Problems and Exercises Collection - Block 2: Classification

Most of the problems and exercises of this collection have been taken from previous years exams. The topics covered by each exercise are shown next to the exercise number according to:

2.1. Multiclass classification.

2.2. Binary decision.

2.3. Binary decision with Gaussian likelihoods.

2.4. Characterization of classifiers using ROC curves.

2.5. Other classification rules: Neyman-Pearson and minimax.

2.6. Discriminant functions.

Notation:

- ML decider: Maximum likelihood decider $[\phi_{\mathrm{ML}}(\mathbf{x})]$.
- MAP decider: Maximum *a posteriori* decider $[\phi_{\mathrm{MAP}}(\mathbf{x})]$.
- LRT: Likelihood ratio test.
- P_{e}: Probability of error.
- P_{FA}: Probability of false alarm.
- P_{M}: Probability of missing.
- P_{D}: Probability of detection.
- ROC curve: Operating characteristic curve.

Exercise 1 (2.2; 2.4; 2.6)

Consider the binary decision problem characterized by likelihoods

$$p_{X|H}(x|1) = \frac{3}{4}(1 - x^2), \qquad |x| \leq 1,$$
$$p_{X|H}(x|0) = \frac{15}{16}\left(1 - x^2\right)^2, \qquad |x| \leq 1,$$

and prior probability $P_H(1) = \frac{1}{3}$.

(a) Find the decision regions of the MAP decision maker.

(b) Obtain the detection probability of the MAP decision maker.

(c) Considering cost parameters $c_{00} = c_{11} = 0$, $c_{10} = c$, and $c_{01} = 1$, determine for which values of c the associated Bayesian decision maker always decides $D = 1$.

Solution:

(a) $|x| \underset{D=0}{\overset{D=1}{\gtrless}} \sqrt{\frac{3}{5}}$

(b) $P_{\mathrm{D}} = 1 - \frac{6}{5}\sqrt{\frac{3}{5}}$

(c) $c \leq \frac{2}{5}$

Exercise 2 (2.2; 2.4; 2.6)

Consider a binary classification problem where observations are distributed according to:

$$\begin{aligned} p_{X|H}(x|0) &= \exp(-x), & x > 0 \\ p_{X|H}(x|1) &= a\exp(-ax), & x > 0 \end{aligned}$$

with $a > 1$. For the decision, K independent observations, taken under the same hypothesis, are available: $\left\{X^{(k)}\right\}_{k=1}^{K}$.

(a) Obtain the ML decider based on the set of observations $\left\{X^{(k)}\right\}_{k=1}^{K}$ and check, using such a classifier, that $T = \sum_{k=1}^{K} X^{(k)}$ is a sufficient statistic for the decision.

Consider $K = 2$ for the rest of the exercise.

(b) Find the likelihoods in terms of the sufficient statistic T, $p_{T|H}(t|0)$ and $p_{T|H}(t|1)$.

(c) Calculate P_{FA} and P_{M} for the following threshold decider, as a function of η:

$$t \underset{D=1}{\overset{D=0}{\gtrless}} \eta$$

(d) Provide an approximate plot of the ROC curve for the previous decider, indicating:
- How the operation point moves when increasing η.
- How the ROC curve would change if we had access to a larger number of observations K.
- How the ROC curve changes as the value of a increases.

Solution:

(a) $t \underset{D=1}{\overset{D=0}{\gtrless}} \frac{K\ln a}{a-1}$

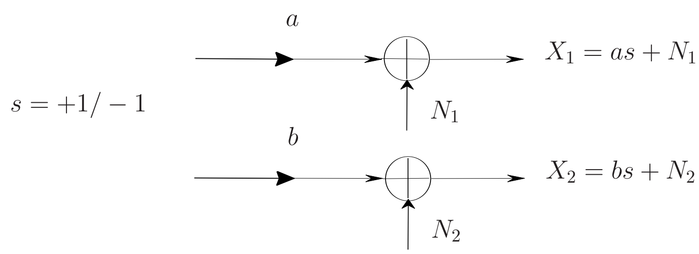
(b)
$$\begin{aligned} p_{T|H}(t|0) &= t\exp(-t), & t > 0 \\ p_{T|H}(t|1) &= a^2 t\exp(-at), & t > 0 \end{aligned}$$

(c) $P_{\mathrm{FA}} = 1 - (\eta + 1)\exp(-\eta) \quad P_{\mathrm{M}} = (a\eta + 1)\exp(-a\eta)$

(d)
- For $\eta = 0$, $P_{\mathrm{FA}} = P_{\mathrm{D}} = 0$; for $\eta \to \infty$, $P_{\mathrm{FA}} = P_{\mathrm{D}} = 1$.
- If the number of observations increases, then necessarily the performance of the classifier should improve (the area below the ROC curve increases).
- The same occurs if the value of a is increased. A rigorous demonstration would be: $\frac{\partial P_{\mathrm{M}}}{\partial a} = -a\eta^2 exp(-a\eta) < 0$, thus P_M decreases as the value of a is increased.

Exercise 3 (2.3; 2.6)

Consider a communication system in which one of the symbols, "+1" or "−1", is simultaneously transmitted through two noisy channels, as illustrated in the figure:



with a and b being two unknown positive constants which characterize the channels, and where N_1 and N_2 are two Gaussian noises with joint pdf

$$\begin{pmatrix} N_1 \\ N_2 \end{pmatrix} \sim G\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right].$$

It is also known that both symbols can be transmitted with equal *a priori* probabilities.

(a) If we wish to design a decider for discriminating the transmitted symbol using just one of the two available observations, X_1 or X_2, indicate which of the two variables you would use, justifying your answer as a function of the values of constants a and b. Provide the analytical expression for the corresponding ML decider.

(b) Obtain now the binary classifier with a minimum probability of error, based on the joint observation of X_1 and X_2, expressing it as a function of a, b, and ρ. Simplify your expression as much as possible.

(c) For $\rho = 0$, calculate the probability of error of the decider obtained in b). Express your result by means of function:

$$F(x) = 1 - Q(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

Solution:

(a) If $a > b$: $\quad x_1 \underset{D=0}{\overset{D=1}{\gtrless}} 0 \qquad$ If $a < b$: $\quad x_2 \underset{D=0}{\overset{D=1}{\gtrless}} 0$

(b) $(a - \rho b)x_1 + (b - \rho a)x_2 \underset{D=0}{\overset{D=1}{\gtrless}} 0$

(c) $P_{\mathrm{e}} = F\left(-\sqrt{a^2 + b^2}\right)$

Exercise 4 (2.2; 2.4)

The following likelihoods characterize a bidimensional binary decision problem with $P_H(0) = 3/5$:

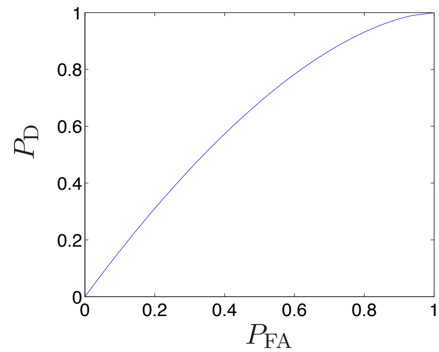
$$p_{X_1,X_2|H}(x_1, x_2|0) = \begin{cases} 2, & 0 < x_1 < 1 \quad 0 < x_2 < 1 - x_1 \\ 0, & \text{otherwise} \end{cases}$$

$$p_{X_1,X_2|H}(x_1, x_2|1) = \begin{cases} 3\,(x_1 + x_2)\,, & 0 < x_1 < 1 \quad 0 < x_2 < 1 - x_1 \\ 0, & \text{otherwise} \end{cases}$$

Consider a generic LRT decider with threshold η,

(a) Calculate P_{FA} as a function of η.

