Bayesian Detection

Aleksandar Dogandžić

May 14, 2017

Contents

Bayesian Detection **1**

0–1 loss and MAP rule 4

Example: DC level in AWGN 4

Bayes Risk **6**

Average error probability (0–1 loss) 7

Bayesian Detection for Simple Hypotheses **7**

Bayes risk and average error probability 8

Nuisance Parameters **10**

Reading: [Hero 2015, §7], [Parmigiani 2002, §2.6], [Parmigiani and Inoue 2009, §7], [Van Trees et al. 2013, §2], [Johnson 2013, §5].

✻ Reminder. A decision rule $\phi(x)$ maps the *measurement space* $\mathcal{X}$ to $\{0, 1\}$:

$$\phi(x) = \begin{cases} 1, & \text{decide } \mathbb{H}_1, \\ 0, & \text{decide } \mathbb{H}_0 \end{cases}$$

and $\mathcal{X}_i = \{x \mid \phi(x) = i\}$.

Bayesian Detection

Posterior expected loss:

$$\begin{aligned} \rho(\text{action} \mid x) &= \int_\Theta \mathbb{L}(\theta, \text{action})\, f_{\Theta|X}(\theta \mid x)\, \mathrm{d}\theta \\ &= \mathrm{E}_{\Theta|X}\big[\mathbb{L}(\Theta, \text{action}) \,\big|\, x\big]. \end{aligned}$$

In *point estimation*, we used loss functions of the form $\mathbb{L}(\widehat{\theta}(x) - \Theta)$, see handout `best`. In *hypothesis testing* discussed here, our action space consists of only two choices. In this case, a popular loss function $\mathbb{L}(\theta, \text{action})$ is piecewise constant in θ, specified by the loss table:

- $\mathbb{L}(1 \mid 0)$ quantifies loss due to a false alarm,
- $\mathbb{L}(0 \mid 1)$ quantifies loss due to a miss,

	$\mathrm{sp}_\Theta(1)$	$\mathrm{sp}_\Theta(0)$
$x \in \mathcal{X}_1$	$\mathbb{L}(1 \mid 1) = 0$	$\mathbb{L}(1 \mid 0)$
$x \in \mathcal{X}_0$	$\mathbb{L}(0 \mid 1)$	$\mathbb{L}(0 \mid 0) = 0$

Table 1: Loss table for binary hypothesis testing.

- losses of correct decisions $\mathbb{L}(1 \mid 1)$ and $\mathbb{L}(0 \mid 0)$ are normally set to zero [Parmigiani 2002].[1]

[1] [Hero 2015, §7.2.2] keeps $\mathbb{L}(1 \mid 1)$ and $\mathbb{L}(0 \mid 0)$ nonzero and different, in general

In particular,

$$\begin{aligned}
\mathbb{L}(\theta, \text{say } \mathbb{H}_0) &= \underbrace{\mathbb{L}(0 \mid 0)}_{0} \mathbb{1}_{\mathrm{sp}_\theta(0)}(\theta) + \mathbb{L}(0 \mid 1) \mathbb{1}_{\mathrm{sp}_\theta(1)}(\theta) \\
&= \mathbb{L}(0 \mid 1) \mathbb{1}_{\mathrm{sp}_\Theta(1)}(\theta) \qquad (1a)\\
\mathbb{L}(\theta, \text{say } \mathbb{H}_1) &= \mathbb{L}(1 \mid 0) \mathbb{1}_{\mathrm{sp}_\theta(0)}(\theta) + \underbrace{\mathbb{L}(1 \mid 1)}_{0} \mathbb{1}_{\mathrm{sp}_\theta(1)}(\theta) \\
&= \mathbb{L}(1 \mid 0) \mathbb{1}_{\mathrm{sp}_\Theta(0)}(\theta) \qquad (1b)
\end{aligned}$$

or, in general,

$$\begin{aligned}
\mathbb{L}(\theta, \text{say } \mathbb{H}_m) &= \sum_{i=0}^{1} \mathbb{L}(m \mid i) \mathbb{1}_{\mathrm{sp}_\theta(i)}(\theta) \\
&= \mathbb{L}(m \mid 0) \mathbb{1}_{\mathrm{sp}_\theta(0)}(\theta) + \mathbb{L}(m \mid 1) \mathbb{1}_{\mathrm{sp}_\theta(1)}(\theta). \qquad (2)
\end{aligned}$$

