Decision Theory: Problems

Notation:

- ML: Maximum Likelihood decision-maker [$\phi_{\mathrm{ML}}(\mathbf{x})$].
- MAP: Maximum *a posteriori* decision-maker [$\phi_{\mathrm{MAP}}(\mathbf{x})$].
- 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: Receiver Operating Characteristic.

Índice

1. Binary ML and MAP decision

1.1. One-dimensional observations

DT1

Consider the binary decision problem given by likelihoods

$$p_{X|H}(x|1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(x - 4\sqrt{2\pi}\right)^2\right),$$

$$p_{X|H}(x|0) = \sqrt{2\pi} \exp\left(-\sqrt{2\pi}x\right), \qquad x \geq 0$$

(a) Compute the decision regions of the ML classifier based on x.

(b) Compute the missing probability of the ML classifier.

(c) Compute the false alarm probability of the ML classifier.

When it was appropriate, express the 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) $D = 1$ si $x \in [-\infty, 0] \cup [A, B]$, where $A = 5\sqrt{2\pi} - \sqrt{18\pi - 2\ln(2\pi)}$, $B = 5\sqrt{2\pi} + \sqrt{18\pi - 2\ln(2\pi)}$.

(b) $P_{\mathrm{M}} = F(A - 4\sqrt{2\pi}) - F(-4\sqrt{2\pi}) + 1 - F(B - 4\sqrt{2\pi})$

(c) $P_{\mathrm{FA}} = \exp\left(-\sqrt{2\pi}A\right) - \exp\left(-\sqrt{2\pi}B\right)$

DT2

Consider a binary detection problem ($H \in \{0, 1\}$) and observations $X \in \mathbb{R}$. The likelihoods are

$$p_{X|H}(x|0) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right),$$

$$p_{X|H}(x|1) = \begin{cases} \frac{1}{2a}, & -a < x < a, \\ 0, & \text{otherwise}, \end{cases}$$

and the hypotheses are equally likely. Derive:

(a) The decision regions of the detector that minimizes the probability of error for an arbitrary value of a, with $a > 0$.

Solution: The detector that minimizes the probability of error is given by

$$\frac{p_{X|H}(x|1)}{p_{X|H}(x|0)} \underset{D=0}{\overset{D=1}{\gtrless}} \frac{P_H(0)}{P_H(1)} = 1,$$

which is the maximum likelihood detector for equally likely hypotheses. Before proceeding, as always, it is convenient to plot these likelihoods. However, we need to consider two different cases:

A) The largest value of $p_{X|H}(x|0)$ is larger than that of $p_{X|H}(x|1)$, i.e.,

$$\frac{1}{2a} > \frac{1}{\sqrt{2\pi}} \Rightarrow a < \sqrt{\frac{\pi}{2}} \approx 1.25.$$

B) The largest value of $p_{X|H}(x|0)$ is smaller (or equal) than that of $p_{X|H}(x|1)$, i.e.,

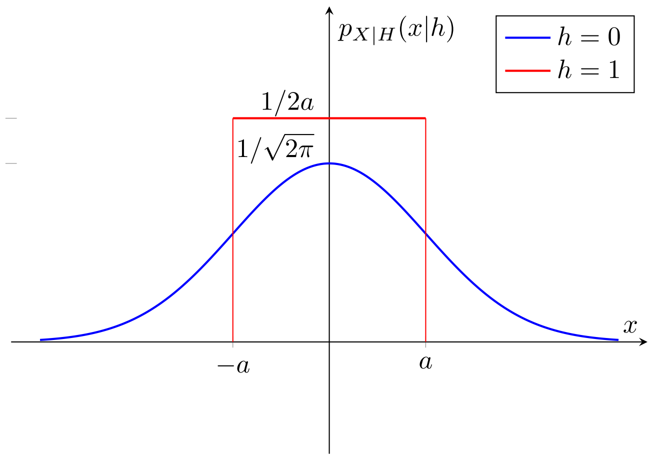
$$\frac{1}{2a} \leq \frac{1}{\sqrt{2\pi}} \Rightarrow a \geq \sqrt{\frac{\pi}{2}} \approx 1.25.$$

Then, for Case A) the likelihoods are shown in the following figure. From this figure, it is easy to see that

$$|x| \underset{D=1}{\overset{D=0}{\gtrless}} a,$$

and the decision regions are

$$\mathcal{X}_0 = \{x \in \mathbb{R} \mid |x| \geq a\},$$
$$\mathcal{X}_1 = \{x \in \mathbb{R} \mid -a < x < a\}.$$



For Case B), the likelihoods are shown in the next figure, where we can see that

$$\mathcal{X}_0 = \{x \in \mathbb{R} \mid -b < x < b\} \cup \{x \mid |x| \geq a\},$$
$$\mathcal{X}_1 = \{x \in \mathbb{R} \mid b \leq |x| < a\},$$

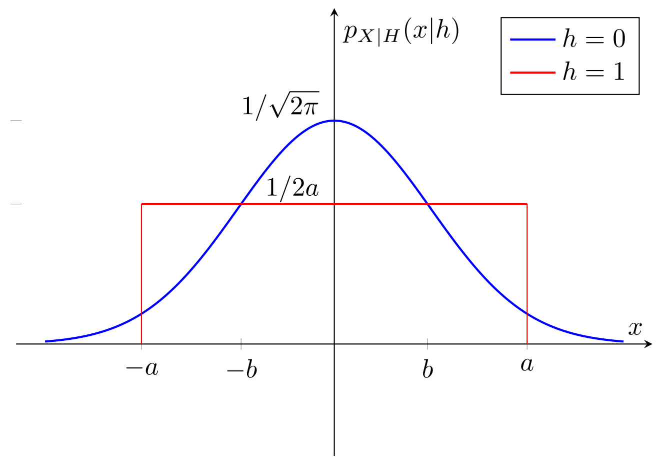
where b is obtained as the positive solution to

$$p_{X|H}(b|0) = p_{X|H}(b|1), \Rightarrow b = \sqrt{2\log\left(\sqrt{\frac{2}{\pi}}a\right)}.$$

Then, we have

$$\mathcal{X}_0 = \left\{x \in \mathbb{R} \,\middle|\, -\sqrt{2\log\left(\sqrt{\frac{2}{\pi}}a\right)} < x < \sqrt{2\log\left(\sqrt{\frac{2}{\pi}}a\right)}\right\} \cup \{x \mid |x| \geq a\},$$

$$\mathcal{X}_1 = \left\{x \in \mathbb{R} \,\middle|\, \sqrt{2\log\left(\sqrt{\frac{2}{\pi}}a\right)} \leq |x| < a\right\}.$$



(b) The probability of detection, P_{D}, as a function of a, with $a > 0$. Sketch a plot of P_{D} vs. a for $a \in (0, 50)$.

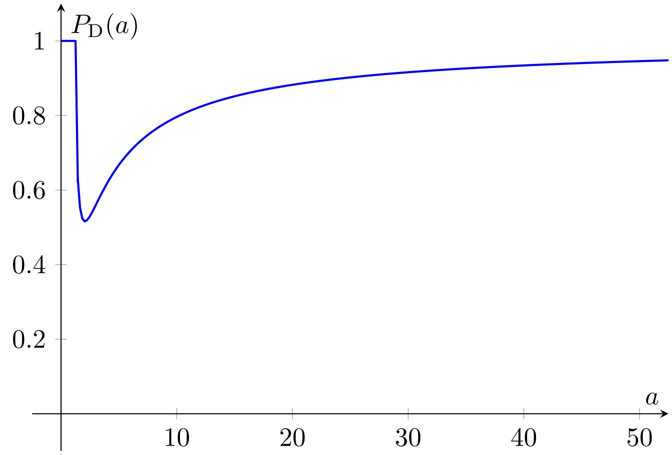
Solution: Let us start again with Case A). In this case, the probability of detection is

$$P_{\text{D}} = P(D = 1|H = 1) = \int_{\mathcal{X}_1} p_{X|H}(x|1)dx = 1,$$

regardless of the value of a, with $a < \sqrt{\frac{\pi}{2}}$. That is, for Case A) we are integrating the whole likelihood under $H = 1$. When we consider Case B), it becomes a bit more involved. Concretely, for $a \geq \sqrt{\frac{\pi}{2}}$, we have

$$P_{\text{D}} = P(D = 1|H = 1) = \int_{\mathcal{X}_1} p_{X|H}(x|1)dx = \int_{-a}^{-b} \frac{1}{2a}dx + \int_{b}^{a} \frac{1}{2a}dx = 2\frac{a-b}{2a}$$
$$= 1 - \frac{b}{a} = 1 - \frac{1}{a}\sqrt{2\log\left(\sqrt{\frac{2}{\pi}}a\right)}.$$

the plot of P_{D} is shown in the following figure



(c) The probability of error for $a = 1$.

Solution: Since $a = 1 < \sqrt{\pi/2}$, we are in Case A), for which we already know that $P_{\mathrm{D}} = 1$. Then, since

$$P_{\mathrm{e}} = P_{\mathrm{FA}} \cdot P_H(0) + P_{\mathrm{M}} \cdot P_H(1) = \frac{1}{2} P_{\mathrm{FA}} + \frac{1}{2}(1 - P_{\mathrm{D}}) = \frac{1}{2} P_{\mathrm{FA}},$$

it only remains to compute P_{FA}. For Case A), the probability of false alarm is given by

$$P_{\mathrm{FA}} = P(D = 1 | H = 0) = \int_{\mathcal{X}_1} p_{X|H}(x|0) dx = \int_{-a}^{a} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx.$$

Taking now into account the symmetry of the likelihood, P_{FA} simplifies to

$$P_{\mathrm{FA}} = 2 \int_0^a \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = 2\left[\frac{1}{2} - \int_a^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx\right] = 1 - 2Q(a),$$

which yields

$$P_{\mathrm{e}} = \frac{1}{2} - Q(a).$$

DT3

Consider the binary decision problem characterized by an observation $X \in [0, 2]$ and likelihoods

$$p_{X|H}(x|1) = \frac{1}{2} x$$

$$p_{X|H}(x|0) = \frac{3}{4} x(2 - x)^2,$$

with $P_H(1) = \frac{2}{5}$.

(a) Find the MAP classifier.

(b) Obtain the probability of missing of the MAP classifier.

(c) Assume now that the same decision maker that was designed in subsection (a) is applied to a scenario in which the likelihood of $H = 1$ is

$$p'_{X|H}(x|1) = \frac{7}{8} p_{X|H}(x|1) + \frac{1}{16},$$

whereas the likelihood of $H = 0$ remains unchanged. Obtain the increment in the probability of error that takes place as a consequence of this different scenario.

Solution:

(a) $x \underset{D=0}{\overset{D=1}{\gtrless}} \frac{4}{3}$

(b) $P_{\mathrm{M}} = \frac{4}{9}$

(c) Since P_{FA} does not change and $P'_{\mathrm{M}} = \frac{17}{36}$, the increment of the probability of error is $P_H(1)(P'_{\mathrm{M}} - P_{\mathrm{M}}) = \frac{1}{90}$

DT4

The random variables X, Y and Z are statistically independent and follow a uniform distribution:

$$\begin{aligned} p_X(x) &= 1, && 0 \le x \le 1 \\ p_Y(y) &= 1, && 0 \le y \le 1 \\ p_Z(z) &= 1, && 0 \le z \le 1 \end{aligned}$$

Consider the following decision problems. In all of them, X is observed, but neither Y nor Z are known.

Problem 1: given by the hypotheses:

$$\begin{aligned} H = 1: &\quad X > 0.2 \\ H = 0: &\quad X \le 0.2 \end{aligned}$$

Problem 2: given by the hypotheses:

$$\begin{aligned} H = 1: &\quad X > Y \\ H = 0: &\quad X \le Y \end{aligned}$$

Problem 3: given by the hypotheses:

$$\begin{aligned} H = 1: &\quad (X > Y) \quad \text{and} \quad (X > Z) \\ H = 0: &\quad (X \le Y) \quad \text{or} \quad (X \le Z) \end{aligned}$$

(a) Determine the MAP decision-maker for problem 1.

(b) Compute the error probability of the MAP decision-maker for problem 1.

(c) Determine the MAP decision-maker for problem 2.

(d) (hard) Compute the error probability of the MAP decision-maker for problem 2.

(e) Determine the MAP decision-maker for problem 3.

(f) (very hard) Compute the error probability of the MAP decision-maker for problem 3.

Solution:

(a) Observing $X = x$ implies knowing the true hypothesis: if $x > 0.2$, the true hypothesis is 1, and it will be 0 otherwise. Thus, the MAP decision-maker is given by

$$x \underset{D=0}{\overset{D=1}{\gtrless}} 0.2$$

Formally, this can also be verified by applying the MAP decision rule in a straightforward way:

$$\begin{aligned} &P_{H|X}(1|x) \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2} \\ \Leftrightarrow &\quad P\{X > 0.2 | X = x\} \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2} \\ \Leftrightarrow &\quad \begin{bmatrix} 1 & \text{si } x > 0.2 \\ 0 & \text{si } x \le 0.2 \end{bmatrix} \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2} \\ \Leftrightarrow &\quad x \underset{D=0}{\overset{D=1}{\gtrless}} 0.2 \end{aligned}$$

(b) Since observing x implies knowing H deterministically, the error probability will be zero. Formally, this can be determined by knowing that:

$$P_{\mathrm{FA}} = P\{D=1|H=0\} = P\{X > 0.2|X \leq 0.2\} = 0$$
$$P_{\mathrm{M}} = P\{D=0|H=1\} = P\{X \leq 0.2|X > 0.2\} = 0$$

thus $P_e = P_H(0)P_{\mathrm{FA}} + P_H(1)P_{\mathrm{M}} = 0$

(c) Knowing that

$$P_{H|X}(1|x) = P\{X > Y|X=x\} = P\{Y < x\} = \int_0^x 1 dx = x$$

the MAP classifier will be

$$x \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2}$$

(d) The error probability is

$$\begin{aligned}
P_e &= P\{D=1, H=0\} + P\{D=0, H=1\} \\
&= P\left\{X > \frac{1}{2}, X \leq Y\right\} + P\left\{X \leq \frac{1}{2}, X > Y\right\} \\
&= \int_0^1 P\left\{X > \frac{1}{2}, X \leq Y|X=x\right\} p_X(x)dx + \int_0^1 P\left\{X \leq \frac{1}{2}, X > Y|X=x\right\} p_X(x)dx \\
&= \int_0^1 P\left\{x > \frac{1}{2}, Y \geq x|X=x\right\} dx + \int_0^1 P\left\{x \leq \frac{1}{2}, Y < x|X=x\right\} dx \\
&= \int_{\frac{1}{2}}^1 P\left\{Y \geq x|X=x\right\} dx + \int_0^{\frac{1}{2}} P\left\{Y < x|X=x\right\} dx \\
&= \int_{\frac{1}{2}}^1 (1-x)dx + \int_0^{\frac{1}{2}} xdx = \frac{1}{4}
\end{aligned}$$

(e) Knowing that

$$P_{H|X}(1|x) = P\{X > Y, X > Z|X=x\} = P\{Y < x, Z < x\} = P\{Y < x\} \cdot P\{Z < x\} = x^2$$

the MAP decision-maker is

$$x \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{\sqrt{2}}$$

(f) The error probability is

$$P_e = P\{D=1, H=0\} + P\{D=0, H=1\}$$

The first term can be computed in a similar way to (c):

$$\begin{aligned}
P\{D=0,H=1\} &= P\left\{X<\frac{1}{\sqrt{2}}, X\geq Y, X\geq Z\right\}\\
&= \int_0^1 P\left\{X<\frac{1}{\sqrt{2}},\ X\geq Y,\ X\geq Z \mid X=x\right\} p_X(x)dx\\
&= \int_0^{\frac{1}{\sqrt{2}}} P\left\{x<\frac{1}{\sqrt{2}},\ Y\leq x,\ Z\leq x \mid X=x\right\} dx\\
&= \int_0^{\frac{1}{\sqrt{2}}} P\left\{Y\leq x\right\} P\left\{Z\leq x\right\} dx\\
&= \int_0^{\frac{1}{\sqrt{2}}} x^2 dx = \left[\frac{1}{3}x^3\right]_0^{\frac{1}{\sqrt{2}}}\\
&= \frac{1}{6\sqrt{2}}
\end{aligned}$$

The second term can be computed as follows:

$$P\{D=1,H=0\} = P\{D=1\} - P\{D=1,H=1\}$$

The first term in this new equation is

$$P\{D=1\} = P\left\{X>\frac{1}{\sqrt{2}}\right\} = \int_{\frac{1}{\sqrt{2}}}^1 p_X(x)dx = 1-\frac{1}{\sqrt{2}}$$

The second term is

$$\begin{aligned}
P\{D=1,H=1\} &= P\left\{X>\frac{1}{\sqrt{2}}, X\geq Y, X\geq Z\right\}\\
&= \int_0^1 P\left\{X>\frac{1}{\sqrt{2}},\ X\geq Y,\ X\geq Z \mid X=x\right\} p_X(x)dx\\
&= \int_{\frac{1}{\sqrt{2}}}^1 P\left\{Y\leq x\right\} P\left\{Z\leq x\right\} dx\\
&= \int_{\frac{1}{\sqrt{2}}}^1 x^2 dx = \left[\frac{1}{3}x^3\right]_0^{\frac{1}{\sqrt{2}}}\\
&= \frac{1}{3}-\frac{1}{6\sqrt{2}}
\end{aligned}$$

Joining all the above term, we get

$$\begin{aligned}
P_e &= P\{D=1\} - P\{D=1,H=1\} + P\{D=0,H=1\}\\
&= 1-\frac{1}{\sqrt{2}}-\frac{1}{3}+\frac{1}{6\sqrt{2}}+\frac{1}{6\sqrt{2}}\\
&= \frac{2}{3}\left(1-\frac{1}{\sqrt{2}}\right)
\end{aligned}$$

DT5

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

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

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

$$\begin{aligned} H = 0: &\quad Z \leq 0.6 \\ H = 1: &\quad Z > 0.6 \end{aligned}$$

(a) Find $p_{Z|X}(z|x)$.

(b) Obtain the *a posteriori* probabilities of both hypotheses.

(c) Find the MAP classifier 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 classifier.

(f) Determine the ML classifier based on the observation of X.

