



## **Model-Based Clustering of Single and Multiple Samples**

Geoff McLachlan\* and Sharon Lee  
University of Queensland

In this talk we consider the modelling of multivariate observations in a random sample from a population composed of one or more components with asymmetric distributions. The distribution for each component is modelled by the skew normal and various skew t-distributions. For an individual sample, we discuss the fitting of these distributions by maximum likelihood. In the case where there are multiple samples taken on objects with significant inter-object variation, we consider the alignment problem of matching the sample-specific clusters across the available samples. We discuss the adoption of a framework with random effects terms that allows the clustering of the samples and the matching of the clusters to be carried out simultaneously. The development of this methodology has been motivated by the aim of providing an automated analysis of data on cells in flow cytometry and examples from this field will be given to illustrate the proposed methods.