# Graph Theoretic Clustering

# Renyi entropy

- Free parameter: $\alpha$ or q
- Denoted as $H_{R\alpha}$ or $H_{Rq}$

$$H_\alpha(X) = \frac{1}{1-\alpha}\log_2\left(\sum_{i=1}^{n} p_i^\alpha\right)$$

## $H_{R2}$ (i.e. $\alpha$ =2)

- $H_{R2} = -\log \sum_i p_i^2$

# Information potential

$$H_\alpha(X) = \frac{1}{1-\alpha}\log\left(\sum_{k=1}^{N} p_k^\alpha\right) = -\log\boxed{\left(\sum_{k=1}^{N} p_k^\alpha\right)}^{\frac{1}{\alpha-1}}$$

$$Information\ Potential = V^\alpha(x) = \sum_{k=1}^{N} p_k^\alpha$$

Information potential is
Sum of powered probabilities

# Information potential is Sum of powered probabilities

- How to get PDF?
- Parzen windowing

# Example

Given a set of five data points x1 = 2, x2 = 2.5, x3 = 3, x4 = 1 and x5 = 6

Find **Parzen probability density function** (pdf) estimates at x = 3, using the Gaussian function with $\sigma = 1$ as window function

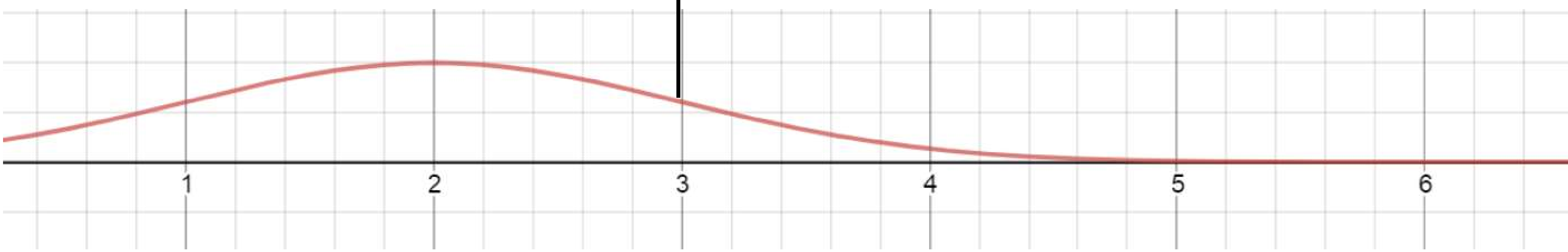
# Algorithm

**x1 = 2, x2 = 2.5, x3 = 3, x4 = 1 and x5 = 6**

1. Place a Gaussian at x=2 i.e. $\mu = 2$
2. Find its value @ x=3
3. Place a Gaussian at x=2.5 i.e. $\mu = 2.5$
4. Find its value @ x=3
5. ..
6. ..
7. ..
8. ..
9. Place a Gaussian at x=6 i.e. $\mu = 6$
10. Find its value @ x=3

# Gaussian with $\mu = 2$

Find value x=3

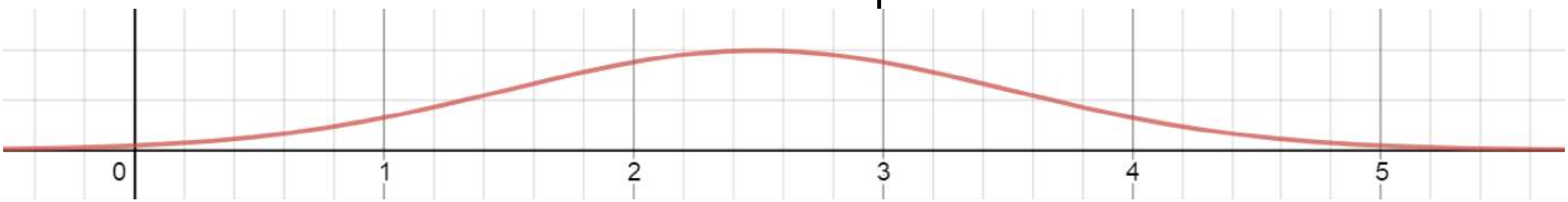



$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_1 - x)^2}{2}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(2-3)^2}{2}\right) = 0.2420$$

