(a) We have the priors $P(\omega_1)$ and $P(\omega_2) = 1 - P(\omega_1)$. The Bayes risk is given by Eqs. 12 and 13 in the text:

$$R(P(\omega_1)) = P(\omega_1) \int\limits_{\mathcal{R}_2} p(x|\omega_1)dx + (1 - P(\omega_1)) \int\limits_{\mathcal{R}_1} p(x|\omega_2)dx.$$

To obtain the prior with the minimum risk, we take the derivative with respect to $P(\omega_1)$ and set it to 0, that is

$$\frac{d}{dP(\omega_1)} R(P(\omega_1)) = \int\limits_{\mathcal{R}_2} p(x|\omega_1)dx - \int\limits_{\mathcal{R}_1} p(x|\omega_2)dx = 0,$$

which gives the desired results:

$$\int\limits_{\mathcal{R}_2} p(x|\omega_1)dx = - \int\limits_{\mathcal{R}_1} p(x|\omega_2)dx.$$

(b) This solution is not always unique, as shown in this simple counterexample. Let $P(\omega_1) = P(\omega_2) = 0.5$ and

$$\begin{aligned} p(x|\omega_1) &= \begin{cases} 1 & -0.5 \leq x \leq 0.5 \\ 0 & \text{otherwise} \end{cases} \\ p(x|\omega_2) &= \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

It is easy to verify that the decision regions $\mathcal{R}_1 = [-0.5, 0.25]$ and $\mathcal{R}_1 = [0, 0.5]$ satisfy the equations in part (a); thus the solution is not unique.

4. Consider the minimax criterion for a two-category classification problem.

(a) The total risk is the integral over the two regions $\mathcal{R}_i$ of the posteriors times their costs:

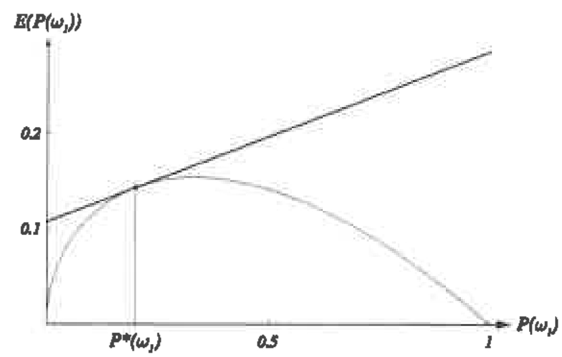
$$\begin{aligned} R &= \int\limits_{\mathcal{R}_1} [\lambda_{11} P(\omega_1) p(\mathbf{x}|\omega_1) + \lambda_{12} P(\omega_2) p(\mathbf{x}|\omega_2)] \; d\mathbf{x} \\ &\quad + \int\limits_{\mathcal{R}_2} [\lambda_{21} P(\omega_1) p(\mathbf{x}|\omega_1) + \lambda_{22} P(\omega_2) p(\mathbf{x}|\omega_2)] \; d\mathbf{x}. \end{aligned}$$

We use $\int\limits_{\mathcal{R}_2} p(\mathbf{x}|\omega_2)\; d\mathbf{x} = 1 - \int\limits_{\mathcal{R}_1} p(\mathbf{x}|\omega_2)\; d\mathbf{x}$ and $P(\omega_2) = 1 - P(\omega_1)$, regroup to find:

$$\begin{aligned} R &= \lambda_{22} + \lambda_{12} \int\limits_{\mathcal{R}_1} p(\mathbf{x}|\omega_2)\; d\mathbf{x} - \lambda_{22} \int\limits_{\mathcal{R}_1} p(\mathbf{x}|\omega_2)\; d\mathbf{x} \\ &\quad + P(\omega_1) \Bigg[(\lambda_{11} - \lambda_{22}) + \lambda_{11} \int\limits_{\mathcal{R}_2} p(\mathbf{x}|\omega_1)\; d\mathbf{x} - \lambda_{12} \int\limits_{\mathcal{R}_1} p(\mathbf{x}|\omega_2)\; d\mathbf{x} \\ &\quad + \lambda_{21} \int\limits_{\mathcal{R}_2} p(\mathbf{x}|\omega_1)\; d\mathbf{x} + \lambda_{22} \int\limits_{\mathcal{R}_1} p(\mathbf{x}|\omega_2)\; d\mathbf{x} \Bigg] \end{aligned}$$

