CHAPTER 7

Decision theory

7.1 Introduction

In section 6.2 the detection of baseband binary symbols was described and analysed for the special case of Gaussian noise and equiprobable symbols. The decision reference voltage, against which detected symbols were tested, was set intuitively (and correctly) mid-way between those voltages which would have been detected, if the binary symbols were received without noise. In this chapter a more systematic analysis of symbol decision theory is given, applying the descriptions of random signals and noise developed in Chapter 3. This allows the optimum levels for receiver decision voltages to be found in the more general case of arbitrary noise distributions and unequal symbol probabilities.

For a binary data stream there are two principal types of decision:

(a) Soft (multi-level) decisions.
(b) Hard (2 level) decisions.

Soft decision receivers, Figure 7.1(a), quantise the decision instant signal plus noise voltage using several allowed levels, each represented by a decision word of a few binary bits. Figure 7.1(c) illustrates an 8-level (3-bit) soft decision process which is typical. Each soft decision contains information not only about the most likely transmitted symbol (000 to 011 indicating a likely 0 and 100 to 111 indicating a likely 1) but also information about the confidence or likelihood which can be placed on this decision. The soft decisions are converted to final or hard decisions using an algorithm which inspects a sequence of several PCM words and makes decisions accounting for the confidence levels they represent, in conjunction with the error control decoding rules. The algorithm tends to be heuristic and requires tailoring to the particular application. Although such algorithms can now be implemented using high speed VLSI decoders, these techniques will not be discussed further here.

Hard decisions (Figure 7.1(b)) are more common than soft decisions. The two major decision criteria used in this case are Bayes and Neyman-Pearson. The Bayes decision criterion is used extensively in binary communications while the Neyman-Pearson criterion is used more in radar applications. The principal difference between the two is

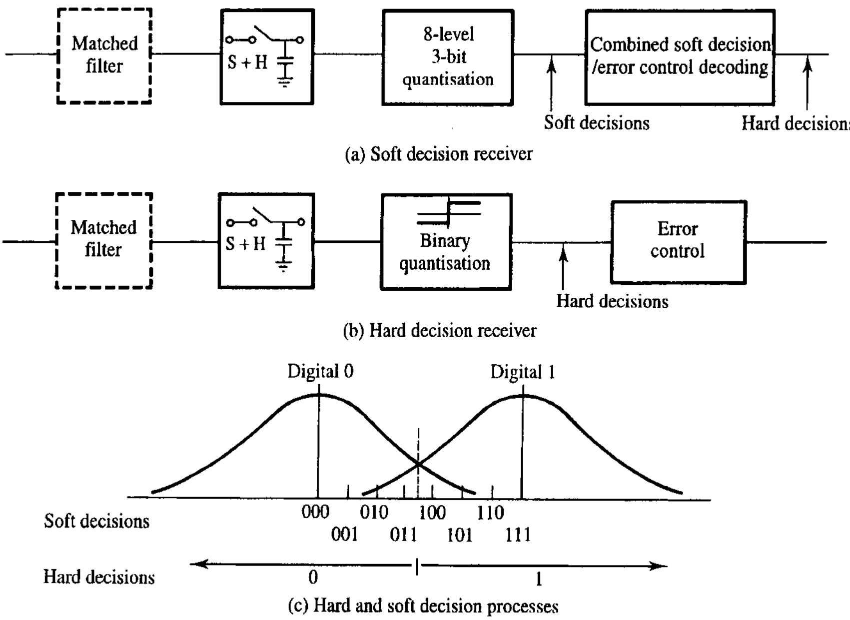

Figure 7.1 *Comparison of hard and soft decision receivers.*

that the Bayes decision rule assumes known a priori source statistics for the occurrence of digital ones and zeros while the Neyman-Pearson criterion makes no such assumption. This is appropriate to radar applications because the a priori probability of the appearance of a target is not normally known.

7.2 A priori, conditional and a posteriori probabilities

There are four probabilities and two probability density functions, see Chapter 3, associated with symbol transmission and reception. These are:

1. $P(0)$ – a priori probability of transmitting the digit 0.
2. $P(1)$ – a priori probability of transmitting the digit 1.
3. $p(v|0)$ – conditional probability density function of detecting voltage v given that a digital 0 was transmitted.
4. $p(v|1)$ – conditional probability density function of detecting voltage v given that a digital 1 was transmitted.
5. $P(0|v)$ – a posteriori probability that a digital 0 was transmitted given that voltage v was detected.
6. $P(1|v)$ – a posteriori probability that a digital 1 was transmitted given that voltage v was detected.