# Gaussian with $\mu$ = 2.5

Find value x=3

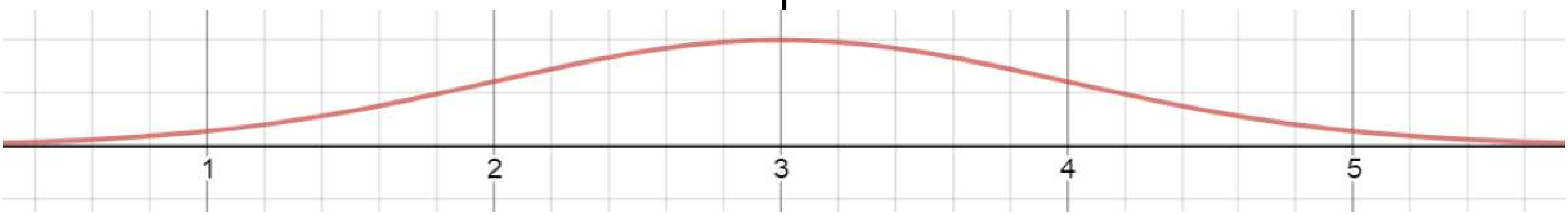



$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_2 - x)^2}{2}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(2.5 - 3)^2}{2}\right) = 0.3521$$

# Gaussian with $\mu$ = 3

Find value x=3  (we get maximum value)

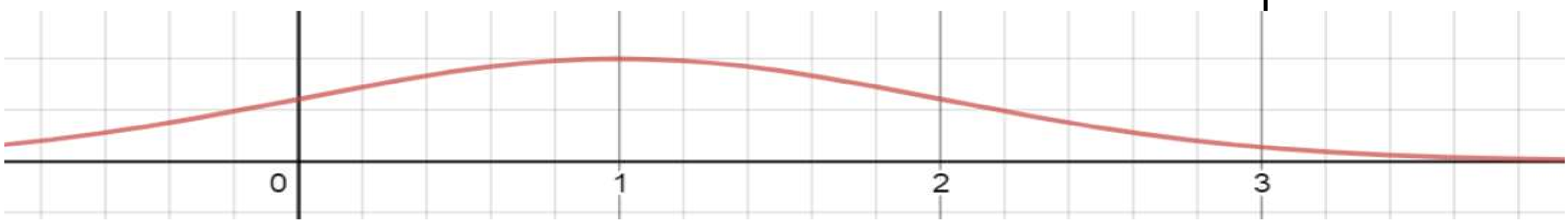



$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(x\;3 - x)^2}{2}\right)$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(3 - 3)^2}{2}\right) = 0.3989$$

# Gaussian with $\mu$ = 1

Find value x=3

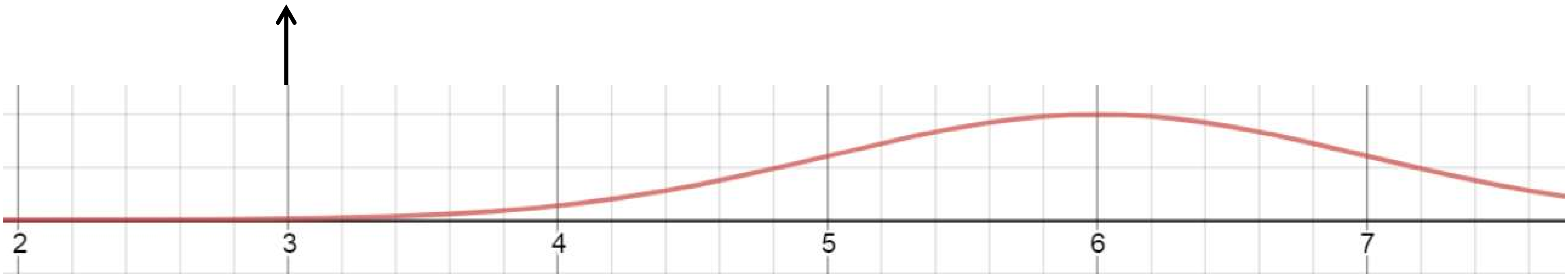



$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(x_4 - x)^2}{2}\right)$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(1 - 3)^2}{2}\right) = 0.054$$