Now, our posterior expected loss takes two values:

$$\begin{aligned}
\underbrace{\rho_0(x)}_{\rho(\text{say } \mathbb{H}_0 \mid x)} &= \mathrm{E}_{\Theta|X}\left[\mathbb{L}(\Theta, \text{say } \mathbb{H}_0) \mid x\right] \\
&= \mathrm{E}_{\Theta|X}\left[\sum_{i=0}^{1} \mathbb{L}(0 \mid i) \mathbb{1}_{\mathrm{sp}_\theta(i)}(\Theta) \,\middle|\, x\right] \\
&= \sum_{i=0}^{1} \mathbb{L}(0 \mid i) \mathrm{E}_{\Theta|X}\left[\mathbb{1}_{\mathrm{sp}_\theta(i)}(\Theta) \,\middle|\, x\right] \\
&= \sum_{i=0}^{1} \mathbb{L}(0 \mid i) \Pr(\mathbb{H}_i \mid x) \\
&= \mathbb{L}(0 \mid 0) \Pr(\mathbb{H}_0 \mid x) + \mathbb{L}(0 \mid 1) \Pr(\mathbb{H}_1 \mid x) \\
&= \mathbb{L}(0 \mid 1) \Pr(\mathbb{H}_1 \mid x)
\end{aligned}$$

and, similarly,

$$\underbrace{\rho_1(x)}_{\rho(\text{say } \mathbb{H}_1 | x)} = \mathbb{L}(1 \mid 0) \Pr(\mathbb{H}_0 \mid x).$$

☞ Notation: As [Hero 2015], we use

$$\begin{aligned}
\Pr(\mathbb{H}_i) &\triangleq \Pr(\Theta \in \mathrm{sp}_\Theta(i)), \qquad (3a)\\
\Pr(\mathbb{H}_i \mid x) &\triangleq \Pr_{\Theta|X}(\Theta \in \mathrm{sp}_\Theta(i) \mid x) \qquad (3b)
\end{aligned}$$

for $i = 0, 1$.

☞ In general, for a decision rule $\phi(x)$,

$$\begin{aligned}
\rho_{\phi(x)}(x) &= \mathrm{E}_{\Theta\,|\,X}\big[\mathbb{L}(\Theta, \text{say } \mathbb{H}_{\phi(x)}) \,|\, x\big] \\
&= \mathrm{E}_{\Theta\,|\,X}\Big[\sum_{i=0}^{1} \mathbb{L}(\phi(x)\,|\,i)\, \mathbb{1}_{\mathrm{sp}_\theta(i)}(\Theta) \,|\, x\Big] \qquad \text{see (2)} \\
&= \sum_{i=0}^{1} \mathbb{L}(\phi(x)\,|\,i)\, \mathrm{E}_{\Theta\,|\,X}\big[\mathbb{1}_{\mathrm{sp}_\theta(i)}(\Theta) \,|\, x\big] \\
&= \sum_{i=0}^{1} \mathbb{L}(\phi(x)\,|\,i)\, \Pr(\mathbb{H}_i \,|\, x) \\
&= \mathbb{L}(\phi(x)\,|\,0)\, \Pr(\mathbb{H}_0 \,|\, x) + \mathbb{L}(\phi(x)\,|\,1)\, \Pr(\mathbb{H}_1 \,|\, x).
\end{aligned} \tag{4}$$

The *Bayes' decision rule* minimizes the posterior expected loss and corresponds to the following measurement-space partitioning:

$$\mathcal{X}_1 = \{x : \rho_1(x) \le \rho_0(x)\}$$

or

$$\mathcal{X}_1 = \big\{x : \mathbb{L}(1\,|\,0)\, \Pr(\mathbb{H}_0 \,|\, x) \le \mathbb{L}(0\,|\,1)\, \Pr(\mathbb{H}_1 \,|\, x)\big\}$$

i.e.,

$$\frac{\Pr(\mathbb{H}_1 \,|\, x)}{\Pr(\mathbb{H}_0 \,|\, x)} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \frac{\mathbb{L}(1\,|\,0)}{\mathbb{L}(0\,|\,1)} \tag{5}$$

or, equivalently, upon applying the Bayes' rule:

$$\frac{f(x\,|\,\mathbb{H}_1)\, \Pr(\mathbb{H}_1)/f(x)}{f(x\,|\,\mathbb{H}_0)\, \Pr(\mathbb{H}_0)/f(x)} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \frac{\mathbb{L}(1\,|\,0)}{\mathbb{L}(0\,|\,1)}$$

i.e.,

$$\underbrace{\frac{f(x\,|\,\mathbb{H}_1)}{f(x\,|\,\mathbb{H}_0)}}_{\text{likelihood ratio}} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \frac{\mathbb{L}(1\,|\,0)}{\mathbb{L}(0\,|\,1)} \frac{\Pr(\mathbb{H}_0)}{\Pr(\mathbb{H}_1)}. \tag{6}$$

Bayes' rule expressed in terms of likelihood and prior distributions

☞ Note:

$$f(x\,|\,\mathbb{H}_i) = \frac{\int_{\mathrm{sp}_\Theta(i)} f_{X\,|\,\Theta}(x\,|\,\theta)\, f_\Theta(\theta)\, \mathrm{d}\theta}{\Pr(\mathbb{H}_i)}, \qquad i = 0, 1.$$

0–1 loss and MAP rule

	$\mathrm{sp}_\Theta(1)$	$\mathrm{sp}_\Theta(0)$
$x \in \mathcal{X}_1$	$\mathbb{L}(1\,|\,1) = 0$	$\mathbb{L}(1\,|\,0) = 1$
$x \in \mathcal{X}_0$	$\mathbb{L}(0\,|\,1) = 1$	$\mathbb{L}(0\,|\,0) = 0$

Table 2: 0-1 loss table for binary hypothesis testing.

CHOOSING $\mathbb{L}(1\,|\,0) = \mathbb{L}(0\,|\,1) = 1$ yields the 0–1 loss in Table 2. The corresponding Bayes' decision rule for the 0–1 loss is called the maximum *a posteriori* (MAP) rule:

$$\frac{\Pr(\mathbb{H}_1\,|\,x)}{\Pr(\mathbb{H}_0\,|\,x)} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} 1 \tag{7a}$$

see (5)

i.e.,

$$\frac{f(x\,|\,\mathbb{H}_1)}{f(x\,|\,\mathbb{H}_0)} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \frac{\Pr(\mathbb{H}_0)}{\Pr(\mathbb{H}_1)}. \tag{7b}$$

see (6)

✻ ML rule. For equiprobable hypotheses:

$$\Pr(\mathbb{H}_0) = \Pr(\mathbb{H}_1) = 0.5 \tag{8a}$$

the MAP rule (7b) is known as the maximum-likelihood (ML) rule. Substituting (8a) into (7b) yields

$$f_{X|\Theta}(x\,|\,\mathbb{H}_1) \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} f_{X|\Theta}(x\,|\,\mathbb{H}_0). \tag{8b}$$

ML rule

Example: DC level in AWGN

WE collect conditionally independent, identically distributed (i.i.d.) measurements $(X[n])_{n=0}^{N-1}$ given μ following $\mathcal{N}(x[n]\,|\,\mu,\sigma^2)$, where the variance σ^2 is known.

Assign the following mixture prior on μ:

$$f_\mu(\mu) = \pi_0\delta(\mu) + \underbrace{(1-\pi_0)}_{\triangleq \pi_1}\mathcal{N}\big(\mu\,|\,\mu_0,\tau_0^2\big)$$

spike-and-slab prior

where $\pi_0 \in (0,1)$ is a known constant. We wish to test

$$\begin{aligned}&\mathbb{H}_0 : \mu \in \mathrm{sp}_\mu(0) = \{0\} \quad \text{versus}\\ &\mathbb{H}_1 : \mu \in \mathrm{sp}_\mu(1) = \mathbb{R}\backslash\{0\}.\end{aligned}$$

In this case,

$$\begin{aligned}\Pr(\mathbb{H}_0) &= \pi_0\\ \Pr(\mathbb{H}_1) &= \pi_1 = 1-\pi_0.\end{aligned}$$

