Statistical Decision Theory

G V V Sharma*

Contents

***Abstract*—This manual provides an introduction to statistical decision theory.**

1 Introduction

1.1 Let

$$Y = f(\mathbf{X}) \tag{1.1}$$

The *mean square error* (MSE) is defined as

$$MSE(f) = E\left[Y - f(\mathbf{X})\right]^2 \tag{1.2}$$

Show that

$$MSE(f) = E_{\mathbf{X}}\left\{E_Y\left[Y - f(\mathbf{x})\right]^2 |\mathbf{X} = \mathbf{x}\right\} \tag{1.3}$$

1.2 Let

$$c = f(\mathbf{x}) \tag{1.4}$$

Using (1.3)

$$\begin{aligned}\min MSE(f) &= \min MSE(f)|X \\ &= \min_c E_Y\left\{\left[Y - c\right]^2 |\mathbf{X} = \mathbf{x}\right\}\end{aligned} \tag{1.5}$$

Show that

$$\begin{aligned}MSE(f)|X &= E_Y\left\{\left[Y - c\right]^2 |\mathbf{X} = \mathbf{x}\right\} \\ &= -2cE_Y\left\{Y|\mathbf{X} = \mathbf{x}\right\} + E_Y\left\{Y^2|\mathbf{X} = \mathbf{x}\right\} + c^2\end{aligned} \tag{1.6}$$

1.3 $MSE(f)$ is minimum when

$$\frac{d}{dc}MSE(f)|X = 0. \tag{1.7}$$

Show that this results in

$$c = f(\mathbf{x}) = E\left[Y|X = \mathbf{x}\right] \tag{1.8}$$

2 Nearest Neighbour

2.1 Find $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$,

$$y_2 = \frac{1}{k}\sum_{i \in N_k(y_1)} x_{2i} \tag{2.1}$$

where

$$N_k(y_1) = \left\{i : |x_{1i} - y_1| < \epsilon < \left|x_{1j} - y_1\right|\right\}, \tag{2.2}$$

$$|N_k(y_1)| = k \tag{2.3}$$

2.2 Plot $\mathbf{y}_m, 1 \le m \le 100$ for $k = 15$.

2.3 Repeat the exercise for $k = 1$.

2.4 Compare with the least squares estimate.

*The author is with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad 502285 India e-mail: gadepall@iith.ac.in.