Isaac Newton Institute for Mathematical Sciences

Design of Experiments

21 July to 15 August 2008

Organisers: Professor RA Bailey (Queen Mary, London), Dr B Bogacka (Queen Mary, London), Dr H Grossmann (Queen Mary, London), Dr D Woods (Southampton)

The experiment has been a cornerstone of progress in science and technology since the days Galileo dropped cannonballs from the Leaning Tower of Pisa. But it was not before the 1930s that rigorous theories for the Design of Experiments (DoE) began to emerge, when it was recognized that a mathematical-statistical approach was required to meet the needs of experimentation in an increasingly complex world. DoE is essential for assuring the integrity and reproducibility of research findings, and design principles such as randomisation, replication and blocking have become standards of best practice in many areas.

As a discipline DoE is closely intertwined with combinatorics and statistical modelling, but has also been influenced by applied sciences like agriculture and chemical engineering among others. New challenges are arising as technology is advancing rapidly. This programme embraces both applications and theory and is organised around three themes in each area. In each of the application areas—genomics and proteomics, computer experiments and clinical trials—experiments are performed to foster the understanding of genetic processes, the climate, diseases and other complex phenomena. The corresponding design problems are varied and often require the development of new theory or interpreting existing results in an entirely new way.

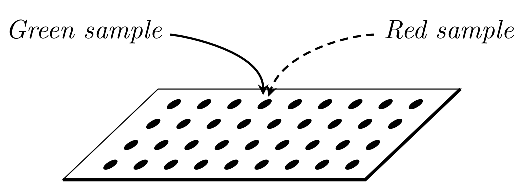

Figure 1: Two-colour microarray experiment

To illustrate some of the issues involved consider Figure 1 which schematically depicts a two-colour microarray experiment in genomics. Put simply, these experiments are undertaken to identify from among the thousands of genes potentially controlling the behaviour of cells those which can be found more (or less) frequently in, for example, normal as compared to cancer tissue. The experimental procedure involves comparing coloured pairs of genetic samples extracted from biological material of interest with different reference genes represented by dots in the figure. Design questions arising for this kind of experiment are: Which biological samples should be paired and how should the colours be assigned? How many microarrays are needed to obtain reliable and precise results? What is an optimal design given the available resources? A solution to the last problem for six samples and nine microarrays is shown in Figure 2 in the form of a directed graph where arrows connect samples that occur on the same microarray. The sample at the tip of each arrow is coloured red and the one at the base green.

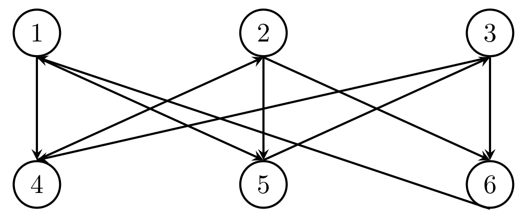

Figure 2: Optimal design

The methodology driven themes in the programme are multi-stratum experiments, experiments under nonlinear models and differential equations, and multitiered experiments. Each of these is already used in several areas of application, but often with different vocabulary. Statisticians from different areas will bring different tools to the table and pool expertise. Because the underlying methodology is the same, it should be possible to learn from each other in a short time and so incorporate the best from each area to produce a more unified theory.