The terms a priori and a posteriori imply reasoning from cause to effect and reasoning from effect to cause respectively. The cause of a communication event, in this context, is the symbol transmitted and the effect is the voltage detected. The a priori probabilities (in communications applications) are usually known in advance. The conditional probabilities are dependent on the data (1 or 0) transmitted. The a posteriori probabilities can only be established after many events (i.e. symbol transmissions and receptions) have been completed. (Note that the lower case notation is used for conditional probability density functions and upper case for a posteriori probabilities. Also, note that the position of the transmitted and detected quantities is interchanged between these cases.)

7.3 Symbol transition matrix

When communications systems operate in the presence of noise (which is always the case) there is the possibility of transmitted symbols being sufficiently corrupted to be interpreted, at the receiver, in error. If the characteristics of the noise are precisely enough known, or many observations of symbol transmissions and receptions are made, then the probability of each transmitted symbol being interpreted at the receiver as any of the other symbols can be found. These *transition* probabilities are usually denoted using the conditional probability notation, $P(i|j)$, where i represents the symbol received and j represents the symbol transmitted. $P(i|j)$ is therefore the probability that a symbol will be interpreted as i given that j was transmitted. The symbol transition probabilities can be arranged as the elements of a matrix to describe the end to end properties of a communications channel. Examples of such matrices are given below.

7.3.1 Binary symmetric channel

In a binary channel four types of communication events can occur. These events are:

0 transmitted and 0 received
0 transmitted and 1 received
1 transmitted and 1 received
1 transmitted and 0 received

Denoting transmitted symbols with subscript *TX* and received symbols with subscript *RX* (for extra clarity) the binary channel can be represented schematically as shown in Figure 7.2(a). The corresponding symbol transition matrix is:

$$\begin{bmatrix} P(0_{RX}|0_{TX}) & P(0_{RX}|1_{TX}) \\ P(1_{RX}|0_{TX}) & P(1_{RX}|1_{TX}) \end{bmatrix}$$

If the probability, p, of a transmitted 0 being received in error as a 1, is equal to the probability of a transmitted 1 being received as a 0, then the binary channel is said to be symmetric and the transition matrix can be written as:

$$\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

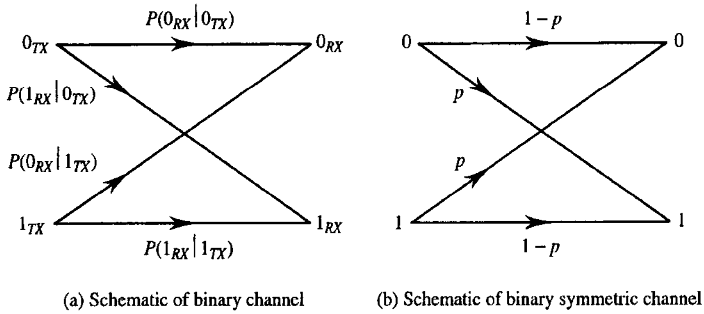

Figure 7.2 *Schematic representations of binary channels.*

Figure 7.2(b) shows the equivalent schematic diagram. The conditional probabilities in the matrix must sum vertically to unity, i.e.:

$$P(0_{RX}|0_{TX}) + P(1_{RX}|0_{TX}) = 1 \tag{7.1}$$

For the binary symmetric case:

$$P(0_{RX}|0_{TX}) = P(1_{RX}|1_{TX}) = 1 - p \tag{7.2}$$

The unconditional probability of receiving a 0 is therefore:

$$P_{0_{RX}} = P(0_{TX})P(0_{RX}|0_{TX}) + P(1_{TX})P(0_{RX}|1_{TX}) \tag{7.3}$$

which can be rewritten for the binary symmetric channel as:

$$P_{0_{RX}} = P(0)(1-p) + P(1)p \tag{7.4}$$

Similarly:

$$P_{1_{RX}} = P(1)(1-p) + P(0)p \tag{7.5}$$

Equations (7.4) and (7.5) relate the observed probabilities of received symbols to the a priori probabilities of transmitted symbols and the error probabilities of the channel.