(b) The following plot represents the ROC curve of the LRT. Justifying your answer:
  - Indicate on the ROC how the operation point moves on the curve when increasing or decreasing the threshold of the test.
  - Place on the ROC the operation points corresponding to the ML decider, to the decider with minimum probability of error, and to the Neyman-Pearson classifier with $P_{\mathrm{FA}} = 0.3$.



Solution:

(a) $x_1 + x_2 \underset{D=0}{\overset{D=1}{\gtrless}} \frac{2}{3}\eta = \eta' \qquad P_{FA} = 1 - \eta'^2$

(b)
- P_{FA} and P_D decrease as the threshold is increased.
- ML decider: $\eta = 1$, $\eta' = \frac{2}{3}$, $P_{FA} = \frac{5}{9}$.
  MAP decider: $\eta = \frac{3}{2}$, $\eta' = 1$, $P_{FA} = 0$.
  N-P decider: $P_{FA} = 0.3$.

Exercise 5 (2.3)

Consider two equally probable hypotheses, with associated observations:

$$\begin{aligned} H = 0: &\quad X = N \\ H = 1: &\quad X = N + aS \end{aligned}$$

where N and S are independent Gaussian random variables, with zero mean and variances v_n and v_s, respectively, and where a is a known positive constant.

(a) Verify that the minimum probability error test can be written down as

$$c_1 \exp\left(c_2 x^2\right) \gtrless \eta$$

and calculate the value of constants c_1 and c_2, indicating the associated criterion for the decision.

(b) Determine the decision regions (over x) induced by the classifier. Note that such regions can be expressed as a function of constants c_1 and c_2.

Solution:

(a) $c_1 \exp\left(c_2 x^2\right) \underset{D=0}{\overset{D=1}{\gtrless}} 1$, where $c_1 = \frac{P_H(0)}{P_H(1)}\sqrt{\frac{v_n}{v_n + a^2 v_s}}$ and $c_2 = \frac{1}{2v_n} - \frac{1}{2\left(v_n + a^2 v_s\right)}$

(b) $|x| \underset{D=0}{\overset{D=1}{\gtrless}} \sqrt{\frac{-\ln c_1}{c_2}}$

Exercise 6 (2.2)

The joint probability density function of random variables X and Z is given by

$$p_{X,Z}(x,z) = x + z, \qquad 0 \leq x, z \leq 1$$

Consider the decision problem based on the observation of X (but not Z), with hypotheses:

$$\begin{array}{ll} H = 0: & Z < 0.6 \\ H = 1: & Z > 0.6 \end{array}$$

(a) Determine $p_{Z|X}(z|x)$.
(b) Obtain the *a posteriori* probabilities of both hypotheses.
(c) Determine the MAP decider based on X.
(d) Applying Bayes' Theorem, find the likelihoods $p_{X|H}(x|0)$ and $p_{X|H}(x|1)$.
(e) Calculate the probability of false alarm of the MAP decider.
(f) Determine the ML decider based on the observation of X.

Solution:

(a) $p_{Z|X}(z|x) = \dfrac{2\,(x+z)}{2x+1}, \quad 0 \leq x, z \leq 1$

(b) $P_{H|X}(0|x) = \dfrac{1.2x + 0.36}{2x+1} \qquad P_{H|X}(1|x) = 1 - \dfrac{1.2x + 0.36}{2x+1}$

(c) $x \underset{D=1}{\overset{D=0}{\gtrless}} 0.7$

(d) $p_{X|H}(x|0) = \dfrac{2x + 0.6}{1.6}$ and $p_{X|H}(x|1) = \dfrac{0.8x + 0.64}{1.04}$

(e) $P_{\text{FA}} = 0.5687$

(f) $x \underset{D=1}{\overset{D=0}{\gtrless}} 0.5$

Exercise 7 (2.2; 2.4; 2.5)

Consider a binary decision problem with $P_H(1) = 2P_H(0)$ and likelihoods:

$$\begin{array}{ll} p_{X|H}(x|0) = 2(1-x), & 0 \leq x \leq 1 \\ p_{X|H}(x|1) = 2x - 1, & \dfrac{1}{2} \leq x \leq \dfrac{3}{2} \end{array}$$

(a) Find the minimum mean cost decider for cost policy $c_{00} = c_{11} = 0$, $c_{10} = 4c_{01}$.
(b) Determine the Neyman-Pearson classifier with $P_{\text{FA}} = 0.04$.
(c) Obtain, as a function of parameter α, the false alarm and detection probabilities of the family of deciders with analytical shape

$$x \underset{D=0}{\overset{D=1}{\gtrless}} \alpha$$

(d) Plot (in an approximate manner) the operating characteristic (ROC) curve, taking α as the free parameter, and illustrating how the operation point of the decider changes as a function of the value of such parameter.

(e) Indicate whether the deciders obtained in (a) and (b) correspond to certain operation points of the previous ROC and, if so, identify it (or them).

Solution:

(a) If $x < \dfrac{1}{2}:\ D = 0$; If $\dfrac{1}{2} < x < 1:\ x \underset{D=0}{\overset{D=1}{\gtrless}} \dfrac{5}{6}$; If $x > 1:\ D = 1$

(b) $\alpha = 0.8$.

(c) $P_{\mathrm{FA}} = \begin{cases} (1-\alpha)^2 & 0 < \alpha < 1 \\ 0 & 1 < \alpha < \dfrac{3}{2} \end{cases}$ $\qquad P_{\mathrm{D}} = \begin{cases} 1 & 0 < \alpha < \dfrac{1}{2} \\ 1 - \left(\alpha - \dfrac{1}{2}\right)^2 & \dfrac{1}{2} < \alpha < \dfrac{3}{2} \end{cases}$

(d) $$\begin{cases} 1 < \alpha < \dfrac{3}{2} & P_{\mathrm{FA}} = 0 & P_{\mathrm{D}} = 1 - \left(\alpha - \dfrac{1}{2}\right)^2 \\ \dfrac{1}{2} < \alpha < 1 & P_{\mathrm{FA}} = (1-\alpha)^2 & P_{\mathrm{D}} = 1 - \left(\alpha - \dfrac{1}{2}\right)^2 \\ 0 < \alpha < \dfrac{1}{2} & P_{\mathrm{FA}} = (1-\alpha)^2 & P_{\mathrm{D}} = 1 \end{cases}$$

(e) (a) $\alpha = 5/6$ (b) $\alpha = 0.8$

Exercise 8 (2.3)

Let the following likelihoods characterize a bidimensional binary decision problem:

$$p_{X_1,X_2|H}(x_1,x_2|0) = G\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

$$p_{X_1,X_2|H}(x_1,x_2|1) = G\left(\mathbf{m}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

Plot in plane $X_1 - X_2$ the decision border given by the MAP decider, if the following conditions hold: $P_H(0) = P_H(1)$, $v_0 = v_1$ and $\rho = 0$. Indicate how that decision border would change if:

(a) The *a priori* probabilities were $P_H(0) = 2P_H(1)$.