$$
\begin{aligned}
&= \lambda_{22} + (\lambda_{12} - \lambda_{22}) \int\limits_{\mathcal{R}_1} p(\mathbf{x}|\omega_2)\, d\mathbf{x} \\
&\quad + P(\omega_1)\Bigg[(\lambda_{11} - \lambda_{22}) + (\lambda_{11} + \lambda_{21}) \int\limits_{\mathcal{R}_2} p(\mathbf{x}|\omega_1)\, d\mathbf{x} \\
&\qquad + (\lambda_{22} - \lambda_{12}) \int\limits_{\mathcal{R}_1} p(\mathbf{x}|\omega_2)\, d\mathbf{x}\Bigg].
\end{aligned}
$$

(b) Consider an arbitrary prior $0 < P^*(\omega_1) < 1$, and assume the decision boundary has been set so as to achieve the minimal (Bayes) error for that prior. If one holds the same decision boundary, but changes the prior probabilities (i.e., $P(\omega_1)$ in the figure), then the error changes *linearly*, as given by the formula above. But the true Bayes error must be less than or equal to that (linearly bounded) value, since one has the freedom to change the decision boundary at each value of $P(\omega_1)$. Moreover, we note that the Bayes error is 0 at $P(\omega_1) = 0$ and at $P(\omega_1) = 1$, since the Bayes decision rule under those conditions is to always decide ω_2 or ω_1, respectively. Thus the curve of Bayes error rate is concave down for all prior probabilities.



(c) According to the general minimax equation in part (a), for our case (i.e., $\lambda_{11} = \lambda_{22} = 0$ and $\lambda_{12} = \lambda_{21} = 1$) the decision boundary is chosen to satisfy

$$\int\limits_{\mathcal{R}_2} p(x|\omega_1)\, dx = \int\limits_{\mathcal{R}_1} p(x|\omega_2)\, dx.$$

We assume that a *single* decision point suffices, and thus we seek to find x^* such that

$$\int\limits_{-\infty}^{x^*} N(\mu_1, \sigma_1^2)\, dx = \int\limits_{x^*}^{\infty} N(\mu_2, \sigma_2^2)\, dx,$$

where, as usual, $N(\mu_i, \sigma_i^2)$ denotes a Gaussian. We assume for definiteness and without loss of generality that $\mu_2 > \mu_1$, and that the single decision point lies between the means. Recall the definition of an error function, given by Eq. 96

in the Appendix of the text, that is,

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

We can rewrite the above as

$$\mathrm{erf}[(x^* - \mu_1)/\sigma_1] = \mathrm{erf}[(x^* - \mu_2)/\sigma_2].$$

If the values of the error function are equal, then their corresponding arguments must be equal, that is

$$(x^* - \mu_1)/\sigma_1 = (x^* - \mu_2)/\sigma_2$$

and solving for x^* gives the value of the decision point

$$x^* = \left(\frac{\mu_2\sigma_1 + \mu_1\sigma_2}{\sigma_1 + \sigma_2}\right).$$

(d) Because the mimimax error rate is independent of the prior probabilities, we can choose a particularly simple case to evaluate the error, for instance, $P(\omega_1) = 0$. In that case our error becomes

$$E = 1/2 - \mathrm{erf}[(x^* - \mu_1)/\sigma_1] = 1/2 - \mathrm{erf}\left[\frac{\mu_2\sigma_1 - \mu_1\sigma_2}{\sigma_1(\sigma_1 + \sigma_2)}\right].$$