# Gaussian with $\mu$ = 6

Find value x=3



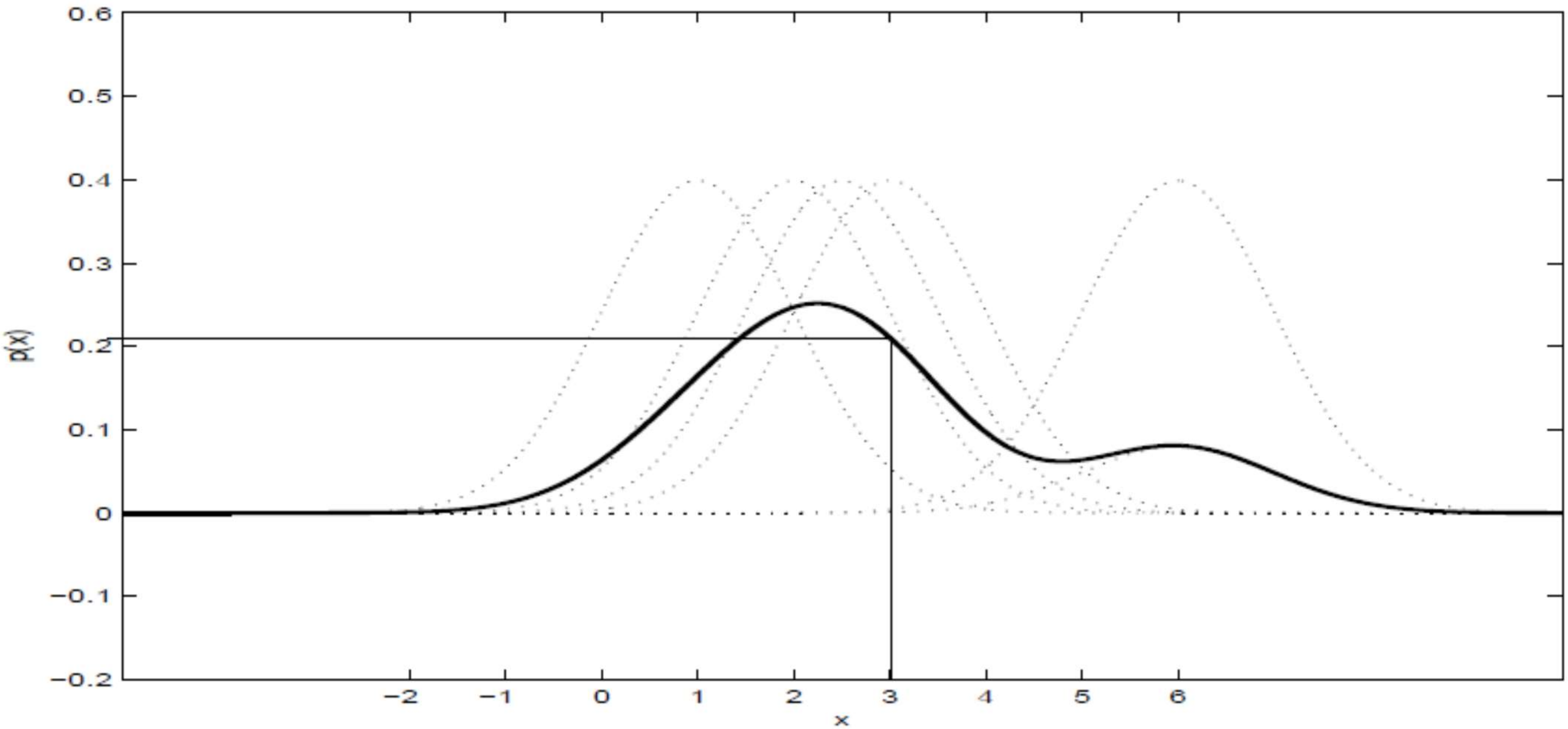

$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_5 - x)^2}{2}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(6 - 3)^2}{2}\right) = 0.0044$$

