

Academy for PhD Training in Statistics (APTS)

A taught course centre for PhD students in the Mathematical Sciences

A: Vision for the taught course centre

The **aim** of this proposal is to improve the quality of PhD training in statistics throughout the UK, by creating a collaboration between nine prominent UK university statistics research groups and departments to deliver a basic core of first-year training for statistics PhD students across the UK. (Here “statistics” refers to statistical methodology, applied statistics, and applied probability.) The collaboration includes all three statistics groups funded to date by EPSRC Science and Innovation awards, and all collaborating groups have considerable experience in training PhD students. Our aim is to meet the concerns about UK PhD training raised by the 2003 EPSRC/CMS International Review of Mathematics, about the breadth and quality of UK PhD training and specifically the chronic shortage of well-trained research statisticians; specifically we plan to provide a course of modules which is loosely comparable in general terms to the training delivered by major statistics departments in the USA. The need for first-year PhD training is particularly pressing in statistics, since few incoming PhD students have taken a first degree in statistics. It is already the case that such courses are being planned by individual institutions in the UK. EPSRC funding will enable us to develop and establish a pattern and expectation of collaboration which will:

- be much more efficient in terms of use of researcher time;
- ensure a broader and more systematic training than can be provided by any one institution acting on its own;
- build and strengthen a community of PhD students across the UK which in the long term will have profound positive effects on collaborative research in UK statistics;
- provide students with the opportunity to form networks with each other which could be useful for them whatever eventual career paths they take, whether in the UK or overseas.

There is already a precedent for collaboration in statistical research education at a higher level: EPSRC/LMS funded advanced courses, the EPSRC/RSS Graduate Training Programme in statistics, and a well-established annual Research Students Conference in statistics run by the research students themselves. In general this training takes place either at a more advanced level than basic first-year PhD training, or is based around specific themes rather than providing a comprehensive training. The **objectives** of the current APTS proposal are:

- to provide a yearly programme of 8 modules, each of 10 lectures plus tutorials and substantial additional preparatory material, to cover the majority of the core statistical knowledge which should form the first-year training for UK statistics PhD students;
- to make this accessible to 30 EPSRC-funded students per year drawn from across the UK, plus up to 40 further students. In the period funded by this grant, the 30 EPSRC-funded places will be supported from the APTS budget in terms of travel, accommodation and subsistence. Further access will be provided to students not able to come to the training weeks by means of low-cost web-casting technology;
- by the end of the funding period, to have established a model for collaboration between statistics groups across the UK which allows this level of training to continue.

B: Course provision

The philosophy of this proposal is to produce training opportunities in first-year PhD statistics which are based on the expectation that students will be able to learn independently in their own time if they are also exposed to regular intensive training periods. Students will register for a course centred on 4 intensive training weeks delivered over the year, each training week being composed of 2 taught modules plus occasional supplementary lectures (for example, on the process of preparation and submission of scientific papers). While statistical research is rather diverse in nature, nevertheless it is possible to identify a core framework of topics drawn from applied statistics, statistical methodology, and applied probability, both in terms of specific tools and capabilities and in terms of broad overviews which provide the student with mental “world maps” of the relevant area and its major literature. An indicative list of possible modules is as follows: statistical computing, statistical asymptotics, matrix and tensor methods for statistics, statistical distribution theory, approaches to statistical modelling, applied stochastic processes, Markov chain Monte Carlo, Bayesian methods.

Statistics researchers come from a variety of backgrounds; the rationale underlying choice of module topics will be to supply a range to cover the major part of statistical training for most first-year PhD statistics students, to cover gaps in knowledge and to reinforce and establish competence. In comparison with major US statistics departments we expect to provide a broadly comparable training framework, somewhat less extensive because of the need to organise the course around 4 intensive training weeks, but exploiting a much wider range of potential lecturers.

A typical intensive week contains 2 modules each of 10 lectures plus practical sessions. In order to make the most efficient use of the week, there will be an expectation of substantial prior preparation by the students using prior reading and exercises as well as web casts of a couple of preparatory lectures,¹ all supplied by the module lecturer. During the module formative assessment will be practiced (for example, in-lecture tests marked by lecturer); normative assessment will be left to the students’ home institutions for resource reasons, and because this will need to comply with the home institution’s regulatory framework. There will also be a follow-up process based on student self-assessment at the end of the module, leading to written suggestions for further study by the module lecturer.

Potential lecturers will be identified and invited by the committees described in the management structure given in the next section. Institutions participating in this grant have committed to supply module lecturers if required, and they include institutions specifically resourced by EPSRC Science and Innovation awards to develop UK statistical research; we therefore do not anticipate any difficulty in finding excellent candidates for lecturers.

This use of distance learning in conjunction with face-to-face teaching is intended to maximize the benefit of the intensive weeks, both by increasing student-preparedness and by allowing routine material to be covered beforehand. We will also experiment with enhancing the webcasts using more recent technology (specifically, weblogs and interactive whiteboards) in as cost-effective a manner as possible.