We know that $\overline{x} \triangleq \frac{1}{N}\sum_{n=0}^{N-1} x[n]$ is a sufficient statistic and

$$\{\overline{X} \mid \mu\} \sim \mathcal{N}\Big(\mu, \frac{\sigma^2}{N}\Big).$$

Hence,

$$\begin{aligned} f(\overline{x} \mid \mathbb{H}_0) &= \mathcal{N}\Big(\overline{x} \,\Big|\, 0, \frac{\sigma^2}{N}\Big) \\ f(\overline{x} \mid \mathbb{H}_1) &= \mathcal{N}\Big(\overline{x} \,\Big|\, \mu_0, \frac{\sigma^2}{N} + \tau_0^2\Big) \end{aligned}$$

yielding

$$\begin{aligned} \ln f(\overline{x} \mid \mathbb{H}_1) - \ln f(\overline{x} \mid \mathbb{H}_0) &= -0.5 \ln\Big(\frac{\sigma^2}{N} + \tau_0^2\Big) - 0.5 \frac{(\overline{x} - \mu_0)^2}{\sigma^2/N + \tau_0^2} + 0.5 \ln \frac{\sigma^2}{N} + 0.5 \frac{\overline{x}^2}{\sigma^2/N} \\ &\underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \ln \frac{\mathbb{L}(1 \mid 0)\,\Pr(\mathbb{H}_0)}{\mathbb{L}(0 \mid 1)\,\Pr(\mathbb{H}_1)} \qquad \text{see (6)} \end{aligned}$$

i.e.,

$$0.5 \frac{\overline{x}^2}{\sigma^2/N} - 0.5 \frac{(\overline{x} - \mu_0)^2}{\sigma^2/N + \tau_0^2} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \underbrace{\ln \frac{\mathbb{L}(1 \mid 0)\,\Pr(\mathbb{H}_0)}{\mathbb{L}(0 \mid 1)\,\Pr(\mathbb{H}_1)} + 0.5 \ln\Big(1 + \frac{N\tau_0^2}{\sigma^2}\Big)}_{\triangleq \eta}. \quad (9)$$

Define

$$a \triangleq \frac{1}{\sigma^2/N} - \frac{1}{\sigma^2/N + \tau_0^2} \geq 0.$$

- ✻ If $\tau_0^2 = 0$ (and, consequently, $a = 0$), this detector reduces to comparing $\overline{x}$ with a threshold because terms in (9) that contain $(\overline{x})^2$ cancel out.
- ✻ $\tau_0^2 > 0$. If $\tau_0^2 > 0$, then

$$0.5 \frac{\overline{x}^2}{\sigma^2/N} - 0.5 \frac{(\overline{x} - \mu_0)^2}{\sigma^2/N + \tau_0^2} = 0.5a(\overline{x} + c)^2 - b \qquad \text{complete the squares}$$

where

$$\begin{aligned} c &\triangleq \frac{\mu_0}{a(\sigma^2/N + \tau_0^2)} \\ &= \frac{\mu_0 \sigma^2/N}{\tau_0^2}. \end{aligned}$$

and b can be determined easily as well. Hence, we can write

$$0.5a(\overline{x} + c)^2 \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \eta + b$$

which can be reduced to

$$(\overline{x} + c)^2 \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \gamma.$$

Here, if $\gamma \le 0$, we decide $\mathbb{H}_1$. If $\gamma > 0$, then

$$|\overline{x} + c| \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \sqrt{\gamma}$$

and decide $\mathbb{H}_0$ if

$$-c - \sqrt{\gamma} < \overline{x} < -c + \sqrt{\gamma}.$$

✻ $\tau_0^2 > 0$ AND $\mu_0 = 0$. For $\tau_0^2 > 0$ (implying $a > 0$) and $\mu_0 = 0$ (implying $c = 0$), our test simplifies to

$$0.5\frac{\overline{x}^2}{\sigma^2/N} - 0.5\frac{(\overline{x})^2}{\sigma^2/N + \tau_0^2} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \eta$$

i.e.,

$$0.5a\overline{x}^2 \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \eta$$

i.e.,

$$\overline{x}^2 \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} 2\eta/a.$$

In summary, $\mathcal{X}_0(\overline{x})$ has the following shapes as function of $\overline{x}$:

$$\mathcal{X}_0(\overline{x}) = \begin{cases} \text{one-sided interval,} & a = 0 \\ \text{two-sided interval,} & a \ne 0, \gamma > 0, . \\ \text{entire real line,} & a \ne 0, \gamma \le 0. \end{cases}$$

