Northeastern University

EECE 5564 Department of Electrical and Computer Engineering

Solution 1

1.1 (10 pts) Let x be a real-valued random variable.

(a) Prove that the variance of $x = \sigma^2 = E[(x-\mu)^2] = E[x^2] - \mu^2$.

(b) Let x be a real-valued random vector. Prove that the covariance matrix of $x = \Sigma = E[xx^T] - \mu\mu^T$.

Solution:

(a)

$$\begin{aligned}
&Var(x) = E[(x-\mu)^2] = E[x^2 + \mu^2 - 2\mu x] \\
&\mu = E[x], linearity \\
&= E[x^2] + \mu^2 - 2\mu E[x] \\
&= E[x^2] + \mu^2 - 2\mu^2 = E[x^2] - \mu^2
\end{aligned}$$

(b)

$$\begin{aligned}
\Sigma &= E[(x-\mu)(x-\mu)^T] = E[(xx^T - x\mu^T - \mu x^T + \mu\mu^T)] \\
&= E[xx^T] - E[x]\mu^T - \mu E[x]^T + \mu\mu^T = E[xx^T] - \mu\mu^T
\end{aligned}$$

1.2 (10 pts) Suppose two equally probable one-dimensional densities are of the form $p(x|\omega_i) \propto e^{-|x-a_i|/b_i}$ for $i = 1, 2$ and $b > 0$.

(a) Write an analytic expression for each density, that is, normalize each function for arbitrary a_i, and positive b_i.

(b) Calculate the likelihood ratio $p(x|\omega_1)/p(x|\omega_2)$ as a function of your four variables.

(c) Plot a graph (using MATLAB) of the likelihood ratio for the case $a_1 = 0$, $b_1 = 1$, $a_2 = 1$ and $b_2 = 2$. Make sure the plots are correctly labeled (axis, titles, legend, etc) and that the fonts are legible when printed.

Solution:

(a) It is given that $p(x|\omega_i) = Ke^{-|x-a_i|/b_i}$, with $K > 0$. Using the normalization axioms, we have that:

$$\int_{-\infty}^{+\infty} p(x|\omega_i)dx = 1$$
$$\int_{-\infty}^{+\infty} Ke^{-|x-a_i|/b_i}dx = 1$$
$$\int_{-\infty}^{a_i} Ke^{(x-a_i)/b_i}dx + \int_{a_i}^{+\infty} Ke^{(a_i-x)/b_i}dx = 1$$
$$Kb_ie^{(x-a_i)/b_i}\Big|_{-\infty}^{a_i} - Kb_ie^{(a_i-x)/b_i}\Big|_{a_i}^{+\infty} = 1$$
$$K = \frac{1}{2b_i}$$

Therefore, $p(x|\omega_i) = \frac{1}{2b_i}e^{-|x-a_i|/b_i}$

(b)

$$\frac{p(x|\omega_1)}{p(x|\omega_2)} = \frac{b_2}{b_1}e^{\frac{|x-a_2|}{b_2}-\frac{|x-a_1|}{b_1}}$$

(c) Plug values into MATLAB and plot. See figure 1. The MATLAB code is attached separately.