$$p(x=3) = (0.2420 + 0.3521 + 0.3989$$

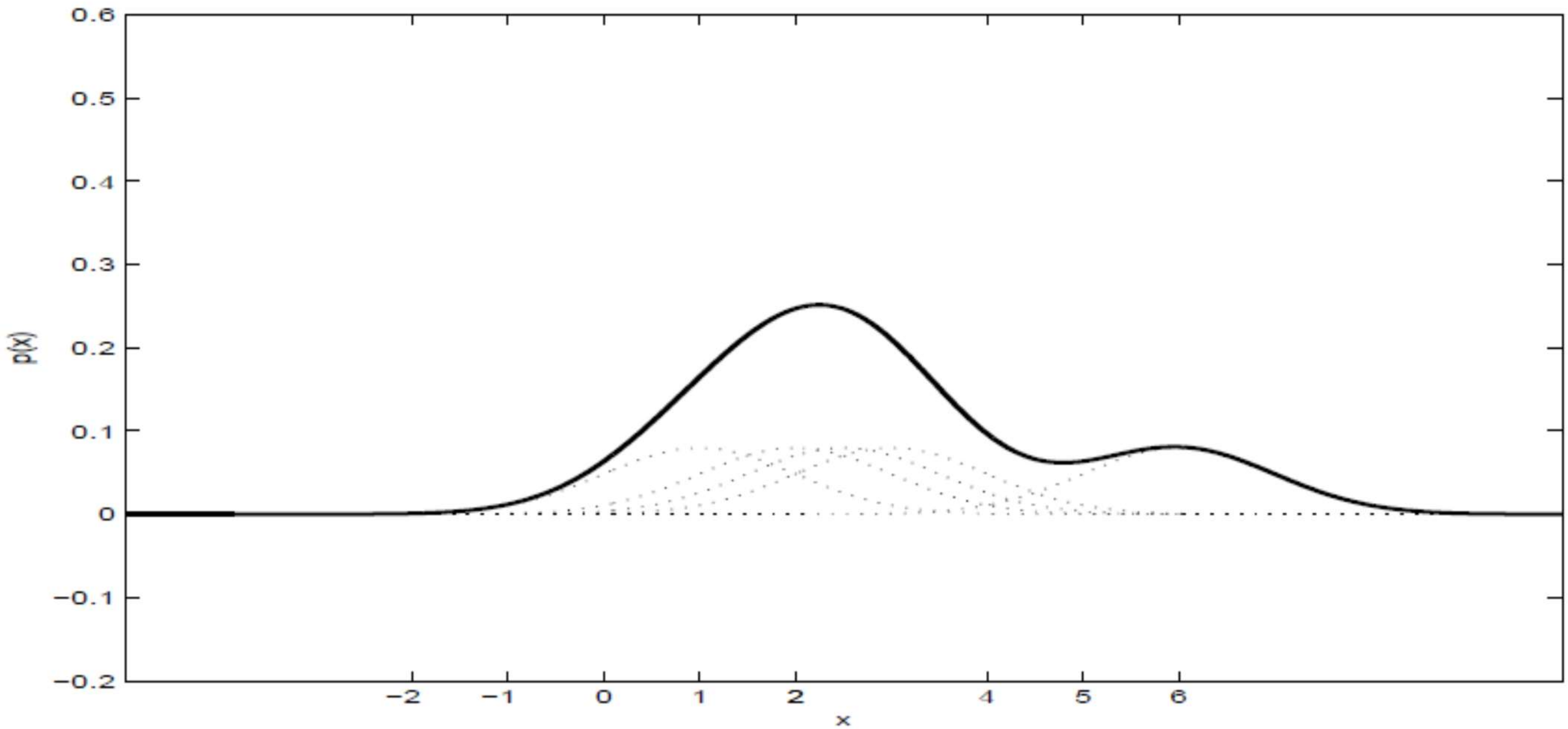
$$+0.0540 + 0.0044)/5 \quad = 0.2103$$

x1 = 2, x2 = 2.5, x3 = 3, x4 = 1 and x5 = 6




What is p(3)? Answer is 0.21

Given: x1 = 2, x2 = 2.5, x3 = 3, x4 = 1 and x5 = 6



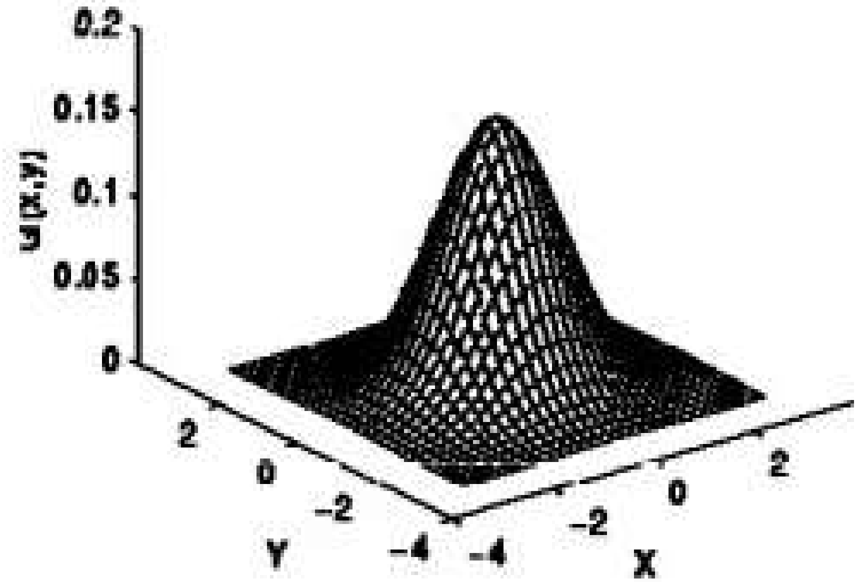

What is p(x) in general?
Answer is the estimated curve

# If data is two dimensional …[1]

- Use two dimensional Gaussian function to estimate PDF

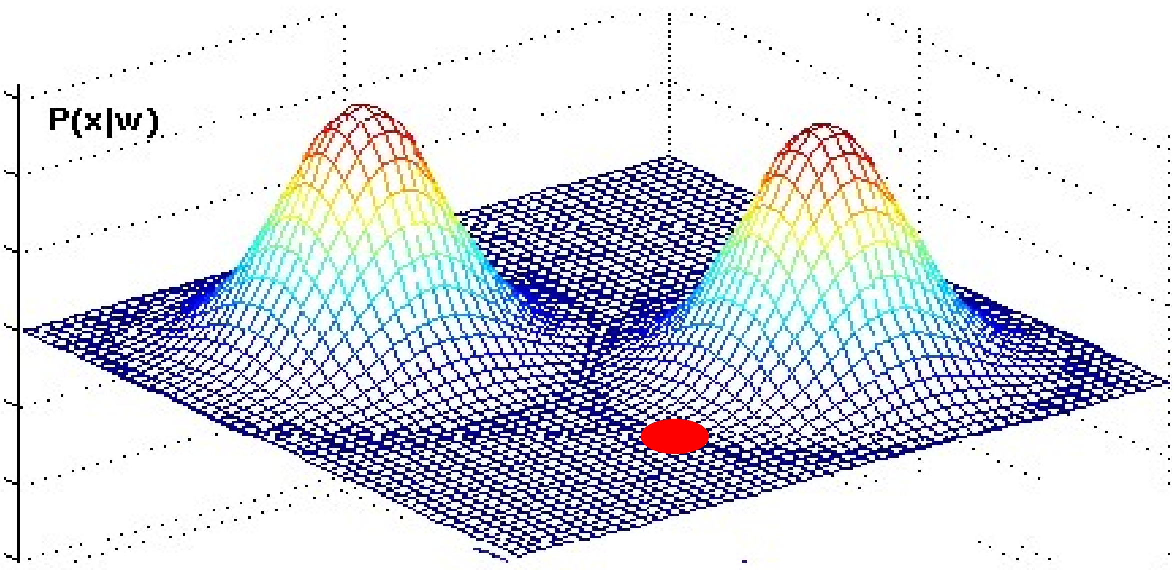
$$G(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}$$

A graphical representation of the 2D Gaussian distribution with mean(0,0) and σ = 1



# If data is two dimensional …²

- Place the Gaussian @ various points
- Find their contributions @ the given point
- Sum up all the contributions
- Take average




- Assume we have
  3 points on a plane
- @ red color point
  the PDF is given by
  average contributions
  from other two Gaussians

# Parzen Window

- Number of data points (N)

**To calculate the density @ one point**

- N Gaussian computations and then summing up
  - One $\Sigma$

## Information potential (IP)

- Need to know PDf @ N points and then summing up
  - One more $\Sigma$
  - i.e. Two $\Sigma$

