* Deal with Zero Probability

→ if $P(x_i / y) = 0$ for some $X_i = x_i$ & $Y = y_i$ it will make posterior 0.

→ To avoid this, we apply "smoothing"

$$P(x_i / y) = \frac{\# x_i \cap y + \alpha}{\# y + \alpha \cdot k}$$

α ← small constant

k ↑ No. of examples

$P(x_i / y) \geq 0$

* Variables are continuous

	M	M	H	M	M
X_1 (Temp)	30.5	32.7	35.8	30.5	33.4
X_2 (Humidity)	80%	70%	80%	65%	75%
Y (Yes/No)	Yes	No	Yes	Yes	No

This does not work

$P(X_1 / Y)$: $P(X_1 = 30.5 / Y = Yes)$

$P(X_1 = 32.7 / Y = Yes)$

$P(X_1 = 35.8 / Y = Yes)$

⋮

Idea 1

Divide variable values into different categories

e.g temp = High if $t > 35^0$
= Med if $30^0 < t < 35^0$
= low if $t < 30^0$

Multinomial Naive Bayes

Variables are multinomial /(binary) can take a few finite labels.

→ Model parameters are

$$P(x_i / y) \; , \; P(y)$$

Gaussian Naive Bayes

→ Variables are continuous. e.g temp value, pressure value

→ Counting probability for all values of a R.V. is not possible

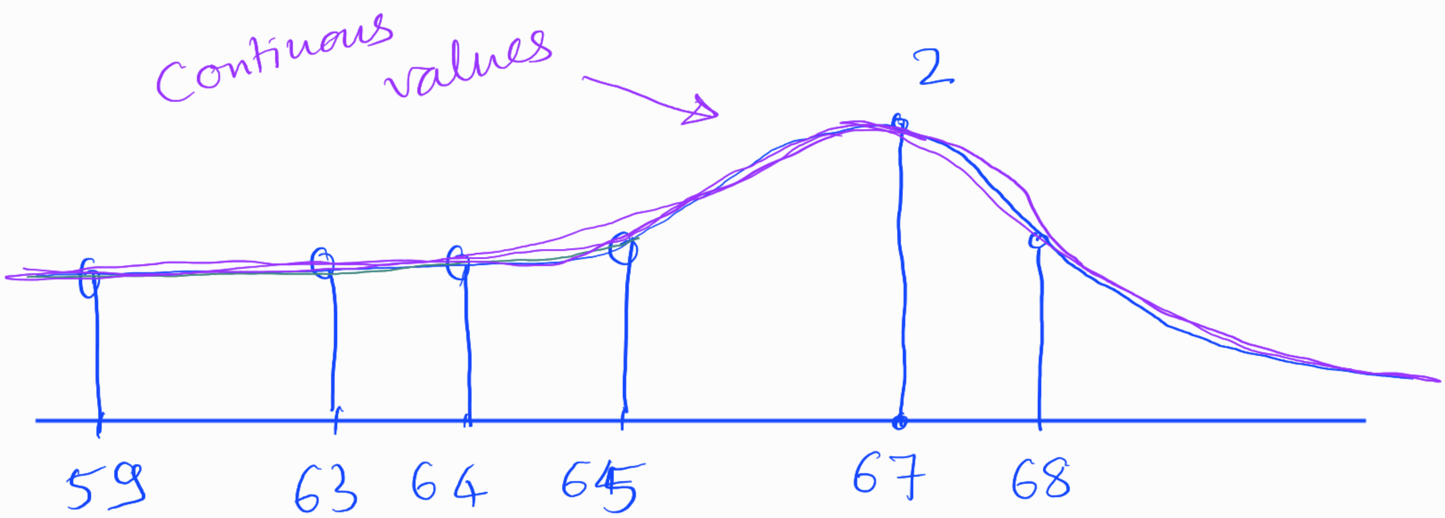
X_1 (Temp)	30.5	32.7	35.8	30.5	33.4
X_2 (Humidity)	80%	70%	80%	65%	75%

Y (Yes/No)   Yes   No   Yes   Yes   No

- X_1 & X_2 are Random Variables with continuous values.
- $P(x_1/y)$ & $P(x_2/y)$ are coming from an underlying Normal Distri-bution.
- Distribution $P(X_i/y)$ is Normal

R 65"   J 67"   N 68"   G 59"   M 63"