Solution:

(a) Knowing that

$$\begin{aligned} p_X(x) &= \int_{-\infty}^{\infty} p_{Z,X}(z,x)dz = \int_0^1 (x+z)dz = \frac{(x+1)^2}{2} - \frac{x^2}{2} \\ &= x + \frac{1}{2}, \qquad 0 \leq x \leq 1 \end{aligned}$$

we obtain

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

(b)

$$\begin{aligned} P_{H|X}(0|x) &= P\{H=0|x\} = P\{Z < 0.6|x\} = \int_{-\infty}^{0.6} p_{Z|X}(z|x)dz \\ &= \int_0^{0.6} \frac{2\,(x+z)}{2x+1}dz = \frac{1.2x + 0.36}{2x+1} \\ P_{H|X}(1|x) &= 1 - P_{H|X}(0|x) = \frac{0.8x + 0.64}{2x+1} \end{aligned}$$

(c) The MAP decision-maker is given by

$$\begin{aligned} P_{H|X}(1|x) \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2} \quad &\Leftrightarrow \quad \frac{0.8x+0.64}{2x+1} \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2} \\ &\Leftrightarrow \quad 0.8x + 0.64 \underset{D=0}{\overset{D=1}{\gtrless}} x + 0.5 \\ &\Leftrightarrow \quad x \underset{D=1}{\overset{D=0}{\gtrless}} 0.7 \end{aligned}$$

(d) Knowing that

$$\begin{aligned} P_H(0) &= \int_{-\infty}^{\infty} P_{H|X}(0|x)p_X(x)dx = \int_0^1 \frac{1.2x+0.36}{2x+1}\left(x + \frac{1}{2}\right)dx \\ &= \int_0^1 (0.6x + 0.18)dx = \frac{1}{2}\left((0.6+0.18)^2 - 0.18^2\right) = 0.48 \\ P_H(1) &= 1 - P_H(0) = 0.52 \end{aligned}$$

we have

$$p_{X|H}(x|0) = \frac{P_{H|X}(0|x)p_X(x)}{P_H(0)} = \frac{1}{0.48}\frac{1.2x+0.36}{2x+1}\left(x+\frac{1}{2}\right) = \frac{10x+3}{8}$$

$$p_{X|H}(x|1) = \frac{P_{H|X}(1|x)p_X(x)}{P_H(1)} = \frac{1}{0.52}\frac{0.8x+0.64}{2x+1}\left(x+\frac{1}{2}\right) = \frac{10x+8}{13}$$

(e)

$$P_{\mathrm{FA}} = P\{D=1|H=0\} = P\{X<0.7|H=0\} = \int_{-\infty}^{0.7} p_{X|H}(x|0)dx$$
$$= \frac{1}{8}\int_0^{0.7}(10x+3)dx = 0.5687$$

(f) The ML decision-maker is

$$p_{X|H}(x|1) \underset{D=0}{\overset{D=1}{\gtrless}} p_{X|H}(x|0) \quad\Leftrightarrow\quad \frac{10x+8}{13} \underset{D=0}{\overset{D=1}{\gtrless}} \frac{10x+3}{8}$$
$$\Leftrightarrow\quad x \underset{D=1}{\overset{D=0}{\gtrless}} 0.5$$

DT6

A test to detect the presence of a certain bacteria in a microbial culture has been developed based on the measure of CO_2 concentration in the culture. The basal level (when the bacteria is not present) for CO_2 concentration is characterized by a gamma distribution:

$$p_T(t) = (0.15)^2\ t\ \exp(-0.15t), \qquad t>0.$$

In contaminated samples (the bacteria is present), the concentration level increases 20 units with respect to the basal level. Therefore, the two hypotheses to consider are:

$$\begin{array}{lcl} H=0 & : & X=T \\ H=1 & : & X=T+20 \end{array} \tag{1}$$

It is also known that the *a priori* probability of contaminated samples is 0.2.

(a) Obtain the expressions for the likelihoods of both hypotheses, expressing them in terms of random variable X.

(b) Find the decision regions of the likelihood ratio test (LRT), as a function of parameter η.

(c) Particularize the decision regions for the ML classifier, as well as for the decision maker that minimizes the probability of error.

(d) Obtain general expressions for P_{FA} and P_{D} as functions of the LRT threshold. Simplify your expressions as much as you can, so that the provided solutions do not imply the evaluation of any integrals.

(e) Find the minimum P_{FA} that can be achieved, if the test has to be adjusted with the goal that no contaminated cultures can remain undetected.

Hint: Simplify your expressions using approximation $\exp(3) \approx 20$.

Solution:

(a) $p_{X|0} = (0.15)^2\ x\ \exp(-0.15x)$,
$p_{X|0} = (0.15)^2\ (x-20)\ \exp(-0.15(x-20))$

(b) $X \underset{D=0}{\overset{D=1}{\gtrless}} \dfrac{400}{20-\eta} = \eta'$

(c) $P_{\text{FA}} = (0.15\eta' + 1)\exp(-0.15\eta')$,
$P_{\text{FA}} = 20(0.15\eta' - 2)\exp(-0.15\eta')$

(d) $P_{\text{FA}} = \dfrac{1}{5}$

DT7

Three random variables are characterized by the following likelihoods:

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

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

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

Considering the following three hypotheses:

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

obtain:

(a) The Bayesian decision-maker that minimizes the overall risk 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:

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

obtain:

(c) The corresponding ML classifier.

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

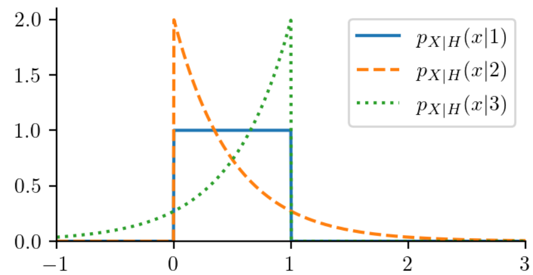
Solution:

(a) Since the costs for all types of error are equal and the cost of hitting is 0, the Bayesian decision-maker is MAP. Also, since the hypothesis are equally likely, the MAP decision-maker is also ML.

According to the statement, the likelihoods of the hypothesis are

$$p_{X|H}(x|1) = p_{X_1}(x)$$
$$p_{X|H}(x|2) = p_{X_2}(x)$$
$$p_{X|H}(x|3) = p_{X_3}(x)$$

and they are represented in the figure.



The cutting point of the likelihoods for hypotheses 1 and 2 is given by the solution of

$$\begin{aligned} p_{X|H}(x|1) &= p_{X|H}(x|2) \\ \Leftrightarrow 1 &= 2\exp(-2x) \\ \Leftrightarrow x &= \frac{\ln(2)}{2} \approx 0.35 \end{aligned}$$

in the same way (an also because of the symmetry of the distributions) it is straightforward to verify that the cut point of the likelihoods for hypotheses 1 and 3 is

$$x = 1 - \frac{\ln(2)}{2} \approx 0.66.$$

Thus, the Bayesian decision rule is

$$D = \left[\begin{array}{ll} 1, & x \in (0.5\ln(2),\, 1 - 0.5\ln(2)) \\ 2, & x \in [0,\, 0.5\ln(2)] \cup [1,\, \infty) \\ 3, & x \in (-\infty,\, 0] \cup [1 - 0.5\ln(2),\, 1] \end{array} \right.$$

(b)

$$\begin{aligned} P\{D = 1|H = 1\} &= P\{x \in (0.5\ln(2),\, 1 - 0.5\ln(2))|H = 1\} \\ &= \int_{0.5\ln(2)}^{1-0.5\ln(2)} 1 \cdot dx \\ &= 1 - \ln(2) \approx 0.31 \end{aligned}$$

Analogamente

$$\begin{aligned} P\{D = 2|H = 2\} &= \int_0^{0.5\ln(2)} 2\exp(-2x)dx + \int_1^{\infty} 2\exp(-2x)dx \\ &= [-\exp(-2x)]_0^{0.5\ln(2)} + [-\exp(-2x)]_1^{\infty} \\ &= \frac{1}{2} + e^{-2} \approx 0.64 \end{aligned}$$

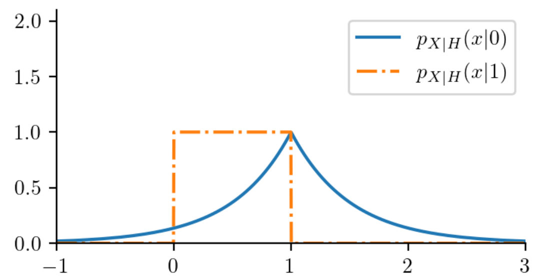
y, por simetría

$$P\{D = 3|H = 3\} = \frac{1}{2} + e^{-2} \approx 0.64$$

(c) The likelihoods for the new hypotheses are:

$$p_{X|H}(x|0) = (p_{X_2} * p_{X_3})(x) = \exp(-2|x-1|)$$
$$p_{X|H}(x|1) = p_{X_1}(x)$$

(where $*$ denotes the convolution operator), and they are represented in the figure



Thus, the ML decision-maker is

$$D = \left[\begin{array}{ll} 0, & x \notin [0,\ 1] \\ 1, & x \in [0,\ 1] \end{array} \right.$$

(d)

$$\begin{aligned} P_{\mathrm{FA}} &= P\{D=1|H=0\} = P\{0 \le x \le 1|H=0\} \\ &= \int_0^1 \exp(2x-2)dx = \frac{1}{2}(1-e^{-2}) \approx 0.4323 \end{aligned}$$

$$P_{\mathrm{M}} = P\{D=0|H=1\} = P\{x \notin [0,1]|H=1\} = 0$$

DT8

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 modeled 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 modeled 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 decision maker 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 decision maker in the light of the new information. Compare this decision maker 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) $P_{\mathrm{FA}} = 8\,\%$ (this is the percentage of prudent clients that will leave the company).
$P_{\mathrm{D}} = 76.2\,\%$ (this is the percentage of reckless clients that are correctly identified as such)

(c) $k \underset{D=0}{\overset{D=1}{\gtrless}} 4.16.$ $P_{\mathrm{FA}} = 0.37\,\%$. $P_{\mathrm{M}} = 37.11\,\%$ and $P_{\mathrm{e}} = 4\,\%$.
For the ML classifier, $P_{\mathrm{e}} = 8.9\,\%$.

DT9

Consider the binary hypotheses

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

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

$$p_N(n) = \frac{1}{s}\left(1 - \frac{|n|}{s}\right), \qquad |n| < s$$

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

(a) Design the MAP classifier.

(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 decision maker but designed under the assumption that N is Gaussian with the same variance of the noise actually present (and zero mean).

Solution:

(a) $$\begin{aligned} D = 0 &: \quad -s < x < \frac{s}{3} \\ D = 1 &: \quad \frac{s}{3} < x < 2s \end{aligned}$$

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

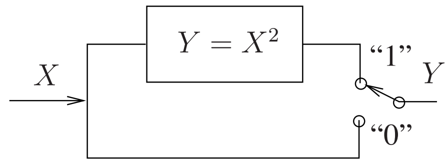
(c) $P_{\mathrm{FA}} = \frac{1}{8}\left(1 + \frac{\ln 2}{3}\right)^2 \approx 0.1894$ (decreases)

$P_{\mathrm{M}} = \frac{1}{8}\left(1 - \frac{\ln 2}{3}\right)^2 \approx 0.0739$ (increases)

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

DT10

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 decision maker to predict which is the position of the switch. The cost policy is $c_{00} = c_{11} = 0$, $c_{10} = 2c_{01}$.

(a) State the decision problem in the usual formulation.

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

(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 $P > 4/5$: $P_{\text{FA}} = 1;\ P_{\text{M}} = 0$

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

1.2. Multi-dimensional observations

DT11

Consider the decision problem given by observation $\mathbf{X} = (x_1, x_2)$, likelihoods

$$\begin{aligned} p_{\mathbf{X}|H}(\mathbf{x}|1) &= x_1 + x_2, \quad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1, \\ p_{\mathbf{X}|H}(\mathbf{x}|0) &= \frac{6}{5}\left(x_1^2 + x_1\right), \quad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1, \end{aligned} \tag{2}$$

and $P_H(1) = \dfrac{6}{11}$.

(a) Determine the decision regions of the MAP classifier, and sketch the corresponding decision boundary on the plane $x_1 - x_2$.

(b) Calculate the missing probability of such classifier.

(c) Obtain the decision regions of the MAP classifier which is based just on variable x_2.

Solution:

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

(b) $P_{\mathrm{M}} = \frac{7}{20}$

(c) $x_2 \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{3}$.

DT12

Consider a binary decision problem characterized by observations $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, and likelihoods

$$
\begin{aligned}
p_{\mathbf{X}|H}(\mathbf{x}|1) &= \exp(-x_1 - x_2), && x_1 \geq 0, \quad x_2 \geq 0 \\
p_{\mathbf{X}|H}(\mathbf{x}|0) &= 2, && x_1 \geq 0, \quad x_2 \geq 0, \quad x_1 + x_2 \leq 1
\end{aligned}
$$

It is also known that $P_H(1) = 4/5$.

(a) Design the ML classifier.

(b) Design the MAP classifier.

(c) Calculate the probability of error of the ML classifier.

(d) Calculate the probability of false alarm of the MAP classifier.

Solution:

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

(b) The MAP classifier decides $D = 0$ if $\ln(2) < x_1 + x_2 < 1$, and $D = 1$ otherwise.

(c) $P_e = (1 - 2e^{-1})/5$

(d) $P_{\mathrm{FA}} = \ln(2)^2$

DT13

Consider the decision problem given by equally likely hypothesis and observations X_1, X_2, X_3 that are independent under any of the hypothesis, and identically distributed, with likelihoods

$$
\begin{aligned}
p_{X_n|H}(x|1) &= \exp(-x), && x \geq 0, \quad n = 1, 2, 3 \\
p_{X_n|H}(x|0) &= 2\exp(-2x), && x \geq 0, \quad n = 1, 2, 3
\end{aligned}
$$

Three MAP classifiers are applied, one for each variable, in such a way that decision D_n of the n-th decision maker is based in observation X_n only (for $n = 1$, 2 or 3).

(a) Determine the false alarm, missing and error probability of each decision maker.

(b) Determine the probability that all decision makers take the same decision, given $H = 0$.

(c) Let $Z = (D_1, D_2, D_3)$ the vector containing the three decisions. Consider the MAP classifier based on observation $\mathbf{Z}$ (that is, the decision maker does not observe X_1, X_2 or X_3, and its only input is $\mathbf{Z}$). Determine its decision when $\mathbf{Z} = (1, 1, 0)$.

Solution:

(a) $P_{\mathrm{M}} = \frac{1}{2}$, $P_{\mathrm{FA}} = \frac{1}{4}$, $P_e = \frac{3}{8}$.

(b) $P = \frac{7}{64}$.

(c) $D_1 + D_2 + D_3 \underset{D=0}{\overset{D=1}{\gtrless}} \frac{3}{2}$.

DT14

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 classifier based on X_1, and calculate its probability of error.

Let DMAP1 be the decision maker 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 decision maker DMAP1.

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

Solution:

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

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

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

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

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

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

DT15

Consider a binary decision problem characterized by:

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

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

(a) Design the corresponding MAP classifier 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 decision maker designed in section (a).

Solution:

(a) $x_1 + x_2 \underset{D=1}{\overset{D=0}{\gtrless}} \ln(4R)$

(b) $t \underset{D=1}{\overset{D=0}{\gtrless}} \ln(4R)$

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

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

DT16

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

$$p_{X_1,X_2|H}(x_1, x_2|0) = K_0 x_1(1 - x_2), \qquad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1$$
$$p_{X_1,X_2|H}(x_1, x_2|1) = K_1 x_1 x_2, \qquad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1$$

(a) Compute 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) Compute 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, X_2 and H; i.e., we observe $Y = X_2 + N$. Design the optimal decision-maker for this situation when the noise pdf is:

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

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

Solution:

(a) Given that the likelihoods are probability density functions, their integrals must be unity

$$\int_0^1 \int_0^1 p_{X_1,X_2|H}(x_1, x_2|0)dx_1dx_2 = 1$$
$$\Rightarrow \quad K_0 \int_0^1 \int_0^1 x_1(1 - x_2)dx_1dx_2 = 1$$
$$\Rightarrow \quad K_0 \int_0^1 x_1dx_1 \int_0^1 (1 - x_2)dx_2 = 1$$
$$\Rightarrow \quad K_0 = 4$$

In a similar way

$$\int_0^1 \int_0^1 p_{X_1,X_2|H}(x_1,x_2|1)dx_1dx_2 = 1 \quad \Rightarrow \quad K_1 = 4$$

(b) The classifier with minimum error probability is MAP, dado por

$$\begin{aligned} P_H(1)p_{X_1,X_2|H}(x_1,x_2|1)dx_1dx_2 &\underset{D=0}{\overset{D=1}{\gtrless}} P_H(0)p_{X_1,X_2|H}(x_1,x_2|0)dx_1dx_2 \\ \Leftrightarrow \quad 4x_1x_2 &\underset{D=0}{\overset{D=1}{\gtrless}} 4x_1(1-x_2) \\ \Leftrightarrow \quad x_2 &\underset{D=0}{\overset{D=1}{\gtrless}} (1-x_2) \\ \Leftrightarrow \quad x_2 &\underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2} \end{aligned}$$

Variable X_1 is irrelevant for the decision and X_2 is a sufficient statistic.

(c) It is easy to see that both likelihood can be factorized as the product of two probability density functions, one per each variable:

$$\begin{aligned} p_{X_1,X_2|H}(x_1,x_2|0) &= (2x_1)\cdot(2(1-x_2)), \qquad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1 \\ p_{X_1,X_2|H}(x_1,x_2|1) &= (2x_1)\cdot(2x_2), \qquad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1 \end{aligned}$$

Therefore, X_1 and X_2 are independent, the marginal distributions are given by these factors

$$\begin{aligned} p_{X_1|H}(x_1|0) &= 2x_1, \qquad 0 \le x_1 \le 1 \\ p_{X_2|H}(x_2|0) &= 2(1-x_2), \qquad 0 \le x_2 \le 1; \\ p_{X_1|H}(x_1|1) &= 2x_1, \qquad 0 \le x_1 \le 1 \\ p_{X_2|H}(x_2|1) &= 2x_2, \qquad 0 \le x_2 \le 1 \end{aligned}$$

(In any case, we can compute the marginal distributions following the general procedure. For instance,

$$p_{X_1|H}(x_1|0) = \int_0^1 p_{X_1,X_2|H}(x_1,x_2|0)dx = 4x_1\int_0^1 (1-x_2)dx_2 = 2x_1, \qquad 0 \le x_1 \le 1$$

which matches the result previously obtained. We could proceed with the other distributions in the same way).

