Naïve Bayes Classifier

Implementation Notes

Python Scikit learn

- Gaussian Naive Bayes
- Multinomial Naive Bayes
- Categorical Naive Bayes
  (Attribute values are different categories)

$$P(Y/X) = \frac{P(X/Y) \cdot P(Y)}{P(X)}$$

X_1	0	1	2	1	0	2	2	0	1
Y	0	1	0	1	1	1	0	0	0

$X_1 = \{0, 1, 2\}$

$P(x_1/y)$

X_1 \ Y	0	1
0	2/5	
1	1/5	
2	2/5	

X	0	1	0	1	1	0
Y	1	0	1	1	0	0

$\boxed{P(Y) = 3/5}$

What is probability of an unbiased coin toss (observing head)

$P(H) = \frac{1}{2}$

$P(T) = 1 - \frac{1}{2} = \frac{1}{2}$

Binomial (Bernaulli)

$f(x) = p^{x} \times (1-p)^{1-x}$

$P(x) = p$ if $x=1$

$= 1-p$ if $x=0$

$P(x=H) = 1/2$

$P(x=T) = 1/2$

$\Rightarrow$ H H H H T

4/5

No. of trials

Binomial Distribution := ${}^{n}C_{r}\, p^{r}(1-p)^{n-r}$

$r \rightarrow$ No. times $x=1$

$P(x=H) = {}^{5}C_{4}\,(1/2)^{4}\cdot(1/2)^{1}$

$(x=1)$

$= 5\times\left(\frac{1}{2}\right)^{5} = 5/32$

Multinomial Distribution $\frac{n!}{x_1!\, x_2! \ldots x_k!} P_1^{x_1} \times P_2^{x_2} \ldots P_n^{x_k}$

No. of times $X = x_i$

$x_1 + x_2 + \ldots x_k = n$

Gaussian Distribution

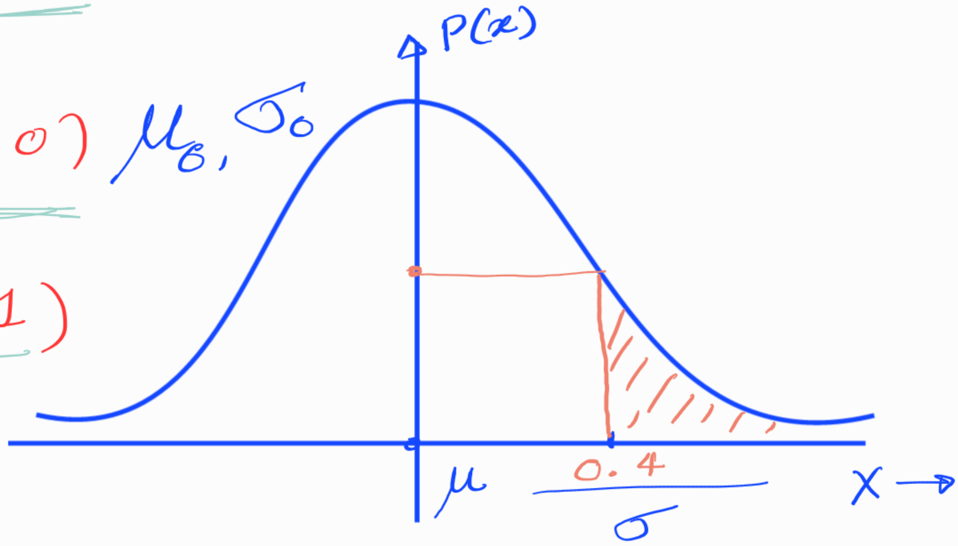
X_1	0.2	0.4	0.5	0.3	0.7
Y	1	0	1	0	1

$P(x_1/y)$ $P(y)$

$P(x_1/y=0)$ μ_0, σ_0

$P(x_1/y=1)$

μ_1, σ_1



$$P(x) = \boxed{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x-\mu}{\sigma}}}$$