(e) We substitute the values given in the problem into the formula in part (c) and find

$$x^* = \frac{\mu_2\sigma_1 + \mu_1\sigma_2}{\sigma_1 + \sigma_2} = \frac{1/2 + 0}{1 + 1/2} = 1/3.$$

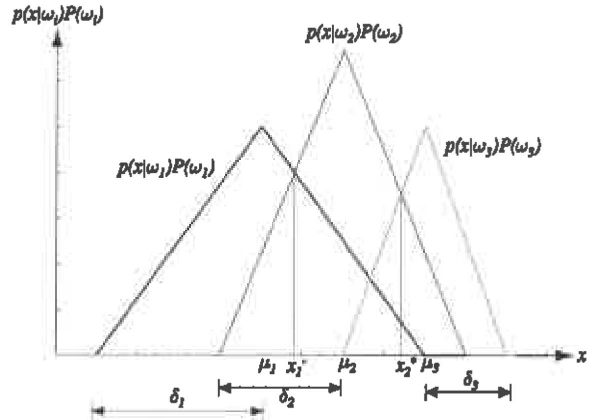
The error is then

$$E = 1/2 - \mathrm{erf}\left[\frac{1/3 - 0}{1 + 0}\right] = 1/2 - \mathrm{erf}[1/3] = 0.1374.$$

(f) Note that the distributions have the same form (in particular, the same variance). Thus, by symmetry the Bayes error for $P(\omega_1) = P^*$ (for some value P^*) must be the same as for $P(\omega_2) = P^*$. Because $P(\omega_2) = 1 - P(\omega_1)$, we know that the curve (analogous to the one shown above) is symmetric around the point $P(\omega_1) = 0.5$. Because the curve is concave down, therefore it must *peak* at $P(\omega_1) = 0.5$, i.e., equal priors. The tangent to the graph of the error versus $P(\omega_1)$ is thus horizontal at $P(\omega_1) = 0.5$. For this case of equal priors, the Bayes decision point for this problem can be stated simply: it is the point midway between the means of the two distributions, i.e., $x^* = 5.5$.

5. We seek to generalize the notion of minimax criteria to the case where *two* independent prior probabilities are set.

(a) We use the triangle distributions and conventions in the figure. We solve for the



decision points as follows (being sure to keep the signs correct, and assuming that the decision boundary consists of just two points):

$$P(\omega_1)\left(\frac{\delta_1-(x_1^*-\mu_1)}{\delta_1^2}\right)=P(\omega_2)\left(\frac{\delta_2-(\mu_2-x_1^*)}{\delta_2^2}\right),$$

which has solution

$$x_1^*=\frac{P(\omega_1)\delta_2^2\delta_1+P(\omega_1)\delta_2^2\mu_1-P(\omega_2)\delta_1^2\delta_2+P(\omega_2)\mu_2\delta_1^2}{P(\omega_1)\delta_2^2+P(\omega_2)\delta_1^2}.$$

An analogous derivation for the other decision point gives:

$$P(\omega_2)\left(\frac{\delta_2-(x_2^*-\mu_2)}{\delta_2^2}\right)=P(\omega_3)\left(\frac{\delta_3-(x_2^*-\mu_3)}{\delta_3^2}\right),$$

which has solution

$$x_2^*=\frac{-P(\omega_2)\delta_3^2\mu_2+P(\omega_2)\delta_3^2\delta_2+P(\omega_3)\delta_2^2\delta_3+P(\omega_3)\delta_2^2\mu_3}{P(\omega_2)\delta_3^2+P(\omega_3)\delta_2^2}.$$