(d)

$$\begin{aligned} P_{\text{FA}} &= P\{D=1|H=0\} = P\left\{X_2 > \frac{1}{2}|H=0\right\} = \int_{\frac{1}{2}}^1 p_{X_2|H}(x_2|0)dx_2 \\ &= \int_{\frac{1}{2}}^1 2(1-x_2)dx_2 = \frac{1}{4} \\ P_{\text{M}} &= P\{D=0|H=1\} = P\left\{X_2 < \frac{1}{2}|H=1\right\} = \int_0^{\frac{1}{2}} p_{X_2|H}(x_2|1)dx_2 \\ &= \int_0^{\frac{1}{2}} 2x_2dx_2 = \frac{1}{4} \\ P_{\text{e}} &= P_H(0)P_{\text{FA}} + P_H(1)P_{\text{M}} = \frac{1}{4} \end{aligned}$$

(e) The MAP decisión-maker based on X_1 and Y is given by

$$p_{X_1,Y|H}(x_1,y|0) \underset{D=0}{\overset{D=1}{\gtrless}} p_{X_1,Y|H}(x_1,y|1)$$

Since Y depends on X_2 and N only, and knowing that these variables are independent of X_1, we conclude that Y is also independent of X_1. Thus, the MAP decision-maker can be written, also, as

$$\begin{aligned} p_{X_1|H}(x_1|1)p_{Y|H}(y|1) &\underset{D=0}{\overset{D=1}{\gtrless}} p_{X_1|H}(x_1|0)p_{Y|H}(y|0) \\ \Leftrightarrow \quad p_{Y|H}(y|1) &\underset{D=0}{\overset{D=1}{\gtrless}} p_{Y|H}(y|0) \end{aligned}$$

Since Y is the sum of two independent random variables, its pdf will be the convolution of the pdf's of each one of them. Therefore,

$$\begin{aligned} p_{Y|H}(y|0) &= p_{X_2|H}(y|0) * p_{N|H}(y|0) = p_{X_2|H}(y|0) * p_N(y) \\ &= \left[\begin{array}{ll} 2(1-y), & y \in [0,1] \\ 0, & y \notin [0,1] \end{array}\right] * \left[\begin{array}{ll} 1, & y \in [0,1] \\ 0, & y \notin [0,1] \end{array}\right] \\ &= \left[\begin{array}{ll} 2y - y^2, & 0 \le y \le 1 \\ 4 - 4y + y^2, & 1 < y \le 2 \\ 0, & y \notin [0,2] \end{array}\right] \end{aligned}$$

In the same way,

$$\begin{aligned} p_{Y|H}(y|1) = p_{X_2|H}(y|0) * p_N(y) &= \left[\begin{array}{ll} 2y, & y \in [0,1] \\ 0, & y \notin [0,1] \end{array}\right] * \left[\begin{array}{ll} 1, & y \in [0,1] \\ 0, & y \notin [0,1] \end{array}\right] \\ &= \left[\begin{array}{ll} y^2, & 0 \le y \le 1 \\ 2y - y^2, & 1 < y \le 2 \\ 0, & y \notin [0,2] \end{array}\right] \end{aligned}$$

Thus, the MAP decision-maker will be

$$\begin{aligned} &\left[\begin{array}{ll} y^2, & 0 \le y \le 1 \\ 2y - y^2, & 1 < y \le 2 \\ 0, & y \notin [0,2] \end{array}\right] \underset{D=0}{\overset{D=1}{\gtrless}} \left[\begin{array}{ll} 2y - y^2, & 0 \le y \le 1 \\ 4 - 4y + y^2, & 1 < y \le 2 \\ 0, & y \notin [0,2] \end{array}\right] \\ \Leftrightarrow &\left[\begin{array}{ll} y^2 \underset{D=0}{\overset{D=1}{\gtrless}} 2y - y^2 & 0 \le y \le 1 \\ 2y - y^2 \underset{D=0}{\overset{D=1}{\gtrless}} 4 - 4y + y^2, & 1 < y \le 2 \end{array}\right] \\ \Leftrightarrow &\left[\begin{array}{ll} y \underset{D=0}{\overset{D=1}{\gtrless}} 1 & 0 \le y \le 1 \\ 0 \underset{D=0}{\overset{D=1}{\gtrless}} (2-y)(1-y), & 1 < y \le 2 \end{array}\right] \\ \Leftrightarrow &\left[\begin{array}{ll} D = 0 & 0 \le y \le 1 \\ D = 1, & 1 < y \le 2 \end{array}\right] \\ \Leftrightarrow & \; y \underset{D=0}{\overset{D=1}{\gtrless}} 1 \end{aligned}$$

(f)

$$P'_{\mathrm{FA}} = P_{D|H}(1|0) = P\{Y > 1|H = 0\} = \int_1^2 p_{Y|H}(y|0)dy = \int_1^2 \left(4 - 4y + y^2\right) dy = \frac{1}{3}$$
$$P'_{\mathrm{M}} = P_{D|H}(0|1) = P\{Y < 1|H = 1\} = \int_0^1 p_{Y|H}(y|1)dy = \int_0^1 2y - y^2 dy = \frac{1}{3}$$
$$P_{\mathrm{e}} = P_H(0)P_{\mathrm{FA}} + P_H(1)P_{\mathrm{M}} = \frac{1}{3}$$

2. Multiclass ML and MAP decision making

DT17

Consider a detection problem with three hypothesis ($H \in \{0, 1, 2\}$), observation $\mathbf{X} = (X_1, X_2)^T \in \mathbb{R}^2$ and likelihoods

$$\begin{aligned} p_{\mathbf{X}|H}(\mathbf{x}|0) &= \frac{1}{\pi}, & & x_1^2 + x_2^2 < 1, \\ p_{\mathbf{X}|H}(\mathbf{x}|1) &= \frac{1}{4}, & & 0 < x_1 < 2, \quad 0 < x_2 < 2, \\ p_{\mathbf{X}|H}(\mathbf{x}|2) &= 1, & & 1 < x_1 < 2, \quad 1 < x_2 < 2, \end{aligned}$$

The a priori probabilities are $P_H(0) = 1/8$, $P_H(1) = 1/2$, and $P_H(2) = 3/8$. Find:

(10 %) (a) The decision regions of the detector that minimizes the probability of error.

Solution: The detector that minimizes the probability of error is the maximum a posteriori detector, which is given by

$$d = \arg\max_h P_{H|\mathbf{X}}(h|\mathbf{x}),$$

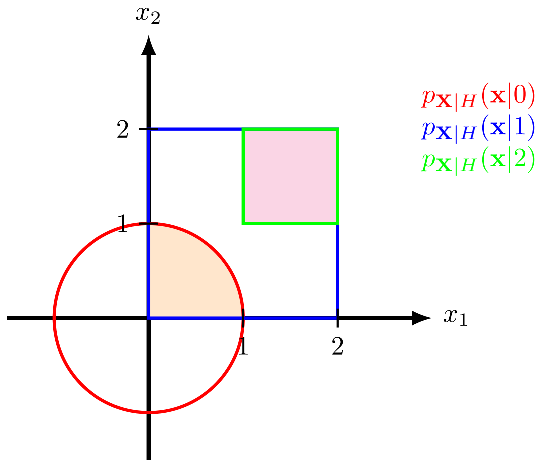
and can be rewritten as

$$d = \arg\max_h p_{\mathbf{X}|H}(\mathbf{x}|h)P_H(h).$$

Hence, the decision regions are

$$\mathcal{X}_d = \{\mathbf{x}|d = \arg\max_h p_{\mathbf{X}|H}(\mathbf{x}|h)P_H(h)\}.$$

To compute these decision regions, it is convenient to plot the supports of the likelihoods as shown in the next figure (each colored line corresponds to the support boundary)



From this plot, we can see that the supports of the likelihoods only overlap in two regions, which are shaded. Then, we only need to see which $p_{X|H}(x|h)P_H(h)$ is larger in these region. In the orange-shaded area, it is easy to see that

$$p_{X|H}(x|0)P_H(0) = \frac{1}{\pi} \cdot \frac{1}{8} < p_{X|H}(x|1)P_H(1) = \frac{1}{4} \cdot \frac{1}{2},$$

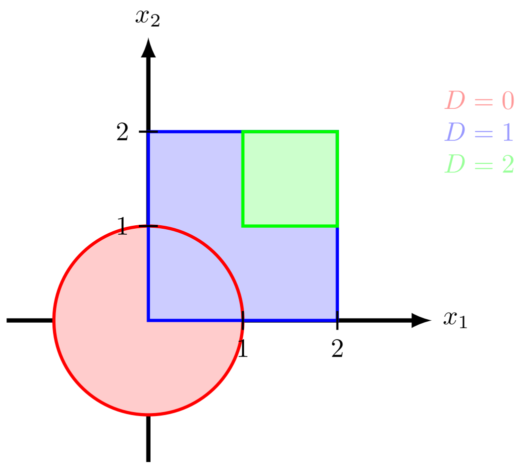
and, therefore, in this region we should decide $D = 1$. In the magenta-shaded area, we have

$$p_{X|H}(x|1)P_H(1) = \frac{1}{4} \cdot \frac{1}{2} < p_{X|H}(x|2)P_H(2) = 1 \cdot \frac{3}{8},$$

which implies that in this region we should decide $D = 2$. Hence, the decision regions are

$$\begin{aligned}
\mathcal{X}_0 &= \{(x_1, x_2)^T \mid x_1^2 + x_2^2 < 1, x_1 \leq 0\} \cup \{(x_1, x_2)^T \mid x_1^2 + x_2^2 < 1, x_1 > 0, x_2 \leq 0\}, \\
\mathcal{X}_1 &= \{(x_1, x_2)^T \mid 0 < x_1 \leq 1, 0 < x_2 < 2\} \cup \{(x_1, x_2)^T \mid 1 < x_1 < 2, 0 < x_2 \leq 1\} \\
\mathcal{X}_2 &= \{(x_1, x_2)^T \mid 1 < x_1 < 2, 1 < x_2 < 2\},
\end{aligned}$$

which are the shaded areas shown in the following figure



(10 %) (b) The conditional probability of correct decision of the derived detector under $H = 0$, $P(D = 0|H = 0)$.

Solution: The requested probability is

$$P(D=0|H=0) = \int_{\mathcal{X}_0} p_{\mathbf{X}|H}(\mathbf{x}|0) d\mathbf{x}.$$

That is, we need to integrate the constant $p_{\mathbf{X}|H}(\mathbf{x}|0) = 1/\pi$ in the region $\mathcal{X}_0$. Since we know that $p_{\mathbf{X}|H}(\mathbf{x}|0) = 1/\pi$ integrates to 1 in the region $\{(x_1, x_2)^T \mid x_1^2 + x_2^2 < 1\}$, and we are leaving out one quarter of that region, $P(D=0|H=0)$ becomes

$$P(D=0|H=0) = \frac{3}{4}.$$

DT18

Consider a classification problem with three hypotheses and likelihoods given by

$$\begin{aligned}
p_{\mathbf{X}|H}(\mathbf{x}|0) &= 1, & 0 \le x_1 \le 1, \quad & 0 \le x_2 \le 1, \\
p_{\mathbf{X}|H}(\mathbf{x}|1) &= \frac{4}{9}, & \frac{1}{2} \le x_1 \le 2, \quad & \frac{1}{2} \le x_2 \le 2, \\
p_{\mathbf{X}|H}(\mathbf{x}|2) &= \frac{1}{4}, & 1 \le x_1 \le 3, \quad & 1 \le x_2 \le 3
\end{aligned}$$

(a) Obtain the decision regions of the maximum likelihood classifier.

(b) Find the condition relating $P_H(1)$ and $P_H(2)$ that guarantees that the MAP classifier selects hypothesis $H=2$ for any x in the domain of $p_{\mathbf{X}|H}(\mathbf{x}|2)$.

(c) Knowing that $P_H(0) = \frac{1}{2}$ and $P_H(2) = 2P_H(1)$, calculate the Probability of error given $\mathbf{x}$ incurred by the MAP classifier.

(d) For the *a priori* probabilities given in the previous section, find the decision regions of the MAP classifier based just on the observation of X_1, and obtain the probability of error of such classifier.

(e) We define a binary classification problem with hypotheses:

$$\begin{aligned}
H' = 0 \quad & \text{if} \quad H \in \{0, 2\} \\
H' = 1 \quad & \text{if} \quad H = 1
\end{aligned}$$

Obtain the decision regions of the MAP classifier based just on observation X_1, and calculate its probability of error.

Solution:

(a) Denoting as $\mathcal{R}_0$, $\mathcal{R}_1$, and $\mathcal{R}_2$ the domains of $p_{\mathbf{X}|H}(\mathbf{x}|0)$, $p_{\mathbf{X}|H}(\mathbf{x}|1)$, and $p_{\mathbf{X}|H}(\mathbf{x}|2)$, respectively, the ML criterion results in the following regions:

$$\begin{cases}
D = 0, & \text{if } \mathbf{x} \in \mathcal{R}_0 \\
D = 1, & \text{if } \mathbf{x} \in \mathcal{R}_1 \setminus \mathcal{R}_0 \text{ (i.e., } \mathbf{x} \in \{\mathbf{z} | \mathbf{z} \in \mathcal{R}_1 \text{ and } \mathbf{z} \notin \mathcal{R}_0\}) \\
D = 2, & \text{if } \mathbf{x} \in \mathcal{R}_2 \setminus \mathcal{R}_1
\end{cases}$$

(b) $P_H(2) > \frac{16}{9} P_H(1)$

(c) MAP criterion:

$$\begin{cases}
D = 0, & \text{if } \mathbf{x} \in \mathcal{R}_0 \\
D = 1, & \text{if } \mathbf{x} \in \{\mathbf{z} | \mathbf{z} \in \mathcal{R}_1 \text{ and } \mathbf{z} \notin \mathcal{R}_0 \text{ and } \mathbf{z} \notin \mathcal{R}_2\}) \\
D = 2, & \text{if } \mathbf{x} \in \mathcal{R}_2
\end{cases}$$

$$P_e(\mathbf{x} \in \mathcal{R}_0 \cap \mathcal{R}_1) = \frac{4}{31};\; P_e(\mathbf{x} \in \mathcal{R}_1 \cap \mathcal{R}_2) = \frac{8}{17};\; P_e(\text{otro } \mathbf{x}) = 0$$

(d) MAP criterion:

$$\begin{cases} D = 0, & \text{if } x_1 \in (0,1) \\ D = 2, & \text{if } x_1 \in (1,3) \end{cases}$$

$$P_e = P_H(1) = \frac{1}{6}$$

(e) MAP criterion: $D' = 0 \quad \forall x_1 \; (\in (0,3))$

$$P_e = P_{H'}(1) = \frac{1}{6}$$

DT19

Consider a detection problem with three hypothesis ($H \in \{0,1,2\}$) and observation $\mathbf{X} = (X_1, X_2)^T \in \mathbb{R}^2$. Moreover, we know that hypotheses are equally likely, also that

$$p_{X_1|X_2,H}(x_1|x_2,0) = p_{X_1|X_2,H}(x_1|x_2,1) = p_{X_1|X_2,H}(x_1|x_2,2),$$

and

$$p_{X_2|H}(x_2|0) = \begin{cases} 1/3, & |x_2| < 1.5, \\ 0, & \text{otherwise}, \end{cases}$$

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

$$p_{X_2|H}(x_2|2) = \begin{cases} -x_2/2, & -2 < x_2 < 0, \\ 0, & \text{otherwise}. \end{cases}$$

Derive the decision regions of the detector that minimizes the probability of error.

Solution: The detector that minimizes the probability of error is the MAP detector, which is given by

$$d = \operatorname*{argmax}_h P_{H|\mathbf{X}}(h|\mathbf{x}),$$

and can be rewritten as

$$d = \operatorname*{argmax}_h p_{\mathbf{X}|H}(\mathbf{x}|h) P_H(h) = \operatorname*{argmax}_h p_{\mathbf{X}|H}(\mathbf{x}|h),$$

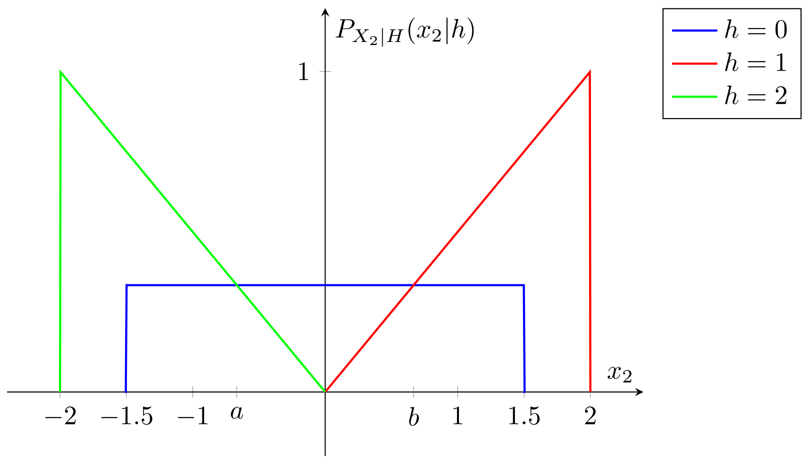
where the last step follows from $P_H(h) = 1/M$. However, we do not have the joint likelihood $p_{\mathbf{X}|H}(\mathbf{x}|h)$, but only $p_{X_1|X_2,H}(x_1|x_2,h)$ and $p_{X_2|H}(x_2|h)$. Using Bayes's theorem, the joint likelihood becomes

$$p_{\mathbf{X}|H}(\mathbf{x}|h) = p_{X_1,X_2|H}(x_1,x_2|h) = p_{X_1|X_2,H}(x_1|x_2,h) p_{X_2|H}(x_2|h),$$

which yields

$$\begin{aligned} d &= \operatorname*{argmax}_h p_{\mathbf{X}|H}(\mathbf{x}|h) = \operatorname*{argmax}_h p_{X_1|X_2,H}(x_1|x_2,h) p_{X_2|H}(x_2|h) \\ &= \operatorname*{argmax}_h p_{X_2|H}(x_2|h), \end{aligned}$$

where, in the last step, we have taken into account that $p_{X_1|X_2,H}(x_1|x_2,h)$ does not depend on h. Then, the decision regions can only depend on x_2 and, to derive them, we need $p_{X_2|H}(x_2|h)$, which is plotted in the following figure



Hence, the decision regions, which are defined as

$$\mathcal{X}_d = \{\mathbf{x} | d = \underset{h}{\operatorname{argmax}}\, p_{X_2|H}(x_2|h)\},$$

are given by

$$\mathcal{X}_0 = \{x_2 \in \mathbb{R} \mid a < x_2 < b\},$$
$$\mathcal{X}_1 = \{x_2 \in \mathbb{R} \mid b \le x_2 < 2\}, \mathcal{X}_2 = \{x_2 \in \mathbb{R} \mid -2 < x_2 \le a\}.$$

Hence, it remains to find the decision boundaries, which are the solution to

$$P_{X_2|H}(b|0) = P_{X_2|H}(b|1) \Rightarrow b = \frac{2}{3},$$
$$P_{X_2|H}(a|0) = P_{X_2|H}(a|2) \Rightarrow a = -\frac{2}{3},$$

yielding

$$\mathcal{X}_0 = \{x_2 \in \mathbb{R} \mid -2/3 < x_2 < 2/3\},$$
$$\mathcal{X}_1 = \{x_2 \in \mathbb{R} \mid 2/3 \le x_2 < 2\},$$
$$\mathcal{X}_2 = \{x_2 \in \mathbb{R} \mid -2 < x_2 \le -2/3\}.$$

DT20

A unidimensional classification problem involves three (*a priori*) equally probable hypotheses, which are characterized by the following likelihoods:

$$p_{X|H}(x|0) = 2\left(1 - 2\left|x - \frac{1}{2}\right|\right), \qquad 0 < x < 1$$
$$p_{X|H}(x|1) = 1, \qquad 0 < x < 1$$
$$p_{X|H}(x|2) = 2x, \qquad 0 < x < 1$$

(a) Determine the decision maker that provides a minimum probability of error.

