Now lets consider the Bayes Classifier with M classes & any loss func

$Y \in \{c_0, c_1, \ldots, c_{M-1}\}$ True labels.

$h(x) \in \{\alpha_0, \alpha_1, \ldots, \alpha_{k-1}\}$ Decisions from the hypothesis func

in general $M \neq k$

$$L : h(x) \times Y \rightarrow \mathbb{R}^+$$

$L(\alpha_j, c_t)$ loss when $h(x)$ is α_j & true label is c_t

Risk of the classifier is

$$R(h) = \mathbb{E}_{XY}\left[L(h(x), Y)\right] \qquad h \in \mathcal{H}$$

We want a classifier with least risk value.

$$R(h) = \mathbb{E}_{XY}\left[L(h(x), Y)\right]$$

$$= \mathbb{E}_{X}\Big[\underbrace{\mathbb{E}_{Y|X}\left[L(h(x), Y) \mid X\right]}_{\text{Conditional Risk}}\Big]$$

$$= \int_{X} R(h(x) \mid x)\, p_X(x)\, dx$$

Consider x

$h(x)$ is α_i

$R(\alpha_i | x)$ denotes expected loss when classifier says α_i

$$R(\alpha_i | x) = \mathbb{E}_{Y|X} \left[L(h(x), Y) \mid h(x) = \alpha_i, x \right]$$

$$= \mathbb{E}_{Y|X} \left[L(\alpha_i, Y) \mid X = x \right]$$

$$= \sum_{j=0}^{M-1} L(\alpha_i, c_j) \, p_{Y|X}(Y = c_j | x)$$

$p_{Y|X}$: posterior

p_Y : prior

$p_{X|Y}$: Class conditional

In general we have

$$R(h(x) | x) = \sum_{j=0}^{M-1} L(h(x), c_j) \, p_{Y|X}(Y = c_j \mid X = x)$$

$\therefore$ optimal classifier

for each x, $h(x)$ should be such that

$$R(h(x)|x) \leq R(h'(x)|x) \quad \forall h' \in \mathcal{H}$$

The Bayes classifier h_B for M-class is

$h_B(x) = \alpha_i$ if

$$\sum_{j=0}^{M-1} L(\alpha_i, c_j)\, p_{Y|X}(y=c_j|x) \leq \sum_{j=0}^{M-1} L(\alpha_t, c_j)\, p_{Y|X}(y=c_j|x) \quad \forall t$$

Break ties arbitrarily.

Lets consider $M=2$

$Y \in \{c_0, c_1\}$

$h(x) \in \{\alpha_0, \alpha_1\}$

$h_B(x) = \alpha_0$ if

$$L(\alpha_0, c_0)\, p_{Y|X}(y=c_0|x) + L(\alpha_0, c_1)\, p_{Y|X}(y=c_1|x) \leq$$

$$L(\alpha_1, c_0)\, p_{Y|X}(y=c_0|x) + L(\alpha_1, c_1)\, p_{Y|X}(y=c_1|x)$$

Consider 0-1 Loss

$$L(\alpha_0, c_0) = L(\alpha_1, c_1) = 0$$

$$L(0,1)\, p_{y=1|x} \leq L(1,0)\, p_{y=0|x}$$

$$\frac{p_{y=0|x}}{p_{y=1|x}} \geq \frac{L(0,1)}{L(1,0)}$$

Now lets generalize this to M-Class Case

& 0-1 Loss func

$$R(\alpha_i|x) = \sum_{j=0}^{M-1} L(\alpha_i, c_j)\, p_{y|x}(y=c_j|x)$$

$$= \sum_{j\neq i} p_{y|x}(y=c_j|x)$$

$$= 1 - p_{y|x}(y=c_i|x)$$

$\therefore$ Bayes Classifier $h_B(x) = \alpha_i$

$$\left(1 - p_{y|x}(y=c_i|x)\right) \leq \left(1 - p_{y|x}(y=c_j|x)\right) \quad \forall j$$

$$p_{y|x}(y=c_i|x) \geq p_{y|x}(y=c_j|x)$$

This is the M-class classifier for 0-1 loss func

Minimizer of the prob of misclassification.