(b) The value of ρ were increased.

Solution: The decision border is the bisector of the segment joining the means of both Gaussian distributions.

(a) If $P_H(0)$ gets larger, then the decision border is shifted towards the likelihood of hypothesis $H = 1$, i.e., towards $\mathbf{m}$.

(b) The decision border does not change.

Exercise 9 (2.2; 2.3; 2.6)

It is known that in a binary decision problem the observations follow discrete Bernouilli distributions with parameters p_0 and p_1 ($0 < p_0 < p_1 < 1$):

$$P_{X|H}(x|0) = \begin{cases} p_0 & x = 1 \\ 1 - p_0 & x = 0 \\ 0 & \text{otherwise} \end{cases} \qquad P_{X|H}(x|1) = \begin{cases} p_1 & x = 1 \\ 1 - p_1 & x = 0 \\ 0 & \text{otherwise} \end{cases}$$

We have access to a set of K independent observations taken under the same hypothesis for the decision process: $\left\{X^{(k)}\right\}_{k=1}^{K}$. Let T be a statistic defined as the following function of the observations: $T = \sum_{k=1}^{K} X^{(k)}$, i.e., random variable T is the number of observations which are equal to one.

(a) Obtain the ML decider based on the set of observations $\left\{X^{(k)}\right\}_{k=1}^{K}$, expressing it as a function of r.v. T.

(b) Taking into consideration that the mean and variance of a Bernouilli distribution with parameter p are given by p and $1-p$, respectively, find the means and variances of statistic T conditioned on both hypotheses: m_0 and v_0 (for $H = 0$) and m_1 and v_1 (for $H = 1$).

Consider for the rest of the exercise $p_0 = 1 - p_1$.

For K large enough, the distribution of T can be approximated by means of a Gaussian distribution, using the previously calculated means and variances.

(c) Calculate P_{FA} and P_{M} for the threshold decider

$$t \underset{D=0}{\overset{D=1}{\gtrless}} \eta$$

as a function of η. Express your result using function:

$$F(x) = 1 - Q(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

(d) Provide an approximate representation of the ROC curve of the previous decider, indicating:
- How the operation point moves when increasing η.
- How the ROC curve would be modified if the number of available observations (K) increased.
- How the ROC curve would change if the value of p_1 gets larger (keeping condition $p_0 = 1 - p_1$).

Solution:

(a) $$t \underset{D=0}{\overset{D=1}{\gtrless}} \frac{K \ln \dfrac{1-p_1}{1-p_0}}{\ln \dfrac{1-p_1}{1-p_0} - \ln \dfrac{p_1}{p_0}} = \eta$$

(b) $m_0 = Kp_0$ $\quad$ $m_1 = Kp_1$

$v_0 = Kp_0\,(1-p_0)$ $\quad$ $v_1 = Kp_1\,(1-p_1)$

(c) $$P_{\mathrm{FA}} = F\left(\frac{\eta - K(1-p_1)}{\sqrt{Kp_1(1-p_1)}}\right) \qquad P_{\mathrm{M}} = 1 - F\left(\frac{\eta - Kp_1}{\sqrt{Kp_1(1-p_1)}}\right)$$

(d) If $\eta \to -\infty$, $P_{\mathrm{FA}} = 0$ and $P_{\mathrm{D}} = 0$; $\eta \to \infty$ implies $P_{\mathrm{FA}} = 1$ and $P_{\mathrm{D}} = 1$.
The area below the ROC curve increases when K gets larger.
The area below the ROC curve increases if p_1 is reduced.

Exercise 10 (2.2)

Consider a binary decision problem with hypotheses $H = 0$ and $H = 1$, and observation X. A particular classifier decides $D = 1$ if X falls within region R_1, and $D = 0$ otherwise, obtaining false alarm and detection probabilities P_{FA} and P_{D}, respectively.

The complementary classifier decides $D = 0$ if X is situated inside R_1 and $D = 1$ otherwise, P'_{FA} and P'_{D} being the associated probabilities of false alarm and detection, respectively. Find the existing relationship between the probabilities of false alarm and detection of both deciders.

Solution:

$$P'_{\text{FA}} = 1 - P_{\text{FA}} \qquad P'_{\text{D}} = 1 - P_{\text{D}}$$

Exercise 11 (2.3; 2.6)

We have a binary decision problem with likelihoods:

$$p_{X_1,X_2|H}(x_1, x_2|0) = G\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

$$p_{X_1,X_2|H}(x_1, x_2|1) = G\left(\mathbf{m}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

with $\mathbf{m} = [m, m]^T$, where $m > 0$, and $|\rho| < 1$.

(a) Knowing that $P_H(0) = P_H(1)$, obtain the Bayes' decider incurring in a minimum probability of error. Plot the obtained decision boundary on the plane $X_1 - X_2$.

(b) For the classifier obtained in a), verify that $Z = X_1 + X_2$ is a sufficient statistic for the decision. Obtain the likelihoods of hypotheses $H = 0$ and $H = 1$ over random variable Z, $p_{Z|H}(z|0)$ and $p_{Z|H}(z|1)$.

(c) Calculate the false alarm, missing, and error probabilities of the previous decider, expressing them in terms of function

$$F(x) = 1 - Q(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

(d) Analyze how the probability of error changes with ρ; in order to do so, consider cases $\rho = -1$, $\rho = 0$, and $\rho = 1$. Indicate, for each of these values of ρ, how the likelihoods and decision boundary look like on the plane with coordinate axis $X_1 - X_2$.

Solution:

(a) $x_1 + x_2 \underset{D=0}{\overset{D=1}{\gtrless}} m$

(b) $z \underset{D=0}{\overset{D=1}{\gtrless}} m$

$$p_{Z|H}(z|0) = G\left(0, 2(1+\rho)\right) \qquad p_{Z|H}(z|1) = G\left(2m, 2(1+\rho)\right)$$

(c) $P_{\text{FA}} = P_{\text{M}} = P_{\text{e}} = 1 - F\left(\dfrac{m}{\sqrt{2(1+\rho)}}\right)$

(d) If $\rho \to -1: \ P_{\text{e}} = 0$ $\qquad$ If $\rho = 0: \ P_{\text{e}} = 1 - F\left(\dfrac{m}{\sqrt{2}}\right)$ $\qquad$ If $\rho \to 1: \ P_{\text{e}} = 1 - F\left(\dfrac{m}{2}\right)$

Exercise 12 (2.3)

Consider an N-dimensional binary (and Gaussian) decision problem, where observation vectors $\mathbf{X}$ are distributed according to likelihoods

$$p_{\mathbf{X}|H}(\mathbf{x}|0) = G(\mathbf{0}, v\mathbf{I})$$

$$p_{\mathbf{X}|H}(\mathbf{x}|1) = G(\mathbf{m}, v\mathbf{I})$$

where $\mathbf{0}$ and $\mathbf{m}$ are N-dimensional vectors with components 0 and $\{m_n\}$, respectively, and $\mathbf{I}$ is the NxN unitary matrix.