EXAMPLE 7.1 – Multisymbol transmission

For an M symbol communication alphabet the 2×2 transition matrix of the binary channel must be extended to an $M \times M$ matrix. Consider a source with the $M = 6$ symbols, $A\ B\ C\ D\ E\ F$, and a 6-ary receiver which can distinguish these symbols. The transition matrix, $\mathbf{T}$, for this system contains 6×6 transition probabilities which are:

$$\mathbf{T} = \begin{bmatrix} \frac{1}{2} & \frac{0}{1} & \frac{1}{24} & \frac{0}{1} & \frac{0}{1} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{12} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{4} & \frac{1}{6} & \frac{0}{1} & \frac{1}{12} \\ \frac{0}{1} & \frac{1}{3} & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{8} & \frac{0}{1} & \frac{1}{6} & \frac{1}{12} & \frac{1}{2} & \frac{1}{8} \\ \frac{0}{1} & \frac{1}{24} & \frac{1}{8} & \frac{1}{12} & \frac{0}{1} & \frac{1}{3} \end{bmatrix}$$

(For a transmitted symbol B the probabilities of receiving symbols $A, \cdots, F$ are given by the $M = 6$ entries in the matrix's second column.)

If the a priori transmission probabilities of the six symbols are:

$$P(A) = 0.1$$
$$P(B) = 0.15$$
$$P(C) = 0.4$$
$$P(D) = 0.05$$
$$P(E) = 0.2$$
$$P(F) = 0.1$$

then find the probability of error when (i) only the symbol D is transmitted and (ii) when a random string of symbols is transmitted. Also find, for this latter case, the probability of receiving the symbol C.

The probability of error, P_e, when the symbol D is transmitted is given by the sum of all probabilities in the fourth (i.e. D transmitted) column, excluding the fourth (i.e. D received) element:

$$P_e = \sum_{m=A}^{F} P(m_{RX}|D_{TX}) \quad \text{(excluding the } m = D \text{ term)}$$

$$= 0 + 1/3 + 1/6 + 1/12 + 1/12 = 2/3$$

This is identical to $1 - P(D_{RX}|D_{TX})$, i.e.:

$$P_e = 1 - 1/3 = 2/3$$

(Note that the leading diagonal elements, $P(m|m)$, in the transition matrix are the probabilities of correct symbol reception and $1 - P(m|m)$ are therefore the probabilities of symbol error.)

The probability of error when a (long) random string of symbols $(A, \cdots, F)$ is transmitted is a weighted sum of the probabilities of error for each transmitted symbol where the weighting factors are the a priori probabilities of transmission, i.e.:

$$P_e = \sum_{m=A}^{F} P(m)\,[1 - P(m_{RX}|m_{TX})]$$

$$= 0.1 \times \frac{1}{2} + 0.15 \times \frac{1}{2} + 0.4 \times \frac{3}{4} + 0.05 \times \frac{2}{3} + 0.2 \times \frac{1}{2} + 0.1 \times \frac{2}{3}$$

$$= 0.625$$

In addition to probabilities of error, the various probabilities of symbol reception can also be found from the transition matrix. The probability of receiving symbol C when symbol D is transmitted is given by inspection of the appropriate matrix element, i.e.:

$$P(C_{RX}|D_{TX}) = 1/6$$

The (unconditional) probability of receiving symbol C when a (long) random string of symbols $(A, \cdots, F)$ is transmitted is the a priori weighted sum of all elements in the third (i.e. C received) row of the transition matrix:

$$P(C_{RX}) = \sum_{m=A}^{F} P(C_{RX}|m_{TX})\, P(m)$$

$$= 0.1 \times \frac{1}{8} + 0.15 \times \frac{1}{8} + 0.4 \times \frac{1}{4} + 0.05 \times \frac{1}{6} + 0.2 \times 0 + 0.1 \times \frac{1}{12}$$

$$= 0.148$$

7.4 Bayes's decision criterion

This is the most widely applied decision rule in communications systems and, as a consequence, it will be discussed in detail. In essence it operates so as to minimise the average *cost* (in terms of errors or lost information) of making decisions.

7.4.1 Decision costs

In binary transmission there are two ways to lose information:

1. Information is lost when a transmitted digital 1 is received in error as a digital 0.
2. Information is lost when a transmitted digital 0 is received in error as a digital 1.

The cost in the sense of lost information due to mechanisms (1) and (2) is denoted here by C_0 and C_1 respectively. There is no cost (i.e. no information is lost) when correct decisions are made.

7.4.2 Expected conditional decision costs

The expected conditional cost, $C(0|v)$, incurred when a detected voltage v is interpreted by a decision circuit as a digital 0 is given by:

$$C(0|v) = C_0\, P(1|v) \tag{7.6(a)}$$

where C_0 is the cost if the decision is in error and $P(1|v)$ is the (a posteriori) probability that the decision is in error. By symmetry the corresponding equation can be written:

$$C(1|v) = C_1\, P(0|v) \tag{7.6(b)}$$

This is the expected conditional cost incurred when v is interpreted as a digital 1.