(b) Discuss whether the previous decision maker is equivalent or not to a second decision maker operating in two stages: The first stage classifier decides, with minimum probability of error, between $H = 0$ and $\{H = 1 \cup H = 2\}$; then, if hypothesis $\{H = 1 \cup H = 2\}$ is selected, a second decision maker is applied to discriminate, again minimizing the probability of error, between $H = 1$ and $H = 2$.

Solution:

(a) $\begin{cases} D = 1: & 0 < x < 1/4 \\ D = 0: & 1/4 < x < 2/3 \\ D = 2 & 2/3 < x < 1 \end{cases}$

(b) $\begin{cases} D = 1: & 0 < x < 1/2 \\ D = 2 & 1/2 < x < 1 \end{cases}$ Different and worse than the classifier in Part (a).

3. Bayesian decision making

DT21

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 classifier.

(b) Obtain the detection probability of the MAP classifier.

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

DT22

Consider a binary decision problem characterized by the following likelihoods:

$$p_{X|H}(x|0) = \exp(-x), \qquad x > 0$$
$$p_{X|H}(x|1) = \sqrt{\tfrac{2}{\pi}}\exp(-\tfrac{x^2}{2}), \qquad x > 0$$

It is also known that $P_H(0) = \sqrt{\frac{2}{\pi}} P_H(1)$ and $c_{00} = c_{11} = 0$, $c_{10} = \exp\left(\frac{1}{2}\right) c_{01}$:

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

(b) Calculate the probability of error of the MAP classifier. Express your result by means of function:

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

(c) Determine the decision regions of the Bayesian classifier with minimum risk.

(d) Calculate the probability of error of the decision maker obtained in the previous subsection.

Solution:

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

$$P_e = \frac{1}{1+\sqrt{\frac{2}{\pi}}}\left[\sqrt{\frac{2}{\pi}}\,(1-\exp(-2)) + 2 - 2F(2)\right]$$

(b) This classifier always decides $D = 0$, $P_e = \dfrac{1}{\sqrt{\frac{2}{\pi}}+1}$

DT23

Consider a unidimensional binary classification problem with likelihoods

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

It is easy to check that, for this particular case, the likelihood ratio test is equivalent to the application of a threshold over x:

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

(a) Obtain the probability of detection and the probability of false alarm as a function of η.

(b) Plot the ROC curve, and place on it the operation point corresponding to the MAP classifier for $P_H(1) = \dfrac{3}{4}$.

(c) Knowing that $c_{00} = c_{11} = 0$, $c_{01} = 1$ y $c_{10} = 3$, express the risk of the classifier as a function of η and find the optimum threshold minimizing such risk.

Solution:

(a) $P_{\text{FA}} = \eta$; $P_{\text{D}} = 1 - (1-\eta)^3$

(b) $P_{\text{D}} = 1 - (1 - P_{\text{FA}})^3$. For the MAP classifier: $P_{\text{FA}} = \frac{2}{3}$ and $P_{\text{D}} = \frac{26}{27}$

(c) $\bar{C} = \frac{3}{4}\left[(1-\eta)^3 + \eta\right]$; $\eta^\star = 1 - \frac{\sqrt{3}}{3}$

DT24

The ship of a certain treasure hunters company is looking for Spanish galleon sunken in the eighteenth century. From sensor measurements taken at a secret location in the ocean, they have obtained a variable X correlated with the presence of the sunken galleon. The likelihoods of hypotheses $H = 1$ ("there is a sunken galleon") and $H = 0$ ("there is not a sunken galleon") are given by

$$p_{X|H}(x|1) = 4x^3, \qquad 0 \geq x \geq 1$$

$$p_{X|H}(x|0) = 4(1-x)^3, \qquad 0 \geq x \geq 1$$

From other evidence, it is estimated that $P_H(1) = 0.1$. Depending on a decision about whether the galleon has been located or not, the captain of the ship will initiate an underwater scanning operation ($D = 1$) or leave the area unexplored ($D = 0$).

It is known that

- The cost of the underwater operation is 100 MM$(million dollars).
- The galleon hides a treasure worth 1000 MM$.

Suppose that other costs and benefits of the operation (e.g, cost of leaving the area, extraction of the treasure, selling the treasure, etc.) are negligible compared to the figures above.

(a) Determine for which values of x the underwater operation should be carried out according to a minimum risk (risk) criterion.

(b) Determine the risk of the decision maker obtained in the previous section.

(c) The cost of the underwater operation is so high that the company would go bankrupt if the Spanish galleon is not found in that location. For this reason, it is preferred to use a decision-maker that maximizes the probability of detection while maintaining bounded the probability of false alarm in $P_{\text{FA}} \leq 10^{-4}$. Determine for which values of x the underwater operation must be addressed in this case.

(d) The treasure hunters company knows that a rival company may have anticipated their plans. They estimate the probability that the sunken galleon no longer contains any treasure is 0.2. Find the risk of the decision-maker obtained in paragraph a) under these conditions.

Solution:

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

(b) $r = -\frac{315}{4} = -78.75$

(c) $x \underset{D=0}{\overset{D=1}{\gtrless}} \frac{9}{10}$

(d) $r' = -60$

DT25

Consider a binary classification problem characterized by $P_H(0) = P_H(1) = 1/2$, $c_{00} = c_{11} = 0$, $c_{01} = 9$, $c_{10} = 8$, and likelihoods

$$p_{X|H}(x|0) = 1 - \frac{x}{2}; \qquad 0 \leq x \leq 2$$

$$p_{X|H}(x|1) = \frac{2}{3}; \qquad 0 \leq x \leq 3/2$$

(a) Consider a generic LRT classifier:

$$\frac{p_{X|H}(x|0)}{p_{X|H}(x|1)} \underset{D=1}{\overset{D=0}{\gtrless}} \eta$$

Illustrate the decision regions of such a classifier for interval $x \in [0, 2]$, explaining how these regions change when modifying the threshold of the test.

(b) Obtain P_{FA} and P_{D} for the LRT classifier, expressing them as a function of η.

(c) Design the ML classifier for the problem under consideration, and obtain its P_{FA} and P_{M}.

(d) Consider now the following threshold classifier:

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

Obtain, as a function of η', the values of P_{FA} and P_{D}. Fill in the following table particularizing your expressions for the indicated values of the threshold.

η'	0	0.5	1	1.5	2
P_{FA}					
P_{D}					

(e) Provide, as a function of η', an expression for the risk of the threshold classifier considered in the previous subsection. Find the value of η' that minimizes such risk.

Solution:

(a) If $x > \frac{3}{2}$ the classifier always decides $D = 0$. If $x > \frac{3}{2}$ the LRT classifier is:

$$x \underset{D=0}{\overset{D=1}{\gtrless}} 2 - \frac{4\eta}{3} = \mu$$

So we can find the following situations:

- If $\eta > \frac{3}{2}$ ($\mu < 0$) it always decides $D = 1$.
- If $\eta < \frac{3}{8}$ ($\mu > \frac{3}{2}$) it always decides $D = 0$.
- If $\frac{3}{2} < \eta < \frac{3}{8}$, the classifier decides $D = 0$ for $0 < x < \mu$ and $D = 1$ for $\mu < x < \frac{3}{2}$

(b)
- If $\eta > \frac{3}{2}$ ($\mu < 0$), $P_{\text{FA}} = P_{\text{D}} = 1$.
- If $\eta < \frac{3}{8}$ ($\mu > \frac{3}{2}$), $P_{\text{FA}} = P_{\text{D}} = 0$.
- If $\frac{3}{2} < \eta < \frac{3}{8}$, $P_{\text{FA}} = \dfrac{15}{16} - \mu + \dfrac{\mu^2}{4}$, $P_{\text{D}} = 1 - \dfrac{2\mu}{3}$

(c) ML Decider ($\eta = 1$ and $\mu = \dfrac{2}{3}$): $P_{\text{M}} = \dfrac{4}{9}$ and $P_{\text{FA}} = \dfrac{55}{144}$

(d) If $0 < \eta' < \frac{3}{2}$: $P_{\text{FA}} = 1 - \eta' + \dfrac{\eta'^2}{4}$ and $P_{\text{D}} = 1 - \dfrac{2\eta'}{3}$

If $\frac{3}{2} < \eta' < 2$: $P_{\text{FA}} = 1 - \eta' + \dfrac{\eta'^2}{4}$ and $P_{\text{D}} = 0$

η'	0	0.5	1	1.5	2
P_{FA}	1	$\frac{9}{16}$	$\frac{1}{4}$	$\frac{1}{16}$	0
P_{D}	1	$\frac{2}{3}$	$\frac{1}{3}$	0	0

(e) $\mathbb{E}\{c_{DH}\} = [\eta' - 2]^2 + 3\eta'$, if $0 < \eta' < \dfrac{3}{2}$

$\mathbb{E}\{c_{DH}\} = [\eta' - 2]^2 + \dfrac{9}{2}$, if $\eta' > \dfrac{3}{2}$

$\eta'^* = \dfrac{1}{2}$

DT26

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

$$p_{X|H}(x|0) = \lambda_0 \exp(-\lambda_0 x) \qquad x \geq 0$$
$$p_{X|H}(x|1) = \lambda_1 \exp(-\lambda_1 x) \qquad x \geq 0$$

where $\lambda_0 = 2\lambda_1$.

(a) Assuming that $P_H(1) = 1/2$ design the classifier that minimizes the risk.

(b) Calculate P_{FA} and P_{M} for the decision maker 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 decision maker as a function of P.

(d) The previous decision maker 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) For the given costs, the optimal decision maker is MAP, that is:

$$\begin{aligned} P_H(1)p_{X|H}(x|1) &\underset{D=0}{\overset{D=1}{\gtrless}} P_H(0)p_{X|H}(x|0) \\ \Leftrightarrow \lambda_1 \exp(-\lambda_1 x) &\underset{D=0}{\overset{D=1}{\gtrless}} 2\lambda_1 \exp(-2\lambda_1 x) \\ \Leftrightarrow -\lambda_1 x &\underset{D=0}{\overset{D=1}{\gtrless}} \ln(2) - 2\lambda_1 x \\ \Leftrightarrow x &\underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{\lambda_1}\ln(2) \end{aligned}$$

(b)

$$\begin{aligned} P_{\mathrm{FA}} &= \int_{\mathcal{X}_1} p_{X|H}(x|0)dx = \int_{\frac{1}{\lambda_1}\ln(2)}^{\infty} 2\lambda_1 \exp(-2\lambda_1 x)\, dx \\ &= \left[-\exp(-2\lambda_1 x)\right]_{\lambda_1^{-1}\ln(2)}^{\infty} = \exp(-2\ln(2)) = 2^{-2} = \frac{1}{4} \\ P_{\mathrm{M}} &= \int_{\mathcal{X}_0} p_{X|H}(x|1)dx = \int_0^{\frac{1}{\lambda_1}\ln(2)} \lambda_1 \exp(-\lambda_1 x)\, dx \\ &= \left[-\exp(-\lambda_1 x)\right]_0^{\lambda_1^{-1}\ln(2)} = 1 - \exp(-\ln(2)) = \frac{1}{2} \end{aligned}$$

(c) For the given costs, the risk is equivalent to the error probability:

$$R = P_{\mathrm{e}} = P_H(0)P_{\mathrm{FA}} + P_H(1)P_{\mathrm{M}} = (1-P)\frac{1}{4} + P\frac{1}{2} = \frac{1}{4}(1+P)$$

(d) Let $E_i = 1$ if there is an error in decision i, and 0 otherwise. Since the observations are independent, so the errors, therefore

$$\begin{aligned} P\{2 \text{ errors}\} &= P\{E_1 = 1, E_2 = 1\} = P\{E_1 = 0\}P\{E_2 = 0\} = \frac{1}{16}(1+P)^2 \\ P\{0 \text{ errors}\} &= P\{E_1 = 0, E_2 = 0\} = \left(1 - \frac{1}{4}(1+P)\right)^2 = \frac{1}{16}(3-P)^2 \\ P\{1 \text{ error}\} &= P\{E_1 = 0, E_2 = 1\} + P\{E_1 = 1, E_2 = 0\} \\ &= 2 \cdot \frac{1}{4}(1+P)\left(1 - \frac{1}{4}(1+P)\right) = \\ &= \frac{1}{8} \cdot (1+P)(3-P) \end{aligned}$$

(e) The risk associated to the two decisions is

$$\begin{aligned} R' &= 0 \cdot P\{0 \text{ errors}\} + 1 \cdot P\{1 \text{ error}\} + 18 \cdot P\{2 \text{ errors}\} \\ &= \frac{1}{8} \cdot (1+P)(3-P) + \frac{18}{16}(1+P)^2 = P^2 + \frac{5}{2}P + \frac{3}{2} \end{aligned}$$

DT27

Consider a binary decision problem described by

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

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 decision maker.

(d) Design the Neyman-Pearson detector 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) $$\begin{array}{ll} D = 1: & |x| < 0.43 \text{ and } |x| > 1 \\ D = 0: & 0.43 < |x| < 1 \end{array}$$

(c) $P_e = 0.184$.

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

DT28

Consider a binary decision problem with equally likely hypothesis, based on the observation of a random variable X with likelihoods

$$p_{X|H}(x|0) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

$$p_{X|H}(x|1) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) Calculate the probability of error of the MAP classifier.

(b) Design the Neyman-Pearson detector satisfying $P_{\text{FA}} \leq 1/4$.

(c) Assume now that H can take a third value $H = 2$. The likelihood of this hypothesis is

$$p_{X|H}(x|2) = \begin{cases} 2(1-x), & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

If all three hypotheses have the same *a priori* probability, and the cost policy is

$$c_{00} = c_{11} = c_{22} = 0, c_{02} = c_{10} = c_{12} = c_{20} = 1, c_{01} = c_{21} = 2$$

where c_{dh} is the cost of decision $D = d$ when $H = h$ is the true hypothesis, obtain the risk of each possible decision as a function of X, i.e., calculate

$$\mathbb{E}\{c_{0,H}|x\},\ \mathbb{E}\{c_{1,H}|x\} \text{ and } \mathbb{E}\{c_{2,H}|x\}$$

(d) Plot the risks calculated in the previous section as a function of the observation x, and determine the decision regions of the minimum risk classifier.

Solution:

(a) $P_{\mathrm{e}} = \dfrac{3}{8}$

(b) $x \underset{D=0}{\overset{D=1}{\gtrless}} \dfrac{3}{4}$

(c) $\mathbb{E}\{c_{0,H}|x\} = \dfrac{2}{3}x + \dfrac{2}{3} \qquad \mathbb{E}\{c_{1,H}|x\} = 1 - \dfrac{2}{3}x \qquad \mathbb{E}\{c_{2,H}|x\} = \dfrac{4}{3}x + \dfrac{1}{3}$

(d) $\begin{cases} D = 2 & 0 \le x \le \dfrac{1}{3} \\ D = 1 & \dfrac{1}{3} \le x \le 1 \end{cases}$

DT29

Consider a binary decision problem where the hypotheses have the same *a priori* probabilities and where the likelihoods are given by

$$p_{X_1|H}(x_1|0) = \begin{cases} 2x_1, & 0 \le x_1 \le 1 \\ 0, & \text{otherwise} \end{cases}$$

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

It is also known that that costs of right decisions is zero, and the cost of errors is one (i.e., $c_{00} = c_{11} = 0$, $c_{10} = c_{01} = 1$).

(a) Obtain the familitiy of LRT decision makers

$$\frac{p_{X_1|H}(x_1|0)}{p_{X_1|H}(x_1|1)} \underset{D=1}{\overset{D=0}{\gtrless}} \eta$$

and calculate their false alarm and missing probabilities, P_{FA} and P_{M}, as functions of η.

(b) Using the result of the previous subsection, find the probabilities of false alarm and missing of the Bayes' classifier, as well as the probability of missing for a Neyman-Pearson detector with $P_{\mathrm{FA}} = 0.01$.

(c) We wish to improve the performance of the Bayes' classifier based on the observation of X_1 by recurring to a second variable X_2 which follows, under each of the hypotheses, the following distribution:

$$p_{X_2|H}(x_2|0) = \begin{cases} 3x_2^2, & 0 \le x_2 \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$p_{X_2|H}(x_2|1) = \begin{cases} 3(1-x_2)^2, & 0 \le x_2 \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Obtain P_{FA} and P_{M} for the Bayes' decision maker based on X_2.

(d) We wish to analyze the overall risk of implementing each of the two Bayes' classifiers considered in the exercise, defined as the sum of the risk of the decision maker, (r_{ϕ_i}), and the cost C_i associated to measuring the observation, X_i, i.e.:

$$R_{\mathrm{TOTi}} = r_{\phi_i} + C_i.$$

Knowing that the cost of measuring X_1 is zero, but the cost of measuring X_2 is given by a constant a, indicate for which values of a each of the two schemes, the one based on X_1 or the one based on X_2, incurs in a smaller overall risk.

Solution:

(a) $x_1 \overset{D=0}{\underset{D=1}{\gtrless}} \dfrac{\eta}{1+\eta} = \eta'$

$P_{\text{FA}} = \eta'^2$ and $P_{\text{M}} = (1-\eta')^2$

(b) Bayes' decision maker: $P_{\text{FA}} = \dfrac{1}{4}$ and $P_{\text{M}} = \dfrac{1}{4}$

N-P decision maker: $P_{\text{FA}} = 0.01$ and $P_{\text{M}} = 0.81$

(c) $x_2 \overset{D=0}{\underset{D=1}{\gtrless}} \dfrac{1}{2}$

$P_{\text{FA}} = \dfrac{1}{8}$ and $P_{\text{M}} = \dfrac{1}{8}$

(d) $R_{\text{TOT1}} = \dfrac{1}{4}$ and $R_{\text{TOT2}} = \dfrac{1}{8} + a$

If $a < \dfrac{1}{8}$, $R_{\text{TOT2}} < R_{\text{TOT1}}$. On the contrary, if $a > \dfrac{1}{8}$, $R_{\text{TOT2}} > R_{\text{TOT1}}$.

DT30

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.

DT31

Consider a multiple-choice exam where each question has 3 possible options and only one is correct. In each question, each student may select as many options as they wish. The scoring policy for these questions is as follows:

- Selecting the correct answer adds 1 point.
- Each incorrect answer subtracts half a point.