(a) Design the ML classifier.

(b) If $P_H(0) = 1/4$, design the minimum probability of error classifier.

(c) Obtain P_{FA} and P_{M} for the ML decider. What behavior would be observed if the number of observations grows with $\{m_n\} \neq 0$?

(d) In practice, we just have access to random variable Z, which is related to $\mathbf{X}$ via

$$Z = \mathbf{m}^T\mathbf{X} + N$$

where N is $G(m', v_n)$ and independent of $\mathbf{X}$. What would the new ML classifier based on the observation of Z be like?

(e) Calculate P'_{FA} and P'_M for the design in part d). How do they change with respect to P_{FA} and P_{M}?

Indication: When, convenient, express your result using function:

$$F(x) = 1 - Q(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

Solution:

(a) $\mathbf{m}^T\mathbf{X} \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2}||\mathbf{m}||_2^2$

(b) $\mathbf{m}^T\mathbf{X} \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2}||\mathbf{m}||_2^2 - v\ln 3$

(c) $P_{\mathrm{FA}} = P_{\mathrm{M}} = F\left(\frac{||\mathbf{m}||_2}{2\sqrt{v}}\right)$, which goes to 0 as N increases towards infinity.

(d) $z \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2}||\mathbf{m}||_2^2 + m'$

(e) $P'_{FA} = P'_M = F\left(\frac{||\mathbf{m}||_2}{2\sqrt{v + \frac{v_n}{||\mathbf{m}||_2^2}}}\right)$; they increase with $\frac{v_n}{||\mathbf{m}||_2^2}$.

Exercise 13 (2.2; 2.4)

Consider a binary decision problem characterized by:

$$p_{X_1,X_2|H}(x_1,x_2|0) = \begin{cases} \alpha x_2 & 0 < x_1 < \frac{1}{4} \quad 0 < x_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{X_1,X_2|H}(x_1,x_2|1) = \begin{cases} \beta x_1 & 0 < x_1 < 1 \quad 0 < x_2 < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

(a) After finding the values of constants α and β, provide a graphic representation of the decision regions corresponding to an LRT classifier. Indicate how those regions change as a function of the classifier threshold. Can the threshold be set so that the resulting classifier is linear?

(b) Obtain the marginal probability density functions of x_1 and x_2 conditioned on both hypotheses ($H = 0$ and $H = 1$). What is the existing statistical relationship between X_1 and X_2?

(c) For simplicity, we opt to use a threshold classifier based in just one variable: X_1 or X_2:

$$\text{DEC1:} \quad x_1 \underset{D=0}{\overset{D=1}{\gtrless}} \eta_1 \qquad\qquad \text{DEC2:} \quad x_2 \underset{D=1}{\overset{D=0}{\gtrless}} \eta_2$$

Calculate the probabilities of false alarm and detection of classifiers DEC1 and DEC2, expressing them as functions of the thresholds of such classifiers, η_1 and η_2, respectively.

(d) Plot the ROC curves (i.e., the curves that represent P_{D} as a function of P_{FA}), corresponding to deciders DEC1 and DEC2. Discuss how the operation points of both classifiers change when modifying the corresponding threholds.

(e) In the light of the obtained results, can it be concluded that one of the two proposed classifiers, DEC1 or DEC2, always outperforms the other one?

Solution:

(a) $\alpha = 8$ and $\beta = 4$.
We decide the only plausible hypothesis where $p_{X_1,X_2|H}(x_1,x_2|0)$ or $p_{X_1,X_2|H}(x_1,x_2|1)$ are zero. In the region where both likelihoods overlap, considering the LRT given by $\frac{p_{X_1,X_2|H}(x_1,x_2|0)}{p_{X_1,X_2|H}(x_1,x_2|1)} \underset{D=1}{\overset{D=0}{\gtrless}} \eta$, the decider is:

$$2x_2 - \eta x_1 \underset{D=1}{\overset{D=0}{\gtrless}} 0$$

For $\eta = 4$ a linear border is obtained.

(b) Observations X_1 and X_2 are independent under both hypotheses.

$p_{X_1|H}(x_1|0) = 4, \quad 0 < x_1 < \frac{1}{4} \qquad p_{X_2|H}(x_2|0) = 2x_2, \quad 0 < x_2 < 1$

$p_{X_1|H}(x_1|1) = 2x_1, \quad 0 < x_1 < 1 \qquad p_{X_2|H}(x_2|1) = 2, \quad 0 < x_2 < \frac{1}{2}$

(c) DEC1: $\begin{cases} P_{\text{FA}} = \begin{cases} 1 - 4\eta_1, & 0 < \eta_1 < 1/4 \\ 0, & 1/4 < \eta_1 < 1 \end{cases} \\ P_{\text{D}} = 1 - \eta_1^2, \quad 0 < \eta_1 < 1 \end{cases}$ DEC2: $\begin{cases} P_{\text{FA}} = \eta_2^2, \quad 0 < \eta_2 < 1 \\ P_{\text{D}} = \begin{cases} 2\eta_2, & 0 < \eta_2 < 1/2 \\ 1, & 1/2 < \eta_2 < 1 \end{cases} \end{cases}$

(d) DEC1: When $\eta_1 = 1$ the operation point is $P_{\mathrm{FA}} = 0$ and $P_{\mathrm{D}} = 0$; for $\eta_1 = 0$ the operation point is $P_{\mathrm{FA}} = 1$ and $P_{\mathrm{D}} = 1$.
DEC2: When $\eta_2 = 1$ the operation point is $P_{\mathrm{FA}} = 1$ and $P_{\mathrm{D}} = 1$; for $\eta_2 = 0$ the operation point is $P_{\mathrm{FA}} = 0$ and $P_{\mathrm{D}} = 0$.

(e) None of the classifiers can be stated to always outperform the other.

Exercise 14 (2.1)

Three random variables are characterized by the following likelihoods:

$$p_{X_1}(x_1) = \begin{cases} 1, & 0 < x_1 < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$p_{X_2}(x_2) = 2\exp(-2x_2), \quad x_2 > 0$$

$$p_{X_3}(x_3) = 2\exp(2(x_3 - 1)), \quad x_3 < 1$$

Considering the following three hypotheses:

$$\begin{aligned} H = 1: &\quad X = X_1 \\ H = 2: &\quad X = X_2 \\ H = 3: &\quad X = X_3 \end{aligned}$$

obtain:

(a) The Bayes' decider that minimizes the overall mean cost when all hypotheses are *a priori* equally probable, and the cost policy is $c_{ii} = 0$, $i = 1, 2, 3$ and $c_{ij} = c$ with $i \neq j$.

(b) Probabilities of deciding $D = i$ given hypothesis $H = i$, i.e., $P(D = i|H = i)$ for $i = 1, 2, 3$.