7.4.3 Optimum decision rule

A rational decision rule to adopt is to interpret each detected voltage, v, as either a 0 or a 1, so as to minimise the expected conditional cost, i.e.:

$$C(1|v) \mathop{\lessgtr}_{0}^{1} C(0|v) \tag{7.7}$$

The interpretation of inequality (7.7) is 'if the upper inequality holds then decide 1, if the lower inequality holds then decide 0'. Substituting equations (7.6) into the inequality (7.7) and cross dividing gives:

$$\frac{P(0|v)}{P(1|v)} \mathop{\lessgtr}_{0}^{1} \frac{C_0}{C_1} \tag{7.8}$$

If the costs of both types of error are the same, or unknown (in which case the rational assumption must be $C_0 = C_1$), then (7.8) represents a *maximum a posteriori* probability (MAP) decision criterion. Inequality (7.8) is one form of Bayes's decision criterion. It uses a posteriori probabilities, however, which are not usually known. The criterion can be transformed to a more useful form by using Bayes's theorem, equation (3.4), i.e.:

$$P(0|v) = \frac{p(v|0)\, P(0)}{p(v)} \tag{7.9(a)}$$

and:

$$P(1|v) = \frac{p(v|1)\, P(1)}{p(v)} \tag{7.9(b)}$$

Figure 7.3 illustrates the conditional probability density functions for the case of zero mean Gaussian noise. Dividing equation (7.9(a)) by (7.9(b)):

$$\frac{P(0|v)}{P(1|v)} = \frac{p(v|0)\, P(0)}{p(v|1)\, P(1)} \tag{7.10}$$

and substituting equation (7.10) into (7.8):

$$\frac{p(v|0)\, P(0)}{p(v|1)\, P(1)} \mathop{\lessgtr}_{0}^{1} \frac{C_0}{C_1} \tag{7.11}$$

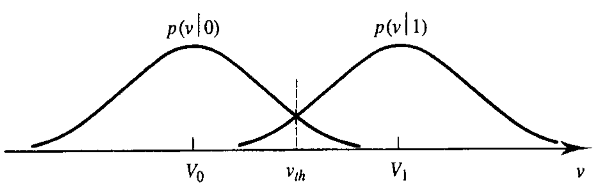

Figure 7.3 *Probability distributions for binary transmissions V_0 or V_1.*

Rearranging the inequality (7.11) gives Bayes's criterion in a form using conditional probability density functions and the a priori source probabilities which are usually known:

$$\frac{p(v|0)}{p(v|1)} \underset{0}{\overset{1}{\lessgtr}} \frac{C_0\, P(1)}{C_1\, P(0)} \tag{7.12}$$

The left hand side of the above inequality is called the *likelihood ratio* (which is a function of v) and the right hand side is a likelihood threshold which will be denoted by L_{th}. If $C_0 = C_1$ and $P(0) = P(1)$ (or, more generally, $C_0 P(1) = C_1 P(0)$), or if neither costs nor a priori probabilities are known (in which case $C_0 = C_1$ and $P(0) = P(1)$ are the most rational assumptions) then $L_{th} = 1$ and equation (7.12) is called a *maximum likelihood* decision criterion. This type of hypothesis testing is used by the receivers described in Chapter 11, e.g. Figure 11.7. Table 7.1 illustrates the relationship between maximum liklihood, MAP and Bayes decision criteria.

Table 7.1 *Comparison of receiver types*

Receiver	*A priori probabilities known*	*Decision costs known*	*Assumptions*	*Decision criterion*
Bayes	Yes	Yes	None	$\frac{p(v\|0)}{p(v\|1)} \underset{0}{\overset{1}{\lessgtr}} \frac{C_0\, P(1)}{C_1\, P(0)}$
MAP	Yes	No	$C_0 = C_1$	$\frac{p(v\|0)}{p(v\|1)} \underset{0}{\overset{1}{\lessgtr}} \frac{P(1)}{P(0)}$
Max. likelihood	No	No	$C_0 P(1) = C_1 P(0)$	$\frac{p(v\|0)}{p(v\|1)} \underset{0}{\overset{1}{\lessgtr}} 1$