To follow an optimal option-marking strategy, we translate the problem into a statistical decision scenario. Since each question has three options and only one is correct, we consider 3 hypotheses: $H = 1$, $H = 2$, and $H = 3$, where $H = h$ means that h is the correct option. Since one can select as many options as desired, there are 7 possible decisions (not selecting any option is equivalent to selecting all of them): $D = 1$, $D = 2$, $D = 3$, $D = (1,2)$, $D = (1,3)$, $D = (2,3)$ and $D = (1,2,3)$ ($D = (i,j)$ means selecting options i and j).

First, we adopt a cost policy such that if M is the score obtained on each question and $c_{d,h}$ is the cost of deciding $D = d$ when the correct option is $H = h$, then the score of each question is computed as:

$$m_{d,h} = 1 - c_{d,h}$$

For example, if we set the cost of getting it right when only the correct option is selected to be 0, $c_{h,h} = 0$, the score obtained if the decision is to select only the correct answer $D = h$ would be 1 point:

$$m_{h,h} = 1 - c_{h,h} = 1 - 0 = 1$$

Analogously, the cost of selecting all options should be $c_{(1,2,3),h} = 1$. This is because by selecting all options, you add one point for selecting the correct option and subtract half a point for each incorrect one, resulting in a score of zero points:

$$m_{(1,2,3),h} = 1 - c_{(1,2,3),h} = 1 - 1 = 0$$

(a) Complete the table that defines the cost policy needed to compute the score of each question as a function of the options selected by the student.

	$H = 1$	$H = 2$	$H = 3$
$D = 1$	0		
$D = 2$		0	
$D = 3$			0
$D = 1, 2$			
$D = 1, 3$			
$D = 2, 3$			
$D = 1, 2, 3$	1	1	1

(b) A student reads a question and estimates that the *a posteriori* probabilities that each of the first two options is correct are: $P_{H|X}(1|x) = 0.5$ and $P_{H|X}(2|x) = 0.3$. Which option will the student select?

(c) In another question, the student is certain that option $H = 3$ is not correct and that option $H = 1$ is more likely than option $H = 2$, so they are deciding between selecting only the first option (deciding $D = 1$) or selecting options 1 and 2 (deciding $D = (1, 2)$). For which values of $p = P_{H|X}(1|x)$ should the student select options 1 and 2 instead of only option 1?

Solution:

(a)

	$H = 1$	$H = 2$	$H = 3$
$D = 1$	0	1.5	1.5
$D = 2$	1.5	0	1.5
$D = 3$	1.5	1.5	0
$D = 1, 2$	0.5	0.5	2
$D = 1, 3$	0.5	2	0.5
$D = 2, 3$	2	0.5	0.5
$D = 1, 2, 3$	1	1	1

(b) The student should select the option with the smallest *a posteriori* expected cost. If we compute the cost of each decision:

$$\mathbb{E}\{c_{dh}|x\} = P_{H|X}(1|x)c_{d,1} + P_{H|X}(2|x)c_{d,2} + P_{H|X}(3|x)c_{d,3}$$

	$\sum_{h=1}^{3} P_{H|X}(h|x)c_{d,h}$	Cost
$D = 1$	$0 \cdot 0.5 + 1.5 \cdot 0.3 + 1.5 \cdot 0.2$	0.75
$D = 2$	$1.5 \cdot 0.5 + 0 \cdot 0.3 + 1.5 \cdot 0.2$	1.05
$D = 3$	$1.5 \cdot 0.5 + 1.5 \cdot 0.3 + 0 \cdot 0.2$	1.2
$D = 1, 2$	$0.5 \cdot 0.5 + 0.5 \cdot 0.3 + 2 \cdot 0.2$	0.8
$D = 1, 3$	$0.5 \cdot 0.5 + 2 \cdot 0.3 + 0.5 \cdot 0.2$	0.95
$D = 2, 3$	$2 \cdot 0.5 + 0.5 \cdot 0.3 + 2 \cdot 0.2$	1.25
$D = 1, 2, 3$	$0.5 + 0.3 + 0.2$	1

The student should select option $D = 1$, which has the lowest expected cost.

(c) Since $H = 3$ is not correct, if $p = P_{H|X}(1|x)$ then $P_{H|X}(2|x) = 1 - p$, and the costs of the two decisions the student is considering are:

- $D = 1$: $0 \cdot p + 1.5 \cdot (1 - p) + 1.5 \cdot 0 = 1.5 \cdot (1 - p)$
- $D = 1, 2$: $0.5 \cdot p + 0.5 \cdot (1 - p) + 2 \cdot 0 = 0.5$

The correct decision is the one with the minimum expected cost, which is equivalent to solving the inequality

$$1.5(1 - p) > 0.5 \Rightarrow p < 2/3$$

That is, if the student believes that $P_{H|X}(1|x) < 2/3$, they should select options 1 and 2, whereas if they believe that $P_{H|X}(1|x) \geq 2/3$, they should select only option 1.

DT32

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

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

$a \geq 1$ 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 risk, 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 risk 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 $a > 1$.

4. Non-Bayesian decision making

DT33

Consider the binary decision problems given by likelihoods

$$p_{X|H}(x|1) = \frac{3}{8}x^2, \qquad 0 \le x \le 2,$$
$$p_{X|H}(x|0) = \frac{3}{4} - \frac{3}{16}x^2 \qquad 0 \le x \le 2$$

(a) Compute the decision regions of the LRT decision maker with threshold η

(b) Compute the Neyman-Pearson detector with $P_{\mathrm{FA}} \le \frac{1}{8}$

(c) Compute the detection probability of the Neyman-Pearson detector

Solution:

(a)

$$\begin{aligned}
p_{X|H}(x|1) &\underset{D=0}{\overset{D=1}{\gtrless}} \eta p_{X|H}(x|0) \\
\Leftrightarrow \quad \frac{3}{8}x^2 &\underset{D=0}{\overset{D=1}{\gtrless}} \eta\left(\frac{3}{4} - \frac{3}{16}x^2\right) \\
\Leftrightarrow \quad \left(\frac{3}{8} + \frac{3}{16}\eta\right)x^2 &\underset{D=0}{\overset{D=1}{\gtrless}} \frac{3}{4}\eta \\
\Leftrightarrow \quad x &\underset{D=0}{\overset{D=1}{\gtrless}} \frac{\frac{3}{4}\eta}{\frac{3}{8} + \frac{3}{16}\eta} \\
\Leftrightarrow \quad x &\underset{D=0}{\overset{D=1}{\gtrless}} \sqrt{\frac{4\eta}{2+\eta}} = \mu
\end{aligned}$$

(b)

$$\begin{aligned}
P_{\mathrm{FA}} &= \int_{\mu}^{2}\left(\frac{3}{4} - \frac{3}{16}x^2\right)dx = \left.\frac{3}{4}x - \frac{1}{16}x^3\right]_{\mu}^{2} \\
&= \frac{3}{2} - \frac{1}{2} - \frac{3}{4}\mu + \frac{1}{16}\mu^3 = 1 - \frac{3}{4}\mu + \frac{1}{16}\mu^3
\end{aligned}$$

The Neyman-Pearson detector is the LRT whose threshold is the solution of

$$1 - \frac{3}{4}\mu + \frac{1}{16}\mu^3 = \frac{1}{8}$$

or, equivalently,

$$\mu^3 - 12\mu + 14 = 0$$

(c)

$$P_{\mathrm{D}} = \int_{\mu}^{2}\frac{3}{8}x^2 dx = \left.\frac{1}{8}x^3\right]_{\mu}^{2} = 1 - \frac{\mu^3}{8}$$

DT34

Consider the decision problem given by the likelihoods:

$$p_{X|H}(x|1) = \frac{n+1}{n}\left(1 - x^n\right), \qquad 0 \le x \le 1$$
$$p_{X|H}(x|0) = (n+1)\, x^n, \qquad 0 \le x \le 1$$

where $n > 0$.

(a) Determine the decision regions of the LRT of threshold $\lambda > 0$.

(b) Determine the false alarm and missing probabilities.

(c) Determine the detection probability of the Neyman-Pearson detector with $P_{\mathrm{FA}} \le 0.1$

Solution:

(a) The LRT decision maker will be given by

$$\begin{aligned} p_{X|H}(x|1) &\underset{D=0}{\overset{D=1}{\gtrless}} \lambda p_{X|H}(x|0) \\ \Leftrightarrow \quad \frac{n+1}{n}\left(1-x^n\right) &\underset{D=0}{\overset{D=1}{\gtrless}} \lambda(n+1)x^n \\ \Leftrightarrow \quad 1 - x^n &\underset{D=0}{\overset{D=1}{\gtrless}} \lambda n x^n \\ \Leftrightarrow \quad x &\underset{D=1}{\overset{D=0}{\gtrless}} \frac{1}{(\lambda n+1)^{\frac{1}{n}}} \end{aligned}$$

Thus, the LRT decision maker decides $D = 0$ for $x \in [\mu, 1]$ and $D = 1$ for $x \in [0, \mu]$, where $\mu = (\lambda n + 1)^{-\frac{1}{n}}$

(b)

$$\begin{aligned} P_{\mathrm{M}} &= P\{D=0|H=1\} = P\{X > \mu|H=1\} = \int_{\mu}^{1} \frac{n+1}{n}\left(1-x^n\right)dx \\ &= \frac{n+1}{n}\left(x - \frac{x^{n+1}}{n+1}\right)_{\mu}^{1} \\ &= 1 - \frac{n+1}{n}\mu + \frac{1}{n}\mu^{n+1} \end{aligned}$$

$$\begin{aligned} P_{\mathrm{FA}} &= P\{D=1|H=0\} = P\{X < \mu|H=0\} = \int_{0}^{\mu} (n+1)x^n dx \\ &= \mu^{n+1} \end{aligned}$$

(c) Taking $P_{\mathrm{FA}} = 10^{-1}$, we get

$$\mu = 10^{-\frac{1}{n+1}}$$

so

$$P_{\mathrm{D}} = 1 - P_{\mathrm{M}} = \frac{n+1}{n}10^{-\frac{1}{n+1}} + \frac{0.1}{n}$$

DT35

Consider a binary decision problem with equally probable hypotheses and likelihoods

$$p_{x|H}(x|1) = x\exp(-x), \qquad x \ge 0 \tag{3}$$

$$p_{x|H}(x|0) = \exp(-x), \qquad x \geq 0 \tag{4}$$

(a) Determine, as a function of η, the decision regions of an LRT decision maker with parameter η.

(b) Obtain, as a function of η, the false alarm and missing probabilities of an LRT decision maker.

(c) Calculate the probability of detection of a Neyman-Pearson detector with $P_{\text{FA}} \leq e^{-1}$.

(d) Obtain the probability of error conditioned on the observation, $P\{D \neq H|x\}$, for an LRT decision maker with parameter η.

Solution:

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

(b) $P_{\text{FA}} = e^{-\eta}$, $P_{\text{FA}} = 1 - (1+\eta)e^{-\eta}$

(c) $P_{\text{D}} = 2e^{-1}$

(d) $P\{D \neq H|x\} = \left[\begin{array}{ll} \frac{x}{1+x}, & \text{if } x < \eta \\ \frac{1}{1+x}, & \text{if } x > \eta \end{array} \right.$

DT36

Consider a binary decision problem characterized by the following likelihoods

$$p_{X|H}(x|0) = n(1-x)^{n-1}, \qquad 0 \leq x \leq 1$$

$$p_{X|H}(x|1) = nx^{n-1}, \qquad 0 \leq x \leq 1$$

with $n \geq 2$ a natural number.

(a) Determine the decision regions of an LRT decision maker, as a function of the threshold of the test, η.

(b) Obtain, as a function of n and η, the false alarm and missing probabilities.

(c) Determine the minimax decision maker.

Solution:

(a) $x \underset{D=0}{\overset{D=1}{\gtrless}} \dfrac{\eta^{\frac{1}{n-1}}}{1+\eta^{\frac{1}{n-1}}}$

(b) $P_{\text{FA}} = \left(\frac{1}{1+\eta^{\frac{1}{n-1}}}\right)^n$

$P_{\text{M}} = \left(\frac{\eta^{\frac{1}{n-1}}}{1+\eta^{\frac{1}{n-1}}}\right)^n$

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

DT37

Consider the binary decision problem given by observation $X \in [0,4]$ and likelihoods

$$p_{X|H}(x|0) = \frac{1}{8}x, \qquad 0 \leq x \leq 4$$

$$p_{X|H}(x|1) = cx\exp(-x), \qquad 0 \leq x \leq 4,$$

where $c = (1 - 5\exp(-4))^{-1}$.

(a) Find the decision regions of the LRT decision maker with parameter η.

(b) Find the values of η for which $P\{D=0\}=1$.

(c) Find the Neyman-Pearson detector with $P_{\text{FA}} \leq 0.1$.

Solution:

(a) $x \underset{D=1}{\overset{D=0}{\gtrless}} \ln\left(\frac{8c}{\eta}\right)$

(b) $\eta \geq 8c$

(c) $x \underset{D=1}{\overset{D=0}{\gtrless}} \sqrt{1.6}$

DT38

Consider the binary decision problem given by equally probable hypothesis and likelihoods

$$p_{X|H}(x|1) = \frac{1}{(1+x)^2}, \qquad x \geq 0$$

$$p_{X|H}(x|0) = \frac{2x}{(1+x)^3}, \qquad x \geq 0$$

(a) Compute the decision regions of the LRT decision maker with parameter η.

(b) Sketch the ROC of LRT approximately.

(c) Compute the decision regions of the minimax decision maker.

(d) Compute the decision regions of the Neyman-Pearson detector with $P_{\text{FA}} \leq \dfrac{1}{16}$

Hint: the probability distribution functions corresponding to the given likelihoods are:

$$F_{X|H}(x|1) = \frac{x}{(1+x)}, \qquad x \geq 0$$

$$F_{X|H}(x|0) = \frac{x^2}{(1+x)^2}, \qquad x \geq 0$$

Solution:

(a) Si $\eta > \dfrac{1}{2}$, $x \underset{D=1}{\overset{D=0}{\gtrless}} \dfrac{1}{2\eta - 1}$.

If $\eta < \frac{1}{2}$, $D = 1$ for any x.

(b) $P_{\text{D}} = \sqrt{P_{\text{FA}}}$

(c) $x \underset{D=1}{\overset{D=0}{\gtrless}} \dfrac{1}{2}(1+\sqrt{5})$

(d) $x \underset{D=1}{\overset{D=0}{\gtrless}} \dfrac{1}{3}$

(e) $x \underset{D=1}{\overset{D=0}{\gtrless}} \dfrac{1}{2}(1+\sqrt{5})$

DT39

Consider a binary decision problem characterized by likelihoods

$$p_{X|H}(x_1, x_2|1) = 4\exp(-2(x_1 + x_2)), \qquad x_1 \geq 0, \quad x_2 \geq 0,$$
$$p_{X|H}(x_1, x_2|0) = 1, \qquad 0 \leq x_1 \leq 1, \quad 0 \leq x_2 \leq 1, \tag{5}$$

(a) Find the decision regions of the ML classifier. Plot your result in the plane $x_1 - x_2$.

(b) Obtain the Neyman-Pearson detector with False Alarm Probability 0.005.

(If you find it useful, consider $\ln(2) = 0.7$).

Solution:

(a) $x_1 \geq 1$ or $x_2 \geq 1 : D = 1$
$x_1, x_2 \leq 1 \quad y \quad x_1 + x_2 < 0.7 : D = 1$
$x_1, x_2 \leq 1 \quad y \quad x_1 + x_2 > 0.7 : D = 0$

(b) $x_1, x_2 \geq 1 : D = 1$
$x_1, x_2 \leq 1 \quad y \quad x_1 + x_2 < \eta : D = 1$
$x_1, x_2 \leq 1 \quad y \quad x_1 + x_2 > \eta : D = 0$
with $\eta = 0.1$

DT40

Consider the binary decision problem given by likelihood functions

$$p_{X|H}(x|1) = 2x, \qquad 0 \leq x \leq 1$$
$$p_{X|H}(x|0) = 1, \qquad 0 \leq x \leq 1$$

(a) Obtain the decision regions of the Neyman-Pearson (NP) decision maker with $P_{\text{FA}} \leq 0.1$.

(b) In this and the following questions, assume that n independent observations are given, $X_1, \ldots, X_n$, all of them driven by the same likelihoods than X. Let $Y = \text{máx}\{X_1, \ldots, X_n\}$. Compute $P\{Y \leq y|H = 1\}$ y $P\{Y \leq y|H = 0\}$, as a function of $y > 0$. (Hints: (I) try to express the probability of event $Y \leq y$ as a function of the probability of events $X_i \leq y$, taking advantage of the independence between observations, (II) the correct answer has the form $P\{Y \leq y|H = h\} = y^{a_h n}$, where a_0 y a_1 are constant values that must be computed).

(c) Compute the likelihood functions $p_{Y|H}(y|1)$ y $p_{Y|H}(y|0)$

(d) Compute the NP decision maker based on Y with $P_{\text{FA}} < 0.19$

(e) Compute the detection probability of the NP decision maker obtained in the previous question.

Solution:

(a) $x \underset{D=0}{\overset{D=1}{\gtrless}} 0.9.$

(b) $P\{Y \geq y|H = 1\} = y^{2n}$, $P\{Y \geq y|H = 0\} = y^n$

(c) $p_{Y|H}(y|1) = 2ny^{2n-1}$, $p_{Y|H}(y|0) = ny^{n-1}$

(d) $x \underset{D=1}{\overset{D=0}{\gtrless}} 0.81^{1/n}$

(e) $P_{\text{D}} = 0.3439$

5. ROC

DT41

Consider the binary decision problems given by likelihoods

$$p_{X|H}(x|1) = \frac{1}{2}x, \qquad 0 \le x \le 2,$$
$$p_{X|H}(x|0) = 2(1-x) \qquad 0 \le x \le 1$$

(a) Compute the decision regions of the LRT with threshold η

(b) Represent, approximately, the ROC of the LRT.

(c) Compute the minimax classifier.

Solution:

(a) For $1 \le x \le 2$, $\quad D = 1$.

