Workshops and training sessions

What is? Workshop – Monday 9 September, 1.30pm

Chair: Paul Norman (University of Leeds)

What is …?

This session is aimed at those who want to know more about getting skilled up. The audience might be PhD students, postdocs, local or national government researchers, or just anybody intrigued about knowing about something new to them. Each (c. 10 minute) presentation will introduce in an accessible manner a quantitative method and will explain what you can do with it and with what kinds of data. Where possible there will be research examples and pointers to materials which can help people learn the method.

The methods: Logistic regression & survival; Iterative Proportional Fitting; Spatial Interaction; K-means classification; Principal Components Analysis; Microsimulation; Curve fitting / smoothing.

The presenters: Fran Darlington-Pollock, Nik Lomax, Adam Dennett, Dan Exeter & Paul Norman. All are fully house-trained in presenting complicated material in an accessible manner.

Logistic regression and survival

Logistic regression and survival analysis allow us to analyse binary outcomes. For example, we might be interested in a particular event such as moving house or a particular characteristic such as health status. Both methods model the relationship between the categorical response variable of interest and one or more predictor variables yielding either odds ratios or hazard ratios. However, while the predictor variables in logistic regression are either categorical or continuous, survival analysis can also account for time. Both methods are useful in population geography and demographic research.

Iterative Proportional Fitting

Iterative Proportional Fitting (IPF) is a technique that can be used to adjust a distribution reported in one data set by totals reported in others. There are various data situations in population research when values for population attributes might be missing due to being unknown, unreliable, outdated, or a sample. IPF provides a tool for estimating these missing data.

Spatial Interaction Models

Spatial Interaction Models have a well-established history in population research and are used to understand, analyse, explain and estimate movements of people (typically migration or commuting flows) between geographic locations and infer properties associated with origins, destinations, and their physical separation, such as their attractiveness or accessibility. Where flow data are available, the fitting of models and the estimation of parameters associated with origins, destinations and separation can reveal important information about the system being studied.

K-means classification

K-means classification uses a range of socio-demographic data as inputs to group geographic areas based on the similarity of their characteristics. Also known as clustering, this technique groups areas by whether places are like each other rather

than necessarily near each other. The relative presence or absence of social (or other) attributes are used to devise a 'label' for each group of areas. These categories of places can be used for targeting resources or as an analytical variable.

Principal Components Analysis

Principal Components Analysis (PCA) is a dimension reduction technique used widely in population statistics. Through PCA, 'latent' underlying relationships between a large number of highly correlated variables can be captured by fewer factors which represent sets of the original variables. PCA has contemporary use in the production of deprivation indexes. In the New Zealand Deprivation Index (NZDep), census variables representing different dimensions of deprivation for small area 'meshblocks' are reduced down to one Principal Component. The values and signs of the coefficients for the NZDep scores are consistent with the expectation for a measure of socio-economic circumstances.

Microsimulation

Microsimulation is an approach used to estimate the characteristics of individuals within a population from a range of attribute-rich data sources and for modelling those individuals over time. The creation of a synthetic population of individuals is often called spatial microsimulation because spatial identifiers are added to the individuals. Dynamic models deal with time steps; they are used to age the individuals and alter their characteristics by introducing the probability of transitioning between different states. For example, dynamic models can be used in health policy to assess outcomes for individuals in the model under different intervention scenarios.

Curve fitting/smoothing

In demographic analysis we often have data which are grouped by age or are spiky. Non-linear curve-fitting models can be used for the estimation of single-year-of-age information from five-year grouped data. The same approach can be applied to smooth rates where the original data are ragged. Curve fitting is a useful skill to have within the toolbox of population geographers and demographers. The scary equations are a barrier which can be overcome fairly readily (though not necessarily easily!).

Email: P.D.Norman@leeds.ac.uk

Training session: UK Census Longitudinal Studies – Tuesday 10 September, 9.00am

Lee Williamson[1], Tom Clemens[1], Oliver Duke-Williams[2], Ian Shuttleworth[3]

[1]SLS, [2]CeLSIUS, [3]NILS

An introduction to using the UK Census Longitudinal Studies.

This training workshop session is designed to introduce people unfamiliar with the analysis of longitudinal data and the unique social science that can be undertaken with microdata that tracks individuals over time to the kinds of analyses that can be carried out. The session will provide a general introduction to the UK national Census Longitudinal Studies (LS): the England and Wales ONS LS, the Scotland LS (SLS) and the Northern Ireland LS (NILS).

This brief talk will be followed by an opportunity for delegates to have a hands-on session to:

• explore which variables are held by each LS in the data dictionary and use test data;

• have help completing an application to use LS data (main example from the SLS);

• meet with Support Unit staff from all three studies and discuss the development of new research projects.

Please bring your own laptop for the hands-on part of the session.

No previous experience of microdata or statistical analysis techniques is required.

Early career panel: Grant applications – Tuesday 10 September, 9.00am

Panel members: Professor Jakub Bijak (University of Southampton), Professor Jane Falkingham (University of Southampton), Professor Hill Kulu (University of St Andrews), Professor Melinda Mills (University of Oxford)

Chair: Dr. Julia Mikolai (University of St Andrews)

This panel session will provide valuable, otherwise not easily accessible, information to early career academics on the decision-making process, what grant reviewers look for, what grant assessment panel members look for, and tips and tricks on writing a successful grant application.

Training Workshop: An introduction to social science genomic data analysis – Wednesday 11 September, 9.00am

Melinda Mills[1] and Nicola Barban[2]

[1]University of Oxford, [2]University of Essex

Within the past few years, there has been a growth in genetic discoveries related to key demographic topics, such as fertility behaviour (age at first birth, number of children ever born), educational attainment, risk, well-being and many more. This has been coupled with increased availability of data that contain genetic information. For demographers it is often difficult to understand how and if this approach can be implemented into their own research. This workshop is aimed at beginners who are interested in integrating genetics into their demographic research. We will introduce some of the fundamental concepts and misnomers, followed by several hands-on computer-based exercises using sample data. The workshop is open to all career stages. The expectation is that you have a basic understanding of statistics, but there are no other prerequisites.