7.4.4 Optimum decision threshold voltage

Bayes's decision criterion represents a general solution to setting the optimum reference or threshold voltage, v_{th}, in a receiver decision circuit. The threshold voltage which minimises the expected conditional cost of each decision is the value of v which satisfies:

$$\frac{p(v|0)}{p(v|1)} = \frac{C_0\, P(1)}{C_1\, P(0)} = L_{th} \tag{7.13}$$

This is illustrated for two conditional pdfs in Figure 7.4. (These pdfs actually correspond to those obtained for envelope detection of OOK rectangular pulses, see Figure 6.11, $p(v|0)$ being a Rayleigh distribution (see section 4.7.1) and $p(v|1)$ being Rician.) If $L_{th} = 1.0$ such as would be the case for statistically independent, equiprobable symbols with equal error costs, then the voltage threshold would occur at the intersection of the two conditional pdfs. (This is exactly the location of the decision threshold selected intuitively for centre point detection of equiprobable, rectangular binary symbols in section 6.2.1.)

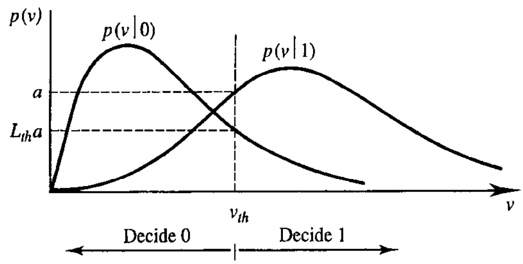

Figure 7.4 *Optimum decision threshold voltage for envelope detected OOK signal.*

In practice, a single threshold, such as that shown in Figure 7.4, is normally adequate against which to test detected voltages. It is possible in principle, however, for the conditional pdfs to intersect at more than one point. This situation is illustrated in Figure 7.5. Assuming the symbols are equiprobable the decision thresholds are then set at the intersection of the condition pdfs. Here, however, this implies more than one threshold, with several decision regions delineated by these thresholds. It is important to appreciate that despite their unconventional appearance the conditional pdfs in equation (7.13) behave in the same way as any other function. Equation (7.13), which can be rewritten in the form:

$$p(v|0) = L_{th}\, p(v|1) \tag{7.14}$$

can therefore be solved using any of the normal techniques including, where necessary, numerical methods.

7.4.5 Average unconditional decision cost

It is interesting to note (and perhaps self evident) that as well as minimising the cost of each decision, Bayes's criterion also minimises the average cost of decisions (i.e. the

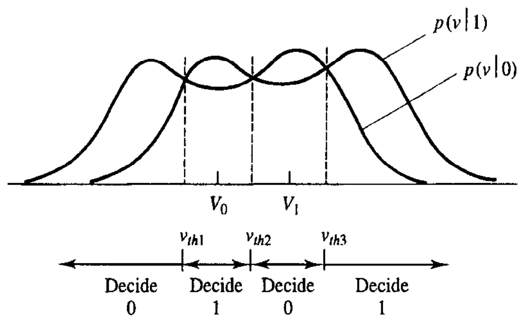

Figure 7.5 *Optimum decision thresholds for equiprobable binary voltages in noise with multimodal pdf.*

costs averaged over all decisions irrespective of type). This leads to the following alternative derivation of Bayes's criterion.

The average cost, $\bar{C}$, of making decisions is given by:

$$\bar{C} = C_1 \; P(0)P(1_{RX}|0_{TX}) + C_0 \; P(1)P(0_{RX}|1_{TX}) \tag{7.15}$$

where $P(0)P(1_{RX}|0_{TX})$ is the probability that any given received symbol is a digital 1 which has been detected in error and $P(1)P(0_{RX}|1_{TX})$ is the probability that any given received symbol is a digital 0 also detected in error. From Figure 7.3 it can be seen that:

$$P(1_{RX}|0_{TX}) = \int_{v_{th}}^{\infty} p(v|0) \; dv \tag{7.16(a)}$$

$$P(0_{RX}|1_{TX}) = \int_{-\infty}^{v_{th}} p(v|1) \; dv \tag{7.16(b)}$$

where v_{th} is the receiver decision threshold voltage. Substituting equations (7.16) into equation (7.15) the average cost of making decisions can be written as:

$$\bar{C} = C_1 \; P(0) \int_{v_{th}}^{\infty} p(v|0) \; dv + C_0 \; P(1) \int_{-\infty}^{v_{th}} p(v|1) \; dv \tag{7.17}$$