For $0 \le x \le 1$,

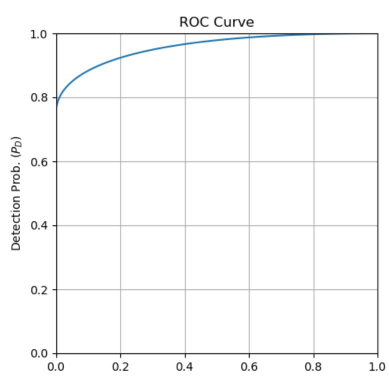
$$\begin{aligned} p_{X|H}(x|1) &\underset{D=0}{\overset{D=1}{\gtrless}} \eta \cdot p_{X|H}(x|0) \\ \Leftrightarrow \quad \frac{1}{2}x &\underset{D=0}{\overset{D=1}{\gtrless}} \eta 2(1-x) \\ \Leftrightarrow \quad (1+4\eta)x &\underset{D=0}{\overset{D=1}{\gtrless}} 4\eta \\ \Leftrightarrow \quad x &\underset{D=0}{\overset{D=1}{\gtrless}} \frac{4\eta}{1+4\eta} = \mu \end{aligned}$$

(b)

$$P_{\text{FA}} = \int_{\mu}^{1} 2(1-x)dx = \left[-(1-x)^2\right]_{\mu}^{1} = (1-\mu)^2$$
$$P_{\text{D}} = \int_{\mu}^{2} \frac{1}{2}x dx = \left[\frac{x^2}{4}\right]_{\mu}^{2} = 1 - \frac{1}{4}\mu^2$$

therefore

$$P_{\text{D}} = 1 - \frac{1}{4}(1 - \sqrt{P_{\text{FA}}})^2$$



(c)

$$\begin{aligned} P_{\mathrm{FA}} = P_{\mathrm{M}} \Leftrightarrow \quad & (1-\mu)^2 = \frac{1}{4}\mu^2 \\ \Leftrightarrow \quad & 1-\mu = \frac{1}{2}\mu \\ \Leftrightarrow \quad & \mu = \frac{2}{3} \end{aligned}$$

DT42

Consider the decision problem given by observation $\mathbf{X} = (x_1, x_2)$, likelihoods

$$\begin{aligned} p_{X|H}(x|1) &= 2x, \quad 0 \le x \le 1, \\ p_{X|H}(x|0) &= 6x(1-x), \quad 0 \le x \le 1, \end{aligned} \tag{6}$$

and $P_H(1) = \dfrac{3}{5}$.

(a) Determine the decision regions of the LRT decision-maker with parameter η.

(b) Compute and plot (approximately) the ROC of the LRT decision-maker.

(c) Compute the coordinates in the ROC of the MAP decision-maker.

Solution:

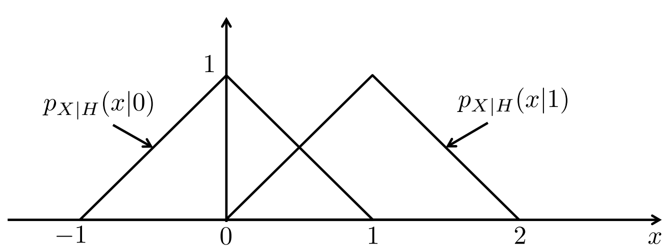
(a) $x \underset{D=0}{\overset{D=1}{\gtrless}} 1 - \dfrac{1}{3\eta} = \mu$

(b) If $\mu \geq 0$, $P_{\mathrm{FA}} = 1 - 3\mu^2 + 2\mu^2$, $P_{\mathrm{D}} = 1 - \mu^2$.
If $\mu < 0$, $P_{\mathrm{FA}} = 1$, $P_{\mathrm{D}} = 1$.

(c) $(P_{\mathrm{FA}}, P_{\mathrm{D}}) = \left(\dfrac{1}{2}, \dfrac{3}{4}\right)$.

DT43

We have a binary decision problem characterized by the likelihoods depicted in the following figure:



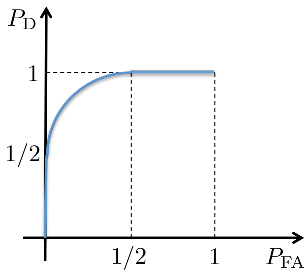
(a) Find an analytical expression for the decision regions of a generic LRT.

(b) Obtain the probabilities of false alarm and missing, and plot the ROC curve.

Solution:

(a) $\begin{cases} -1 \leq x \leq 0 & D = 0 \\ 0 \leq x \leq 1 & x \underset{D=0}{\overset{D=1}{\gtrless}} \dfrac{\eta}{1+\eta} = \nu \\ 1 \leq x \leq 2 & D = 1 \end{cases}$

(b) $\begin{cases} -1 \leq \nu \leq 0 & P_{\mathrm{M}} = 0 \\ 0 < \nu < 1 & P_{\mathrm{FA}} = \frac{1}{2}(1-\nu)^2, \quad P_{\mathrm{M}} = \frac{1}{2}\nu^2 \\ 1 \leq \nu \leq 2 & P_{\mathrm{FA}} = 0 \end{cases}$



This concludes.

DT44

Consider a decision problem characterized by the following likelihoods:

$$p_{X|H}(x|0) = \frac{2}{a^2}x, \qquad 0 < x < a$$
$$p_{X|H}(x|1) = \frac{1}{a}, \qquad 0 < x < a$$

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 decision maker.

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

DT45

Consider the binary decision problem given by observation $X \in \left[0, \frac{\pi}{2}\right]$ and likelihoods

$$p_{X|H}(x|0) = \cos(x), \qquad 0 \leq x \leq \frac{\pi}{2}$$

$$p_{X|H}(x|1) = \sin(x), \qquad 0 \leq x \leq \frac{\pi}{2},$$

(a) Compute the decision regions of an LRT decision maker with parameter $\eta \geq 0$.

(b) Compute the ROC.

(c) Compute the decision regions of a minimax decision maker.

Hint: for any $\alpha \in \mathbb{R}$, $\cos(\arctan(\alpha)) = \dfrac{1}{\sqrt{\alpha^2 + 1}}$

Solution:

(a) $x \underset{D=0}{\overset{D=1}{\gtrless}} \arctan(\eta)$

(b) $P_{\mathrm{D}} = \sqrt{P_{\mathrm{FA}}(1 - P_{\mathrm{FA}})}$

(c) $x \underset{D=0}{\overset{D=1}{\gtrless}} \frac{\pi}{4}$

DT46

Consider a binary decision problem with likelihoods:

$$p_{X|H}(x|1) = \frac{\pi}{2} \sin\left(\frac{\pi}{2}x\right), \qquad 0 < x < 1$$
$$p_{X|H}(x|0) = \frac{\pi}{2} \cos\left(\frac{\pi}{2}x\right), \qquad 0 < x < 1$$

(a) Find the decision regions of an LRT decision maker with parameter η:

$$\frac{p_{X|H}(x|1)}{p_{X|H}(x|0)} \underset{D=0}{\overset{D=1}{\gtrless}} \eta.$$

(b) Provide an approximate plot of the ROC of the LRT classifier.

(c) Indicate which point of the ROC corresponds to the operating point of the ML classifier.

(d) Indicate which point of the ROC corresponds to the operating point of the minimax decision maker.

(e) Indicate which point of the ROC corresponds to the operating point of Neyman Pearson detector with $P_{\mathrm{FA}} \leq 0.4$.

Solution:

(a) $x \underset{D=0}{\overset{D=1}{\gtrless}} \frac{2}{\pi} \arctan(\eta)$,

(b) The ROC is a circumference arch of radius 1, and centered in (1,0).

(c) $(P_{\mathrm{FA}}, P_{\mathrm{M}}) = \left(1 - \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$

(d) $(P_{\mathrm{FA}}, P_{\mathrm{M}}) = \left(1 - \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$

(e) $(P_{\mathrm{FA}}, P_{\mathrm{M}}) = (0.4, 0.8)$

DT47

Consider a binary decision problem characterized by the following likelihoods:

$$p_{X|H}(x|0) = 2\exp(-2x) \quad x > 0$$

$$p_{X|H}(x|1) = 1 \quad 0 < x < 1$$

(a) Obtain the likelihood ratio test for a generic value of threshold η.

$$\frac{p_{X|H}(x|1)}{p_{X|H}(x|0)} \underset{D = 0}{\overset{D = 1}{\gtrless}} \eta$$

(b) Calculate the false alarm and missing probabilities of the previous decision maker as a function of η.

(c) Plot the operating characteristic curve (ROC) of the decision maker, indicating in your representation the operation points of:

- The maximum likelihood decision maker
- The maximum *a posteriori* decision maker, for $P_H(0) = 2P_H(1)$
- The Neyman-Pearson detector with $P_{FA} \leq 0.1$

(d) Consider now a second decision maker consisting on imposing a threshold on the observation x

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

Obtain the false alarm and missing probabilities of this classifier as a function of η_u.

(e) Plot the ROC of the new decision maker, and compare it with the ROC of the LRT decision maker. Which decision scheme (the one based on the LRT or the one based on a threshold over x) offers a better performance? Justify your answer.

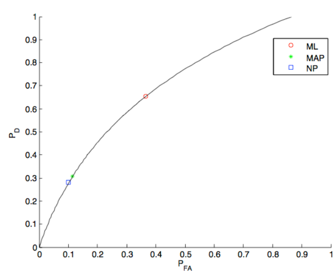
Solution:

(a) $\begin{cases} D = 1: & \eta' < x < 1 \\ D = 0: & 0 < x < \eta' \quad \text{and} \quad x > 1 \end{cases}$

where $\eta' = \frac{1}{2}\ln 2\eta$ and $\eta' > 0$

(b) $P_{\mathrm{FA}} = \begin{cases} \exp(-2\eta') - \exp(-2) & 0 < \eta' < 1 \\ 0 & \eta' > 1 \end{cases}$ $\qquad P_{\mathrm{M}} = \begin{cases} \eta' & 0 < \eta' < 1 \\ 1 & \eta' > 1 \end{cases}$

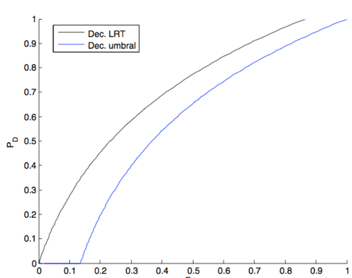
(c)



$$\begin{array}{lll} \mathrm{ML}: & P_{\mathrm{FA}} = \frac{1}{2} - \exp(-2) & P_D = 1 - \frac{1}{2}\ln 2 \\ \mathrm{MAP}: & P_{\mathrm{FA}} = \frac{1}{4} - \exp(-2) & P_D = 1 - \ln 2 \\ \mathrm{N-P}: & P_{\mathrm{FA}} = 0.1 & \end{array}$$

(d) $P_{\mathrm{FA}} = \exp(-2\eta_u)$ $\qquad P_{\mathrm{M}} = \begin{cases} \eta_u & 0 < \eta_u < 1 \\ 1 & \eta_u > 1 \end{cases}$

(e)



As expected, the ROC corresponding to the LRT is above the ROC of the based on thresholding x; we confirm that the LRT decision makers achieve better performance.

DT48

Consider a binary decision problem characterized by the observation vector $\mathbf{X} = (x_1, x_2)$ and likelihoods

$$p_{\mathbf{X}|H}(\mathbf{x}|1) = a^2 \exp\left[-a(x_1 + x_2)\right], \qquad x_1, x_2 > 0,$$
$$p_{\mathbf{X}|H}(\mathbf{x}|0) = b^2 \exp\left[-b(x_1 + x_2)\right], \qquad x_1, x_2 > 0, \tag{7}$$

for b and a two positive constants with $b > a$.

(a) Show that the likelihood ratio test of this problem can be expressed as

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

where we have defined random variable $T = X_1 + X_2$. Obtain the threshold value corresponding to the ML classifier.

(b) Determine the likelihood of both hypotheses expressed in terms of random variable T, i.e., $p_{T|H}(t|i), \quad i = 0, 1$.

(c) Express the missing and false alarm probabilities of the LRT as a function of the threshold η.

(d) Sketch in an approximate manner the OC curve, and place on this curve the operation points corresponding to $\eta = 0$, $\eta = \infty$, the Neyman-Pearson detector with false alarm probability $P_{\rm FA} = 0.1$, and the ML test for the particular case $b = 3a$.

(e) If both hypotheses have the same *a priori* probability, calculate the average risk of the decision maker for the following cost policy: $c_{00} = 0$, $c_{11} = 0.5$, and $c_{01} = c_{10} = 1$. Obtain the threshold value that minimizes this average risk.

Solution:

(a) $\eta_{\rm ML} = \frac{2\ln(b/a)}{b-a}$

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

(c) $P_{\rm M} = 1 - (1 + a\eta)\exp(-a\eta)$ and $P_{\rm FA} = (1 + b\eta)\exp(-b\eta)$

(d) For $\eta = 0$, we have $P_{\rm FA} = P_{\rm D} = 1$; whereas for $\eta = \infty$, we obtain $P_{\rm FA} = P_{\rm D} = 0$

(e) $\bar{r} = \frac{\eta}{2}\left[\frac{a^2}{2}\exp(-a\eta) - b^2 exp(-b\eta)\right]$
$\eta^* = \frac{\ln 2 + 2\ln(b/a)}{b-a}$

DT49

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

$$p_{X|H}(x|0) = 2(1-x), \quad 0 \le x \le 1$$
$$p_{X|H}(x|1) = 2x - 1, \quad \frac{1}{2} \le x \le \frac{3}{2}$$

(a) Find the Bayesian decision-maker for cost policy $c_{00} = c_{11} = 0$, $c_{10} = 4c_{01} > 0$.

(b) Determine the Neyman-Pearson detector with $P_{\rm FA} = 0.04$.

(c) Obtain, as a function of parameter α, the false alarm and detection probabilities for the family of decision-makers with analytic 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 decision-maker changes as a function of the value of such parameter.

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

Solution:

(a) The Bayesian decision-maker for the given costs is given by the decision rule

$$\begin{aligned}
&(c_{01}-c_{11})P_H(1)p_{X|H}(x|1) \underset{D=0}{\overset{D=1}{\gtrless}} (c_{10}-c_{00})P_H(0)p_{X|H}(x|0) \\
\Leftrightarrow\quad & c_{01}2P_H(0)p_{X|H}(x|1) \underset{D=0}{\overset{D=1}{\gtrless}} 4c_{01}P_H(0)p_{X|H}(x|0) \\
\Leftrightarrow\quad & p_{X|H}(x|1) \underset{D=0}{\overset{D=1}{\gtrless}} 2p_{X|H}(x|0) \\
\Leftrightarrow\quad & \begin{bmatrix} D=0, & \text{if } 0 \le x \le \frac{1}{2} \\ 2x-1 \underset{D=0}{\overset{D=1}{\gtrless}} 4(1-x), & \text{if } \frac{1}{2} \le x \le 1 \\ D=1, & \text{if } 1 \le x \le \frac{3}{2} \end{bmatrix} \\
\Leftrightarrow\quad & \begin{bmatrix} D=0, & \text{if } 0 \le x \le \frac{1}{2} \\ x \underset{D=0}{\overset{D=1}{\gtrless}} \frac{5}{6}, & \text{if } \frac{1}{2} \le x \le 1 \\ D=1, & \text{if } \frac{3}{2} \le x \le 1 \end{bmatrix} \\
\Leftrightarrow\quad & x \underset{D=0}{\overset{D=1}{\gtrless}} \frac{5}{6}
\end{aligned}$$

(b) The LRT for threshold $\lambda \ge 0$ has the form

$$\begin{aligned}
p_{X|H}(x|1) \underset{D=0}{\overset{D=1}{\gtrless}} \lambda p_{X|H}(x|0) \Leftrightarrow\quad & \begin{bmatrix} D=0, & \text{if } 0 \le x \le \frac{1}{2} \\ 2x-1 \underset{D=0}{\overset{D=1}{\gtrless}} 2\lambda(1-x), & \text{if } \frac{1}{2} \le x \le 1 \\ D=1, & \text{if } 1 \le x \le \frac{3}{2} \end{bmatrix} \\
\Leftrightarrow\quad & \begin{bmatrix} D=0, & \text{if } 0 \le x \le \frac{1}{2} \\ x \underset{D=0}{\overset{D=1}{\gtrless}} \frac{2\lambda+1}{2(1+\lambda)}, & \text{if } \frac{1}{2} \le x \le 1 \\ D=1, & \text{if } 1 \le x \le \frac{3}{2} \end{bmatrix} \\
\Leftrightarrow\quad & x \underset{D=0}{\overset{D=1}{\gtrless}} \alpha
\end{aligned}$$

where $\alpha = \frac{2\lambda+1}{2(1+\lambda)} \in [\frac{1}{2}, 1]$. The false alarm probability is

$$\begin{aligned}
P_{\mathrm{FA}} &= P_{D|H}(1|0) = P\{x \ge \alpha | H=0\} \\
&= \int_\alpha^\infty p_{X|H}(x|0)dx = \int_\alpha^1 2(1-x)dx \\
&= (1-\alpha)^2
\end{aligned}$$

Taking $P_{\mathrm{FA}} \le 0.04$, we get $(1-\alpha)^2 = 0.04$, thus $\alpha = 0.8$ and the NP decision-maker is

$$x \underset{D=0}{\overset{D=1}{\gtrless}} 0.8$$

(c) According to part (b), the false alarm probability, for any $\alpha \in [0, \frac{3}{2}]$ is

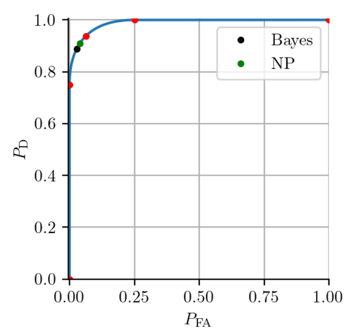
$$P_{\text{FA}} = \left[\begin{array}{ll} (1-\alpha)^2 & 0 < \alpha < 1 \\ 0 & 1 < \alpha < \frac{3}{2} \end{array} \right.$$

In a similar way, the detection probability is

$$\begin{aligned} P_{\text{D}} = P_{D|H}(1|1) = P\{x \leq \alpha | H = 1\} &= \int_{\alpha}^{\infty} p_{X|H}(x|1)dx \\ &= \left[\begin{array}{ll} 1 & 0 < \alpha < \frac{1}{2} \\ \int_{\alpha}^{\frac{3}{2}} (2x-1)dx & \frac{1}{2} < \alpha < \frac{3}{2} \end{array} \right. \\ &= \left[\begin{array}{ll} 1 & 0 < \alpha < \frac{1}{2} \\ 1 - \left(\alpha - \frac{1}{2}\right)^2 & \frac{1}{2} < \alpha < \frac{3}{2} \end{array} \right. \end{aligned}$$

(d) Obtaining some sample points and plotting the ROC in an approximate manner will suffice. For instance:

$$\begin{array}{llll} \alpha = 0 & \Rightarrow & P_{\text{FA}} = 1, & P_{\text{D}} = 1 \\ \alpha = \frac{1}{2} & \Rightarrow & P_{\text{FA}} = \frac{1}{4}, & P_{\text{D}} = 1 \\ \alpha = \frac{3}{4} & \Rightarrow & P_{\text{FA}} = \frac{1}{4}, & P_{\text{D}} = \frac{15}{16} \\ \alpha = 1 & \Rightarrow & P_{\text{FA}} = 0, & P_{\text{D}} = \frac{3}{4} \\ \alpha = \frac{3}{2} & \Rightarrow & P_{\text{FA}} = 0, & P_{\text{D}} = 0 \end{array}$$



(e) The Bayesian decision maker is equivalent $\alpha = \frac{5}{6}$. The NP decision-maker is equivalent to $\alpha = 0.8$. Their respective locations in the ROC are shown in the figure.