It should be noted that a major value of the intensive weeks will be the opportunities given for students in a given cohort to mix informally and to interact with more senior and established researchers, and thus to gain a sense of the potential for collaboration at national levels. The well-established UK Research Students Conference in statistics already provides some limited opportunity for this; we expect it to be greatly enhanced by contacts made in the intensive training weeks.

The course will be marketed by means of:

- advertisements in the Royal Statistical Society’s *RSS News*;
- posters distributed to all UK university statistics departments and groups;

¹For example, see the INI lectures such as <http://www.newton.cam.ac.uk/webseminars/mondays/2005/1205/bartnik/>; we have identified means of cheaply delivering lectures which include synchronization of slides and voice.

- the UK `allstat` web bulletin board;
- a course website including facilities for web registration;
- personal contacts between committee members (see below) and UK statistics PhD supervisors.

We anticipate the last of these to be of major importance; representatives of institutions playing a part in the committee structure described below will be expected to play their part in ensuring national coverage of personal contacts to encourage supervisors to commit their students to this course.

A modest registration fee will be charged (of the order of £50 per student per training week). In the event of over-subscription for a course, preference will be shown first for EPSRC-funded students, then for students from UK institutions, then (if applicable) for students from institutions outside the UK. Should it be necessary to show further preference, this will be done by allocating places to institutions on a quota basis with some preference given to students from participating institutions. In the event of over-subscription for EPSRC-funded places in a given year, we will offer reduced support *pro-rata*. Students who are unable to attend will be given access to lecture slides and preparatory material including webcasts, at a reduced registration fee.

C: Management strategy

The centre is intended to have a national reach and its management structure is designed to reflect and to enable this by involving UK statistics departments across the country. We propose a three-tier structure of (1) centre Director (and associated staff) based at Warwick; (2) Executive Committee involving representatives of institutions named as collaborating on the grant² and some others; and (3) an Advisory Committee composed of representatives of “member institutions” of the centre, being institutions which commit to requiring their first-year statistics PhD students to register and to use the programme.

At an operational level the centre will be managed at Warwick by a Director (one of Firth and Kendall), an academic administrative assistant, and secretarial and technical staff. The Director will be responsible for EPSRC funding for the centre, for managing the process of registration of PhD students, and for liaising between the students and the institutions running the training weeks and the module lecturers.

The *Executive Committee* will be chaired by the Director and will be composed of representatives of all nine institutions collaborating in the grant, and 2 further representatives serving 2-year terms and nominated by the Advisory Committee as below. It will meet 3 times a year to monitor and to establish training weeks, to solicit and to evaluate applications to run training weeks and to deliver course modules, and to ensure quality and balance of activities. Each member will assist the Director in advertising the course, and their institutions will commit to be available to run training weeks up to the rate of 1 training week per year, and to be available to supply lecturers to deliver modules up to the rate of 2 modules per year. (These last do not exclude other institutions from running training weeks or supplying lecturers.)

The *Advisory Committee* is composed of the Executive Committee together with representatives of a number of other “member institutions”. Member institutions commit to requiring their statistics PhD students to register with and to use the course. The Advisory Committee will meet once per year (though much business will be transacted by email), will provide feedback on modules, will advise on topic and lecturer choice, and will contribute modules to training weeks where appropriate and cost-effective. Member institutions will be recruited during the first year of operation of the grant.

²Prof S.N. Wood (Bath), Dr E.J. Collins (Bristol), Prof S.P. Brooks (Cambridge), Prof A.W. Bowman (Glasgow), Prof G.O. Roberts (Lancaster), Prof A.J. Wood (Nottingham), Prof S.L. Lauritzen (Oxford), Prof. J.J. Forster (Southampton), Prof D. Firth (Warwick)

The normal expectation for student registration is that students will register for and attend all 4 training weeks and participate in all 8 modules; however it is expected that some variation will occur where individual students have already acquired relevant training through previous post-graduate education.

D: Quality control strategy

A key driver for the long-term success of APTS is a keen commitment to continuous quality improvement, in relation to every aspect of its activities. Specific quality improvement (QI) action plans will be developed and implemented at four distinct levels:

- individual modules (content and delivery)
- individual training events/locations (organisation, facilities)
- online materials (general information, course materials and registration)
- overall APTS direction and management

The full set of QI action plans will be critically reviewed by the Director and academic administrative assistant after each training week, and by the Executive Committee at each of its meetings. The Director, with the aid of the academic administrative assistant, will be responsible for ensuring that agreed QI actions are duly implemented.

The QI action plans will themselves will be developed and updated in response to systematic quality monitoring activities, which will include:

- student feedback on module content and delivery, on course organisation, etc.
- student self-assessment returns for each module
- a critical report by the academic administrative assistant on each training week
- feedback from lecturers
- feedback solicited regularly from the Advisory Committee

The physical presence of the academic administrative assistant during training weeks will be important in regard to this activity.

