

IT 542: Pattern Recognition and Machine Learning  
Assignment – 5  
Kernel Density Estimation

(1) Draw 1000 samples from  $p(x) = P1.p1(x) + P2.p2(x)$   
Where,  $P1 = 0.6$ ,  $P2 = 0.4$ ,  $p1(x) \sim N(5,10)$  and  $p2(x) \sim N(10,15)$ .  
Draw a histogram for the data.

(2) Use the KDE to estimate pdf of given samples using

$$KDE \left( \frac{1}{nh} \sum_{i=1}^n K \left( \frac{x - x_i}{h} \right), K \sim Normal \right)$$

$h$  is bandwidth,  $n$  is number samples.

$h$  can be used

$$h = \left( \frac{4\hat{\sigma}^5}{3n} \right)^{\frac{1}{5}} \approx 1.06\hat{\sigma}n^{-1/5},$$

→ Std. Dev.  
of samples