Bayes's decision criterion sets the decision threshold v_{th} so as to minimise $\bar{C}$. To find this optimum threshold $\bar{C}$ must be differentiated with respect to v_{th} and equated to zero. This involves differentiating the integrals in equation (7.17) which is most easily done using [Dwight, equations 69.1 and 69.2], i.e.:

$$\frac{d}{dx} \int_{c}^{x} f(t) \; dt = f(x) \tag{7.18(a)}$$

$$\frac{d}{dx} \int_{x}^{c} f(t) \; dt = -f(x) \tag{7.18(b)}$$

We therefore have:

$$\frac{d\bar{C}}{dv_{th}} = C_1 \; P(0) \; [-p(v_{th}|0)] + C_0 \; P(1) \; [p(v_{th}|1)] \tag{7.19}$$

Setting equation (7.19) equal to zero for minimum (or maximum) $\bar{C}$:

$$C_0 \; P(1) \; p(v_{th}|1) = C_1 \; P(0) \; p(v_{th}|0) \tag{7.20(a)}$$

and rearranging, gives:

$$\frac{P(v_{th}|1)}{P(v_{th}|0)} = \frac{C_1 \; P(0)}{C_0 \; P(1)} \tag{7.20(b)}$$

Solution of equation (7.20) for v_{th} gives the Bayes's criterion (i.e. optimum) decision voltage.

EXAMPLE 7.2 – Binary transmission

Consider a binary transmission system subject to additive Gaussian noise which has a mean value of 1.0 V and a standard deviation of 2.5 V. (Note that, unusually, the standard deviation of the noise is not the same as its RMS value in this example.) A digital 1 is represented by a rectangular pulse with amplitude 4.0 V and a digital 0 by a rectangular pulse with amplitude –4. 0 V. The costs (i.e. information lost) due to each type of error are identical (i.e. $C_0 = C_1$) but the a priori probabilities of symbol transmission are different, $P(1)$ being twice $P(0)$, as the symbols are not statistically independent. Find the optimum decision threshold voltage and the resulting probability of symbol error.

Using Bayes's decision criterion the optimum decision threshold voltage is given by solving equation (7.14), i.e.:

$$p(v_{th}|0) = L_{th}\, p(v_{th}|1)$$

where:

$$L_{th} = \frac{C_0}{C_1}\frac{P(1)}{P(0)}$$

Since $C_1 = C_0$ and $P(1) = 2P(0)$ then:

$$L_{th} = 2.0$$

The noise is Gaussian and its mean value adds to the symbol voltages. An equivalent signalling system therefore has symbol voltages of –3 V and 5 V, and noise with zero mean.

The conditional probability density functions are:

$$p(v|0) = \frac{1}{\sigma_0\sqrt{(2\pi)}}\, e^{\frac{-(v-V_0)^2}{2\sigma_0^2}}$$

$$p(v|1) = \frac{1}{\sigma_1\sqrt{(2\pi)}}\, e^{\frac{-(v-V_1)^2}{2\sigma_1^2}}$$

where V_0 and V_1 are the signal voltages representing digital 0s and 1s respectively. Substituting these equations into the above version of equation (7.14) with $\sigma_0 = \sigma_1 = 2.5$, $V_0 = -3$ and $V_1 = 5$:

$$\frac{1}{2.5\sqrt{(2\pi)}}\, e^{\frac{-(v+3)^2}{2(2.5)^2}} = 2.0\, \frac{1}{2.5\sqrt{(2\pi)}}\, e^{\frac{-(v-5)^2}{2(2.5)^2}}$$

Cancelling and taking logs:

$$\frac{-(v+3)^2}{2(2.5)^2} = \ln(2.0) - \frac{(v-5)^2}{2(2.5)^2}$$

$$\frac{-(v^2+6v+9)+(v^2-10v+25)}{2(2.5)^2} = \ln(2.0)$$

$$16 - 16v = 2(2.5)^2 \ln(2.0)$$

or:

$$v = \frac{16 - [2(2.5)^2 \ln(2.0)]}{16} = 0.46 \text{ V}$$

Figure 7.6 illustrates the *un*conditional probability density functions of symbols plus noise and the location of the optimum threshold voltage at $v_{th} = 0.46$ V for this example. (The conditional probability density functions would both enclose unit area.)