$$Information\ Potential = V^{\alpha}(x) = \sum_{k=1}^{N} p_k^{\alpha}$$

- Quadratic information potential with Gaussian kernel

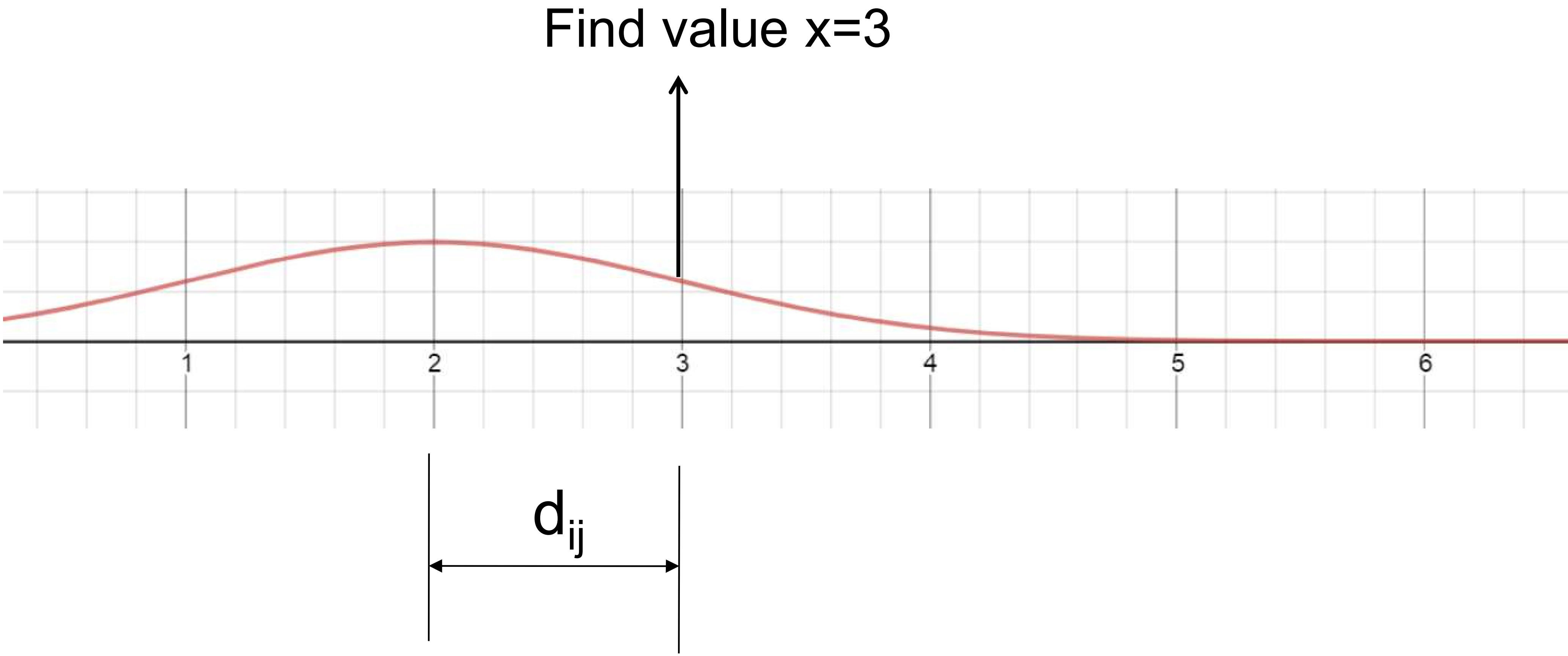
$$\left( \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{\sigma\sqrt{2}}(x_j - x_i) \right)$$

$$\boxed{\hat{V}(X) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \hat{V}_{i,j}}$$

$$d_{i,j} = x_i - x_j$$

$$V_{ij} = G(d_{ij})$$

# What is $G(d_{ij})$?



# Information force

Information force on sample xj due to sample xi

$$\hat{F}(i) = \frac{-1}{2N\sigma^2} \sum_{j=1}^{N} \boxed{\hat{V}_{i,j} d_{i,j}}$$

where

$$d_{i,j} = x_i - x_j$$

$$V_{ij} = G(d_{ij})$$

Total information force acting on sample

# Information force and information potential

$$\hat{F}(i) = \frac{-1}{2N\sigma^2} \sum_{j=1}^{N} \hat{V}_{i,j} d_{i,j}$$

where

$$d_{i,j} = x_i - x_j$$

$$V_{ij} = G(d_{ij})$$

$$\hat{V}(X) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \hat{V}_{i,j}$$

# $G(d_{ij})$

- $G(d_{ij})$ – Gaussian placed @ $x_j$ point, and evaluated @ $x_i$ point
- $G(d_{ij})$ is always positive
- Both IP and IF are functions of $G(d_{ij})$