(b) Note that from our normalization condition, $\sum_{i=1}^{3} P(\omega_i)=1$, we can express all priors in terms of just *two* independent ones, which we choose to be $P(\omega_1)$ and $P(\omega_2)$. We could substitute the values for x_i^* and integrate, but we choose instead to go directly to the calculation of the error, E, as a function of priors $P(\omega_1)$ and $P(\omega_2)$, by considering the four contributions:

$$\begin{aligned}
E &= P(\omega_1)\frac{1}{2\delta_1^2}\left[\mu_1+\delta_1-x_1^*\right]^2 \\
&\quad +P(\omega_2)\frac{1}{2\delta_2^2}\left[\delta_2-\mu_2+x_1^*\right]^2 \\
&\quad +P(\omega_2)\frac{1}{2\delta_2^2}\left[\mu_2+\delta_2-x_2^*\right]^2 \\
&\quad +\big[\underbrace{1-P(\omega_1)-P(\omega_2)}_{P(\omega_3)}\big]\frac{1}{2\delta_3^2}\left[\delta_3-\mu_3+x_2^*\right]^2
\end{aligned}$$

To obtain the minimax solution, we take the two partial and set them to zero. The first of the derivative equations,

$$\frac{\partial E}{\partial P(\omega_1)} = 0,$$

yields the equation

$$\begin{aligned}\left(\frac{\mu_1+\delta_1-x_1^*}{\delta_1}\right)^2 &= \left(\frac{\delta_3-\mu_3+x_2^*}{\delta_3}\right)^2 \quad \text{or}\\ \frac{\mu_1+\delta_1-x_1^*}{\delta_1} &= \frac{\delta_3-\mu_3+x_2^*}{\delta_3}.\end{aligned}$$

Likewise, the second of the derivative equations,

$$\frac{\partial E}{\partial P(\omega_2)} = 0,$$

yields the equation

$$\left(\frac{\delta_2-\mu_2+x_1^*}{\delta_2}\right)^2 + \left(\frac{\mu_2+\delta_2-x_2^*}{\delta_2}\right)^2 = \left(\frac{\delta_3-\mu_3+x_2^*}{\delta_3}\right)^2.$$

These simultaneous quadratic equations have solutions of the general form:

$$x_i^* = \frac{b_i+\sqrt{c_i}}{a_i} \qquad i=1,2.$$

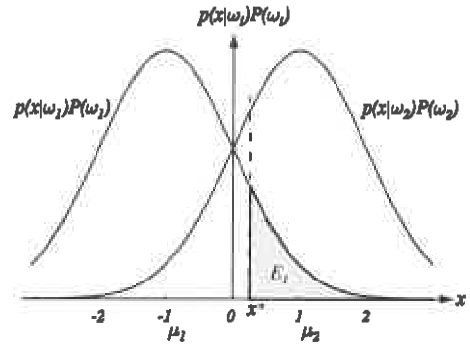
After a straightforward, but very tedious calculation, we find that:

$$\begin{aligned}
a_1 &= \delta_1^2-\delta_2^2+\delta_3^2,\\
b_1 &= -\delta_1^2\delta_2-\delta_1\,\delta_2^2-\delta_1\delta_2\;\delta_3-\delta_2^2\mu_1+\delta_3^2\mu_1+\delta_1^2\mu_2-\delta_1\delta_3\mu_2+\delta_1\delta_3\mu_3,\\
c_1 &= \delta_1^2\big(2\delta_1\delta_2^3+2\delta_2^4+2\delta_1\delta_2^2\delta_3+2\delta_2^3\delta_3+\delta_1\delta_2^2\mu_1+2\delta_2^3\mu_1\\
&\quad +2\,\delta_1\,\delta_2\,\delta_3\,\mu_1-2\,\delta_2\,{\delta_3}^2\,\mu_1+{\delta_2}^2\,{\mu_1}^2-{\delta_3}^2\,{\mu_1}^2-2\,{\delta_1}^2\,\delta_2\,\mu_2-2\,\delta_1\,{\delta_2}^2\,\mu_2\\
&\quad +2\,{\delta_2}^2\,\delta_3\,\mu_2+2\,\delta_2\,{\delta_3}^2\,\mu_2-2\,{\delta_2}^2\,\mu_1\,\mu_2+2\,\delta_1\,\delta_3\,\mu_1\,\mu_2+2\,{\delta_3}^2\,\mu_1\,\mu_2\\
&\quad -{\delta_1}^2\,{\mu_2}^2+2\,{\delta_2}^2\,{\mu_2}^2-2\,\delta_1\,\delta_3\,{\mu_2}^2-{\delta_3}^2\,{\mu_2}^2+2\,{\delta_1}^2\,\delta_2\,\mu_3-2\,{\delta_2}^3\,\mu_3\\
&\quad -2\,\delta_1\,\delta_2\,\delta_3\,\mu_3-2\,{\delta_2}^2\,\delta_3\,\mu_3-2\,\delta_1\,\delta_3\,\mu_1\,\mu_3+2\,{\delta_1}^2\,\mu_2\,\mu_3-2\,{\delta_2}^2\,\mu_2\,\mu_3\\
&\quad +2\,\delta_1\,\delta_3\,\mu_2\,\mu_3-{\delta_1}^2\,{\mu_3}^2+{\delta_2}^2\,{\mu_3}^2\big).
\end{aligned}$$