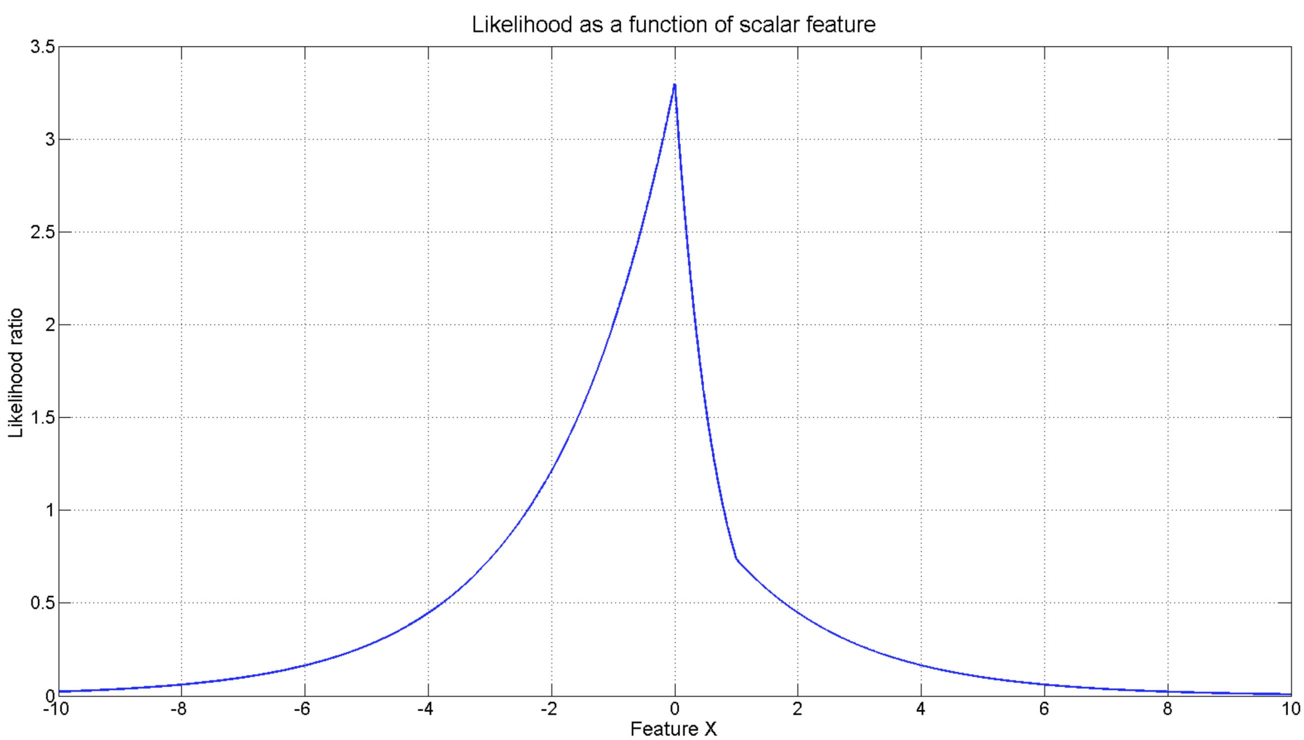

Figure 1: Likelihood plot for problem 2.c

1.3 (10 pts) Consider a two-class problem, with classes $c1$ and $c2$ where $P(c1) = P(c2) = 0.5$. There is a one-dimensional feature variable x. Assume that the x data for class one is uniformly distributed between a and b, and the x data for class two is uniformly distributed between r and t. Assume that $a < r < b < t$. Derive a general expression for the Bayes error rate for this problem. (Hint: a sketch may help you think about the solution.) **Solution:**

$$P(x|w_1) = \frac{1}{b-a} and P(x|w_2) = \frac{1}{t-r}$$

for the case where $\frac{1}{b-a} > \frac{1}{t-r}$

$$P(error|x) = \int_r^b \frac{1}{t-r}\frac{1}{2}dx = \frac{b-r}{2(t-r)}$$

and for the case where $\frac{1}{b-a} < \frac{1}{t-r}$

$$P(error|x) = \int_r^b \frac{1}{b-a}\frac{1}{2}dx = \frac{b-r}{2(b-a)}$$

1.4 (12 pts) Consider a two-class, one-dimensional problem where $P(\omega_1) = P(\omega_2)$ and $p(x|\omega_i) \sim N(\mu_i, \sigma_i^2)$. Let $\mu_1 = 0$, $\sigma_1^2 = 1$, $\mu_2 = \mu$, and $\sigma_2^2 = \sigma^2$.

(a) Derive a general expression for the location of the Bayes optimal decision boundary as a function of μ and σ^2.

(b) With $\mu = 1$ and $\sigma^2 = 2$, make two plots using MATLAB: one for the class conditional pdfs $p(x|\omega_i)$ and one for the posterior probabilities $p(\omega_i|x)$ with the location of the optimal decision regions. Make sure the plots are correctly labeled (axis, titles, legend, etc) and that the fonts are legible when printed.

(c) Estimate the Bayes error rate p_e.

(d) Comment on the case where $\mu = 0$, and σ^2 is much greater than 1. Describe a practical example of a pattern classification problem where such a situation might arise.