The primary mechanism for both the sharing of information on best practice and for the commissioning of suitably high-quality course materials will of course be the Executive/Advisory Committee structure already described. The formal quality improvement activities listed here will serve to inform the agendas of both committees.

Partly for the purposes of tracking long-term impact on the PhD experiences and future careers of APTS students, we will maintain — with the students' consent — a database of APTS alumni. This will allow occasional sample surveys of former students to solicit their perceptions of the impact of APTS, as well as providing potential opportunities for (as yet unplanned) APTS-organised activities in continuing professional development, for occasional alumnus presentations during APTS weeks, etc.

We will of course seek to collaborate and to share resources both with other centres funded by this call and with the existing CETL for postgraduate education in statistics (at Lancaster), so as to establish best practice and to share experience.

E: Sustainability strategy

We attach a letter from the University of Warwick affirming its commitment to the proposed taught course centre. As stated in the objectives above, by the end of the funding period we plan to establish a model for collaboration which is seen to be cost-effective and which attracts sufficiently many PhD students to be broadly sustainable by charging registration fees. We believe this to be realistic: according to figures obtained from EPSRC there are of the order of 40 starts per year of EPSRC-funded statistics PhD students; our objective of attracting 30 EPSRC-funded students annually is based on this figure and is born out by our estimate that the collaborating institutions expect to recruit substantially more than half that number each year. Since non-EPSRC students will also participate on the course, it seems achievable to aim for a total participation of 60 – 70 students per year, particularly if good use is made of the management structure to involve a wide range of UK institutions. After the end of the five-year funding period we plan to increase the registration fee to the order of £100 per training week per student; this will provide a firm basis for reimbursing lecturers and funding administration costs for training weeks and the programme as a whole, while sending departments would take responsibility for funding travel, accommodation and subsistence for all their students. During the grant we aim to run training weeks at a variety of locations across the UK: the resulting model of collaboration would therefore also adapt well to economize on student travel and accommodation by running several instances of the same training week at different locations (albeit with increased costs for instructors and a reduced payoff in terms of national community-building).

F: Evaluation of impact of the course

The main impacts of APTS will be on:

- the preparedness and PhD progress of individual research students
- the recruitment of statistics PhD students in the UK, and the competitiveness of UK-trained doctorates in academic and research job markets
- the long-term health of statistical research in the UK (and thus, given the UK’s prominence internationally in statistics, worldwide).

We will evaluate these annually at the level of the Executive Committee as follows: by systematically canvassing and exploiting the views and experiences of the Advisory Committee, by measuring the extent of continued participation of member institutions, and by reviewing records derived from our alumni database of the frequency of completed PhD for academy students, and of the pattern of appointments to academic positions in the UK and abroad. Note that the third impact is very long-term (time-scale 10–15 years), and will surely be affected by a wide variety of other factors.

The extent to which APTS meets its stated objectives (see section A) should be rather more easily assessed. Simple metrics include the number of modules delivered each year, the number of students attending each module, and the number of member institutions each year; and the corresponding figures for years beyond the initially-funded five.

G: Justification of resources

1. Costs of the Director’s time (estimated at 5% of a professorial post) will be met by Warwick. Secretarial and technical-support costs (10% of each, rising to 50% during course weeks, total 5-year cost £28K) are essential for running the registration process and the website. A fraction of the cost of a post-doctoral position (10%, rising to 100% during course weeks and 50% in the week either side: total 5-year cost

£48K) is needed in order to supply the academic administrative assistant, who will liaise intensively on practical and academic issues both with module lecturers and with institutions running training weeks, will provide on-site monitoring and support at training weeks, and will implement in detail the quality improvement strategy. This combination of management levels and skills will be essential for the efficient running of the taught course centre.

2. Costs of accommodation, subsistence and travel for EPSRC-funded PhD students are based on median figures for the nine collaborating institutions. We estimate 30 EPSRC-funded student places per week; this is based on EPSRC's own figures. We expect the course eventually to establish sufficient value to achieve considerably greater numbers, in which case the arrangements for over-subscription specified in B above may need to be applied.
3. Course preparation and update costs take into account a substantial amount of work required to build in good preparatory material as well as the lectures themselves, and to update modules in the light of further experience. No provision is made for assessment beyond a modest amount of formative assessment at the end of each module, as this would drive actual costs (in terms of lecturer time) much higher and so jeopardize recruitment of module lecturers: it is expected that member institutions will make their own arrangements for any assessment required beyond this.
4. Equipment costs are justified as follows. An important component of course provision is the period of prior student preparation for each module, involving preparatory web-cast lectures. We have identified the most suitable and cost-effective software and equipment to facilitate these lectures (tablet PC, equipment to record from whiteboards, and screen capture software). Each collaborating institution is expected to be equally involved in module presentation activity; we therefore need funding for one set of equipment/software per institution.

H: Diagrammatic plan

(attached)