Bayesian Classification Part-2

2.2 Bayes Decison Theory:

* Let us first focus on two-class classification problem. Let ω_1, ω_2 represent the two possible classes.

* **Bayes Rule:**

$$P(\omega_i|\vec{x})\, P(\vec{x}) = P(\vec{x}|\omega_i)\, P(\omega_i)$$

$$\Rightarrow P(\omega_i|\vec{x}) = \frac{P(\vec{x}/\omega_i)\, P(\omega_i)}{P(\vec{x})} \quad \text{——} (1)$$

- $P(\vec{x})$ is the probability density function (pdf) of $\vec{x}$

$$P(\vec{x}) = \sum_{i=1}^{2} P(\vec{x}|\omega_i)\, P(\omega_i) \quad \text{——} (2)$$

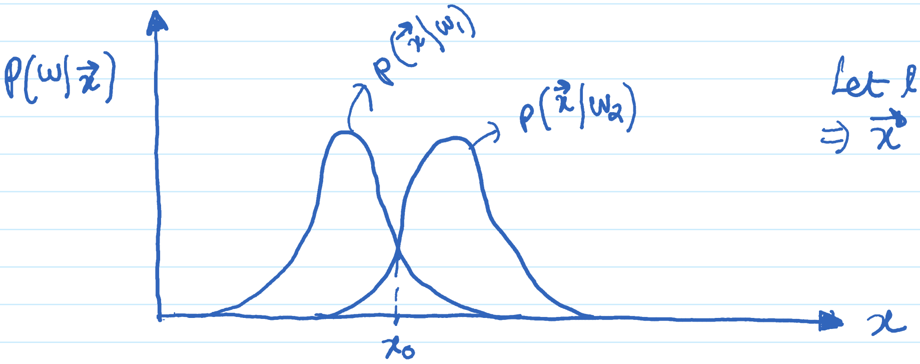
- $P(\omega_1), P(\omega_2)$ are the a priori probabilities of class-1 and class-2.

  It is assumed that these values are known.
  Even if not known, they can be computed from the training data.
  Let N_1 and N_2 be the number of samples in the training data that belong to ω_1 and ω_2

$$P(\omega_1) \approx \left(\frac{N_1}{N_1+N_2}\right); \quad P(\omega_2) \approx \left(\frac{N_2}{N_1+N_2}\right)$$

- $P(\vec{x}|\omega_i)$: It describes the distribution of feature vectors of each class.

  It can also be estimated from the training data.
  It is referred to as likelihood function of ω_i with respect to $\vec{x}$.



Let $l = 1$
$\Rightarrow \vec{x}$ is of size 1×1.

- **Bayes Classification Rule:**

If $P(\omega_1|\vec{x}) > P(\omega_2|\vec{x}) \Rightarrow \vec{x}$ is assigned to ω_1 class.
else, $\vec{x}$ is assigned to ω_2 class. } — (3)

substituting Eq-(3) in Eq-(1) $\Rightarrow$

If $P(\vec{x}|\omega_1) \cdot P(\omega_1) > P(\vec{x}|\omega_2)\, P(\omega_2) \Rightarrow \vec{x} \in \omega_1$
else, $\vec{x} \in \omega_2$ } — (4)

Special Case: prior probabilities of all classes are equal

$\Rightarrow P(\omega_1) = P(\omega_2)$

substituting it in Eq-(4) $\Rightarrow$

If $P(\vec{x}|\omega_1) > P(\vec{x}|\omega_2) \Rightarrow \vec{x} \in \omega_1$.
else, $\vec{x} \in \omega_2$. } — (5)