Model fitting and Time series approaches to the analysis of menstrual bleeding data

by

Nayana Weerasooriya

Department of Statistics and Computer Science

University of Colombo

Thesis submitted for the Degree of Master of Philosophy

University of Colombo

Sri Lanka

January 1991

Abstract

This thesis is based on data made available by the World Health Organisation. The data base consists of daily menstrual bleeding records of 4817 women using one of ten different fertility regulating methods. After introductory chapters that discuss the data base and review previous work on the analysis of menstrual bleeding patterns, two new approaches to the analysis of data are considered in detail. An approach based on techniques used in the analysis of time series data is discussed in Chapter 4. The major part of the thesis, Chapters 5 to 8, describe the use of generalised linear models to model the lengths of bleeding and non-bleeding events.

A time series approach to the analysis of menstrual bleeding data has been introduced by Pochobradsky (1970). In this thesis, his approach is extended by using principal component analysis to analyse the periodicity of vaginal bleeding patterns of women. Statistics that summarise the shape of the periodogram (a powerful tool in the study of rhythmic processes) of daily bleeding data are used to describe the variation in the bleeding patterns among different contraceptive users and to summarise the bleeding experience of women under different contraceptive methods.

In modelling menstrual bleeding data, interest is centered on the relationship between a woman's bleeding pattern behaviour and her age, height, weight, number of pregnancies, outcome of last pregnancy and contraceptive history. To model the lengths of bleeding episodes, the shifted Poisson distributional assumption is made and to model the lengths of non-bleeding intervals, the shifted negative binomial distributional assumption is made. Algorithms for fitting these models within the framework of generalised linear models are developed. Even after isolating fixed effects due to the measured variables such as age, number of pregnancies etc., it was noticed that there is still a considerable amount of variation left unexplained by the model. Due to the fact that data come from different women it is possible to exist

a between women variation in data. In order to improve the model in this respect this women effect is also introduced in to the model. However it is natural to consider the variation due to this women effect as a random component rather than a fixed component in the model, and so a random effects modelling approach is introduced to allow for this between women variation. The merits of this approach are described.