X, Y : random variables

unconditionally independent

$$X \perp Y \iff p(X,Y) = p(X)p(Y)$$

conditionally independent

$$X \perp Y \mid Z \iff p(X,Y \mid z) = p(X \mid z) \cdot p(Y \mid z)$$

Naive Bayes Classifiers

$$p(x \mid y=c, \theta_c) = \prod_{j=1}^{D} p(x_j \mid y=c, \theta_{jc})$$

- If features are binary

Ex: $x = [1, 0, 1, 0, 0, 0]^T$

$x_1 \in \{0, 1\}$ $\quad x_4 = \{0, 1\}$

$x_2 \in \{0, 1\}$ $\quad x_5 = \{0, 1\}$

$x_3 \in \{0, 1\}$ $\quad x_6 = \{0, 1\}$

- If features are categorical

Ex: $x = [3, 0, 1, 2, 4, 3]^T$

$x_1 \in \{1, 2, 3\}$ $\quad x_4 \in \{1, 2\}$

$x_2 \in \{0, 1\}$ $\quad x_5 \in \{-4, 4\}$

$x_3 \in \{-1, 0, 1\}$ $\quad x_6 \in \{0, 1, 2, 3\}$

- If features are real-valued

Ex $x = [1.2, -0.1, -0.03, 10.2, 1.1, -1.1]^T$

$x_j \in \mathbb{R}$

For binary features

$$p(x \mid y=c, \theta) = \prod_{j=1}^{D} \text{Ber}(x_j \mid \mu_{jc})$$

Training Data

	x_{i1}	x_{i2}	x_{i3}	x_{i4}	y_i
	← input				← output
$i=1$	1	0	0	1	0
$i=2$	1	0	1	0	1
$i=3$	1	1	1	0	0
$i=4$	1	1	1	1	1
$i=5$	1	0	1	0	1

$\mu_{10} = \frac{2}{2} = 1$ $\quad \mu_{11} = \frac{3}{3} = 1$

$\mu_{20} = 1/2$ $\quad \mu_{21} = 1/3$

$\mu_{30} = 1/2$ $\quad \mu_{31} = 2/3$

$\mu_{40} = 1/2$ $\quad \mu_{41} = 2/3$

Using NBC for Prediction

Ex: $N = 5$ $x_i \in \{0,1\}^3$ $y_i \in \{0,1\}$

	x_{i1}	x_{i2}	x_{i3}	y_i
$i=1$	1	0	1	0
$i=2$	1	0	0	1
$i=3$	1	1	0	0
$i=4$	0	1	1	0
$i=5$	0	0	0	1

Given this dataset and using Naive Bayes classifier, how would $x = [0,1,0]^T$ be classified?

Compute priors

$p(y=0|D) = 3/5$ $p(y=1|D) = 2/5$

$$p(y=c|D) \propto p(y=c|D) \cdot \prod_{j=1}^{3} p(x_j | y=c, D)$$

$$p(y=0 \mid x, D) \propto \frac{3}{5} \cdot p(x_1=0 \mid y=0, D) \cdot p(x_2=1 \mid y=0, D) \cdot p(x_3=0 \mid y=0, D)$$

$$= \frac{3}{5} \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{45}$$

$$p(y=1 \mid x, D) \propto = \frac{2}{5} \cdot p(x_1=0 \mid y=1, D) \cdot p(x_2=1 \mid y=1, D) \cdot p(x_3=0 \mid y=1, D)$$

$$= \frac{2}{5} \cdot \frac{1}{2} \cdot \frac{0}{2} \cdot \frac{2}{2} = \emptyset$$

NBC would assign $x = [0, 1, 0]^T$ to class 0.

Ex: $N = 5 \quad x_i \in \{0,1\}^3 \quad y \in \{0,1\}$

	x_{i1}	x_{i2}	x_{i3}	y_i
$i=1$	0	0	1	0
$i=2$	1	0	0	0
$i=3$	0	1	0	0
$i=4$	0	1	0	1
$i=5$	1	0	1	1

Given this dataset and using Naive Bayes classifier, compute $p(y=0|x,D)$ for $x = [0,1,1]^T$

$$p(y=0|D) = 3/5$$

$$p(y=1|D) = 2/5$$

$$p(y=c|x,D) \propto p(y=c|D) \cdot \prod_{j=1}^{3} p(x_j|y=c,D)$$

$$p(y=0|x_1 D) \propto 3/5 \cdot p(x_1=0|y=0,D) \cdot p(x_2=1|y=0,D) \cdot p(x_3=1|y=0,D)$$

$$= 3/5 \cdot \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{2}{45}$$

$$p(y=1|x_1 D) \propto \frac{2}{5} \cdot p(x_1=0|y=1,D) \cdot p(x_2=1|y=1,D) \cdot p(x_3=1|y=1,D)$$

$$= \frac{2}{5} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{20}$$

$$p(y=0|x_1 D) \propto \frac{2}{45} = \frac{8}{180} \qquad p(y=1|x_1 D) = \frac{1}{20} = \frac{9}{180}$$

$$p(y=0|x_1 D) + p(y=1|x_1 D) = 1$$

$\uparrow$ $8/17$ $\qquad$ $\uparrow$ $9/17$

$$p(y=0|x_1 D) = 8/17$$