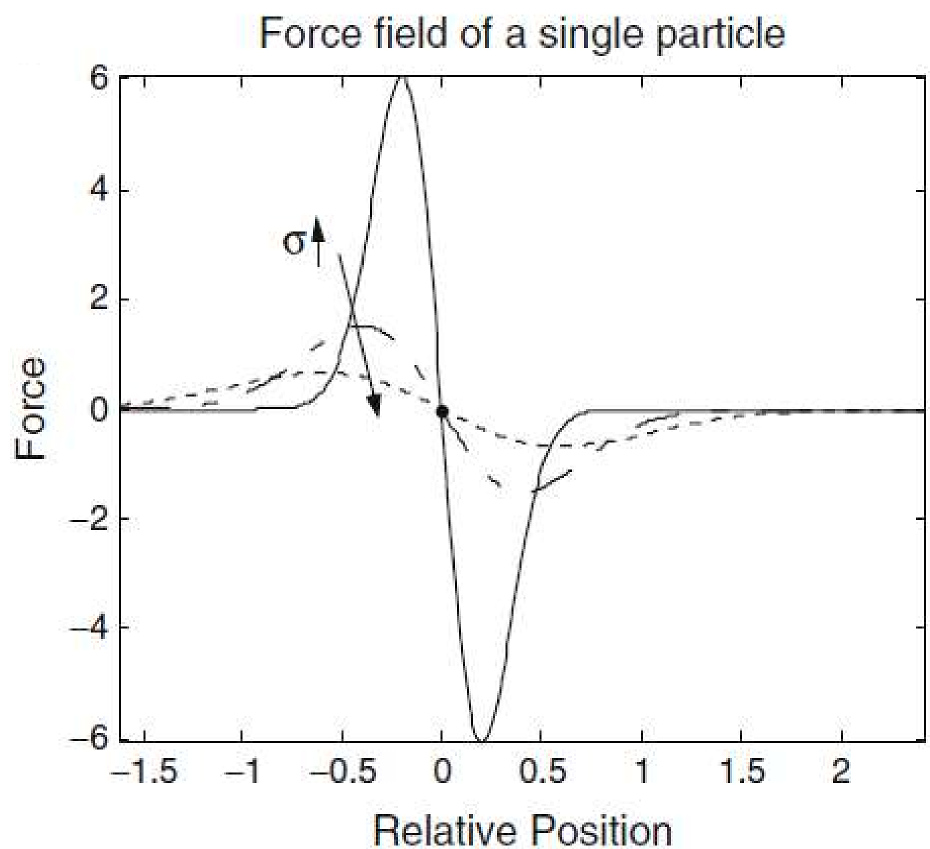
# $d_{ij}$

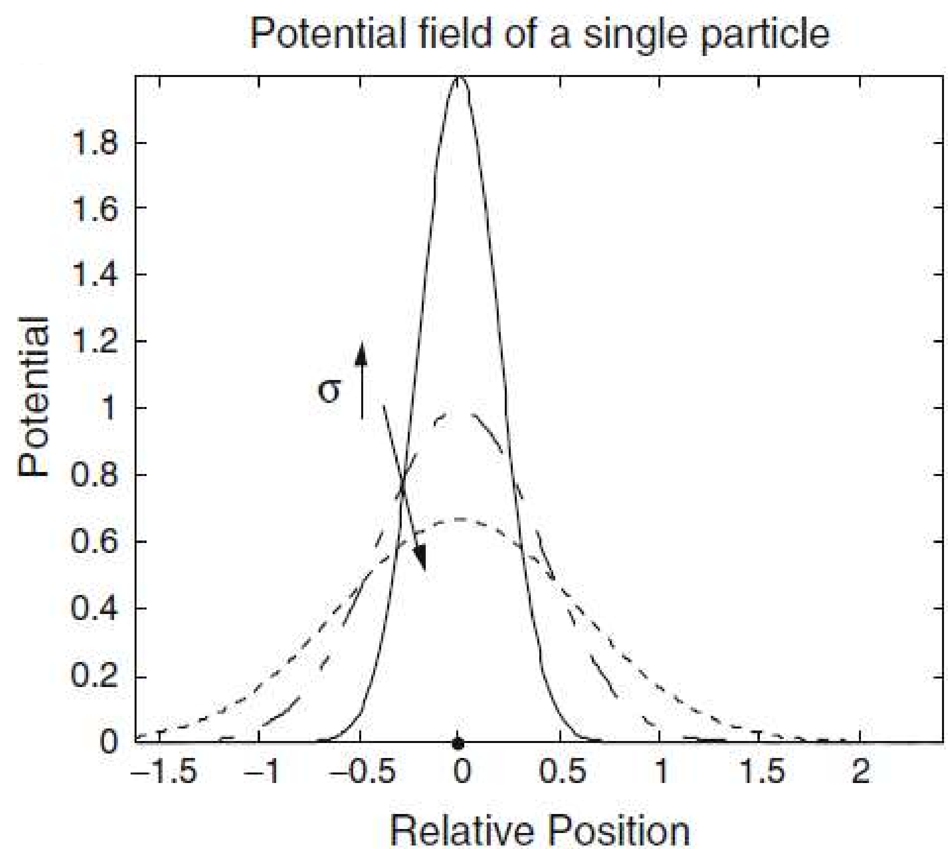
- $d_{ij}=x_i-x_j$
- Can be positive or negative
- IF = $G(d_{ij})$ x $d_{ij}$
- IF could be +ve or -ve