DT50

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

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

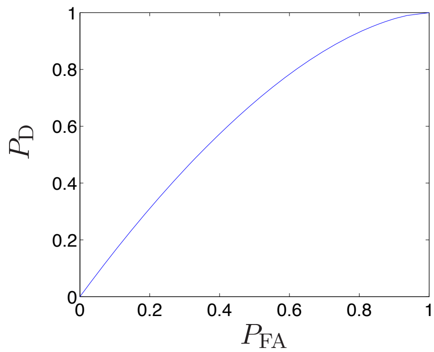
$$p_{X_1,X_2|H}(x_1, x_2|1) = \left\{ \begin{array}{ll} 3\,(x_1 + x_2)\,, & 0 < x_1 < 1 \quad 0 < x_2 < 1 - x_1 \\ 0, & \text{otherwise} \end{array} \right.$$

Consider a generic LRT decision maker with threshold η,

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

(b) The figure 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 classifier, to the decision maker with minimum probability of error, and to the Neyman-Pearson detector with $P_{FA} = 0.3$.

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 classifier: $\eta = 1$, $\eta' = \frac{2}{3}$, $P_{FA} = \frac{5}{9}$.
  MAP classifier: $\eta = \frac{3}{2}$, $\eta' = 1$, $P_{FA} = 0$.
  N-P decision maker: $P_{FA} = 0.3$.

DT51

DT52

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

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

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 classifier 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 decision maker

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

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 decision maker, indicating:

- How the ROC curve would change if the value of p_1 gets larger (keeping condition $p_0 = 1 - p_1$).

Solution:

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

(b) $m_0 = Kp_0 \qquad m_1 = Kp_1$
$v_0 = Kp_0\,(1-p_0) \qquad v_1 = Kp_1\,(1-p_1)$

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

(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.

DT53

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:

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

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 decision makers 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 $\dfrac{p_{X_1,X_2|H}(x_1,x_2|0)}{p_{X_1,X_2|H}(x_1,x_2|1)} \underset{D=1}{\overset{D=0}{\gtrless}} \eta$, the decision maker 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.

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

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

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

(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.

DT54

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 decision maker 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* decision maker 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 decision maker 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 classifier 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 classifier 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 classifier corresponds to an operation point with $0.75 < \eta < 2.5$.

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

DT55

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

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

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 classifier 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 decision maker, 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 decision maker, 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) $t \underset{D=1}{\overset{D=0}{\gtrless}} \dfrac{K\ln a}{a-1}$

(b) Since $T = X_1 + X_2$, and X_1 and X_2 are independent, the pdf of T is the convolution between those of X_1 and X_1, that is

$$\begin{aligned}
p_{T|H}(t|1) &= p_{X|H}(t|1) * p_{X|H}(t|1)\\
&= \int_{-\infty}^{\infty} p_{X|H}(\tau|1)p_{X|H}(t-\tau|1)d\tau, \qquad ,t \geq 0\\
&= \int_{0}^{t} a\exp(-a\tau)a\exp(-a(t-\tau))d\tau \qquad t \geq 0\\
&= a^2\exp(-at)\int_0^t dt \qquad t \geq 0\\
&= a^2 t\exp(-at) \qquad t \geq 0
\end{aligned}$$

The conditional distribution for $H = 0$ is formally equivalent and can be computed from the above result by taking $a = 0$.

$$p_{T|H}(t|0) = t\exp(-t) \qquad t \geq 0$$

(c)

$$P_{\mathrm{FA}} = \int_{\mathcal{X}_1} p_{T|H}(t|0)dt = \int_0^{\eta} t\exp(-t)dt = 1-(\eta+1)\exp(-\eta)$$
$$P_{\mathrm{M}} = \int_{\mathcal{X}_0} p_{T|H}(t|0)dt = \int_{\eta}^{\infty} a^2 t\exp(-at)dt = (a\eta+1)\exp(-a\eta)$$

(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.

6. Gaussian models

DT56

Consider the binary decision problem given by observation $X \in \mathbb{R}$, $P_H(1) = q$ and likelihoods

$$p_{X|H}(x|0) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{8}$$
$$p_{X|H}(x|1) = \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{x^2}{8}\right) \tag{9}$$

(a) Compute the.MAP classifier

(b) Compute the probability of error. Express the result as a function of

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

Solution:

(a) The MAP classifier is given by

$$\begin{aligned}
P_H(1)p_{X|H}(x|1) &\underset{D=0}{\overset{D=1}{\gtrless}} P_H(0)p_{X|H}(x|0) \\
\Leftrightarrow\quad \frac{q}{\sqrt{8\pi}}\exp\left(-\frac{x^2}{8}\right) &\underset{D=0}{\overset{D=1}{\gtrless}} \frac{1-q}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) \\
\Leftrightarrow\quad -\frac{x^2}{8} &\underset{D=0}{\overset{D=1}{\gtrless}} -\frac{x^2}{2} + \ln\left(\frac{2(1-q)}{q}\right) \\
\Leftrightarrow\quad \left(1-\frac{1}{4}\right)x^2 &\underset{D=0}{\overset{D=1}{\gtrless}} 2\ln\left(\frac{2(1-q)}{q}\right) \\
\Leftrightarrow\quad |x| &\underset{D=0}{\overset{D=1}{\gtrless}} \sqrt{\frac{8}{3}\ln\left(\frac{2(1-q)}{q}\right)} = \mu
\end{aligned}$$

(b)

$$P_{\text{FA}} = 2\int_{-\infty}^{-\mu} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = 2F(-\mu)$$
$$P_{\text{M}} = 1 - 2\int_{-\infty}^{-\mu} \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{x^2}{8}\right) dx = 1 - 2F\left(-\frac{\mu}{2}\right)$$
$$P_{\text{e}} = q\left(1 - 2F\left(-\frac{\mu}{2}\right)\right) + (1-q)2F(-\mu)$$

DT57

Consider the binary decision problem given by observation $X \in \mathbb{R}$, $P_H(1) = \frac{1}{4}$, likelihoods

$$p_{X|H}(x|1) = \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{(x-4)^2}{8}\right)$$
$$p_{X|H}(x|0) = \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{x^2}{8}\right)$$

and cost matrix

$$\mathbf{C} = \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 4 & 1 \end{pmatrix}$$

(a) Compute the Bayesian classifier that minimizes the risk

(b) Compute the risk of the Bayesian classifier. Express the result using function

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

Solution:

(a) The Bayesian classifier is

$$(c_{01} - c_{11}) P_H(1) p_{X|H}(x|1) \underset{D=0}{\overset{D=1}{\gtrless}} (c_{10} - c_{00}) P_H(0) p_{X|H}(x|0)$$
$$\Leftrightarrow \quad \frac{1}{4\sqrt{8\pi}} \exp\left(-\frac{(x-4)^2}{8}\right) \underset{D=0}{\overset{D=1}{\gtrless}} \frac{3}{\sqrt{8\pi}} \exp\left(-\frac{x^2}{8}\right)$$
$$\Leftrightarrow \quad \frac{x^2}{8} - \frac{(x-4)^2}{8} \underset{D=0}{\overset{D=1}{\gtrless}} \ln(12)$$
$$\Leftrightarrow \quad 8x - 16 \underset{D=0}{\overset{D=1}{\gtrless}} 8\ln(12)$$
$$\Leftrightarrow \quad x \underset{D=0}{\overset{D=1}{\gtrless}} 2 + \ln(12) = \mu$$

(b)

$$
\begin{aligned}
P_{\mathrm{FA}} = P\{D=1|H=0\} &= \int_{\mu}^{\infty} \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{x^2}{8}\right) dx \\
&= \int_{\frac{\mu}{2}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt = F\left(-\frac{\mu}{2}\right) \\
P_{\mathrm{M}} &= \int_{-\infty}^{\mu} \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{(x-4)^2}{8}\right) dx \\
&= \int_{-\infty}^{\frac{\mu-4}{2}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dx = F\left(\frac{\mu-4}{2}\right)
\end{aligned}
$$

therefore, the risk is

$$
\begin{aligned}
R_{\phi} &= (c_{00}P_H(0) + c_{11}P_H(1)) + (c_{10} - c_{00})P_H(0)P_{\mathrm{FA}} + (c_{01} - c_{11})P_H(1)P_{\mathrm{M}} \\
&= \frac{1}{4} + 3F\left(-\frac{\mu}{2}\right) + \frac{1}{4}F\left(\frac{\mu-4}{2}\right)
\end{aligned}
$$

DT58

Consider a binary decision problem with likelihoods

$$
\begin{aligned}
p_{\mathbf{X}|H}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid 0\right) &\sim G\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right), \\
p_{\mathbf{X}|H}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid 1\right) &\sim G\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)
\end{aligned}
$$

(a) Obtain the ML classifier, and check that the knowledge of $T = X_1 + X_2$ is sufficient for taking decisions.

(b) Obtain the conditional probability density functions $p_{T|H}(t|0)$ and $p_{T|H}(t|1)$.

(c) Calculate the false alarm and missing probabilities using the likelihoods of the previous section. Express your result by means of function

$$
F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du
$$

Solution:

(a) The ML classifier is given by

$$
\begin{aligned}
& p_{\mathbf{X}|H}(\mathbf{x} \mid 1) \underset{D=0}{\overset{D=1}{\gtrless}} p_{\mathbf{X}|H}(\mathbf{x} \mid 1) \\
\Leftrightarrow\quad & \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left(\mathbf{x} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^{\top}\left(\mathbf{x} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)\right) \underset{D=0}{\overset{D=1}{\gtrless}} \frac{1}{2\pi} \exp\left(-\frac{1}{2}\mathbf{x}^{\top}\mathbf{x}\right) \\
\Leftrightarrow\quad & -\left(\mathbf{x} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^{\top}\left(\mathbf{x} - \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) \underset{D=0}{\overset{D=1}{\gtrless}} -\mathbf{x}^{\top}\mathbf{x} \\
\Leftrightarrow\quad & 2\begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\top} \mathbf{x} \underset{D=0}{\overset{D=1}{\gtrless}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\top} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
\Leftrightarrow\quad & x_1 + x_2 \underset{D=0}{\overset{D=1}{\gtrless}} 1
\end{aligned}
$$

Defining $t = x_1 + x_2$ we get the equivalent test

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

(b) Since T is a sum of Gaussian random variables (for any hypothesis), it is Gaussian, too:

$$p_{T|H}(t|0) \sim G(m_0, v_0)$$
$$p_{T|H}(t|1) \sim G(m_1, v_1)$$

where

$$\begin{aligned}
m_0 &= \mathbb{E}\{T \mid H=0\} = \mathbb{E}\{X_1 \mid H=0\} + \mathbb{E}\{X_2 \mid H=0\} = 0 \\
m_1 &= \mathbb{E}\{T \mid H=1\} = \mathbb{E}\{X_1 \mid H=1\} + \mathbb{E}\{X_2 \mid H=1\} = 2 \\
v_0 &= \mathbb{E}\{(T-m_0)^2 \mid H=0\} = \mathbb{E}\{X_1^2 + X_2^2 + 2X_1X_2 \mid H=0\} = 1+1+0 = 2 \\
v_1 &= \mathbb{E}\{(T-m_1)^2 \mid H=1\} = \mathbb{E}\{(X_1-1) + (X_2-1))^2 \mid H=1\} \\
&= \mathbb{E}\{(X_1-1)^2 + (X_2-1)^2 + 2(X_1-1)(X_2-1) \mid H=1\} = 1+1+0 = 1
\end{aligned}$$

Therefore,

$$p_{T|H}(t|0) \sim G(0, 2)$$
$$p_{T|H}(t|1) \sim G(2, 2)$$

(c)

$$\begin{aligned}
P_{\mathrm{M}} &= \int_{-\infty}^{1} p_{T|H}(t \mid 1)dt = \int_{-\infty}^{1} \frac{1}{\sqrt{4\pi}} \exp\left(-\frac{1}{4}(t-2)^2\right) dt \\
&= \int_{-\infty}^{\frac{1}{\sqrt{2}}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{4}z^2\right) dz = F\left(-\frac{1}{\sqrt{2}}\right) \\
P_{\mathrm{FA}} &= \int_{1}^{\infty} p_{T|H}(t \mid 0)dt = \int_{1}^{\infty} \frac{1}{\sqrt{4\pi}} \exp\left(-\frac{1}{4}t^2\right) dt \\
&= \int_{\frac{1}{\sqrt{2}}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt = 1 - F\left(\frac{1}{\sqrt{2}}\right)
\end{aligned}$$

Note that, since $F(-z) = 1 - F(z)$, for any $z \in \mathbb{R}$, we have $P_{\mathrm{FA}} = P_{\mathrm{M}}$.

DT59

A binary decision problem is characterized by Gaussian likelihoods:

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

where $|\rho| < 1$.

(a) Design the maximum likelihood decision maker.

(b) Let $Z = X_1 - \rho X_2$ be a new random variable. Obtain the likelihoods of hypotheses $H=0$ and $H=1$ in terms of the new random variable, $p_{Z|H}(z|0)$ and $p_{Z|H}(z|1)$.

(c) Considering the results of the previous sections, calculate the False Alarm and Missing probabilities of the decision maker designed in Section (a); express your results in terms of function

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

Solution:

(a) $X_1 - \rho X_2 \underset{D=0}{\overset{D=1}{\gtrless}} 0$

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

(c) $P_{\text{FA}} = P_{\text{M}} = F\left(-\frac{1}{\sqrt{1-\rho^2}}\right)$

DT60

Consider the family of Gaussian and statistically independent random variables: $\{U_n,\ n = 0,1,2,3\}$ of means $m_n = n^2$ and variances $v_n = 1 + n$, and the binary decision problem given by the observation $X \in \mathbb{R}$ and hypotheses:

$$\begin{aligned} H = 1: &\quad X = U_0 + U_3 \\ H = 0: &\quad X = U_1 + U_2 \end{aligned}$$

where $P_H(1) = 0.8$.

(a) Determine the ML classifier based on X.

(b) Determine the probability of error of the decision maker. Express the result in terms of the function:

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

Solution:

(a) Under both hypotheses, we have Gaussian likelihoods, of means

$$\begin{aligned} m_1 &= \mathbb{E}\{X|H=1\} = 9 \\ m_0 &= \mathbb{E}\{X|H=0\} = 5 \end{aligned}$$

and variances $v_0 = v_1 = 5$. So

$$\begin{aligned} p_{X|H}(x|1) &= \frac{1}{\sqrt{10\pi}} \exp\left(-\frac{1}{10}(x-9)^2\right) \\ p_{X|H}(x|0) &= \frac{1}{\sqrt{10\pi}} \exp\left(-\frac{1}{10}(x-5)^2\right) \end{aligned}$$

Therefore, the ML classifier will have the form

$$\begin{aligned} & p_{X|H}(x|1) \underset{D=0}{\overset{D=1}{\gtrless}} p_{X|H}(x|0) \\ \Leftrightarrow \quad & -\frac{1}{10}(x-9)^2 \underset{D=0}{\overset{D=1}{\gtrless}} -\frac{1}{10}(x-5)^2 \\ \Leftrightarrow \quad & (x-5)^2 - (x-9)^2 \underset{D=0}{\overset{D=1}{\gtrless}} 0 \\ \Leftrightarrow \quad & x \underset{D=0}{\overset{D=1}{\gtrless}} 7 \end{aligned}$$

(b) The missing probability will be

$$P_{\mathrm{M}} = P\{D=0|H=1\} = P\{X<7|H=0\} = \int_{-\infty}^{7} \frac{1}{\sqrt{10\pi}} \exp\left(-\frac{1}{10}(x-9)^2\right) dx$$
$$= F\left(\frac{-2}{\sqrt{5}}\right)$$

analogously

$$P_{\mathrm{FA}} = P\{D=1|H=0\} = \int_{7}^{\infty} \frac{1}{\sqrt{10\pi}} \exp\left(-\frac{1}{10}(x-5)^2\right) dx = 1 - F\left(\frac{2}{\sqrt{5}}\right) = P_{\mathrm{M}}$$

So,

$$P_e = P_{\mathrm{FA}} = P_{\mathrm{M}} = F\left(\frac{-2}{\sqrt{5}}\right)$$

DT61

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 classifier, 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.

DT62

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' decision maker 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 decision maker, 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) The decision-maker with minimum error probability is MAP, so its decision rule is

$$
\begin{aligned}
&P_H(1)p_{X_1,X_2|H}(x_1,x_2|1) \underset{D=0}{\overset{D=1}{\gtrless}} P_H(0)p_{X_1,X_2|H}(x_1,x_2|0)\\
&\Leftrightarrow \exp\left(-\frac{1}{2}\begin{pmatrix} x_1-m \\ x_2-m \end{pmatrix}^\intercal \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1-m \\ x_2-m \end{pmatrix}\right) \underset{D=0}{\overset{D=1}{\gtrless}} \exp\left(-\frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^\intercal \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right)\\
&\Leftrightarrow -\frac{1}{2(1-\rho^2)}\begin{pmatrix} x_1-m \\ x_2-m \end{pmatrix}^\intercal \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} x_1-m \\ x_2-m \end{pmatrix} \underset{D=0}{\overset{D=1}{\gtrless}} -\frac{1}{2(1-\rho^2)}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^\intercal \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\\
&\Leftrightarrow -(x_1-m)^2 - (x_2-m)^2 + 2\rho(x_1-m)(x_2-m) \underset{D=0}{\overset{D=1}{\gtrless}} -x_1^2 - x_2^2 + 2\rho x_1 x_2\\
&\Leftrightarrow x_1 + x_2 \underset{D=0}{\overset{D=1}{\gtrless}} m
\end{aligned}
$$

(b) Substituting into the decision rule, we obtain the equivalent rule $z \underset{D=0}{\overset{D=1}{\gtrless}} m$.