An analogous calculation gives:

$$\begin{aligned}
a_2 &= \delta_1^2-\delta_2^2+\delta_3^2,\\
b_2 &= \delta_1\,\delta_2\,\delta_3+\delta_2^2\,\delta_3+2\,\delta_2\,\delta_3^2+\delta_1\,\delta_3\,\mu_1-\delta_1\,\delta_3\,\mu_2+\delta_3^2\,\mu_2+\delta_1^2\,\mu_3-\delta_2^2\,\mu_3,\\
c_2 &= (\delta_1^2-\delta_2^2+\delta_3^2)\times\\
&\qquad \big(\delta_2^2\,\delta_3^2+2\,\delta_2\,\delta_3^2\,\mu_1+\delta_3^2\,\mu_1^2-2\,\delta_3^2\,\mu_1\,\mu_2+2\,\delta_3^2\,\mu_2^2+2\,\delta_1\,\delta_2\,\delta_3\,\mu_3\\
&\qquad +2\,\delta_2^2\,\delta_3\,\mu_3+2\,\delta_1\,\delta_3\,\mu_1\,\mu_3-2\,\delta_1\,\delta_3\,\mu_2\,\mu_3+\delta_1^2\,\mu_3^2-\delta_2^2\,\mu_3^2\big).
\end{aligned}$$

(c) For $\{\mu_i,\delta_i\} = \{0,1\},\{.5,.5\},\{1,1\}$, for $i=1,2,3$, respectively, we substitute into the above equations to find $x_1^*=0.2612$ and $x_2^*=0.7388$. It is a simple matter to confirm that indeed these two decision points suffice for the classification problem, i.e., that no more than two points are needed.



5 6. We let x^* denote our decision boundary and $\mu_2 > \mu_1$, as shown in the figure.

(a) The error for classifying a pattern that is actually in ω_1 as if it were in ω_2 is:

$$\int_{\mathcal{R}_2} p(x|\omega_1)P(\omega_1)\ dx = \frac{1}{2}\int_{x^*}^{\infty} N(\mu_1, \sigma_1^2)\ dx \leq E_1.$$

Our problem demands that this error be less than or equal to E_1. Thus the bound on x^* is a function of E_1, and could be obtained by tables of cumulative normal distributions, or simple numerical integration.

