Action $\delta = \{1, \dots, K\}$ of classification

- Zero-One Loss Function

$$I(Y = \delta)$$

- Posterior expected loss

posterior distribution

$$E_{Y|x}[I(Y=\delta)] = \sum_k I(Y=\delta)\, p(Y=k \mid X=x)$$

- Bayes action is an optimal decision rule that minimizes the posterior expected loss

$$\underset{\delta}{\operatorname{argmin}} \sum_k I(\delta = k)\, p(Y=k \mid X=x)$$

$$= \sum_{k \neq \delta} p(Y=k \mid X=x) = \sum_{k \neq \delta} p_k(x)$$

$$= 1 - p_\delta(x)$$

$$\Updownarrow$$

$$\underset{\delta}{\operatorname{argmax}}\ p_\delta(x) = Pr(Y=\delta \mid X=x)$$