathbb{E}[X|Y=y] = \mathbb{E}[2^Y + Z|Y=y] = \mathbb{E}[2^Y|Y=y] + \mathbb{E}[Z|Y=y] = 2^y + \mathbb{E}[Z] = 2^y + 0 = 2^y$

where $\mathbb{E}[Z|Y=y] = \mathbb{E}[Z]$ since Z is independent of Y.

(b) (4 pts) Compute the Mean Squared-Error (MSE) of the MMSE estimator: $\mathbb{E}[(X - \widehat{x}_{MMSE}(Y))^2]$ in terms of σ. Simplify its form as much as you can. Explain your work.

Solution: $\boxed{\mathbb{E}[(X - \widehat{x}_{MMSE}(Y))^2] = \sigma^2.}$

$\mathbb{E}[(X - \widehat{x}_{MMSE}(Y))^2] = \mathbb{E}[(X - 2^Y)^2] = \mathbb{E}[(2^Y + Z - 2^Y)^2] = \mathbb{E}[Z^2] = \sigma^2.$

Note: Since Y only takes two values with positive probability, namely $y = 1$ and $y = 3$, any function $h(y)$ such that $h(1) = 2^1 = 2$ and $h(3) = 2^3 = 8$, will have the same MSE as the MMSE estimator, e.g., the linear function $h(y) = 3y - 1$.

(c) (4 pts) Let $\widehat{x}_{LLSE}(y) = uy + v$ be the LLSE estimator of X based on Y. Compute the exact numerical values of the coefficients u and v and verify that they do not depend on σ. Explain your work.

Solution: $\boxed{u = 3,\ v = -1}$

$$\widehat{x}_{LLSE}(y) = \mathbb{E}[X] + \frac{\mathsf{Cov}[X,Y]}{\mathsf{Var}[Y]}(Y - \mathbb{E}[Y])$$

Thus, $u = \frac{\mathsf{Cov}[X,Y]}{\mathsf{Var}[Y]}$ and $v = \mathbb{E}[X] - u\,\mathbb{E}[Y]$.

$\mathbb{E}[Y] = \frac{(1+3)}{2} = 2,$
$\mathbb{E}[2^Y] = \frac{(2^1+2^3)}{2} = 5,$
$\mathbb{E}[X] = \mathbb{E}[2^Y] + \mathbb{E}[Z] = \mathbb{E}[2^Y] = 5,$

$\mathsf{Var}[Y] = \mathbb{E}[(Y - \mathbb{E}[Y])^2] = 0.5(1-2)^2 + 0.5(3-2)^2 = 1.$
$\mathsf{Cov}[Z,Y] = 0$ because Z and Y are independent.
$\mathsf{Cov}[X,Y] = \mathsf{Cov}[(2^Y + Z), Y] = \mathsf{Cov}[2^Y, Y] + \mathsf{Cov}[Z,Y] = \mathsf{Cov}[2^Y, Y] =$
$= 0.5(2^1 - 5)(1-2) + 0.5(2^3 - 5)(3-2) = 3.$

From this, it follows that
$u = \frac{\mathsf{Cov}[X,Y]}{\mathsf{Var}[Y]} = \frac{3}{1} = 3$ and
$v = \mathbb{E}[X] - u\,\mathbb{E}[Y] = 5 - 3 \times 2 = -1.$

Observe that the values of u and v do not depend on σ.

Note: For $y = 1, 3$, the linear function $h(y) = 3y - 1$ exactly matches the MMSE estimator $\widehat{x}_{MMSE}(y) = 2^y$. Therefore, it must be the LLSE estimator.

(d) (4 pts) Compute the Mean Squared-Error (MSE) of the LLSE estimator: $\mathbb{E}[(X - \widehat{x}_{LLSE}(Y))^2]$ in terms of σ. Simplify its form as much as you can. Explain your work.

Solution: $\boxed{\mathbb{E}[(X - \widehat{x}_{LLSE}(Y))^2] = \sigma^2}$

$\mathbb{E}[(X - \widehat{x}_{LLSE}(Y))^2] = \mathsf{Var}[X] - \frac{(\mathsf{Cov}[X,Y])^2}{\mathsf{Var}[Y]}$.

We have already computed $\mathsf{Cov}[X,Y]$ and $\mathsf{Var}[Y]$ in part (d). Since Z is independent of Y, we have $\mathsf{Var}[X] = \mathsf{Var}[2^Y + Z] = \mathsf{Var}[2^Y] + \mathsf{Var}[Z] = \mathsf{Var}[2^Y] + \sigma^2 = 0.5(2^1 - 5)^2 + 0.5(2^3 - 5)^2 + \sigma^2 = 9 + \sigma^2$.

$\Rightarrow \mathbb{E}[(X - \widehat{x}_{LLSE}(Y))^2] = \mathsf{Var}[X] - \frac{(\mathsf{Cov}[X,Y])^2}{\mathsf{Var}[Y]} = \sigma^2 + 9 - \frac{3^2}{1} = \sigma^2$.