Considering now the binary decision problem characterized by:

$$\begin{aligned} H = 1: &\quad X = X_1 \\ H = 0: &\quad X = X_2 + X_3 \end{aligned}$$

obtain:

(c) The corresponding ML decider.

(d) The false alarm and missing probabilities, $P(D = 1|H = 0)$ and $P(D = 0|H = 1)$, respectively.

Solution:

(a)
$$\begin{aligned} D = 2: &\quad 0 < x < 0.34 \text{ and } x > 1 \\ D = 1: &\quad 0.34 < x < 0.65 \\ D = 3: &\quad 0.65 < x < 1 \text{ and } x < 0 \end{aligned}$$

(b) $P(D = 1|H = 1) = 0.31$, $P(D = 2|H = 2) = 0.6353$ and $P(D = 3|H = 3) = 0.6353$

(c)
$$\begin{aligned} D = 0: &\quad x < 0 \text{ and } x > 1 \\ D = 1: &\quad 0 < x < 1 \end{aligned}$$

(d) $P_{\mathrm{FA}} = P(D = 1|H = 0) = 0.4323$ and $P_{\mathrm{M}} = P(D = 0|H = 1) = 0$

Exercise 15 (2.2)

Consider a decision problem characterized by the following likelihoods:

$$p_{X|H}(x|0) = \begin{cases} \dfrac{2}{a^2}x & 0 < x < a \\ 0 & \text{otherwise} \end{cases} \qquad p_{X|H}(x|1) = \begin{cases} \dfrac{1}{a} & 0 < x < a \\ 0 & \text{otherwise} \end{cases}$$

Plot the characteristic operation curve (P_{D} vs P_{FA}) of the LRT classifier that solves such problem. Place over the curve the operation point corresponding to the maximum likelihood decider.

> **Solution:** The ROC curve is: $P_{\text{FA}} = P_{\text{D}}^2$.
> The operation point of the ML decider is: $P_{\text{D}} = \dfrac{1}{2}$ and $P_{\text{FA}} = \dfrac{1}{4}$

Exercise 16 (2.2)

A bidimensional binary decision probability is characterized by equally probable hypotheses, and likelihoods:

$$p_{X_1,X_2|H}(x_1,x_2|0) = K_0 x_1(1-x_2), \qquad 0 < x_1, x_2 < 1$$

$$p_{X_1,X_2|H}(x_1,x_2|1) = K_1 x_1 x_2, \qquad 0 < x_1, x_2 < 1$$

$(K_0, K_1 > 0)$.

(a) Calculate the values of constants K_0 and K_1.

(b) Find the classifier that minimizes the probability of error, and indicate the importance of X_1 and X_2 in the decision process.

(c) Obtain marginal likelihoods $p_{X_i|H}(x_i|j)$, for $i = 1,2$ and $j = 0,1$. What is the statistical relationship between X_1 and X_2 under each hypothesis?

(d) Calculate P_{FA}, P_{M} y P_{e}.

(e) In practice, X_2 can not be observed directly, but we can just access a version contaminated with an additive noise N independent of X_1 and X_2; i.e., we observe $Y = X_2 + N$. Design the optimal decider for this situation when the noise pdf is:

$$p_N(n) = 1, \quad 0 < n < 1$$

(f) Calculate P'_{FA}, P'_{M} and P'_e for the new situation and the classifier designed in part (e).

> **Solution:**
>
> (a) $K_0 = K_1 = 4$
>
> (b) $x_2 \underset{D=0}{\overset{D=1}{\gtrless}} \dfrac{1}{2}$; X_1 does not provide relevant information for the decision, whereas X_2 is a sufficient statistic.
>
> (c) $p_{X_1|H}(x_1|0) = 2x_1,\ 0 < x_1 < 1$; $\quad p_{X_1|H}(x_1|1) = 2x_1,\ 0 < x_1 < 1$
> $p_{X_2|H}(x_2|0) = 2(1-x_2),\ 0 < x_2 < 1$; $\quad p_{X_2|H}(x_2|1) = 2x_2,\ 0 < x_2 < 1$
> X_1 are X_2 independent under whatever hypothesis.
>
> (d) $P_{\text{FA}} = P_{\text{M}} = P_{\text{e}} = \dfrac{1}{4}$.
>
> (e) With $Y = X_2 + N$, it still holds that: $p_{X_1,Y|H}(x_1,y|i) = p_{X_1|H}(x_1|i)p_{Y|H}(y|i)$, $i = 0,1$ and we just need to work with Y instead of X_2. The likelihoods, expressed as distributions over Y, can be obtained by convolving the distributions of X_2 and N, resulting:
>
> $$p_{Y|H}(y|0) = \begin{cases} 0, & y < 0 \\ 2y - y^2, & 0 < y < 1 \\ 4 - 4y + y^2, & 1 < y < 2 \\ 0, & y > 2 \end{cases} \qquad p_{Y|H}(y|1) = \begin{cases} 0, & y < 0 \\ y^2, & 0 < y < 1 \\ 2y - y^2, & 1 < y < 2 \\ 0, & y > 2 \end{cases}$$
>
> The new decider is: $y = x_2 + n \underset{D=0}{\overset{D=1}{\gtrless}} 1$

(f) $P'_{\mathrm{FA}} = P'_{\mathrm{M}} = P'_{\mathrm{e}} = \dfrac{1}{3}$.

Exercise 17 (2.2; 2.6)

A system generates two observations X_1 and X_2 that, under both hypothesis $H = 0$ and $H = 1$, are independent and identically distributed:

$$\begin{array}{ll} p_{X_i|H}(x_i|1) = 2x_i & 0 < x_i < 1 \\ p_{X_i|H}(x_i|0) = 2(1 - x_i) & 0 < x_i < 1 \end{array}$$

Assume that the *a priori* probability is the same for both hypotheses.

(a) Determine the MAP decider based on X_1, and calculate its probability of error.

Let DMAP1 be the decider of section a), and assume that if $|x_1 - 0.5| < a$ (with $0 < a < 0.5$), X_2 is also observed. When this happens, and with the goal of still applying a threshold classifier, X_1 is discarded (as well as DMAP1 decision, and a second MAP classifier (DMAP2), based on the observation of X_2, is applied.

(b) Plot on plane $X_1 - X_2$, for a generic value a, the decision regions for the joint scheme DMAP1-DMAP2.

(c) Find the probability of error of the joint scheme DMAP1-DMAP2.

(d) Find the maximum reduction of the probability of error that can be achieved using the joint scheme, with respect to the probability of error of decider DMAP1.

(e) Compare the performance of the joint decider DMAP1-DMAP2 with that of the optimum MAP decider based on the joint observation of X_1 and X_2.