$$P(y=1/x_1=0.6)$$

$$= P(x_1=0.6/y=1) \times P(y=1)$$

=

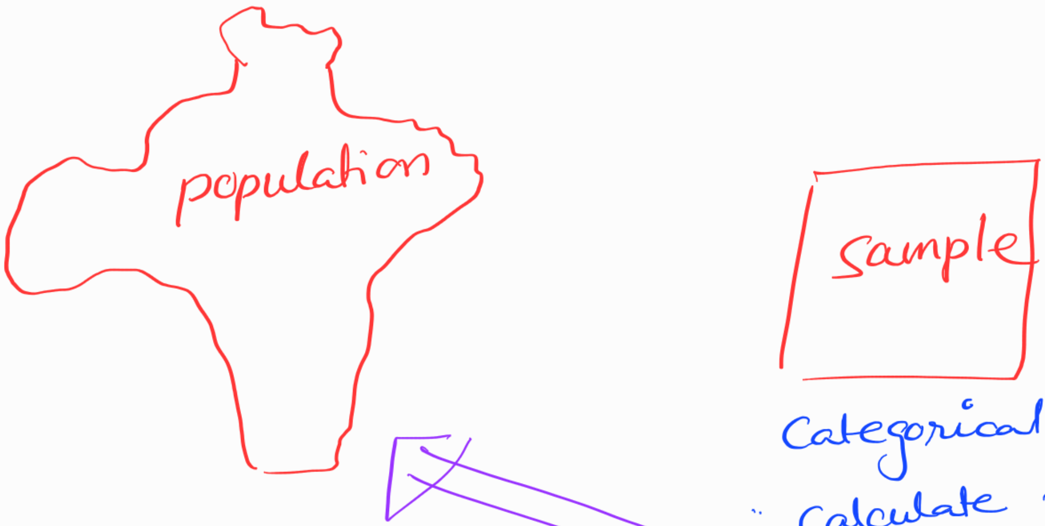
Smoothing : Avoiding 0 probability

- Add a constant in numerator & denominator

$$\left(\frac{c_i + \alpha}{Kc_i + n\alpha}\right)$$

Central Limit Theorem
Z

How the approach of categorial & gaussian different ?



Represents the population

Categorical
: calculate probability from the data

x_1	0.2	0.4	0.5	0.7	0.3	0.8	0.4	0.2
y	0	0	1	1	0	1	0	1

Training

$P(Y) = 1/2$

$P(\bar{Y}) = 1/2$

$P(x_1 / y)$

A $P(x_1 / y = 1)$

- Look at all examples where $y = 1$.
- Calculate μ_1, σ_1

$P(x_1 / y = 0)$

- Look at all examples where $y = 0$
- Calculate μ_0, σ_0

<u>Inference</u> :

$$P(y = 1 / x_1 = 0.1) = \underbrace{P(x_1 = 0.1 / y = 1)}_{?} \cdot P(y = 1)$$

$\cdot \left(\frac{1}{2}\right)$

Use gaussian distri formula to calculate $P(x_1 = 0.1)$

$$P(x) \propto \left(\frac{x - \mu_1}{\sigma_1}\right)$$

<u>Categorical</u> : when each outcome is considere different / each example is an independent experiment

Weather : {sunny, overcast, Rainy)

Multinomial : When all examples are generated by repeating an experiment.

H, H, H, T

Day 1	Day 2	Day 3
Rain	Rain	Sunny

Day 1	Day 2	Day 3	Day 4
W	L	L	W

Check if an email is spam

- each Word is a feature

each Word is a R.V. for the NBC

Word is a ________ R.V.

1. Categorical
2. Multinomial
3. Gaussian

	w_1	w_2	w_3	---	w_n	C.L.
m_1	2	4	1		0	S
						NS
m_2						
⋮						
m_n						S

m_1 | m_2

w_1 w_2 | w_1 w_1 w_2 w_1