i	x_i^1	x_i^2	y_i
1	sunny	car-broken	go-out
2	rainy	car-working	go-out
3	sunny	car-broken	go-out
4	sunny	car-broken	go-out
5	sunny	car-broken	go-out
6	sunny	car-working	stay home
7	rainy	car-working	stay home
8	rainy	car-broken	stay home
9	sunny	car-working	stay home
10	rainy	car-working	stay home

(1) $P(y=\text{go-out}) = 0.5$

(2) $P(y=\text{stay home}) = 0.5$

(3) $$P(y) = \begin{cases} 0.5 & \text{if } y = \text{go-out} \\ 0.5 & \text{if } y = \text{stay home} \end{cases}$$

(4) $$P(x) = \begin{cases} 0.2 & \text{if } x = (\text{sunny}, \text{car-working}) \\ 0.4 & \text{if } x = (\text{sunny}, \text{car-broken}) \\ 0.3 & \text{if } x = (\text{rainy}, \text{car-working}) \\ 0.1 & \text{if } x = (\text{rainy}, \text{car-broken}) \end{cases}$$

(5) $P(x=(\text{rainy}, \text{car-working}) \wedge y = \text{go-out}) = 0.1$

(6) $$P(y=\text{go-out} \mid x=(\text{rainy}, \text{car-working})) = \frac{P(x=(\text{rainy}, \text{car-working}) \wedge y = \text{go-out})}{P(x=(\text{rainy}, \text{car-working}))}$$ (from definition)

$$= \frac{0.1}{0.3} = \frac{1}{3}$$

(7) $$P(x=(\text{rainy}, \text{car-working}) \mid y = \text{go-out})$$

$$= P([x]_1 = \text{rainy} \mid y=\text{go-out}) \times (P([x]_2 = \text{car-working} \mid y = \text{go-out})$$

$$= \frac{P([X]_1 = \text{rainy} \wedge y = \text{go-out})}{P(y=\text{go-out})} \times \frac{P([x]_2 = \text{car-working} \wedge y = \text{go-out})}{P(y=\text{go-out})}$$

$$= \frac{0.1}{0.5} \times \frac{0.1}{0.5} = 0.04$$

(6+) By Bayes rule

$$P(y=\text{go-out} \mid x=(\text{rainy}, \text{car-working}))$$

$$= \frac{P(x=(\text{rainy}, \text{car-working}) \mid y=\text{go-out})\, P(y=\text{go-out})}{P(x=(\text{rainy}, \text{car-working}))}$$

$$= \frac{0.04 \times 0.5}{0.3} = \frac{1}{15}$$

(8) $$P(y=\text{go-out} \mid x=(\text{sunny}, \text{car-broken}))$$

$$= \frac{P(x=(\text{sunny}, \text{car-broken}) \mid y=\text{go-out}) \cdot P(y=\text{go-out})}{P(x=(\text{sunny}, \text{car-broken}))}$$

$$= \frac{\frac{0.4}{0.5} \times \cancel{\frac{0.4}{0.5}} \times \cancel{0.5}}{\cancel{0.4}}$$

$$= 0.8$$

$$\frac{P(x=(\text{sunny}, \text{car-broken}) \mid y=\text{stay home}) \cdot P(y=\text{stay home})}{P(x=(\text{sunny}, \text{car-broken}))}$$

$$= \frac{\cancel{\frac{0.2}{0.5}}^{0.4} \times \frac{0.1}{0.5} \times \cancel{0.5}}{\cancel{0.4}}$$

$$= 0.1$$

i	x_i^1	x_i^2	y_i
1	sunny	car-broken	go-out
2	rainy	car-working	go-out
3	sunny	car-broken	go-out
4	sunny	car-broken	go-out
5	sunny	car-broken	go-out
6	sunny	car-working	stay home
7	rainy	car-working	stay home
8	rainy	car-broken	stay home
9	sunny	car-working	stay home
10	rainy	car-working	stay home

① $P(y = \text{go-out}) = \dfrac{\sum_{i=1}^{10} I(y_i = \text{go-out})}{10} = \dfrac{5}{10} = 0.5$

② $P(y = \text{stay home}) = 0.5$

③ $P(Y) = \begin{cases} 0.5 & \text{if } y = \text{go-out} \\ 0.5 & \text{if } y = \text{stay home} \end{cases}$