Bayes Risk

CHOOSE the rule $\phi(x)$ that minimizes the Bayes risk $\mathrm{E}_{X,\Theta}\big[\mathbb{L}(\Theta, \text{decide } \mathbb{H}_{\phi(X)})\big]$. Now, by the law of iterated expectations,

$$\begin{aligned}\mathrm{E}_{X,\Theta}[\mathbb{L}(\Theta,\ \text{decide } \mathbb{H}_{\phi(X)})] &= \mathrm{E}_X\{\mathrm{E}_{\boldsymbol{\Theta}\,|\,X}\big[\mathbb{L}\big(\Theta,\ \text{decide } \mathbb{H}_{\phi(X)}\big)\,|\,X\big]\} \\ &= \mathrm{E}_X[\rho(\text{decide } \mathbb{H}_{\phi(X)}\,|\,X)].\end{aligned}$$

Hence, the rule $\phi(x)$ that minimizes the posterior expected loss $\rho\big(\text{decide } \mathbb{H}_{\phi(X)}\,|\,x\big)$ also minimizes the Bayes risk.

☞ BAYES' decision rule minimizes the Bayes risk.

Continue:

see (4)

$$
\begin{aligned}
\mathrm{E}_{X,\Theta}\big[\mathbb{L}\big(\Theta,\ \text{say } \mathbb{H}_{\phi(X)}\big)\big] &= \mathrm{E}_X\big[\rho(\text{say } \mathbb{H}_{\phi(X)} \mid X)\big] \\
&= \mathrm{E}_X\Big[\sum_{i=0}^{1} \mathbb{L}(\phi(X) \mid i) \Pr(\mathbb{H}_i \mid X)\Big] \\
&= \int_{\mathcal{X}} \sum_{i=0}^{1} \mathbb{L}(\phi(x) \mid i) \underbrace{\Pr(\mathbb{H}_i \mid x) f_X(x)}_{\text{joint}} \, \mathrm{d}x \\
&= \int_{\mathcal{X}} \sum_{i=0}^{1} \mathbb{L}(\phi(x) \mid i) \underbrace{\Pr(\mathbb{H}_i) f(x \mid \mathbb{H}_i)}_{\text{joint}} \, \mathrm{d}x \\
&= \sum_{m=0}^{1} \int_{\mathcal{X}_m} \sum_{i=0}^{1} \mathbb{L}(m \mid i) \Pr(\mathbb{H}_i) f(x \mid \mathbb{H}_i) \, \mathrm{d}x \\
&= \sum_{m=0}^{1} \sum_{i=0}^{1} \mathbb{L}(m \mid i) \Pr(\mathbb{H}_i) \underbrace{\int_{\mathcal{X}_m} f(x \mid \mathbb{H}_i) \, \mathrm{d}x}_{\Pr(X \in \mathcal{X}_m \mid \mathbb{H}_i)} \\
&= \sum_{m=0}^{1} \sum_{i=0}^{1} \mathbb{L}(m \mid i) \Pr(\mathbb{H}_i) \Pr(X \in \mathcal{X}_m \mid \mathbb{H}_i) \\
&= \mathbb{L}(1 \mid 0) \Pr(\mathbb{H}_0) \Pr(X \in \mathcal{X}_1 \mid \mathbb{H}_0) \\
&\quad + \mathbb{L}(0 \mid 1) \Pr(\mathbb{H}_1) \Pr(X \in \mathcal{X}_0 \mid \mathbb{H}_1). \qquad (10)
\end{aligned}
$$

Average error probability (0–1 loss)

For the 0–1 loss in Table 2, the Bayes risk for rule $\phi(x)$ is

$$
\begin{aligned}
\mathrm{E}_{X,\Theta}\big[\mathbb{L}\big(\Theta,\ \text{decide } \mathbb{H}_{\phi(X)}\big)\big] &= \Pr(\mathbb{H}_0) \Pr(X \in \mathcal{X}_1 \mid \mathbb{H}_0) \\
&\quad + \Pr(\mathbb{H}_1) \Pr(X \in \mathcal{X}_0 \mid \mathbb{H}_1) \qquad (11)
\end{aligned}
$$

total probability

the *average error probability*, with *averaging* performed over the joint distribution of the measurements X and parameter Θ.