# IP and IF

- IF +ve and –ve
- IF depends on both $G(d_{ij})$ & $d_{ij}$

- IP always +ve
- IP depends on $G(d_{ij})$



# IF for two dimensional

**Dimension X**

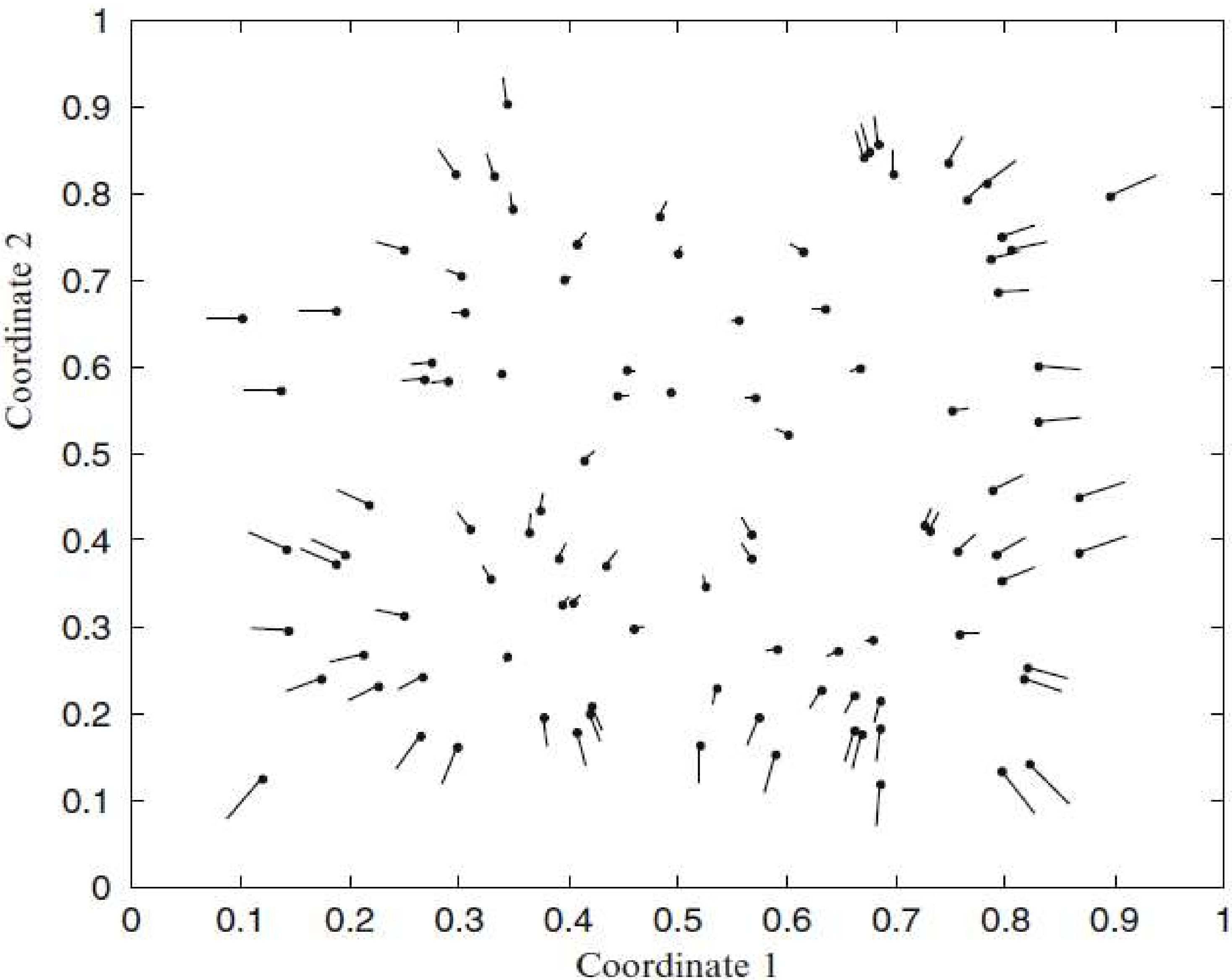
- $d_{ij}=x_i-x_j$
- $IF_x = G(d_{ij}) \times d_{ij}$

**Dimension Y**

- $d_{ij}=y_i-y_j$
- $IF_y = G(d_{ij}) \times d_{ij}$

- Combine $IF_x$ and $IF_y$
- Vector sum
- Produces magnitude as well as direction

# Quadratic information forces on 2 featured data points



# Information forces for clustering

- Some points become center of cluster
- IF of many surrounding points oriented towards them

# Directed tree

- Create graph out of N nodes
- Use their IF as cost
- More the IF better the connection

General constraints

- $d_{ij} < 3\sigma$
- $|F| > 0$