④ $P(X) = \begin{cases} 0.4 & \text{if } x = (\text{sunny}, \text{carbroken}) \\ 0.2 & \text{if } x = (\text{sunny}, \text{carworking}) \\ 0.1 & \text{if } x = (\text{rainy}, \text{car-broken}) \\ 0.3 & \text{if } x = (\text{rainy}, \text{carworking}) \end{cases}$

i	x_i^1	x_i^2	y_i
1	sunny	car-broken	go-out
2	rainy	car-working	go-out
3	sunny	car-broken	go-out
4	sunny	car-broken	go-out
5	sunny	car-broken	go-out
6	sunny	car-working	stay home
7	rainy	car-working	stay home
8	rainy	car-broken	stay home
9	sunny	car-working	stay home
10	rainy	car-working	stay home

⑤ $P(\boxed{x = (\text{rainy}, \text{car-working}) \wedge y = \text{go-out}})$

$$= \frac{\sum_{i=1}^{10} I(x_i = (\text{rainy}, \text{car-working}) \wedge y_i = (\text{go-out}))}{10}$$

$$= \frac{1}{10} = 0.1$$

⑥ $P(\boxed{\underbrace{y = \text{go-out}}_{A} \mid \underbrace{x = (\text{rainy}, \text{car-working})}_{B}})$ directly

$$= \frac{P(y = \text{go-out} \wedge x = (\text{rainy}, \text{car-working}))}{P(x = (\text{rainy}, \text{car-working}))} = \frac{0.1}{0.3} = \frac{1}{3}$$

⑦ $P(\cancel{x} = (\underbrace{\text{rainy}}_{x_1}, \underbrace{\text{car-working}}_{x_2}) \mid y = \text{go-out})$

$$= P(\underbrace{\boxed{[x]_1 = \text{rainy}}}_{A} \mid \underbrace{\boxed{y = \text{go-out}}}_{B}) * P([x]_2 = \text{car-working} \mid y = \text{go-out})$$

$$= \frac{P([x]_1 = \text{rainy} \wedge y = \text{go-out})}{P(y = \text{go-out})} \times \frac{P([x]_2 = \text{car-working} \wedge y = \text{go-out})}{P(y = \text{go-out})}$$

0.1 0.2 0.5

0.1 0.2 0.5

$= \underline{0.04}$

i	x_i^1	x_i^2	y_i
1	sunny	car-broken	go-out
2	rainy	car-working	go-out
3	sunny	car-broken	go-out
4	sunny	car-broken	go-out
5	sunny	car-broken	go-out
6	sunny	car-working	stay home
7	rainy	car-working	stay home
8	rainy	car-broken	stay home
9	sunny	car-working	stay home
10	rainy	car-working	stay home

8 $h(x = (\text{sunny}, \text{car-broken}))$

Naive Bayes assumption: All Feature values are independent

$$P(X=x \mid Y=y) = \prod_{i=1}^{d} P([x]_i = x_i \mid Y=y)$$

Naive Bayes Classifier:

$$h(x) = \arg\max_y P(y \mid x)$$

$$= \arg\max_y \frac{P(x \mid y)\, P(y)}{P(x)}$$

$$= \arg\max_y P(x \mid y)\, P(y)$$

$$= \arg\max_y \prod_{i=1}^{d} P(x_i \mid y)\, P(y)$$

$$= \arg\max_y \log\left(\prod_{i=1}^{d} P(x_i \mid y)\, P(y)\right)$$

$$= \arg\max_y \left(\sum_{i=1}^{d} \log P(x_i \mid y) + \log(P(y))\right)$$

The key to Naive Bayes is to estimate

$$P([x]_\alpha \mid y)$$

Estimating $P([x]_\alpha \mid y)$

- Three notable cases

- Case 1: Categorical features
- Case 2: Multinomial Features
- Case 3: Continuous Features

Categorical Features:

- Each feature α falls into one of k categories

$$[x]_\alpha \in \{\underbrace{c_1, c_2, \ldots, c_k}_{k \text{ options}}\}$$

k-faced dice that results in $P([x]_\alpha | y)$

Ex. {male, female}

{single, married, widowed}

Modeling distribution: $P([x]_\alpha = c \mid y) = \boxed{[\theta_{cy}]_\alpha}$ ← parameter (probability)

Ex. $[\theta_{\text{sunny, goout}}]_1$ $[\theta_{\text{car works, goout}}]_2$

$[\theta_{\text{rainy, goout}}]_1$ $[\theta_{\text{car broken, goout}}]_2$

Constraint: $\sum_{\forall c} [\theta_{cy}]_\alpha = 1$

$[\theta_{cy}]_\alpha$ is the probability of feature α having the value c, given that the label is c.

Parameter estimation:

$$[\theta_{cy}]_\alpha = \frac{\sum_{i=1}^{n} I(y_i = y)\, I([x_i]_\alpha = c)}{\sum_{i=1}^{n} I(y_i = y)} = \frac{\text{\# of label } y \text{ having feature } \alpha \text{ w/ value } j}{\text{\# of samples w/ label } y}$$