NAÏVE BAYES

p features, categorical response $y \in C = \{C_1, \ldots, C_K\}$

$$(x_1, y_1), \ldots, (x_n, y_n), \quad x_i = (x_{i1}, \ldots, x_{ip})$$

Bayes rule:

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} \propto p(x|y)p(y)$$

$$\widehat{p}(y) := \frac{\#[y_i = y]}{n}$$

$$\widehat{p}(x|y) := \frac{\#[(x_i, y_i) = (x, y)]}{\#[(x_i, y_i) = (*, y)]}$$

$$\widehat{p}(y|x) := \frac{\widehat{p}(x|y)\widehat{p}(y)}{\widehat{p}(x)} \propto \widehat{p}(x|y)\widehat{p}(y)$$

Decision rule: Assign x_0 to class i if

$$\widehat{p}(y|x_0) > \widehat{p}(y'|x_0) \quad \text{for all} \quad y' \neq y.$$

Let $\xi \in \mathbb{R}^p$, $c \in C$

$$\widehat{p}(\xi|c) = \frac{\#[(x_i, y_i) = (\xi, c)]}{\#[(x_i, y_i) = (*, c)]} = \frac{\#[(x_{i1}, \ldots, x_{ip}, y_i) = (\xi_1, \ldots, \xi_p, c)]}{\#[(x_i, y_i) = (*, c)]}$$

Suppose $X_1, \ldots, X_p$ are conditionally independent given Y. Then

$$\widehat{p}(\xi|c) = p(\xi_1|c) \cdots p(\xi_p|c)$$

$$\widehat{p}(\xi_j|c) := \frac{\#[(x_{ij}, y_i) = (\xi, c)]}{\#[(x_{ij}, y_i) = (*, c)]}$$