Solution:

(a) Since the priors are equal, the optimal decision boundaries occur when the likelihoods intersect each other (are equal). Let x_o be the optimal decision boundary, then we have that:

$$\begin{aligned}
p(x_o|\omega_1) &= p(x_o|\omega_2) \\
\ln p(x_o|\omega_1) &= \ln p(x_o|\omega_2) \\
\ln G(x_o; \mu_1 = 0, \sigma_1^2 = 1) &= \ln G(x_o; \mu_2 = \mu, \sigma_2^2 = \sigma^2) \\
-\frac{1}{2}x_o^2 &= -\frac{1}{2}\frac{(x_o-\mu)^2}{\sigma^2} - \ln\sigma \\
(1-\sigma^2)x_o^2 - 2x_o\mu + \mu^2 + 2\sigma^2\ln\sigma &= 0
\end{aligned}$$

Solving the quadratic equation, we get:

$$\begin{aligned} x_o &= \frac{2\mu \pm \sqrt{4\mu^2 - 4(1-\sigma^2)(\mu^2 + 2\sigma^2 \ln \sigma)}}{2(1-\sigma^2)} \\ &= \frac{-\mu \pm \sqrt{\mu^2 + (\sigma^2 - 1)(\mu^2 + 2\sigma^2 \ln \sigma)}}{\sigma^2 - 1} \end{aligned}$$

(b) The optimal points are given by $x_o = -2.8402, 0.8402$. Plots are in figure 2:

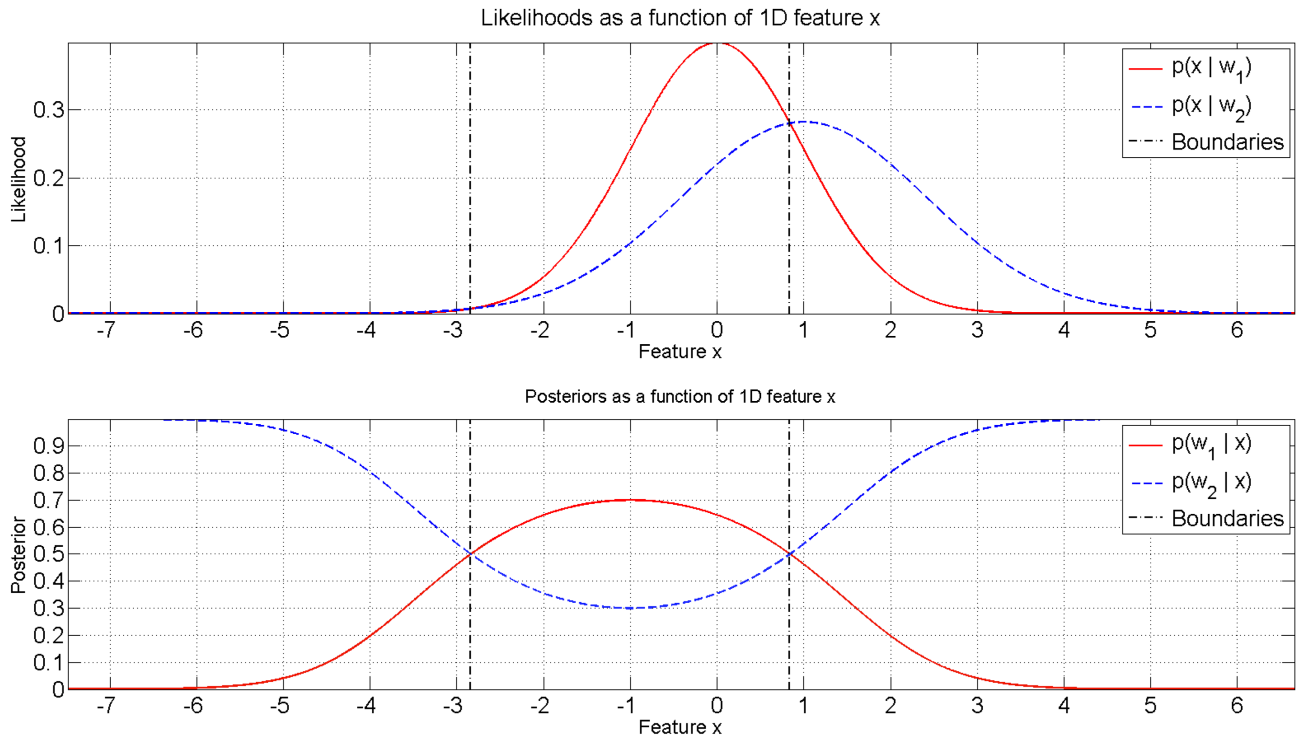

Figure 2: Conditional pdf and posterior probabilities for problem 4.b

(c) With the posteriors, we can perform the integration of $p_e = \int_{-\infty}^{+\infty} \min[p(\omega_1|x), p(\omega_2|x)]p(x)dx$ with MATLAB, obtaining $p_e \approx 0.32718$.

(d) An example for this case (same mean, different variances) would be to classify objects according to their size with respect to a conveyor belt. Or zero mean signals of different variance (check for presence of signal since variance is related to power).