Solution:

(a) $x_1 \underset{D=0}{\overset{D=1}{\gtrless}} \dfrac{1}{2} \qquad P_{\mathrm{e}} = \dfrac{1}{4}$

(b)
$$\begin{array}{lll} D = 0: & x_1 < 1/2 - a \quad \text{and} \quad 1/2 - a < x_1 < 1/2 + a,\ x_2 < 1/2 \\ D = 1: & 1/2 - a < x_1 < 1/2 + a,\ x_2 > 1/2 \quad \text{and} \quad x_1 > 1/2 + a \end{array}$$

(c) $P_{\mathrm{e}} = a^2 - 0.5a + 0.25$

(d) The maximum reduction of the probability of error is $\dfrac{1}{16}$

(e) DMAP(X_1 and X_2): $P_{\mathrm{e}} = \dfrac{1}{6}$

DMAP1- DMAP2: P_{e} changes from $\dfrac{1}{4}$ to $\dfrac{1}{16}$

Exercise 18 (2.2)

Consider a binary decision problem characterized by:

$$p_{X_1,X_2|H}(x_1, x_2|i) = a_i^2 \exp\left[-a_i\left(x_1 + x_2\right)\right] \quad x_1, x_2 > 0 \quad i = 0, 1$$

where $a_0 = 1$ and $a_1 = 2$.

(a) Design the corresponding MAP decider as a function of parameter $R = P_H(1)/P_H(0)$.

(b) Check that $T = X_1 + X_2$ is a sufficient statistic, and calculate the likelihoods expressed as probability density functions of such statistic, $p_{T|H}(t|i)$, $i = 0, 1$.

(c) Calculate the false alarm, missing, and error probabilities of the decider designed in section (a).

Solution:

(a) $\begin{array}{ll} D=1: & x_1+x_2<\ln(4R) \\ D=0: & x_1+x_2>\ln(4R) \end{array}$

(b) $\begin{array}{ll} D=1: & t<\ln(4R) \\ D=0: & t>\ln(4R) \end{array}$

$$p_{T|H}(t|0)=t\exp(-t)\,, \quad t>0 \qquad p_{T|H}(t|1)=4t\exp(-2t)\,, \quad t>0$$

(c) $P_{\mathrm{FA}}=1-\dfrac{1+\ln(4R)}{4R} \qquad P_{\mathrm{M}}=\dfrac{1+2\ln(4R)}{(4R)^2} \qquad P_{\mathrm{e}}=P_H(0)\left(1-\dfrac{3}{16R}-\dfrac{1}{8R}\ln(4R)\right)$

Exercise 19 (2.2; 2.5)

Consider a binary decision problem with $P_H(0)=P_H(1)$ and likelihoods:

$$\begin{array}{ll} p_{X|H}(x|0)=2\,(1-x) & 0<x<1 \\ p_{X|H}(x|1)=1/a & 0<x<a \end{array}$$

$a\geq 1$ being a deterministic parameter.

(a) Design the optimal classifier for cost policy $c_{00}=c_{11}=0$ and $c_{01}=c_{10}=1$, assuming that the value of a is known.

Assume now that the value of a is not known. We opt to apply a minimax strategy, using a threshold x_u^* for the decision process which is selected to minimize the maximum mean cost, i.e.,

$$x_u^*=\arg\left\{\min_{x_u}\left\{\max_a C(x_u,a)\right\}\right\}$$

where x_u is a generic decision threshold

$$x \underset{D=0}{\overset{D=1}{\gtrless}} x_u$$

(b) Obtain x_u^*.

(c) Find the increment of the mean cost that would be produced when applying the minimax strategy over the cost that would be obtained if the value of a were known.

Solution:

(a) $x \underset{D=0}{\overset{D=1}{\gtrless}} 1-\dfrac{1}{2a} \quad 0<x<a$

(b) $x_u^*=\dfrac{1}{2}$

(c) $\Delta P_{\mathrm{e}}=\dfrac{1}{8}-\dfrac{1}{4a}\left(1-\dfrac{1}{2a}\right)$, which is zero for $a=1$, and positive for $a>1$.

Exercise 20 (2.3)

Consider a bidimensional Gaussian decision problem

$$p_{X_1,X_2|H}(x_1,x_2|0) = G\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\right)$$

$$p_{X_1,X_2|H}(x_1,x_2|1) = G\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\right)$$

The *a priori* probabilities of the hypotheses are $P_H(0) = 2/3$ and $P_H(1) = 1/3$, whereas the associated cost policy is $c_{00} = c_{11} = 0$, $c_{01} = c_{10} = 1$.

(a) Establish the expression that provides the corresponding Bayes' decision as a function of the observation vector $\mathbf{X}$.

(b) Show, over a graphic representation, how the decision border changes when varying the value of $P_H(0)$.

Solution:

(a) $x_2 - x_1 \underset{D=0}{\overset{D=1}{\gtrless}} 10 \ln 2$

(b) If $P_H(0)$ increases, the border moves towards point $[0,1]^T$, while a reduction in $P_H(0)$ moves the border towards $[1,0]^T$.

Exercise 21 (2.2; 2.4)

Consider a radar detection problem in which the targets can cause echoes with two different intensity levels:

$$\begin{array}{ll} H = 0 \text{ (no target):} & X = N \\ H = 1 \text{ (target present):} & \begin{cases} H = 1a: & X = s_1 + N \\ H = 1b: & X = s_2 + N \end{cases} \end{array}$$

where s_1 and s_2 are real values associated to the two echo levels for the different targets, and N is a r.v. with distribution $G(0,1)$. It is also known that $P_H(1a|1) = P$ and $P_H(1b|1) = 1 - P$ $(0 < P < 1)$.

(a) Establish the general shape of an LRT which discriminates $H = 0$ and $H = 1$, and justify that such classifier is a threshold classifier when the signs of s_1 and s_2 are the same.

(b) Are there any combination of values of s_1 and s_2 for which a maximum likelihood test decides always in favor of the same hypothesis?

(c) Assuming $s_2 < s_1 < 0$ and the following threshold detector:

$$x \underset{D=1}{\overset{D=0}{\gtrless}} \eta$$

obtain P_{FA} and P_{D} as functions of η, and express your result using function:

$$F(x) = 1 - Q(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

Provide an approximate representation of the classifier's ROC curve (P_{D} vs P_{FA} as function of η), indicating where the points associated to $\eta \to \pm\infty$ would be placed, and how the operation point changes with the threshold.

(d) Explain the effects on the ROC of the following events:
- An increment of s_1.
- A decrement of s_2.
- An increment of P.
- An increment of $P_H(0)$.

Solution:

(a) $$P \exp\left[-\frac{1}{2}\left(s_1^2 - 2s_1 x\right)\right] + (1-P)\exp\left[-\frac{1}{2}\left(s_2^2 - 2s_2 x\right)\right] \underset{D=0}{\overset{D=1}{\gtrless}} \eta$$

(b) No

(c) $P_{\mathrm{FA}} = 1 - F(\eta) \qquad P_{\mathrm{D}} = 1 - PF(\eta - s_1) - (1-P)F(\eta - s_2)$

(d)
- Increasing s_1: reduces the area below the ROC.
- Decreasing s_2: increases the area below the ROC.
- Increasing P: reduces the area below the ROC.
- Increasing $P_H(0)$ does not affect the ROC curve.

Exercise 22 (2.2; 2.5)

Consider a binary decision problem described by

$$\begin{aligned} p_{X|H}(x|0) &= a_0 x^2 && |x| < 1 \\ p_{X|H}(x|1) &= a_1\left(3 - |x|\right) && |x| < 3 \end{aligned}$$

where a_0 and a_1 are constants, with the same *a priori* probabilities for the two hypotheses, and where the following cost policy is used: $c_{00} = c_{11} = 0$, $c_{10} = c_{01} = c$ with $c > 0$.