K = 64"   A 67

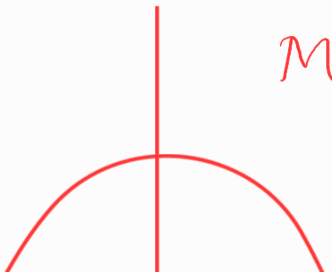


For Normal Distribution the Graph is defined by μ & σ

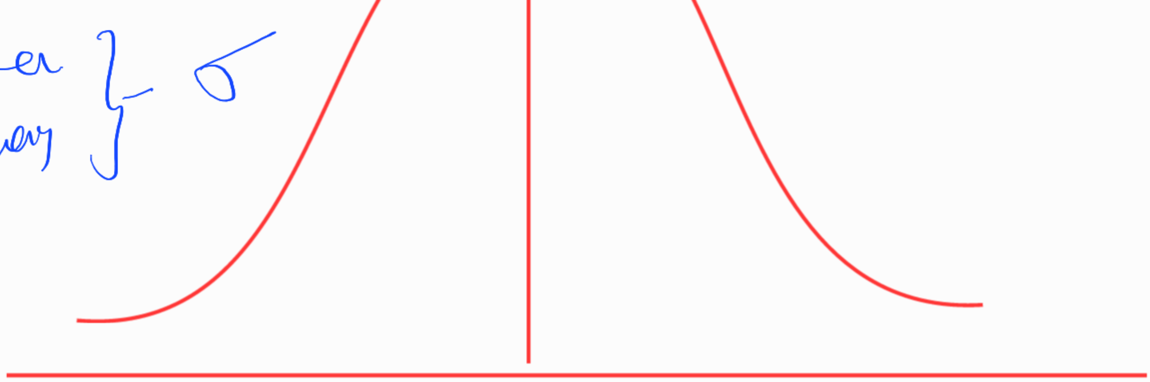
Mean   Standard Deviation

Bell curve

center = μ



Diameter / periphery } – σ



μ – Mean

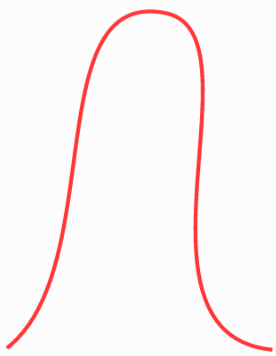
σ Standard Deviation

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \qquad S.D. = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

↳ Average value of x_i

How far are the values from the average

Low SD



High SD

Gaussian Naive Bayes

Parameters $\quad p(x_i / y) \Rightarrow \mu_{x_i y}, \sigma_{x_i y}$

$\mu = 62,$ $\quad \mu = 66.7$

h	59	63	64	65	67	68
g	F	F	F	M	M	M

$p(h/g) \leftarrow$ Gaussian

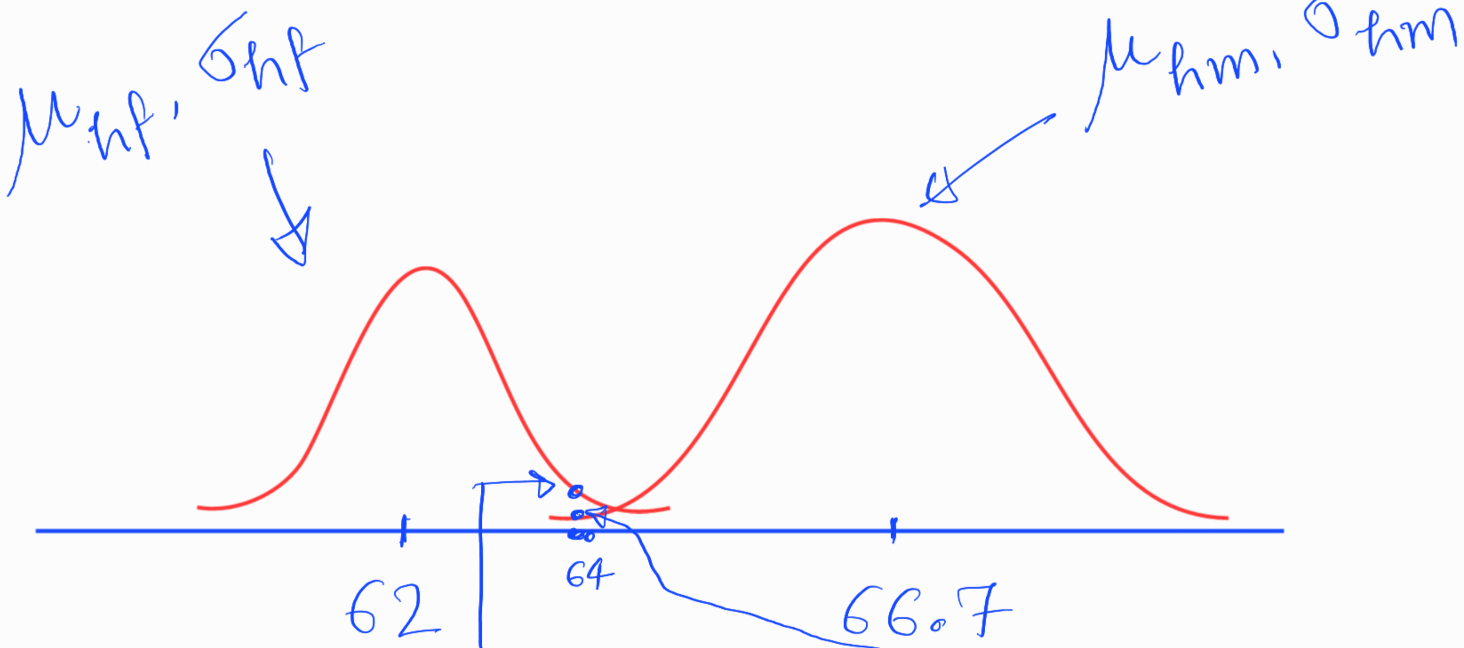
g class label : F/M

h is continuous

Parameters

$$p(h/g=F) = \mu_{hF}, \sigma_{hF}$$

$$p(h/g=M) = \mu_{hM}, \sigma_{hM}$$



$$p(h=64/g=f) > p(h=64/g=m)$$

$h = \underline{64}$