

STATISTICAL PATTERN RECOGNITION

ASSIGNMENT 3

Ali Gholami

Department of Computer Engineering & Information Technology
Amirkabir University of Technology

<https://aligholamee.github.io>
aligholami7596@gmail.com

Abstract

In this paper, we'll review the *parametric* techniques to estimate the *unknown* parameters of data distributions. We'll use, *MLE* and *Bayesian* estimation for *parameter estimation*. Also, we'll delve into the *non-parametric* techniques to estimate the unknown *density* of data distribution. We'll use *Kernel Density Estimation* methods such as *Parzen Windows* and other techniques such as *Histogram* and *k-NN* density estimation.

Keywords. *Parameter Estimation, Density Estimation, Non-parametric Methods, Parametric Methods, Kernel Density Estimation, Maximum Likelihood Estimation, Bayesian Estimation, Histogram Density Estimation, K-NN Density Estimation.*

1 Parameter Estimation

Let x_k , $k = 1, 2, \dots, N$ denote independent training samples from one of the following densities. Obtain the Maximum Likelihood estimate of θ in each case.

(a) $f(x_k; \theta) = \frac{x_k}{\theta^2} \exp(\frac{-x_k^2}{2\theta^2})$ where $x_k \geq 0$ and $\theta \geq 0$

(b) $f(x_k; \theta) = \sqrt{\theta} x_k^{\sqrt{\theta}-1}$ where $0 \leq x_k \leq 1$ and $\theta \geq 0$

Solution

(a) Substituting the given density inside the *MLE* equation yields the following results.

$$\hat{\theta} = \arg \max_{\theta} \{P(D|\theta)\} = \arg \max_{\theta} \left\{ \sum_{k=1}^n \ln P(x_k|\theta) \right\} \quad (1.1)$$

$$\hat{\theta} = \arg \max_{\theta} \left\{ \sum_{k=1}^n \ln \frac{x_k}{\theta^2} \exp(\frac{-x_k^2}{2\theta^2}) \right\}$$

$$\nabla_{\theta} l(\theta) = 0$$

where $l(\theta)$ is $\sum_{k=1}^n \ln \frac{x_k}{\theta^2} \exp(\frac{-x_k^2}{2\theta^2})$ in this case. Performing the gradient on the given equation yields the following results.

$$\sum_{k=1}^n \left(\frac{-2}{\theta} + \frac{x_k^2}{\theta^3} \right) = 0$$

The simplified estimate of unknown θ is given below.

$$\hat{\theta} = \sqrt{\frac{\sum_{k=1}^n x_k^2}{2N}}$$

(b) Substituting the given density in the Maximum Likelihood method yields the following result.

$$\hat{\theta} = \arg \max_{\theta} \left\{ \sum_{k=1}^n \ln \sqrt{\theta} x_k^{\sqrt{\theta}-1} \right\} \quad (1.2)$$

we can obtain the estimate for the unknown θ :

$$\nabla_{\theta} l(\theta) = 0$$

where $l(\theta)$ is $\sum_{k=1}^n \ln \sqrt{\theta} x_k^{\sqrt{\theta}-1}$ in this case. Performing the gradient on the given equation yields the following results.

$$\frac{n}{2\theta} + \frac{1}{2\sqrt{\theta} \sum_{k=1}^n \ln x_k} = 0$$

multiplying the whole equation by θ results in the following equation:

$$\hat{\theta} = \frac{n^2}{(\sum_{k=1}^n \ln x_k)^2}$$