Bayesian Detection for Simple Hypotheses

Binary simple hypotheses: The space of the parameter Θ and its partitions are

$$
\mathrm{sp}_\Theta = \{\theta_0, \theta_1\}, \qquad \mathrm{sp}_\Theta(0) = \{\theta_0\}, \qquad \mathrm{sp}_\Theta(1) = \{\theta_1\}
$$

for testing $\mathbb{H}_0 : \Theta = \theta_0$ versus $\mathbb{H}_1 : \Theta = \theta_1$.

☞ For binary simple hypotheses, the prior probability mass function (pmf) for θ is a Bernoulli pmf:

$$\pi_0 \triangleq p_\Theta(\theta_0) = \Pr(\Theta = \theta_0) = \Pr(\mathbb{H}_0) \tag{12a}$$

$$\begin{aligned}\pi_1 &\triangleq p_\Theta(\theta_1) = \Pr(\Theta = \theta_1) = \Pr(\mathbb{H}_1) \\ &= 1 - \pi_0.\end{aligned} \tag{12b}$$

Bayes risk and average error probability

SUBSTITUTE the simple-hypothesis prior (12) into the general Bayes risk expression for binary hypothesis testing (10):

$$\begin{aligned}\mathrm{E}_{X,\Theta}\big[\mathbb{L}\big(\Theta, \text{ decide } \mathbb{H}_{\phi(X)}\big)\big] = \mathbb{L}(1 \,|\, 0)\pi_0 \underbrace{\Pr(X \in \mathcal{X}_1 \,|\, \theta_0)}_{P_{\mathrm{FA}}} \\ + \mathbb{L}(0 \,|\, 1)\pi_1 \underbrace{\Pr(X \in \mathcal{X}_0 \,|\, \theta_1)}_{P_{\mathrm{M}}}.\end{aligned} \tag{13}$$

✻ AVERAGE error probability. For the 0–1 loss in Table 2, the Bayes risk of a decision rule $\phi(x)$ is equal to the average error probability:

$$P_{\mathrm{av}} = \pi_0 P_{\mathrm{FA}} + \pi_1 P_{\mathrm{M}} \tag{14}$$

total probability

obtained by substituting $\mathbb{L}(1 \,|\, 0) = \mathbb{L}(0 \,|\, 1) = 1$ into (13).

Example 1. Radar

WE revisit the radar example from handout `introdet`. Given the matched-filter output y, we wish to find the MAP rule.

Assume $\tau = 0$ and $s(t) \not\equiv 0$; then

$$y = \int_0^T x(t)s(t)\,\mathrm{d}t$$

and

$$\mathrm{E}(Y \,|\, \mathbb{H}_0) = 0 \tag{15a}$$

$$\mathrm{E}(Y \,|\, \mathbb{H}_1) = \int_0^T |s(t)|^2\,\mathrm{d}t \triangleq \mu_1 > 0 \tag{15b}$$

$$\mathrm{var}(Y \,|\, \mathbb{H}_0) = \frac{\mathcal{N}_0}{2}\int_0^T |s(t)|^2\,\mathrm{d}t \triangleq \sigma_0^2. \tag{15c}$$

✻ MAP rule:

$$\begin{aligned}\frac{f_{Y|\Theta}(y \,|\, \theta_1)}{f_{Y|\Theta}(y \,|\, \theta_0)} &= \frac{\mathcal{N}(y \,|\, \mu_1, \sigma_0^2)}{\mathcal{N}(y \,|\, 0, \sigma_0^2)} \\ &= \frac{\exp\big[-0.5(y-\mu_1)^2/\sigma_0^2\big] \Big/ \sqrt{2\pi\sigma_0^2}}{\exp(-0.5y^2/\sigma_0^2) \Big/ \sqrt{2\pi\sigma_0^2}} \\ &= \exp\big(y\mu_1/\sigma_0^2 - 0.5\mu_1^2/\sigma_0^2\big) \\ &\underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \frac{\pi_0}{\pi_1}\end{aligned} \tag{16}$$

i.e.,

$$\frac{y\mu_1}{\sigma_0^2} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \ln\frac{\pi_0}{\pi_1} + 0.5\frac{\mu_1^2}{\sigma_0^2}.$$

Since $\mu_1 > 0$,[2] we can simplify to

[2] see (15b)

$$y \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \underbrace{\frac{\sigma_0^2}{\mu_1}\ln\frac{\pi_0}{\pi_1} + 0.5\mu_1}_{\gamma} .$$

* ML rule: $\pi_0 = \pi_1 = 0.5$ and our test simplifies to

$$y \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} 0.5\mu_1.$$

In this case, our test does not require knowledge of σ_0^2 or, equivalently, we do not need to know the power spectral density $\mathcal{N}_0/2$ of the additive white Gaussian noise (AWGN) that corrupts our radar returns.