(b) Likewise, the error for categorizing a pattern that is in ω_2 as if it were in ω_1 is:

$$E_2 = \int_{\mathcal{R}_1} p(x|\omega_2)P(\omega_2)\ dx = \frac{1}{2}\int_{-\infty}^{x^*} N(\mu_2, \sigma_2^2)\ dx.$$

(c) The total error is simply the sum of these two contributions:

$$\begin{aligned} E &= E_1 + E_2 \\ &= \frac{1}{2}\int_{x^*}^{\infty} N(\mu_1, \sigma_1^2)\ dx + \frac{1}{2}\int_{-\infty}^{x^*} N(\mu_2, \sigma_2^2)\ dx. \end{aligned}$$

(d) For $p(x|\omega_1) \sim N(-1/2, 1)$ and $p(x|\omega_2) \sim N(1/2, 1)$ and $E_1 = 0.05$, we have (by simple numerical integration) $x^* = 0.2815$, and thus

$$\begin{aligned} E &= 0.05 + \frac{1}{2}\int_{-\infty}^{0.2815} N(\mu_2, \sigma_2^2)\ dx \\ &= 0.05 + \frac{1}{2}\int_{-\infty}^{0.2815} \frac{1}{\sqrt{2\pi}0.05}\exp\left[-\frac{(x-0.5)^2}{2(0.5)^2}\right] dx \\ &= 0.168. \end{aligned}$$

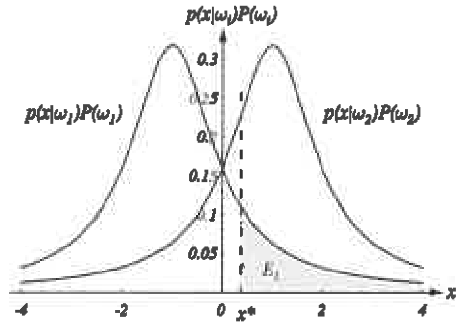
(e) The decision boundary for the (minimum error) Bayes case is clearly at $x^* = 0$. The Bayes error for this problem is:

$$E_B = 2\int_{0}^{\infty} \frac{1}{2}N(\mu_1, \sigma_1^2)\ dx$$

$$= \int_0^{\infty} N(1,1)\, dx = \text{erf}[1] = 0.159,$$

which of course is lower than the error for the Neyman-Pearson criterion case. Note that if the Bayes error were lower than $2 \times 0.05 = 0.1$ in this problem, we would use the Bayes decision point for the Neyman-Pearson case, since it too would ensure that the Neyman-Pearson criteria were obeyed *and* would give the lowest total error.

7. We proceed as in Problem 6, with the figure below.



(a) Recall that the Cauchy density is

$$p(x|\omega_i) = \frac{1}{\pi b} \frac{1}{1 + \left(\frac{x-a_i}{b}\right)^2}.$$

If we denote our (single) decision boundary point as x^*, and note that $P(\omega_i) = 1/2$, then the error for misclassifying a ω_1 pattern as ω_2 is:

$$\begin{aligned} E_1 &= \int_{x^*}^{\infty} p(x|\omega_1)P(\omega_1)\, dx \\ &= \frac{1}{2} \int_{x^*}^{\infty} \frac{1}{\pi b} \frac{1}{1 + \left(\frac{x-a_1}{b}\right)^2}\, dx. \end{aligned}$$

We substitute $(x - a_1)/b = y$, and $\sin\theta = 1/\sqrt{1+y^2}$ to get:

$$\begin{aligned} E_1 &= \frac{1}{2\pi} \int_{\theta=\tilde{\theta}}^{\theta=0} d\theta \\ &= \frac{1}{2\pi} \sin^{-1}\left[\frac{b}{\sqrt{b^2 + (x^* - a_1)^2}}\right], \end{aligned}$$

where $\tilde{\theta} = \sin^{-1}\left[\frac{b}{\sqrt{b^2+(x^*-a_1)^2}}\right]$. Solving for the decision point gives

$$x^* = a_1 + b\sqrt{\frac{1}{\sin^2[2\pi E_1]} - 1} = a_1 + b/\tan[2\pi E_1].$$

(b) The error for the converse case is found similarly:

$$\begin{aligned} E_2 &= \int_{-\infty}^{x^*} \frac{1}{\pi b} \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2} P(\omega_2)\, dx \\ &= \frac{1}{2\pi} \int_{\theta=-\pi}^{\theta=\tilde{\theta}} d\theta \\ &= \frac{1}{2\pi} \left\{ \sin^{-1}\left[\frac{b}{\sqrt{b^2+(x^*-a_2)^2}}\right] + \pi \right\} \\ &= \frac{1}{2} + \frac{1}{\pi} \sin^{-1}\left[\frac{b}{\sqrt{b^2+(x^*-a_2)^2}}\right], \end{aligned}$$

where $\tilde{\theta}$ is defined in part (a).