Note: since the LLSE estimator exactly matches the MMSE estimator over the range of Y, it must have the same MSE as the MMSE estimator, namely, σ^2.

(e) (4 pts) Determine, with proper explanation, whether X and Y are jointly Gaussian. Compare the Mean Squared-Errors (MSEs) of the MMSE and LLSE estimators of X based on Y.

Solution: Since Y only takes the two values $1, 3$ with positive probability, it is not a Gaussian random variable. Therefore X, Y are not jointly Gaussian.

If X, Y were jointly Gaussian, then $\widehat{x}_{MMSE}(y) = \widehat{x}_{LLSE}(y)$ for all values of y and their MSE would be equal. In general, if X, Y are not jointly Gaussian, then the MMSE estimator will be different from the LLSE estimator and the MSE of the MMSE estimator would be strictly smaller (better) than the MSE of the LLSE estimator.

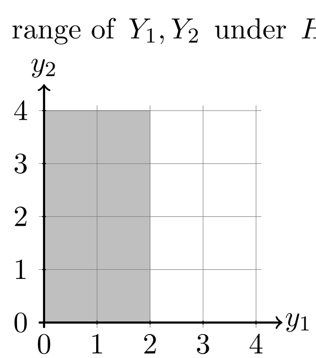
Interestingly, if Y takes only two distinct values (as in this problem), then the MMSE estimator will exactly match the LLSE estimator for all values in the range of Y as you should verify, i.e., for $y = 1, 3$, $\widehat{x}_{MMSE}(y) = 2^y = \widehat{x}_{LLSE}(y) = 3y - 1$, and the MSE of the MMSE and the LLSE estimators will be equal even though X, Y are not jointly Gaussian.

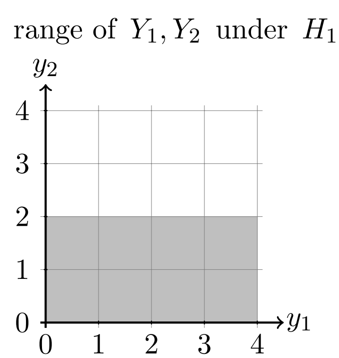
Problem 5 (Detection) *20 points*

Consider the following detection problem with $\mathbb{P}[H_0] = 1/3$ and $\mathbb{P}[H_1] = 2/3$ and two observations Y_1, Y_2.

Under H_0, (Y_1, Y_2) have a joint pdf which is uniform over the range $0 < y_1 < 2$, $0 < y_2 < 4$.

Under H_1, (Y_1, Y_2) have a joint pdf which is uniform over the range $0 < y_1 < 4$, $0 < y_2 < 2$.



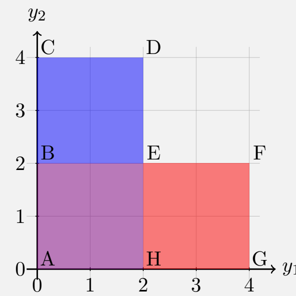


(a) (4 pts) Compute the Maximum Aposteriori Probability (MAP) decision rule $D^{\text{MAP}}(y_1, y_2)$. Simplify its form as much as you can. Explain your work.

Solution: $\boxed{D^{\text{MAP}}(y_1,y_2) = \begin{cases} 1 & (y_1,y_2) \in (0,4)\times(0,2) \\ 0 & (y_1,y_2) \in (0,2)\times(2,4) \\ 0 \text{ or } 1 \text{ equally good} & \text{otherwise.} \end{cases}}$

The pdfs under H_0 and H_1 are both zero outside the shaded region $ACDEFGA$ shown in the figure to the right. Both pdfs are equal to $1/8$ within the square ABEHA. Within the square BCDEB, the pdf under H_0 is $1/8$ and the pdf under H_1 is zero. Within the square HEFGH, the pdf under H_0 is zero and the pdf under H_1 is $1/8$. Finally, $\mathbb{P}[H_0] < \mathbb{P}[H_1]$.

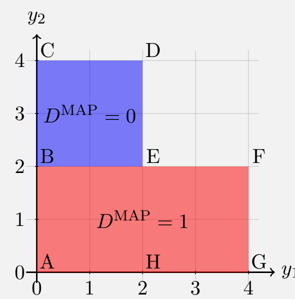
Union of ranges of (Y_1, Y_2) under H_0, H_1



Therefore,
$\mathbb{P}[H_0]\, f_{Y_1,Y_2|H_0}(y_1,y_2) < \mathbb{P}[H_1]\, f_{Y_1,Y_2|H_1}(y_1,y_2)$ within the rectangle ABFGA,
$\mathbb{P}[H_0]\, f_{Y_1,Y_2|H_0}(y_1,y_2) > \mathbb{P}[H_1]\, f_{Y_1,Y_2|H_1}(y_1,y_2)$ within the square BCDEB,
$\mathbb{P}[H_0]\, f_{Y_1,Y_2|H_0}(y_1,y_2) = \mathbb{P}[H_1]\, f_{Y_1,Y_2|H_1}(y_1,y_2) = 0$ outside the region ACDEFGA.