* Probability of false alarm.

$$\begin{aligned}
P_{\text{FA}} &= \Pr(Y > \gamma \mid \mathbb{H}_0) \\
&= \Pr\Big(\underbrace{\frac{Y}{\sigma_0}}_{\mathcal{N}(0,1)} > \frac{\gamma}{\sigma_0} \,\Big|\, \mathbb{H}_0\Big) \\
&= Q\Big(\frac{\gamma}{\sigma_0}\Big) \\
&= 1 - \Phi\Big(\frac{\gamma}{\sigma_0}\Big)
\end{aligned}$$

where $\Phi(x)$ and $Q(x) = 1 - \Phi(x)$ are the cumulative distribution function (cdf) and complementary cdf of a standard normal [$\mathcal{N}(0,1)$] random variable.

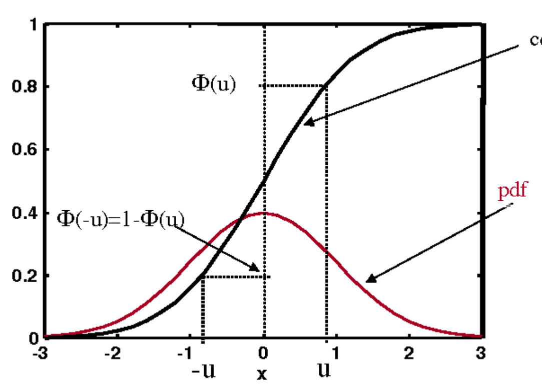

Figure 1: Probability density function (pdf) and cdf of standard normal distribution.

✻ PROBABILITY of miss.

$$\begin{aligned} P_{\mathrm{M}} &= \Pr(Y < \gamma \mid \mathbb{H}_1) \\ &= \Pr\Big(\underbrace{\frac{Y-\mu_1}{\sigma_0}}_{\mathcal{N}(0,1)} < \frac{\gamma-\mu_1}{\sigma_0} \,\Big|\, \mathbb{H}_1\Big) \\ &= \Phi\Big(\frac{\gamma-\mu_1}{\sigma_0}\Big) \\ &= 1 - \Phi\Big(\frac{\mu_1-\gamma}{\sigma_0}\Big). \end{aligned}$$

✻ AVERAGE error probability for the MAP rule:

$$\begin{aligned} P_{\mathrm{av}} &= \pi_0 P_{\mathrm{FA}} + \pi_1 P_{\mathrm{M}} \\ &= \pi_0\Big[1-\Phi\Big(\frac{\gamma}{\sigma_0}\Big)\Big] + \pi_1\Big[1-\Phi\Big(\frac{\mu_1-\gamma}{\sigma_0}\Big)\Big] \end{aligned}$$

which simplifies in the case of the ML rule to:

$$\begin{aligned} P_{\text{av, ML rule}} &= 0.5\Big[1-\Phi\Big(\frac{0.5\mu_1}{\sigma_0}\Big)\Big] + 0.5\Big[1-\Phi\Big(\frac{\mu_1-0.5\mu_1}{\sigma_0}\Big)\Big] \\ &= P_{\mathrm{FA}} = 1-\Phi\Big(\frac{0.5\mu_1}{\sigma_0}\Big). \end{aligned}$$

Nuisance Parameters

INTEGRATE out the nuisance parameters (u, say). Therefore, (5) still holds for testing

$$\begin{aligned} &\mathbb{H}_0 : \Theta \in \mathrm{sp}_\Theta(0) \quad \text{versus} \\ &\mathbb{H}_1 : \Theta \in \mathrm{sp}_\Theta(1) \end{aligned}$$

but $f_{\Theta \mid X}(\theta \mid x)$ is the *marginal posterior*, computed as follows:

$$f_{\Theta \mid X}(\theta \mid x) = \int f_{\Theta, U \mid X}(\theta, u \mid x)\, \mathrm{d}u.$$