The probability of error for each symbol can be found by integrating the error tails of the normalised ($\bar{v} = 0$, $\sigma = 1$) Gaussian probability density function. The probability of a transmitted 0 being received in error is therefore:

$$P_{e1} = \int_{\frac{v_{th}-V_0}{\sigma_0}}^{\infty} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{v^2}{2}} dv = \int_{1.384}^{\infty} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{v^2}{2}} dv$$

Using the substitution $x = v/\sqrt{2}$:

$$P_{e1} = \frac{1}{\sqrt{\pi}} \int_{1.384/\sqrt{2}}^{\infty} e^{-x^2} dx = 0.5 \operatorname{erfc}(1.384/\sqrt{2}) = 0.0831$$

Similarly the probability of a transmitted 1 being received in error is:

$$P_{e0} = \int_{-\infty}^{\frac{v_{th}-V_1}{\sigma_1}} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{v^2}{2}} dv$$

From symmetry P_{e0} can also be written:

$$P_{e0} = \int_{\frac{-(v_{th}-V_1)}{\sigma_1}}^{\infty} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{v^2}{2}} dv = \int_{1.816}^{\infty} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{v^2}{2}} dv$$

Using the same substitution as before this becomes:

$$P_{e0} = 0.5 \operatorname{erfc}(1.816/\sqrt{2}) = 0.0347$$

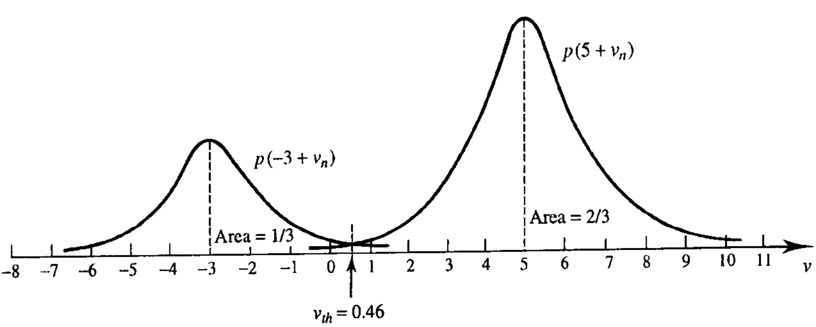

Figure 7.6 *Unconditional symbol plus noise voltage pdf for Example 7.2, $p(V + v_n)$ denotes voltage pdf of symbol voltage V plus noise.*

The overall probability of symbol error is therefore:

$$\begin{aligned} P_e &= P(0)P_{e1} + P(1)P_{e0} \\ &= \frac{1}{3}(0.0831) + \frac{2}{3}(0.0347) = 0.0508 \end{aligned}$$

Optimum thresholding, as discussed above, can be applied to centre point detection processes, such as described in Chapter 6, or after predetection signal processing (e.g. matched filtering), as discussed in Chapter 8. The important point to appreciate is that the sampling instant signal plus noise pdfs, used to establish the optimum threshold voltage(s), are those at the decision circuit input irrespective of whether predetection processing has been applied or not.

7.5 Neyman-Pearson decision criterion

The Neyman-Pearson decision criterion requires only a posteriori probabilities. Unlike Bayes's decision rule it does not require a priori source probability information. This criterion is particularly appropriate to pulse detection in Gaussian noise, as occurs in radar applications, where the source statistics (i.e. probabilities of presence, and absence, of a target) are unknown. It also works well when $C_0 \gg C_1$ or, in radar terms, when the information cost of erroneously deciding a target is present is much less than the information cost of erroneously deciding a target is absent. In this context the important probabilities are those for target detection, P_D, and target false alarm, P_{FA}. The conditional probability density functions for these two decisions and a selected decision threshold voltage are shown in Figure 7.7. The two probabilities are:

$$P_D = \int_{v_{th}}^{\infty} p(v|s+n)\, dv \tag{7.21}$$

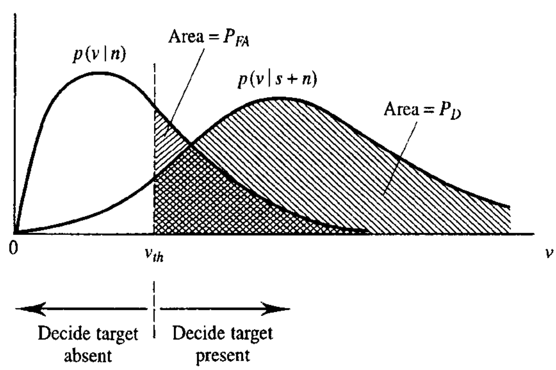

Figure 7.7 *Conditional pdfs and threshold voltage for Neyman-Pearson radar signal detector.*

$$P_{FA} = \int_{v_{th}}^{\infty} p(v|n)\, dv \qquad (7.22)$$

where s denotes signal (arising due to reflections from a target) and n denotes noise. The optimum detector consists of a linear filter followed by a threshold detector [Blahut, 1987] which tests the hypothesis as to whether there is noise only present or a signal pulse plus noise. The filter is matched to the pulse (see section 8.3). In the Neyman-Pearson detector the threshold voltage, v_{th}, is chosen to give an acceptable value of P_{FA} and the detection probability then follows the characteristic shown in Figure 7.8. Note here that detection performance is dependent on both the choice of P_{FA} and the ratio of received pulse energy to noise power spectral density, E/N_0.

