

APTS 2007/08: Spatial and Longitudinal Data Analysis

Module leader: Peter J. Diggle (School of Health and Medicine, Lancaster University)

Assessment exercises (choose one)

1. Analyse the schizophrenia trial data with a view to establishing whether either or both of the active treatments are superior to placebo and which, if either, is the better of the two active treatments, bearing in mind that a 20% reduction in the average PANSS score is considered to be clinically significant.

Summarise your conclusions in a short report suitable for passing to an agency whose task is to decide which, if either, of the two active treatments should be approved for routine clinical use.

The data are available at www.lancs.ac.uk/staff/diggle/APTS_data-sets as the file `PANSS_short_data.txt`

2. Analyse the Galicia lead pollution data with a view to mapping those parts of Galicia in which, with probability at least 0.5, lead pollution exceeds 2.0 units.

Summarise your conclusions in a short report suitable for passing to an environmental protection agency.

The data are available at www.lancs.ac.uk/staff/diggle/APTS_data-sets as the file `lead2000_data.txt`

3. Theoretical exercise on Markov random fields

(a) Prove the result that for a random variable $Y = (Y_1, \dots, Y_n)$, any self-consistent set of full conditional distributions $f_i(y_i | \{y_j : j \neq i\}) : i = 1, \dots, n$ determines the joint distribution of Y .

(b) Show that the autologistic model for a binary-valued random fields on a square lattice, defined by the full conditionals

$$\text{logitP}(Y_{ij} = 1 | \{Y_{kl} : kl \neq ij\}) = \alpha + \beta(Y_{i,j-1} + Y_{i,j+1} + Y_{i-1,j} + Y_{i+1,j})$$

satisfies the Hammersley-Clifford theorem (you may assume that $|\beta| < 0.25$).

(c) Give an example to show that a specification for a binary-valued random field on a square lattice through full conditionals that are individually well-specified may not satisfy the Hammersley-Clifford theorem.

(d) Can you construct a Gaussian Markov random field for $Y = (Y_1, Y_2, Y_3)$ such that all of the off-diagonal elements of $\text{var}(Y)$ are negative?

4. Analyse the macaque retina data with a view to establishing whether they do or do not support either or both of the following developmental hypotheses:

H1 the two types of cell form initially in separate layers which later fuse to form the mature retina;

H2 the two types of cell form initially in a single layer of undifferentiated cells, which later separate into two separate layers.

Summarise your conclusions in a short report suitable for passing to a developmental biologist.

The data are available at `www.lancs.ac.uk/staff/diggle/APTS_data-sets` as the file `wassle_data.txt`

5. The paper by Brown, Diggle, Lord and Young (2001) uses a spatio-temporal Gaussian process to improve radar-based rainfall maps by calibrating the radar images against a network of ground-truth rain-gauges. Write a short summary of the paper (approximately 500 words) and discuss the strengths and weaknesses of the modelling and analysis carried out by the authors (I will not be marking this assignment, so you can be as rude as you like).

Brown, P.E., Diggle, P.J., Lord, M. and Young, P.C. (2001). Space-time calibration of radar-rainfall data. *Applied Statistics*, **50**, 221–242.

PJD, 4 September 2008