(c) The total error is merely the sum of the component errors:

$$E = E_1 + E_2 = E_1 + \frac{1}{2} + \frac{1}{\pi}\sin^{-1}\left[\frac{b}{\sqrt{b^2+(x^*-a_2)^2}}\right],$$

where the numerical value of the decision point is

$$x^* = a_1 + b/\tan\,(2\pi E_1) = 0.376.$$

(d) We add the errors (for $b=1$) and find

$$E = 0.1 + \frac{1}{2} + \frac{1}{\pi}\sin^{-1}\left[\frac{b}{\sqrt{b^2+(x^*-a_2)^2}}\right] = 0.2607.$$

(e) For the Bayes case, the decision point is midway between the peaks of the two distributions, i.e., at $x^*=0$ (cf. Problem 6). The Bayes error is then

$$E_B = 2\int_0^{\infty} \frac{1}{1+\left(\frac{x-a}{b}\right)^2} P(\omega_2)\, dx = 0.2489.$$

This is indeed lower than for the Neyman-Pearson case, as it must be. Note that if the Bayes error were lower than $2 \times 0.1 = 0.2$ in this problem, we would use the Bayes decision point for the Neyman-Pearson case, since it too would ensure that the Neyman-Pearson criteria were obeyed *and* would give the lowest total error.

8. Consider the Cauchy distribution.

(a) We let k denote the integral of $p(x|\omega_i)$, and check the normalization condition, i.e., whether $k=1$:

$$k = \int_{-\infty}^{\infty} p(x|\omega_i)\, dx = \frac{1}{\pi b}\int_{-\infty}^{\infty} \frac{1}{1+\left(\frac{x-a_i}{b}\right)^2}\, dx.$$

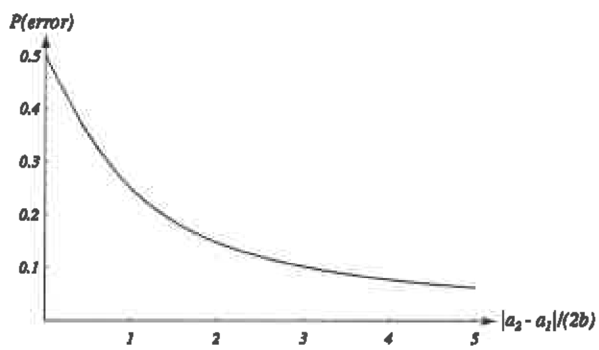
(a) Without loss of generality, we assume that $a_2 > a_1$, note that the decision boundary is at $(a_1 + a_2)/2$. The probability of error is given by

$$
\begin{aligned}
P(error) &= \int\limits_{-\infty}^{(a_1+a_2)/2} p(\omega_2|x)dx + \int\limits_{(a_1+a_2)/2}^{\infty} p(\omega_1|x)dx \\
&= \frac{1}{\pi b} \int\limits_{-\infty}^{(a_1+a_2)/2} \frac{1/2}{1+\left(\frac{x-a_2}{b}\right)^2} dx + \frac{1}{\pi b} \int\limits_{(a_1+a_2)/2}^{\infty} \frac{1/2}{1+\left(\frac{x-a_1}{b}\right)^2} dx \\
&= \frac{1}{\pi b} \int\limits_{-\infty}^{(a_1-a_2)/2} \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2} dx = \frac{1}{\pi} \int\limits_{-\infty}^{(a_1-a_2)/2} \frac{1}{1+y^2} dy,
\end{aligned}
$$

where for the last step we have used the trigonometric substitution $y = (x-a_2)/b$ as in Problem 8. The integral is a standard form for $\tan^{-1}y$ and thus our solution is:

$$
\begin{aligned}
P(error) &= \frac{1}{\pi}\left[\tan^{-1}\left|\frac{a_1-a_2}{2b}\right| - \tan^{-1}[-\infty]\right] \\
&= \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\left|\frac{a_2-a_1}{2b}\right|.
\end{aligned}
$$

(b) SEE FIGURE.