Hence, (5) becomes

$$\begin{aligned} \frac{\Pr(\mathbb{H}_1 \mid x)}{\Pr(\mathbb{H}_0 \mid x)} &= \frac{\int_{\mathrm{sp}_\Theta(1)} f_{\Theta \mid X}(\theta \mid x)\, \mathrm{d}\theta}{\int_{\mathrm{sp}_\Theta(0)} f_{\Theta \mid X}(\theta \mid x)\, \mathrm{d}\theta} \\ &= \frac{\int_{\mathrm{sp}_\Theta(1)} \int f_{\Theta, U \mid X}(\theta, u \mid x)\, \mathrm{d}u\, \mathrm{d}\theta}{\int_{\mathrm{sp}_\Theta(0)} \int f_{\Theta, U \mid X}(\theta, u \mid x)\, \mathrm{d}u\, \mathrm{d}\theta} \\ &\mathop{\gtrless}_{\mathbb{H}_0}^{\mathbb{H}_1} \frac{\mathbb{L}(1 \mid 0)}{\mathbb{L}(0 \mid 1)} \end{aligned} \tag{17a}$$

or, equivalently, upon applying the Bayes' rule:

$$\frac{\int_{\mathrm{sp}_\Theta(1)} \int f_{X\,|\,\Theta,U}(x\,|\,\theta,u)\, f_{\Theta,U}(\theta,u)\,\mathrm{d}u\,\mathrm{d}\theta}{\int_{\mathrm{sp}_\Theta(0)} \int f_{X\,|\,\Theta,U}(x\,|\,\theta,u)\, f_{\Theta,U}(\theta,u)\,\mathrm{d}u\,\mathrm{d}\theta} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \frac{\mathbb{L}(1\,|\,0)}{\mathbb{L}(0\,|\,1)}. \tag{17b}$$

We can also rewrite (6) as

see [Johnson 2013, §5.3.1]

$$\begin{aligned} \frac{f(x\,|\,\mathbb{H}_1)}{f(x\,|\,\mathbb{H}_0)} &= \frac{\int f(x\,|\,u,\mathbb{H}_1)\, f(u\,|\,\mathbb{H}_1)\,\mathrm{d}u}{\int f(x\,|\,u,\mathbb{H}_0)\, f(u\,|\,\mathbb{H}_0)\,\mathrm{d}u} \\ &\underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \frac{\mathbb{L}(1\,|\,0)}{\mathbb{L}(0\,|\,1)} \frac{\Pr(\mathbb{H}_0)}{\Pr(\mathbb{H}_1)}. \end{aligned} \tag{18}$$

If we assume that Θ and U are independent, i.e.,

$$f_{\Theta,U}(\theta,u) = f_\Theta(\theta)\, f_U(u)$$

and, consequently,

$$f(u\,|\,\mathbb{H}_i) = f_U(u)$$

then (18) becomes

$$\frac{\int f(x\,|\,u,\mathbb{H}_1)\, f_U(u)\,\mathrm{d}u}{\int f(x\,|\,u,\mathbb{H}_0)\, f_U(u)\,\mathrm{d}u} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \frac{\mathbb{L}(1\,|\,0)}{\mathbb{L}(0\,|\,1)} \frac{\Pr(\mathbb{H}_0)}{\Pr(\mathbb{H}_1)}. \tag{19}$$

Acronyms

AWGN additive white Gaussian noise. 8

cdf cumulative distribution function. 8, 9

i.i.d. independent, identically distributed. 4

MAP maximum *a posteriori*. 4, 7–9

ML maximum-likelihood. 4, 8, 9

pdf probability density function. 9

pmf probability mass function. 7

References

Hero, Alfred O. (2015). *Statistical Methods for Signal Processing*. Lecture notes. Univ. Michigan, Ann Arbor, MI (cit. on pp. 1, 2).

Johnson, Don H. (2013). *Statistical Signal Processing*. Lecture notes. Rice Univ., Houston, TX (cit. on pp. 1, 11).

Parmigiani, Giovanni (2002). *Modeling in Medical Decision Making: A Bayesian Approach*. New York: Wiley (cit. on pp. 1, 2).

Parmigiani, Giovanni and Lurdes Inoue (2009). *Decision Theory: Principles and Approaches*. New York: Wiley (cit. on p. 1).

Van Trees, Harry L., Kristine L. Bell, and Zhi Tian (2013). *Detection, Estimation, and Modulation Theory, Part I*. 2nd ed. New York: Wiley (cit. on p. 1).