(a) Calculate constants a_0 and a_1.

(b) Determine the Bayes' optimal classifier.

(c) Calculate the probability of error of this decider.

(d) Design the Neyman-Pearson decider that guarantees a P_{FA} not larger than a pre-established value α.

Solution:

(a) $a_0 = 3/2$ and $a_1 = 1/9$.

(b) $$\begin{aligned} D = 1: &\quad |x| < 0.43 \text{ and } |x| > 1 \\ D = 0: &\quad 0.43 < |x| < 1 \end{aligned}$$

(c) $P_{\mathrm{e}} = 0.184$.

(d) $$\begin{aligned} D = 1: &\quad |x| < \alpha^{1/3} \text{ and } |x| > 1 \\ D = 0: &\quad \alpha^{1/3} < |x| < 1 \end{aligned} .$$

Exercise 23 (2.2)

Consider a binary decision problem with cost policy $c_{00} = c_{11} = 0$, $c_{01} = c_{10} = 1$, and likelihoods

$$\begin{aligned} p_{X|H}(x|0) &= \lambda_0 \exp\left(-\lambda_0 x\right) && x \geq 0 \\ p_{X|H}(x|1) &= \lambda_1 \exp\left(-\lambda_1 x\right) && x \geq 0 \end{aligned}$$

where $\lambda_0 = 2\lambda_1$.

(a) Assumming that $P_H(1) = 1/2$ design the classifier that minimizes the mean cost.

(b) Calculate P_{FA} and P_{M} for the decider obtained in (a).

(c) Assuming that the true value of $P_H(1)$ is $P > 0$, but we keep using the classifier designed in part (a). Plot the risk of the decider as a function of P.

(d) The previous decider is applied to two independent observations. Find the probabilities of incurring in exactly 0, 1, and 2 errors, as a function of P.

(e) Assume that the risk associated to two decisions is not the sum of the costs for each decision, but instead:

- If both decisions are correct the associated cost is 0.
- The cost of incurring in just one error is 1.
- The cost incurred by two wrong decisions is $c = 18$.

Plot the mean risk of the two decisions as a function of P.

Solution:

(a) $x \underset{D=0}{\overset{D=1}{\gtrless}} \dfrac{1}{\lambda_1}\ln 2$

(b) $P_{\text{FA}} = 0.25 \qquad P_{\text{M}} = 0.5$

(c) $R = (1+P)/4$

(d) $P\{0 \text{ errors}\} = \left(\frac{1}{4}(3-P)\right)^2$

$P\{1 \text{ error}\} = 2\frac{1}{4}(1-P)\frac{1}{4}(3-P)$

$P\{2 \text{ errors}\} = \left(\frac{1}{4}(1+P)\right)^2$

(e) The risk associated to the two decisions is: $P^2 + \frac{5}{2}P + \frac{3}{2}$.

Exercise 24 (2.2)

A sociological studies institute is working on a project to predict which party will win the next elections. In order to do so, they first evaluate the level of voters turnout. Historically, a low voter turnout favors the PDD party whereas a high voter turnout favors the CSI party. The likelihood of each party winning in each of the two previous scenarios is shown in the following table:

P(voters turnout | Winning party)	low level	high level
PDD	0.7	0.3
CSI	0.4	0.6

The charisma of each candidate also influences the result of the election. This is statistically modelled with the probabilities conditioned on the winning party and the level of voters turnout, provided in the table below:

P(Charisma | voters turnout, winning party)	$-$	$=$	$+$
low, PDD	0.6	0.3	0.1
high, PDD	0.5	0.15	0.35
low, CSI	0.4	0.2	0.4
high, CSI	0.1	0.1	0.8

In this table, − indicates that the PDD candidate is more charismatic than the CSI candidate, + has the opposite meaning, and = denotes that both candidates have the same charisma.

Finally, a survey is taken to predict citizens voting intention (i.e., the output of the survey is a prediction about the winning party). The following table shows the probabilities of the joint distribution of the events ‘winning party’ and ‘survey prediction’.

P(Winning party, survey prediction)	PDD predicted	CSI predicted
PDD	0.35	0.05
CSI	0.2	0.4

Consider in the following that the victory of CSI is the null hypothesis ($h = 0$). Carry out the following tasks to study the relevance of the three measured observations (i.e., voters turnout, charisma, and survey prediction):

(a) Find the maximum likelihood decider that outcomes the winning party using jointly the observations about the level of voters turnout and candidates charisma. Find the probabilities of correctly predicting a victory of both the PDD and the CSI parties with such detector.

(b) Obtain the maximum *a posteriori* decider that outcomes the winning party using jointly the observations about the level of voters turnout and survey predictions. Calculate the probability of error of this detector.

(c) Find the ROC curve for an LRT decider based on the joint observation of voters turnout level and candidates charisma. Place in that curve the maximum likelihood obtained in subsection (a).

(d) Obtain the Neyman-Pearson detector when the three observations are used jointly for a maximum probability of false alarm $P_{\text{FA}} = 0.1$, and its associated probability of detection. In order to do so, you should use the following table of probabilities conditioned on each of the hypotheses:

$P(\text{obs.} \mid H_i)$	PDD low −	PDD low =	PDD low +	PDD high −	PDD high =	PDD high +	CSI low −	CSI low =	CSI low +	CSI high −	CSI high =	CSI high +
PDD	0.3675	0.1837	0.0612	0.1312	0.0525	0.0788	0.0525	0.0262	0.0087	0.0187	0.0075	0.0112
CSI	0.0533	0.0267	0.0533	0.0200	0.0200	0.1600	0.1067	0.0533	0.1067	0.0400	0.0400	0.3200

Solution:

(a) The ML decider is:

Voters turnout \ Charisma	−	=	+
high	PDD	PDD	CSI
low	PDD	CSI	CSI

$P\{D = \text{CSI} | H = \text{CSI}\} = 0.7$ and $P\{D = \text{PDD} | H = \text{PDD}\} = 0.78$

(b) The MAP decider is:

Voters turnout \ Survey Prediction	PDD predicted	CSI predicted
low	PDD	CSI
high	CSI	CSI

$P_{\text{e}} = 0.235$

(c) The ROC curve is characterized by the following operation points

η range	P_{FA}	P_{D}
$\eta < 0.21875$	1	1
$0.21875 < \eta < 0.4375$	0.52	0.895
$0.4375 < \eta < 0.75$	0.36	0.825
$0.75 < \eta < 2.5$	0.3	0.78
$2.5 < \eta < 2.625$	0.24	0.63
$2.625 < \eta$	0	0

The ML decider corresponds to an operation point with $0.75 < \eta < 2.5$.