(c) The maximum value of the probability of error is $P_{max}(\frac{a_2-a_1}{2b}) = 1/2$, which occurs for $|\frac{a_2-a_1}{2b}| = 0$. This occurs when either the two distributions are the same, which can happen because $a_1 = a_2$, or even if $a_1 \neq a_2$ because $b = \infty$ and both distributions are flat.

10. We use the fact that the conditional error is

$$
P(error|x) = \begin{cases} P(\omega_1|x) & \text{if we decide } \omega_2 \\ P(\omega_2|x) & \text{if we decide } \omega_1. \end{cases}
$$

(a) Thus the decision as stated leads to:

$$
P(error) = \int\limits_{-\infty}^{\infty} P(error|x)p(x)dx.
$$

Thus we have

$$
\begin{aligned}
P(error) &= P(x < \theta \text{ and } \omega_1 \text{ is the true state}) \\
&\quad + P(x > \theta \text{ and } \omega_2 \text{ is the true state}) \\
&= P(x < \theta|\omega_1)P(\omega_1) + P(x > \theta|\omega_2)P(\omega_2) \\
&= P(\omega_1)\int_{-\infty}^{\theta} p(x|\omega_1)\, dx + P(\omega_2)\int_{\theta}^{\infty} p(x|\omega_2)\, dx.
\end{aligned}
$$

(b) We take a derivative with respect to θ and set it to zero to find an extremum, that is,

$$\frac{dP(error)}{d\theta} = P(\omega_1)p(\theta|\omega_1) - P(\omega_2)p(\theta|\omega_2) = 0,$$

which yields the condition

$$P(\omega_1)p(\theta|\omega_1) = P(\omega_2)p(\theta|\omega_2),$$

where we have used the fact that $p(x|\omega_i) = 0$ at $x \to \pm\infty$.

(c) No, this condition does not uniquely define θ.

1. If $P(\omega_1)p(\theta|\omega_1) = P(\omega_2)p(\theta|\omega_2)$ over a *range* of θ, then θ would be unspecified throughout such a range.
2. There can easily be multiple values of x for which the condition hold, for instance if the distributions have the appropriate multiple peaks.

(d) If $p(x|\omega_1) \sim N(1,1)$ and $p(x|\omega_2) \sim N(-1,1)$ with $P(\omega_1) = P(\omega_2) = 1/2$, then we have a *maximum* for the error at $\theta = 0$.

11. The deterministic risk is given by Bayes' Rule and Eq. 20 in the text

$$R = \int R(\alpha_i(\mathbf{x})|\mathbf{x})\, d\mathbf{x}.$$

(a) In a random decision rule, we have the *probability* $P(\alpha_i|\mathbf{x})$ of deciding to take action α_i. Thus in order to compute the full probabilistic or randomized risk, R_{ran}, we must integrate over all the conditional risks weighted by their probabilities, i.e.,

$$R_{ran} = \int \left[\sum_{i=1}^{a} R(\alpha_i(\mathbf{x})|\mathbf{x})P(\alpha_i|\mathbf{x})\right] p(\mathbf{x})\, d\mathbf{x}.$$

(b) Consider a fixed point $\mathbf{x}$ and note that the (deterministic) Bayes minimum risk decision at that point obeyes

$$R(\alpha_i(\mathbf{x})|\mathbf{x}) \geq R(\alpha_{max}(\mathbf{x})|\mathbf{x}).$$