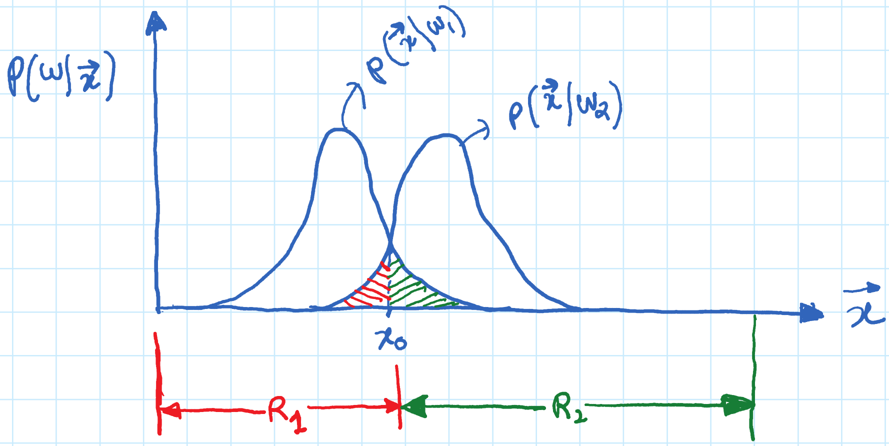
• **Decision Error:**

Let P_e be the total probability of committing a decison error.



For two equiprobable classes, ie, $p(\omega_1) = p(\omega_2) = \frac{1}{2}$,

$$P_e = \frac{1}{2}\int_{-\infty}^{x_0} p(\vec{x}|\omega_2)\, d\vec{x} + \frac{1}{2}\int_{x_0}^{+\infty} p(\vec{x}|\omega_1)\, d\vec{x} \quad (6)$$

P_e is equal to the total shaded area in the above figure.

For any values of $p(\omega_1)$ and $p(\omega_2)$, the above expression can be written as:

$$P_e = \int_{R_2} p(\omega_1)\, p(\vec{x}|\omega_1)\, d\vec{x} + \int_{R_1} p(\omega_2)\, p(\vec{x}|\omega_2)\, d\vec{x} \quad (7)$$

• **Minimizing the Average Risk:**

In practice, there are many applications where falsely classifying a feature vector that belongs to class-A to class-B needs to be penalized differently than falsely classifying a feature vector of class-B to class-A.

For example let ω_1 and ω_2 represent malignant tumor and beningn tumor classes respectively. In such cases, falsely classifying a malignant tumor as a benign tumor has to be penalized more compared to falsely classifying a beningn tumor as malignant tumor.

Let λ_{ij} represent the penalty associated with classifying a feature vector of ith class as jth class.

Then the risk function for $M=2$ is given by:

$$r = \lambda_{12} P(\omega_1) \int_{R_2} p(\vec{x}|\omega_1)\, d\vec{x} + \lambda_{21} P(\omega_2) \int_{R_1} p(\vec{x}|\omega_2)\, d\vec{x} \quad — (8)$$

For the above mentioned scenario of malignant and benign tumors, in order to minimize the risk of falsely classifying a malignant tumor, we can choose $\lambda_{12} > \lambda_{21}$

Generalizing the risk function to M-classes:

let r_k be the risk (penalty) associated with wrongly assigning feature vectors of k^{th} class to other classes.

$$r_k = P(\omega_k) \sum_{i=1}^{M} \int_{R_i} p(\vec{x}|\omega_k)\, d\vec{x} \quad — (9)$$

$\lambda_{ki} = 0$, for $k=i$

Total penalty across all classes is given by:

$$r = \sum_{k=1}^{M} r_k = \sum_{k=1}^{M} P(\omega_k) \left(\sum_{i=1}^{M} \lambda_{ki} \int_{R_i} p(\vec{x}|\omega_k)\, d\vec{x} \right)$$

Interchanging the summations in the above equation $\Rightarrow$

$$r = \sum_{i=1}^{M} \int_{R_i} \left(\sum_{k=1}^{M} \lambda_{ki}\, p(\vec{x}|\omega_k)\, P(\omega_k) \right) d\vec{x} \quad — (10)$$

The goal here is to choose partitionings (R_i) such that the penalty function: r is minimized.

let l_i be defined as: $l_i \triangleq \sum_{k=1}^{M} \lambda_{ki}\, p(\vec{x}|\omega_k)\, P(\omega_k) \quad — (11)$

Minimizing Eq — (10) is equivalent to selecting partitioning region such that

$$\vec{x} \in R_i \quad \text{if} \quad l_i < l_j \,, \quad \forall j \neq i \quad — (12)$$