(d) To obtain the Neyman-Pearson decider, the LRT threshold must be in the interval $(4.92, 6.56)$. In that case, $P_{\text{D}} = 0.6824$

Exercise 25 (2.2)

An insurance company classifies its clients into two groups: prudent and reckless clients ($H = 0$ and $H = 1$, respectively). The probability of a prudent client having k accidents during a year is modelled as a Poisson distribution with unity parameter:

$$P_{K|H}(k|0) = \frac{\exp(-1)}{k!}, \quad k = 0, 1, 2, \ldots$$

In the case of reckless customers, the same probability is modelled as a Poisson distribution with parameter 4:

$$P_{K|H}(k|1) = \frac{4^k \exp(-4)}{k!}, \quad k = 0, 1, 2, \ldots$$

(where it is considered $0! = 1$).

(a) Design a maximum likelihood decider that classifies clients into prudent or reckless based on the number of accidents suffered by the client during its first year in the company.

(b) The performance of the previous classifier can be assessed as a function of these parameters:

- the percentage of prudent clients that will leave the company because they are classified as reckless, and therefore not offered discounts;
- the percentage of reckless clients that are classified as prudent and result in economical losses for the company.

Find the relationship between these quality indicators and the probabilities of False Alarm and Detection, calculating their values (Indication: consider for the calculations 0!=1).

(c) A statistical study paid by the company reflects that just one out of 17 clients is reckless. Find the minimum probability error decider in the light of the new information. Compare this decider with that designed in subsection (a) in terms of probability of error, false alarm, and missing.

Solution:

(a) $k \underset{D=0}{\overset{D=1}{\gtrless}} 2.16.$

(b) $P_{\text{FA}} = 8\,\%$ (this is the percentage of prudent clients that will leave the company).
$P_{\text{D}} = 76.2\,\%$ (this is the percentage of reckless clients that are correctly identified as such)

(c) $k \underset{D=0}{\overset{D=1}{\gtrless}} 4.16$. $P_{\text{FA}} = 0.37\,\%$. $P_{\text{M}} = 37.11\,\%$ and $P_{\text{e}} = 4\,\%$.

For the ML decider, $P_{\text{e}} = 8.9\,\%$.

Exercise 26 (2.2; 2.5)

Bla, bla, bla

Solution: Bla, bla, bla

Exercise 27 (2.2)

Bla, bla, bla

Solution: Bla, bla, bla

Exercise 28 (2.2)

Bla, bla, bla

Solution: Bla, bla, bla

Exercise 29 (2.3)

Bla, bla, bla

Solution: Bla, bla, bla

Exercise 30 (2.3)

Bla, bla, bla

Solution: Bla, bla, bla

Exercise 31 (2.3)

Bla, bla, bla

Solution: Bla, bla, bla

A. Additional Problems

We include in this appendix some additional problems taken from the exercises and problems collection used in courses from other degrees. The resolution of these problems should be feasible for students of the new degrees.

Exercise 2.E2 (2.2)

Consider the binary hypotheses

$$\begin{array}{l} H = 0 : X = N \\ H = 1 : X = s + N \end{array}$$

$s > 0$ being a known constant, and where N is a noise with the following pdf:

$$p_N(n) = \begin{cases} \dfrac{1}{s}\left(1 - \dfrac{|n|}{s}\right), & |n| < s \\ \\ 0, & |n| > s \end{cases}$$

The *a priori* probabilities of the hypotheses are $P_H(0) = 1/3$, $P_H(1) = 2/3$.

(a) Design the MAP decider.

(b) Determine the corresponding P_{FA} and P_{M}, as well as the error probability.

(c) Determine how would these probabilities change if we applied to this situation the same kind of decider but designed under the assumption that N is Gaussian with the same variance of the noise actually present (and zero mean).

Solution:

(a) $$\begin{aligned} D0 &: \quad -s < x < \frac{s}{3} \\ D1 &: \quad \frac{s}{3} < x < 2s \end{aligned}$$

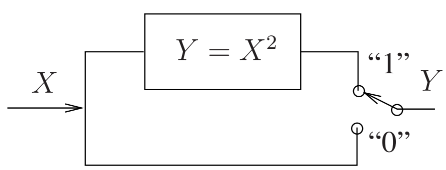
(b) $P_{\mathrm{FA}} = \dfrac{2}{9} \approx 0.2222 \qquad P_{\mathrm{M}} = \dfrac{1}{18} \approx 0.0556 \qquad P_{\mathrm{e}} = \dfrac{1}{9} \approx 0.1111$

(c) $P_{\mathrm{FA}} = \dfrac{\left(1+\dfrac{\ln 2}{3}\right)^2}{8} \approx 0.1894$ (decreases) $\qquad P_{\mathrm{M}} = \dfrac{\left(1-\dfrac{\ln 2}{3}\right)^2}{8} \approx 0.0739$ (increases)

$$P_{\mathrm{e}} = \frac{\left(1-\dfrac{\ln 2}{3}\right)^2}{12} + \frac{\left(1+\dfrac{\ln 2}{3}\right)^2}{24} \approx 0.1124 \qquad \text{(increases)}$$

Exercise 2.E8 (2.2)

The switch shown in the figure is in its upper position ("1") with known probability P. Random variable X has a uniform probability density $U(0, 1)$.



The position of the switch cannot be observed, but the output value Y is available. Based on the observation of this value, we want to apply a Bayesian decider to predict which is the position of the switch. The cost policy is $c_{00} = c_{11} = 0$, $c_{10} = 2c_{01}$.

(a) Pose the problem using the usual equations for an analytical design.

(b) Determine the corresponding test to be used, based on the possible values of P.

(c) Calculate P_{FA} and P_{M}.

(Hint: in order to find $p_Y(y)$, find the relationship that exists between the cumulative distributions of Y and X).

Solution:

(a) $$\begin{aligned} H = 1 &: \quad Y = X^2, \text{ with probability P} \\ H = 0 &: \quad Y = X, \text{ with probability 1-P} \end{aligned}$$

(b) - If $P > 4/5 : \Rightarrow D = 1$ (always)

- If $P < 4/5 :$ $$\begin{cases} 0 < y < \dfrac{1}{16}\left(\dfrac{P}{1-P}\right)^2 & \Rightarrow \quad D = 1 \\ \dfrac{1}{16}\left(\dfrac{P}{1-P}\right)^2 < y < 1 & \Rightarrow \quad D = 0 \end{cases}$$

(c) - If $P > 4/5$: $P_{\mathrm{FA}} = 1;\ P_{\mathrm{M}} = 0$

- If $P < 4/5$: $P_{\mathrm{FA}} = \dfrac{1}{16}\left(\dfrac{P}{1-P}\right)^2;\ P_{\mathrm{M}} = \dfrac{1-\dfrac{5P}{4}}{1-P}$

Exercise 2.E10 (2.1)

A fair dice (with faces from 1 to 6) is thrown and the r.v. X with pdf

$$p_X(x) = \begin{cases} \dfrac{2}{a}\left(1-\dfrac{x}{a}\right), & 0 < x < a \\ 0, & \text{otherwise} \end{cases}$$

is generated so that its mean is given by the result of throwing the dice (i.e., the mean is equal to the number of points in the upper face). Assume that for a given throw we have access to 3 independent measurements of X, with values $x^{(1)} = 2, x^{(2)} = 5, x^{(3)} = 10$. Decide from these values which is the result of throwing the dice according to the maximum likelihood criterion.

Solution: The Maximum Likelihood criterion determines that face '5' should be selected.