Therefore,

$$D^{\text{MAP}}(y_1,y_2) = \begin{cases} 1 & (y_1,y_2) \in (0,4)\times(0,2) \\ 0 & (y_1,y_2) \in (0,2)\times(2,4) \\ 0\ /\ 1 & \text{otherwise.} \end{cases}$$



(b) (4 pts) Compute the probability of error P_e^{MAP} for the MAP decision rule. Explain your work.

Solution: $\boxed{P_e = 1/6}$

Since $D^{\text{MAP}}(y_1,y_2) = 1$ over the entire range of $f_{Y_1,Y_2|H_1}(y_1,y_2)$, i.e., within rectangle ABFGA, $P_{\text{MD}} = 0$.

Since $D^{\text{MAP}}(y_1,y_2) = 1$ over *half* the range of $f_{Y_1,Y_2|H_0}(y_1,y_2)$, i.e., within square ABEHA, and $f_{Y_1,Y_2|H_0}(y_1,y_2)$ is uniform over its entire range, i.e., within rectangle ACDHA, we have $P_{\text{FA}} = 0.5$.

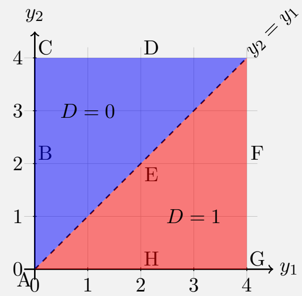
Therefore,

$$P_e = P_{\text{FA}}\mathbb{P}[H_0] + P_{\text{MD}}\mathbb{P}[H_1] = (1/2)(1/3) + (0)(2/3) = 1/6$$

(c) (4 pts) Compute the probability of error P_e when the decision rule is $D(y_1,y_2) = \begin{cases} 1 & y_2 \le y_1 \\ 0 & y_2 > y_1 \end{cases}$.
Explain your work.

Solution: $\boxed{P_e = 1/4}$

$f_{Y_1,Y_2|H_0}$ is uniform within the rectangle ACDHA and the within it, the region $y_2 \le y_1$ is the triangle AEHA which has one-fourth the total area of the rectangle. Thus, $P_{\text{FA}} = 0.25$. In a similar manner, we can deduce that $P_{\text{MD}} = 0.25$.



Therefore,

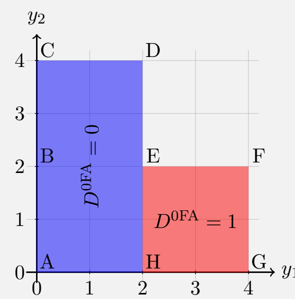
$$P_e = P_{\text{FA}}\mathbb{P}[H_0] + P_{\text{MD}}\mathbb{P}[H_1] = (1/4)(1/3) + (1/4)(2/3) = 1/4.$$

(d) (4 pts) Among all decision rules where the probability of false alarm is zero, determine the rule which has the least probability of error. Simplify its form as much as you can. Explain your work.

Solution:

$$\boxed{D^{0FA}(y_1,y_2) = \begin{cases} 0 & (y_1,y_2) \in \{\text{range of } f_{Y_1,Y_2|H_0}\} = (0,2)\times(0,4) \\ 1 & (y_1,y_2) \in \{\text{range of } f_{Y_1,Y_2|H_1}\} \cap \{\text{range of } f_{Y_1,Y_2|H_0}\}^c = [2,4)\times(0,2) \\ 0/1 \text{ equally good} & \text{otherwise.} \end{cases}}$$

To make $P_{\text{FA}} = 0$ we must decide 0 over the entire range of $f_{Y_1,Y_2|H_0} = (0,2)\times(0,4)$. Outside this region, but within the range of $f_{Y_1,Y_2|H_1}$, we must decide 1 to minimize the error probability. All other observations would be outside the ranges of both pdfs and the decisions we make there do not impact the error probability.



For this rule, $P_{\text{MD}} = 0.5$ since a missed detection occurs only if (y_1,y_2) arise from the pdf $f_{Y_1,Y_2|H_1}$, which is uniform over $(0,4)\times(0,2)$, and $(y_1,y_2) \in (0,2)\times(0,2)$, whose area is exactly half that of the range of $f_{Y_1,Y_2|H_1}$.

Therefore, for this rule,

$P_e = P_{\text{MD}}\mathbb{P}[H_1] = (1/2)(2/3) = 1/3.$

(e) (4 pts) Suppose that we can only observe the sum $Y = Y_1 + Y_2$ and not Y_1 and Y_2 individually. Determine the MAP decision rule based on observing Y, i.e., $D^{\text{MAP}}(y)$. Simplify its form as much as you can. Explain your work.

Solution: $\boxed{D(y) = \begin{cases} 1 & y \in (0,6) \\ 0 \text{ or } 1 \text{ equally good} & \text{otherwise.} \end{cases}}$

Y has the same distribution under the two hypotheses, but $\mathbb{P}[H_1] > \mathbb{P}[H_0]$. Therefore, the MAP decision is 1 over the entire range of Y which is $(0+0=0, 2+4=6)$. Decisions we make outside the range of Y do not impact the error probability.