

Seminar Report - Fitting Mixture of Distributions with R package mixdist

Nik Počuča - 1322223

November 13, 2018

0.1 Introduction

0.1.1 Peter Macdonald

Dr. Peter Macdonald, professor emeritus from McMaster University dives into his lifetime research and life experiences working with finite mixture models in fishery-length frequency analysis. Dr. Macdonald served on numerous United States Environmental Protection Agency FIFRA Scientific Advisory Panels, reviewing methodology proposed by the EPA for pesticide risk assessment. Dr. Macdonald presents the pike measurements taken in 1965 in that of Heming lake and demonstrates how mixture modelling can be used to track length of pike of fish.

0.1.2 Finite Mixture Model

A finite mixture model is a model in which multiple distributions play a role on the sample space. These models often arise when sampling from heterogeneous populations with a specific probability density function on each partition, For example in frequency distributions of animal populations with various age-groups. A finite mixture has a finite number of groups contributing to the density that lives on sample space Ω .

Definition 0.1.1. Finite mixture model Suppose a random variable \mathbf{X} takes on values in a sample space Ω in which it's distribution can be represented by a probability density function (continuous case) of the form

$$p(\mathbf{x}; \boldsymbol{\theta}) = \sum_{k=1}^G \pi_k f_k(\mathbf{x}; \boldsymbol{\theta}_k).$$

Here \mathbf{x} are realizations of \mathbf{X} , π_k is the mixing proportion of group k , where $0 \leq \pi_k \leq 1$, $k = 1, \dots, G$ and $\sum_{k=1}^G \pi_k = 1$. $f_k(\mathbf{x}; \boldsymbol{\theta}_k)$ is the probability density function of partition k and $\boldsymbol{\theta}_k$ as a parameter vector for that group k .