Now lets consider $M = 2$

$$L(0,1)\, p_{y=1|x} \leq L(1,0)\, p_{y=0|x}$$

$$L(0,1)\, p_{x|y=1}\, p_{y=1} \leq L(1,0)\, p_{x|y=0}\, p_{y=0}$$

Taking log on both sides.

$p_{x|y=0}$: Class conditionals be gaussian dist

$$\ln\left(p_{y=0}\, L(1,0)\right) + \ln\left(p_{x|y=0}\right) \leq$$

$$\ln\left(p_{y=1}\, L(0,1)\right) + \ln\left(p_{x|y=1}\right)$$

$$p_{x|y=i}(x|y=i) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right\}$$

taking log

$$\ln\left(\frac{1}{\sigma_i}\frac{1}{\sqrt{2\pi}}\right) + \left\{ \frac{-x^2 - \mu_i^2 + 2x\mu_i}{2\sigma_i^2} \right\}$$

$$= -\ln(\sigma_i) - \frac{1}{2}\ln(2\pi) - \frac{x^2}{2\sigma_i^2} - \frac{\mu_i^2}{2\sigma_i^2} + \frac{2x\mu_i}{2\sigma_i^2}$$

Skipping the Substitution

& Rearranging the terms.

$$\frac{1}{2}x^2\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\right) + x\left(\frac{\mu_0}{\sigma_0^2} - \frac{\mu_1}{\sigma_1^2}\right)$$

$$+ \frac{1}{2}\left(\frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_0^2}{\sigma_0^2}\right) + \ln\left(\frac{\sigma_1}{\sigma_0}\right) + \ln\left(\frac{p_{y=0}\, L(1,0)}{p_{y=1}\, L(0,1)}\right)$$

$$> 0$$

This is of the form

$h_B(x) = 0$ if

$$ax^2 + bx + c > 0$$

$\Rightarrow$ Bayes classifier in this case is a quadratic discriminant func

Some Special Cases :

Case 1:

$$\sigma_0 = \sigma_1 = \sigma$$

$$p_{y=0} = p_{y=1} = \frac{1}{2}$$

$$L(1,0) = L(0,1)$$

then

$$h_B(x) = 0 \quad \text{if}$$

$$\frac{x}{\sigma^2}(\mu_0 - \mu_1) - \frac{1}{2\sigma^2}(\mu_0^2 - \mu_1^2) > 0$$

$$x > \frac{\mu_0 + \mu_1}{2} \quad \text{assuming } \mu_0 > \mu_1$$

Case 2:

$$\mu_0 = \mu_1 = 0, \quad p_{y=0} = p_{y=1}, \quad L(1,0) = L(0,1)$$

$$h_B = 0 \quad \text{if}$$

$$\frac{1}{2} x^2 \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2} \right) - \ln\left(\frac{\sigma_0}{\sigma_1}\right) > 0$$

assume $\sigma_0 > \sigma_1$

$$x^2 > \frac{\sigma_1^2 \sigma_0^2 \ln(\sigma_0/\sigma_1)}{(\sigma_0^2 - \sigma_1^2)}$$

Self help !! Try out for $x \in \mathbb{R}^d$

$$p_{x|y=i} \sim N(\mu_i, \Sigma_i) \qquad x \in \mathbb{R}^d$$

Now Consider the Case of Bayes Classifier

0-1 Loss, $M=2$

wkt

$$R(h_B) = \int_{\mathbb{R}^d} \min\left(p_{Y=0}\, p_{X|Y=0},\ p_{Y=1}\, p_{X|Y=1}\right) dx$$

now lets take a specific case

$x \in \mathbb{R}$

$p_{Y=0} = p_{Y=1} = \frac{1}{2}$

$p_{X|Y=i}$ are Gaussian

$\sigma_0 = \sigma_1 = \sigma$

$\mu_0 < \mu_1$

$h_B(0)$ if

$$x < \frac{\mu_0 + \mu_1}{2}$$

$$R(h_B) = 0.5 \int_{-\infty}^{\frac{\mu_0+\mu_1}{2}} p_{X|Y=1}\, dx + 0.5 \int_{\frac{\mu_0+\mu_1}{2}}^{\infty} p_{X|Y=0}\, dx$$

$$= 0.5\,\phi\left(\frac{\mu_0-\mu_1}{2\sigma}\right) + 0.5\left[1-\phi\left(\frac{\mu_1-\mu_0}{2\sigma}\right)\right] \blacksquare$$