Since Z is the sum of two Gaussian r.v. it is Gaussian too, with conditional means

$$
\begin{aligned}
\mathbb{E}\{Z|1\} &= \mathbb{E}\{X_1|1\} + \mathbb{E}\{X_2|1\} = 2m\\
\mathbb{E}\{Z|0\} &= \mathbb{E}\{X_1|0\} + \mathbb{E}\{X_2|0\} = 0
\end{aligned}
$$

and conditional variances

$$
\begin{aligned}
\mathrm{Var}\{Z|1\} &= \mathbb{E}\{(X_1 - m + X_2 - m)^2|1\}\\
&= \mathrm{Var}\{X_1|1\} + \mathrm{Var}\{X_2|1\} + 2\mathbb{E}\{(X_1-m)(X_2-m)\}\\
&= 2(1+\rho)\\
\mathrm{Var}\{Z|0\} &= 2(1+\rho)
\end{aligned}
$$

Thus

$$
\begin{aligned}
Z|1 &\sim G\left(2m, 2(1+\rho)\right)\\
Z|0 &\sim G\left(0, 2(1+\rho)\right)
\end{aligned}
$$

(c)

$$P_{\mathrm{FA}} = P\{D=1|H=0\} = P\{Z > m|H=0\} = 1 - F\left(\frac{m}{\sqrt{2(1+\rho)}}\right) \tag{10}$$

$$P_{\mathrm{M}} = P\{D=0|H=1\} = P\{Z < m|H=1\} = F\left(\frac{m-2m}{\sqrt{2(1+\rho)}}\right) = P_{\mathrm{FA}} \tag{11}$$

$$P_{\mathrm{e}} = P_H(0)P_{\mathrm{FA}} + P_H(1)P_{\mathrm{M}} = P_{\mathrm{FA}} = P_{\mathrm{M}} \tag{12}$$

(d)
- Si $\rho = -1$: $P_e = 0$
- Si $\rho = 0$: $P_e = 1 - F\left(\frac{m}{\sqrt{2}}\right)$
- Si $\rho = 1$: $P_e = 1 - F\left(\frac{m}{2}\right)$

DT63

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 $[1,0]^T$.

DT64

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

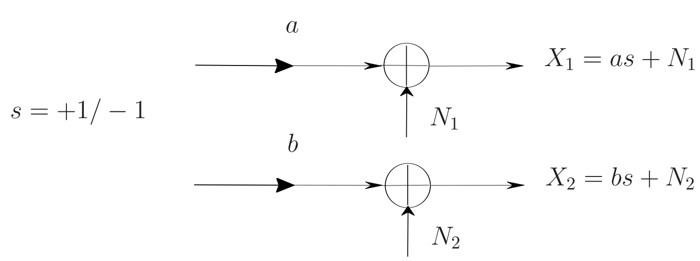
(c) $P_{\text{FA}} = P_{\text{M}} = F\left(\frac{||\mathbf{m}||_2}{2\sqrt{v}}\right)$, which goes to 0 as N increases towards infinity.

(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 $\frac{v_n}{||\mathbf{m}||_2^2}$.

DT65

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 decision maker 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 classifier.

(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 decision maker 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) $P_e = F\left(-\sqrt{a^2 + b^2}\right)$

DT66

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

DT67

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

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

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.

DT68

Consider a binary decision problem with equally probable hypotheses and observations characterized by

$$\begin{aligned} H = 0: &\quad X = N_0 \\ H = 1: &\quad X = a + N_1 \end{aligned}$$

where a is a known constant and N_0 and N_1 are Gaussian random variables with distributions $N_0 \sim G(0, v_0)$ and $N_1 \sim G(0, v_1)$, respectively.

(a) For $a > 0$, provide plots to illustrate the decision regions that would be obtained when $v_0 > v_1$, $v_0 < v_1$, and $v_0 = v_1$.

(b) Consider during the rest of the exercise that $a = 0$, $v_0 = 1$, and $v_1 = 2$. Obtain the decision rule that minimizes the probability of error of the decision maker.

(c) Calculate the incurred probabilities of false alarm and detection when using the previous decision maker. Express your results by means of function

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du$$

(d) Using an approximate representation of the ROC of LRT decision makers

$$\frac{p_{X|H}(x|1)}{p_{X|H}(x|0)} \underset{D=0}{\overset{D=1}{\gtrless}} \eta$$

indicate how would the decision maker operation point move when:

- threshold η is increased.
- the *a priori* probability of hypothesis $H = 1$ grows.

Solution:

(a) If $v_0 = v_1$ we would obtain a classifier based on a single threshold over x; otherwise, there would be two thresholds.

(b) $|x| \underset{D=0}{\overset{D=1}{\gtrless}} \sqrt{2\ln 2} = x_u$

(c) $P_{\mathrm{FA}} = 2F\left(-x_u\right), \qquad P_{\mathrm{D}} = 2F\left(\frac{-x_u}{\sqrt{2}}\right)$

(d) If η increases, then P_{FA} and P_{D} decrease. If $P_H(1)$ increases with η constant, the operation point does not change.

DT69

Let X be a measurement of the instantaneous voltage at a circuit node. Under the null hypothesis $H = 0$, the voltage at the node is characterized by a Gaussian noise with mean 0 and variance v. Under hypothesis $H = 1$, in such node there exists just a sinusoidal signal with mean zero and amplitude $\sqrt{v}$. Since the frequency of the signal is not known, we have that under $H = 1$ the measurement can be probabilistically modeled as $X = \sqrt{v}\cos\Phi$, with Φ a random variable with uniform distribution between 0 and 2π.

(a) Calculate the likelihoods of both hypotheses.

(b) Find the maximum likelihood test to discriminate among them.

(c) Use function $h(a) = a - \log(1-a)$ to express the ML classifier, and calculate the decision regions as functions of v and $h^{-1}(\cdot)$.

(d) Obtain the probability of false alarm of such decision maker as a function of $h^{-1}(\cdot)$ and $Q(z)$.

Hints:

$$\frac{d\cos u}{du} = -\sin u \qquad \frac{d\arccos u}{du} = \frac{-1}{\sqrt{1-u^2}} \qquad \frac{d\sin u}{du} = \cos u \qquad \frac{d\arcsin u}{du} = \frac{1}{\sqrt{1+u^2}}$$

Assume as known function $Q(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\,du$.

Assume as known function $a = h^{-1}(\cdot)$ (reciprocal function of $h(\cdot)$).

Solution:

(a) $p_{X|H}(x|0) = G(x|0,v)$, $p_{X|H}(x|1) = \frac{1}{\pi\sqrt{v-x^2}} \quad \forall_{x\in[-\sqrt{v},\sqrt{v}]}$

(b) $h(\frac{x^2}{v}) \underset{D=0}{\overset{D=1}{\gtrless}} \log\frac{\pi}{2}$ if $x^2 < v$; $D = 0$ otherwise.

(c) $h^{-1}(\log \frac{\pi}{2}) = \frac{x^2}{v} = 0.2126 \approx 0.21 \Rightarrow$
$D_0 : -\infty < x < -\sqrt{v} \quad \cup \quad -\sqrt{0.21v} < x < +\sqrt{0.21v} \quad \cup \quad +\sqrt{v} < x < +\infty$
$D_1 : -\sqrt{v} < x < -\sqrt{0.21v} \quad \cup \quad +\sqrt{0.21v} < x < +\sqrt{v}$

(d) $P_{\mathrm{FA}} = 2(Q(1) - Q(\sqrt{0.21}))$

DT70

Variables Z_1 and Z_2 can only take values, $-m$ or m. Under hypothesis $H = 0$, both variables take the same value. This yields two possible configurations under this hypothesis, both appearing with the same probability. Under hypothesis $H = 1$, both variables take different values. This yields two possible configurations under this hypothesis, both appearing with the same probability. Hypotheses $H = 0$ and $H = 1$ are equiprobable.

Variables z_1 and Z_2 cannot be observed directly. However, we can observe X_1 and X_2, which are noisy measurements of Z_1 and Z_2 respectively, using a device that adds independent zero-mean Gaussian noise of variance one, i. e., $X_i = Z_i + N_i$, where N_1 and N_2 are independent and also independent of Z_1 and Z_2.

(a) Compute $P_{Z_1,Z_2|H}(z_1, z_2|h)$ for all possible values of z_1, z_2 and i.

(b) Compute $P_{X_1,X_2|Z_1,Z_2}(x_1, x_2|z_1, z_2)$.

(c) Without making any computations, reason whether

$$P_{X_1,X_2|Z_1,Z_2}(x_1, x_2|z_1, z_2)$$

is different or identical to $P_{X_1,X_2|Z_1,Z_2,H}(x_1, x_2|z_1, z_2, h)$.

(d) Compute the likelihoods of both hypotheses, $P_{X_1,X_2|H}(x_1, x_2|0)$ and $p_{X_1,X_2|H}(x_1, x_2|1)$.

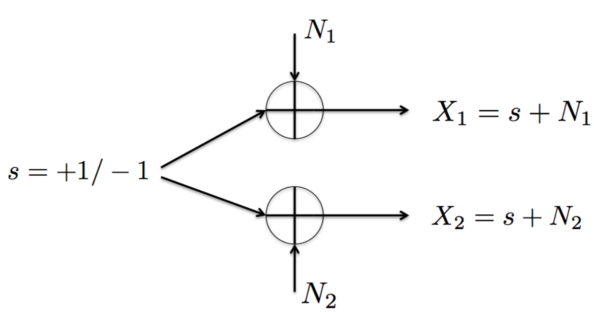
(e) Compute the MAP classifier given observations x_1 and x_2.

Solution:

(a)

DT71

Consider a communication system where the transmitter sends, with equal *a priori* probabilities, the same symbol “+1” or “−1” through two noisy channels, as illustrated in the figure:



where N_1 and N_2 are independent Gaussian random variables, with zero mean and variances λv and $(1 - \lambda)v$, respectively; $v > 0$ and $0 \leq \lambda \leq 1$ are two known constants.

(a) Obtain the binary classifier with a minimum probability of error, based on the joint observation of X_1 and X_2, which allows the receiver to decide whether the transmitted symbol was $+1$ or -1.

(b) Compute the error probability of the above decision maker. Express your result by means of the function

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

(c) Analyze the behaviour of the decision maker (i.e., its decision rule and probability of error) when: $\lambda = 0$ and $\lambda = 1$.

Solution:

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

(b) $P_\mathrm{e} = F(-\frac{1}{\sqrt{\lambda(1-\lambda)v}})$

(c) When $\lambda = 0$, $X_2 = s$ (its variance is zero) and we only consider this observation to make the decision. $P_\mathrm{e} = 0$. When $\lambda = 1$ a similar behaviour happens, but considering observation X_1.

DT72

We wish to find if a certain cell culture grows in a particular liquid environment. In order to do that, we measure the temperature X of the culture (in Celsius degrees) after an elapsed time of $t > 1$ minutes. It is known that, if the culture is growing, the temperature is given by

$$X = 10 \cdot t \exp(-t) + R$$

where R is a noise random variable with zero mean and variance 4.

However, when the cell culture does not evolve, the temperature is given by

$$X = 10 \exp(-t) + R$$

A priori, the probability that the cell culture grows is $P_H(1) = 0.5$. The temperature is measured after t minutes, and we wish to decide whether the culture cell has grown or not.

(a) Find the decision with minimum probability of error.

(b) Find the probability of error of the previous classifier. Express your result in terms of the following normalized distribution function

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

(c) Determine how long should we wait before measuring the temperature in order to minimize the probability of error.

(d) After the time obtained in the previous section, a temperature $x = 10$ degrees is observed. Find and expression for the probability that the cell culture has evolved.

Solution:

(a)

$$X \underset{D=0}{\overset{D=1}{\gtrless}} 5(t+1)\exp(-t)$$

(b)

$$P_e = F\left(\frac{5}{2}(1-t)\exp(-t)\right)$$

(c)

$$t = 2$$

(d)

$$TBD$$

7. General models

DT73

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 decision makers.

Solution:

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

DT74

Consider an M-ary bidimensional classification problem with observations $\mathbf{x} = [x_1,\ x_2]^T$, where it is known that $p_{X_1|X_2,H}(x_1|x_2, H = j)$ does not depend on j (i.e., on the hypothesis). We want to design the ML classifier. If it is further known that $\{P_H(j)\}_{j=1}^{M}$ are different, discuss which of the following classifiers provides the ML one:

(a) $j^* = \arg\text{máx}_j \left\{p_{X_1|H}(x_1|j)\right\}$
(b) $j^* = \arg\text{máx}_j \left\{p_{X_2|H}(x_2|j)\right\}$
(c) $j^* = \arg\text{máx}_j \left\{p_{X_2,H}(x_2, j)\right\}$

Solution: (b)

DT75

Consider a unidimensional binary decision problem with likelihoods $p_{X|H}(x|h)$ and *a priori* probabilities $P_H(h)$, with $h \in \{0, 1\}$ and $P_H(1) = 0.6$.

(a) It is known that $P_{H|X}(h|x) = P_H(h)$, for $h \in \{0, 1\}$, and for all x. Determine the MAP classifier.
(b) Which is the probability of error of the decision maker obtained in the previous section?
(c) Ignore now the condition of section (a). Instead, it is known that the likelihoods are symmetric to each other, i.e., $p_{X|H}(x|1) = p_{X|H}(-x|0)$, and that some decision maker given by

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

verifies $P_{FA} = P_M$. Which is the value of μ?.

(d) Using an equation or a plot, propose and example of symmetric likelihoods (like in the previous section) for which the ML classifier is not a threshold decision maker, i.e., the ML classifier cannot be expressed as

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

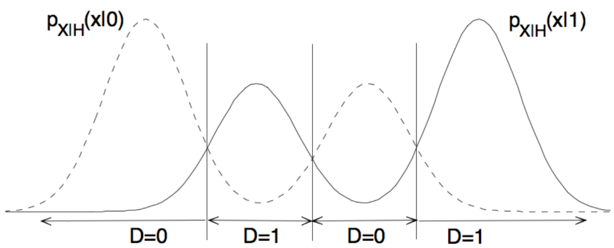
Solution:

(a) The MAP classifier always selects $D = 1$.

(b) $P_e = 0.4$

(c) $\mu = 0$

(d)



8. Sequential decision making

DT76

Most of the time, the returns of a given stock can be modeled as $x[n] = w[n]$, where $w[n]$ is a zero-mean white Gaussian process with variance σ_w^2. However, when there is a significant amount of short sellers (investors that profit from the decline in price of a borrowed asset), the returns can be modeled as $x[n] = s[n] + w[n]$, where $s[n]$ is modeled as a zero-mean white Gaussian process with variance σ_s^2, and independent of $w[n]$.[1]

(10 %) (a) The likelihood ratio test (LRT) when there are N, with $N > 1$, available observations, that is, for $x[n], n = 0, \ldots, N-1$.

Solution: We shall start by defining the vectors

$$\mathbf{x} = (x[0], \ldots, x[N-1])^T, \quad \mathbf{s} = (s[0], \ldots, s[N-1])^T, \quad \mathbf{w} = (w[0], \ldots, w[N-1])^T,$$

which allows us to write

$$H = 0 : \mathbf{x} = \mathbf{w},$$
$$H = 1 : \mathbf{x} = \mathbf{s} + \mathbf{w}.$$

Taking into account that, both, $\mathbf{s}$ and $\mathbf{w}$ are zero-mean Gaussian, white, and independent, it is easy to show that

$$\mathbb{E}\{\mathbf{x}|H = 0\} = \mathbf{0}, \qquad \mathbb{E}\{\mathbf{x}|H = 1\} = \mathbf{0},$$

[1] It is important to notice that $s[n]$ is a random process, *not a deterministic signal.*

and

$$\mathbb{E}\{\mathbf{x}\mathbf{x}^T|H=0\} = \sigma_w^2\mathbf{I}, \qquad \mathbb{E}\{\mathbf{x}\mathbf{x}^T|H=1\} = (\sigma_s^2+\sigma_w^2)\mathbf{I},$$

which yields

$$H=0: \mathbf{x} \sim G(\mathbf{0}, \sigma_w^2\mathbf{I}),$$
$$H=1: \mathbf{x} \sim G(\mathbf{0}, (\sigma_s^2+\sigma_w^2)\mathbf{I}).$$

Once we have the likelihoods, we can compute the likelihood ratio test (LRT) as

$$\frac{p_{\mathbf{X}|H}(\mathbf{x}|1)}{p_{\mathbf{X}|H}(\mathbf{x}|0)} \underset{D=0}{\overset{D=1}{\gtrless}} \eta,$$

which becomes

$$\frac{\frac{1}{(2\pi(\sigma_s^2+\sigma_w^2))^{N/2}} \exp\left(-\frac{1}{2(\sigma_s^2+\sigma_w^2)}\mathbf{x}^T\mathbf{x}\right)}{\frac{1}{(2\pi\sigma_w^2)^{N/2}} \exp\left(-\frac{1}{2\sigma_w^2}\mathbf{x}^T\mathbf{x}\right)} \underset{D=0}{\overset{D=1}{\gtrless}} \eta.$$

Taking logarithms and simplifying the expression, the log-likelihood ratio test (LLRT) is

$$t = \mathbf{x}^T\mathbf{x} = \sum_{n=0}^{N-1} x^2[n] \underset{D=0}{\overset{D=1}{\gtrless}} \mu,$$

where

$$\mu = \frac{\sigma_w^2(\sigma_s^2+\sigma_w^2)}{\sigma_s^2}\left[2\log(\eta) + N\log\left(\frac{\sigma_s^2+\sigma_w^2}{\sigma_w^2}\right)\right].$$

(15 %) (b) The probability of correctly detecting the presence of short sellers of the LRT for an arbitrary threshold. Express your solution in terms of the Q_{χ^2}-function.

Solution: The probability of correctly detecting the presence of short sellers of the LRT for an arbitrary threshold is given by

$$P_{\mathrm{D}} = P(D=1|H=1) = \int_{\mathcal{X}_1} p_{\mathbf{X}|H}(\mathbf{x}|1)d\mathbf{x},$$

where $\mathcal{X}_1 = \{\mathbf{x} \mid \sum_{n=0}^{N-1} x^2[n] > \mu\}$. However, we cannot compute the above multidimensional integral in closed form. To overcome this issue, it can be rewritten as

$$P_{\mathrm{D}} = P(T > \mu|H=1) = \int_{t>\mu} p_{T|H}(t|1)dt.$$

We therefore need the probability density function (PDF) of T under $H=1$. Since T is a the sum of squared Gaussian random variables, we could try to write it as a chi-squared random variable. Nevertheless, they do not have unit variance, which prevents us from using the results below. This is easily overcome by rewriting P_{D} as

$$P_{\mathrm{D}} = P\left(\tilde{T} > \frac{\mu}{\sigma_s^2+\sigma_w^2}|H=1\right) = \int_{\tilde{t}>\mu/(\sigma_s^2+\sigma_w^2)} p_{\tilde{T}|H}(\tilde{t}|1)d\tilde{t},$$

where

$$\tilde{t} = \sum_{n=0}^{N-1}\left(\frac{x[n]}{\sqrt{\sigma_s^2+\sigma_w^2}}\right).$$

Taking into account that $x[n]/\sqrt{\sigma_s^2+\sigma_w^2} \sim G(0,1)$ under $H=1$, it can be shown that

$$\tilde{T} \mid H=1 \sim \chi_N^2.$$

Hence, the sought probability is

$$P_{\mathrm{D}} = \int_{\mu/(\sigma_s^2+\sigma_w^2)}^{\infty} \frac{1}{2^{N/2}\Gamma(N/2)} \tilde{t}^{N/2-1} \exp\left(-\frac{\tilde{t}}{2}\right) d\tilde{t} = Q_{\chi^2}\left(\frac{\mu}{\sigma_s^2+\sigma_w^2}\right).$$