# Unit 3

## Non-parametric Techniques

### * Density Estimation

→ Density estimation is the process of estimating the probability density function PDF $p(x)$ of a random variable based on observed data.

→ Unlike parametric methods, where we assume a specific form for the underlying distribution, non-parametric density estimation techniques make fewer assumptions and instead estimate the density directly from the data.

→ The general idea is to estimate how likely a data point $x$ is to occur by looking at the number of nearby data points.

→ Mathematically, the probability that $x$ falls within a region $R$ can be written as:

$$P = \int_R p(x)\, dx$$

→ $p(x)$ can be approximated using n sample points $x_1, x_2 \ldots x_n$.

The estimate depends on counting the number of sample points that fall within a small region $R_n$ around $x$, and then dividing by both the total number of samples and the volume $V_n$ of that region

i.e $$p_n(x) = \frac{k_n / n}{V_n}$$

$k_n$ = no. of data points inside $R_n$

$n$ = total no. of sample points

$V_n$ is the volume of the region $R_n$.

## Conditions for Convergence

→ To ensure that the estimated density $p_n(x)$ converges to the true density $p(x)$ as the number of samples n increases, certain conditions must be satisfied.

① Shrinking region : volume of region $V_n$ must shrink to 0 as $n \to \infty$

i.e $$\lim_{n \to \infty} V_n = 0$$

② Growing number of points : The no. of points inside the region $k_n$ must increase indefinitely as $n \to \infty$

i.e $\lim_{n \to \infty} k_n = \infty$

③ Small Fraction of Total Points : The fraction of total points inside the region must go to 0.

$$\lim_{n \to \infty} k_n / n = 0.$$

This ensures that the region does not cover too large an area and remains focused on the local structure around $x$.

## * Aspects of Density Estimation

(i) Bias - Variance Trade off ;

## * Parzen Window Method

→ a non-parametric approach to density estimation.

→ It works by placing a window (kernel) function $\varphi$ around each data point and summing up the contributions from all windows to estimate the density at a point $x$.

Formally,

→ define a window (kernel) function $\varphi(u)$, which determines the shape of the region around each data-point.

→ For eg. $\varphi(u)$ could represent a d-dimensional hypercube or a Gaussian kernel.

→ The volume of this region is $V_n = h_n^d$, where $h_n$ is the window width.

→ The density estimate using Parzen's window is given by:

$$P_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{V_n} \varphi\left(\frac{x - x_i}{h_n}\right)$$

$h_n$ = window width

$\varphi\left(\frac{x - x_i}{h_n}\right)$ determines whether a sample point $x_i$ falls inside the window centered at $x$.

→ The value of $h_n$ plays a critical role in determining the smoothness of the estimate:

(i) If $h_n$ is too large → estimate will be overly smooth and might miss important details (underfitting)

(ii) If $h_n$ is too small, the estimate will be too sensitive to noise in the data (overfitting)

* Applications of Parzen window Method

(i) pattern recognition & classification - estimate class-conditional densities of different classes

(ii) probability density estimation

(iii) Data smoothing - smooth noisy data while preserving overall trends

## * Aspects of Density Estimation.

(i) Bias - Variance Tradeoff - choice of window width or num neighbors controls the balance between bias and variance.

small window / kn → high variance ⇒ overfitting

Large window / kn → high bias ⇒ underfitting

(ii) Convergence - The estimate $p_n(x)$ converges to the true density as $n \to \infty$, provided certain conditions are met

(iii) Dimensionality - Non-parametric methods struggle in high-dimensional spaces due to the curse of dimensionality. The no. of samples required grows exponentially w/ the dimensionality of the feature space.