7.6 Summary

Soft and hard decision processes are both used in digital communications receivers but hard decision processes are simpler to implement and therefore more common. Transition probabilities $P(i|j)$ describe the probability with which transmitted symbol j will be corrupted by noise sufficiently to be interpreted at the receiver as symbol i. These probabilities can be assembled into a matrix to describe the end to end error properties of a communications channel.

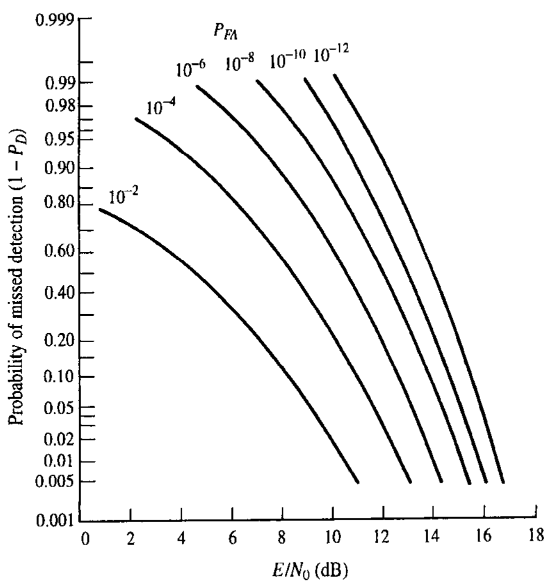

Figure 7.8 *Pulse detection in Gaussian noise using the Neyman-Pearson detection criterion (source: Blahut, 1987, reproduced with the permission of Addison Wesley).*

Bayes's decision criterion is the criterion most often used in digital communications receivers. It is optimum in the sense that it minimises the cost of (i.e. reduces the information lost when) making decisions. Bayes's criterion allows the optimum threshold voltage(s) which delineate(s) decision regions, to be found. Maximum a posteriori probability (MAP) and maximum likelihood detectors are special cases of a Bayes's criterion detector, appropriate when decision costs, or decision costs and a priori probabilities, are unknown.

The Neyman-Pearson decision criterion is normally used in radar applications. It has the advantage over Bayes's criterion that a priori symbol probabilities, whilst known to be very different, need not be quantified. The threshold level of the Neyman-Pearson criterion is set to give acceptable probabilities of false alarm and the probability of detection, given a particular received E/N_0 ratio, follows from this.

7.7 Problems

7.1. Derive the Bayes decision rule from first principles.

7.2. Under what assumptions does the Bayes receiver become a maximum likelihood receiver? Illustrate your answer with the appropriate equations.

7.3. In a baseband binary transmission system a 1 is represented by +5 V and a 0 by –5 V. The channel is subject to additive, zero mean, Gaussian noise with a standard deviation of 2 V. If the a priori probabilities of 1 and 0 are 0.5 and the costs C_0 and C_1 are equal, calculate the optimum position for the decision boundary, and the probability of error for this optimum position. [0 V, 6.23×10^{-3}]

7.4. What degenerate case of Bayes receiver design is being implemented in Problem 7.3? Assume that the decision threshold in Problem 7.3 is incorrectly placed by being moved up from its optimum position by 0.5 V. Calculate the new probability of error associated with this suboptimum threshold. [7.4125×10^{-3}]

7.5. Design a Bayes detector for the a priori probabilities $P(1) = 2/3, P(0) = 1/3$ with costs $C_0 = 1$ and $C_1 = 3$. The conditional pdfs for the received variable v are given by:

$$p(v|0) = \frac{1}{\sqrt{(2\pi)}} e^{-v^2/2}$$

$$p(v|1) = \frac{1}{\sqrt{(2\pi)}} e^{-(v-1)^2/2} \qquad [v = 0.905]$$

7.6. A radar system operates in, zero mean, Gaussian noise with variance 5 volts. If the probability of false alarm is to be 10^{-2} calculate the detection threshold. If the expected return from a target at extreme range is 4 volts, what is the probability that this target will be detected? [5.21, 0.2945]