Bayes Error: (Problem 5a - Marjan Rezvani)

Let's assume we have a vector $\underline{x}$ needs to be classified into on of L classes.

Let $p(C_i)$ denote the prior class probability of class i

$1 \leq i \leq L$

Let $p(x|C_i)$ denote the class likelihood, $\Longrightarrow$ In other words, the conditional probability density of x given it belongs to class i.

Using Bayes Rule: Posterior probability $P(C_i|x)$ is given by

$$p(C_i|x) = \frac{P(x|C_i)\,P(C_i)}{P(x)} \quad \text{where } p(x) = \sum_{i=1}^{L} p(x|C_i)\,p(C_i)$$

The classifier that assigns a vector x to the class with the highest posterior is called the Bayes classifier.

The error associated with this classifier is called Bayes error which can be expressed as

$$\text{error} = 1 - \sum_{i=1}^{L} \int_{C_i} p(C_i)\,p(x|C_i)\,dx \qquad (*)$$

where C_i is the region where class i has the highest posterior.

In our problem, we have:

$L = 2$ (Class A, Class B)

$p(x \mid c_i) \sim N(\mu, \sigma^2)$

To be more specific

$p(x \mid c_A) \sim N(\mu_A, \sigma_A^2)$, $p(x \mid c_B) \sim N(\mu_B, \sigma_B^2)$

Using equation (*)

$$\text{error} = 1 - \sum_{i=1}^{2} \int_{c_i} p(c_i)\, p(x \mid c_i)\, dx$$

$$= 1 - \left(\int_{c_A} p(c_A) * \frac{1}{\sqrt{2\pi\sigma_A^2}} e^{-\frac{(x-\mu_A)^2}{2\sigma_A^2}}\, dx \right.$$

$$\left. + \int_{c_B} p(c_B) * \frac{1}{\sqrt{2\pi\sigma_B^2}} e^{-\frac{(x-\mu_B)^2}{2\sigma_B^2}}\, dx \right) \quad (\square)$$

Regarding prior: In the question, it says " Assume that there are equally many beans of each type (no prior"

So, I assume $p(c_A) = p(c_B) = \frac{1}{2}$

Then equation ($\square$) becomes (Assume if $x \le T \rightarrow$ class A region
$x > T \rightarrow$ 〃 B 〃

$$error = 1 - \left(\int_{-\infty}^{T} \frac{1}{2} \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_A)^2}{2\sigma_A^2}} dx + \int_{T}^{+\infty} \frac{1}{2} \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_B)^2}{2\sigma_B^2}} dx \right)$$