A proposal for a mini-project on

A case study of statistical methods for sustainability

Prof. Adrian Bowman
School of Mathematics & Statistics
The University of Glasgow

Sustainability is an important part of the HEA agenda and statistical methods provide a natural point of connection for the mathematical sciences. This project proposal is built around existing material which has been used for some time in a teaching setting. It has also formed the basis of several school and more general outreach presentations. It therefore has the advantage of being tried and tested as successful teaching material. This proposal is to develop from this a set of resources which will be available to others in a flexible form, applicable to different levels of experience of both teacher and student. The material also has associated software for the R statistical computing system which can be used to create dynamic graphics, allowing student exploration of the data as well as teacher demonstration.

The case study is based on data provided by the Scottish Environment Protection Agency (SEPA) on water quality in the River Clyde. Measurements of dissolved oxygen (DO) from surface water samples are available at several positions on the river over a period of twenty years. The aim is to model changes in the mean level of DO over the years. However, a model also has to allow for seasonal effects and for different patterns at different sampling stations. The basic underlying method is regression but there are several more unusual aspects which can be explored. The four scenarios below give an indication of this. Quite a number of other scenarios are potentially available.

1. Students who are relatively inexperienced in statistical methods can explore the data from a single station. The seasonal effect can be modelled by a trigonometric term. Dynamic graphics using the `rpanel` package are available to help students think through the issues. In fact, the trigonometric terms can be fitted by standard linear methods. Deriving this is in itself a useful and instructive exercise.
   This scenario is likely to be the one of most interest for students at the introductory level. Nonetheless, material from the scenarios below could be usefully discussed in more general, non-technical terms.

2. More experienced students can model the data for the whole river. This introduces issues of repeated measurements, for which convenient and easily implemented models are available in R through the `lme` package.

3. More advanced work would allow effects to be non-linear, potentially using smoothing techniques to model these flexibly. Code in R will be provided to experiment with this.

4. A further interesting aspect to be explored is the effect on DO of the upgrade of a

sewage treatment works at a time in the middle of the period for which data are available.

The work of the project will involve the preparation of teaching materials in the form of:

1. notes for teachers indicating how the material might be used at different levels;
2. background notes and questions for students at different levels;
3. menu-driven dynamic graphics in Rto allow the data to be explored interactively;
4. scripts in R to model the data in different ways.

The timescale of the project would be July-December 2011. There is an expectation that the work might be completed by the end of September. However, for safety, the end of December is proposed as a firm completion date.

The cost of the project is £5500 to cover the time of Prof. Adrian Bowman in preparing the material in this form. In order to reflect the timescale of the project, across two financial years, it is proposed to schedule the payments as £500 for July and £5000 for August to December.

The copyright of the material would remain with Adrian Bowman. In particular, there is a possibility of publishing a paper in a teaching statistics journal. This would have the benefit of making the material available to a wider audience. However, to ensure proper distribution through HEA and MSOR, a permission statement is given below.

Adrian Bowman will disclose all products and materials resulting from this work to the HE Academy Maths, Stats & OR Network. IPR will remain with Adrian Bowman and the University of Glasgow. However, Adrian Bowman grants a royalty-free licence for the HE Academy Maths, Stats & OR Network to use copyright materials for academic purposes and to make such materials freely available to UK Higher Education Institutions for academic purposes.