Institution: UDur: University of Durham

Unit of Assessment: Unit 10: Mathematical Sciences

Title of case study: Industrial impact of Bayes linear analysis

1. Summary of the impact

This study demonstrates how Bayes linear methodologies developed at Durham University have impacted on industrial practice. Two examples are given. The approach has been applied by London Underground Ltd. to the management of bridges, stations and other civil engineering assets, enabling a whole-life strategic approach to maintenance and renewal to reduce costs and increase safety. The approach has won a major award for innovation in engineering and technology. The methodology has also been applied by Unilever and Fera to improve methods of assessing product safety and in particular the risk of chemical ingredients in products causing allergic skin reactions.

2. Underpinning research

Bayesian analysis is a well established approach for combining expert judgements with data to quantify uncertainties about real world outcomes in a probabilistic form appropriate for inference and decision-making. There are two practical problems with this approach, for large and complex problems. Firstly, it requires a level of detail which goes far beyond the ability of the expert to provide meaningful judgements, leading to many arbitrary aspects of the prior formulation. Secondly, the analysis is very computer intensive, typically requiring large-scale numerical simulations which are highly sensitive to certain features of these somewhat arbitrary prior specifications. Therefore, often the analysis is both non-robust and too complex to allow a proper exploration of its sensitivity, particularly in problems of optimal experimental design or sample choice.

Bayes linear analysis has been developed by Michael Goldstein, in Durham, with many collaborators, to address these issues, by both simplifying the specifications required to carry out the analysis and reducing the complexity of the analysis itself. It does this through a geometric approach to statistical inference which takes expectation, rather than probability, as primitive, allowing us to make a limited number of expectation statements, rather than requiring a complete probability specification, and constructing appropriate methodology based on orthogonal projection (which is computationally simpler than full Bayes) for analysing uncertainties based on a partial specification. The foundations and methodology are described in detail in **[1]**, which is the general underpinning research for all of the impact described in this case study, which concerns the ways in which the Bayes linear approach has impacted on industrial practice.

We choose two areas of research and application to demonstrate this impact.

(i) The paper **[2]**, by Goldstein and O'Hagan (U. of Nottingham) considers problems where a decision maker must estimate a set of unknown quantities and receives expert assessments at varying levels of accuracy on samples of the quantities of interest. The paper introduces the general notion of Bayes linear sufficiency, derives its properties and uses the tractability associated with the Bayes linear formulation to underpin a practical methodology for the design and analysis of studies relating to very large systems of assets. We will describe below their impact for London Underground.

(ii) The paper **[3]**, by Goldstein and Shaw (a postdoc at Durham, 1999 - 2002, when this research was carried out) extends the Bayes linear approach by introducing "Bayes linear kinematics" which merges aspects of full Bayes and Bayes linear inferences. (The notion is by analogy with the well-established "probability kinematics".) This allows the construction of "Bayes linear Bayes graphical models", which combine the simplicity of Gaussian graphical models with

the ability to allow full conditioning on marginal distributions of any form. The approach was first developed to address problems in Bayesian reliability testing for complex systems (see **[4]** and **[5]**). The flexibility of the Bayes linear kinematic makes it an appropriate tool for risk assessors who want to quantify their uncertainty about hazards based on disparate sources of information, and we will describe, below, the use of such methods in FERA and Unilever.

3. References to the research

[1] M. Goldstein & D.A. Wooff (2007*) Bayes linear statistics: theory and methods*, Wiley, ISBN: 978-0-470-06567-9.

[2] M. Goldstein and A. O'Hagan (1996) *Bayes linear sufficiency and systems of expert posterior assessments*, Journal of the Royal Statistical Society, series B, 58, 301-316, Stable URL: http://www.jstor.org/stable/2345978.

[3] M. Goldstein and S. Shaw (2004) *Bayes linear kinematics and Bayes linear Bayes Graphical Models*, Biometrika, 91, 425-446, doi:10.1093/biomet/91.2.425.

[4] F.P. Coolen, M. Goldstein & D.A.Wooff (2007) *Using Bayesian statistics to support testing of software systems*, Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 221(1), 85-93, doi:10.1243/1748006XJRR2.

[5] D. Randell, M. Goldstein, G. Hardman and P. Jonathan (2010) *Bayesian linear inspection planning for large-scale physical systems*. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 224(4), 333-345, DOI:10.1243/1748006XJRR322.

Quality of Research: the work contained in **[1]** was supported by a number of EPSRC grants all of which were highly graded in final review. Journal of the Royal Statistical Society, series B (paper **[2]**) and Biometrika (paper **[3]**) are two of the most highly rated statistics journals in the world. Both **[4]** and **[5]** were awarded the Donald Julius Groen Prize by the Safety & Reliability Group of the Institution of Mechanical Engineers. **[5]** is implemented in software within Shell.

Grant referenced in section **2**: 1999 - 2002 High reliability testing for complex software using Bayesian graphical modelling and program comprehension (value £145,729; principal investigator M. Goldstein, EPSRC).

4. Details of the impact

The research **[2]** was conducted by Goldstein and O'Hagan in the context of assessing assets of a regional water company, who sponsored aspects of this work. O'Hagan implemented **[2]** in the inference programme termed ABLE (Assessment with Bayes Linear Estimation - as described in **[2]**, which states that ABLE performs all the calculations in that paper), and applied this approach as a consultant, first for various water companies, then more widely.

ABLE was applied to the assets of London Underground, through a consultancy with Metronet (which was contracted to maintain nine London Underground lines), to achieve a better understanding of long term investment requirements and the sustainability or otherwise of current investment levels in infrastructure. This lead to the development of ESTEEM (Engineering Strategy for Economic and Efficient Management), which applied the Bayes linear methodology of **[1]** and **[2]**, based on the ABLE programme, to Metronet's assets, namely the maintainable items in all of the stations, bridges and other structures, that require an estimated £5 billion investment over a 50 year period. The aim of ESTEEM was to provide the company, through asset management estimates of asset degradation, costs, risks and their probabilities for each maintainable item, with a whole-life cost (WLC) strategic planning process for maintenance and renewal of its civil engineering assets, under varying funding constraints, over a 100 year planning horizon. A 2009 audit report[1] stated that the anticipated benefit of the ESTEEM project was a 20% saving, equalling about £600m (p7 of report). ESTEEM also improved passenger safety by combining safety, business and financial risk factors into a single modelling process (p3). The report (p25) concludes, "ESTEEM, as so far implemented in Metronet, is a clear

demonstration of best practice, leading edge thinking in the areas of civil engineering strategic planning and whole-life cost justification." The ESTEEM project continued throughout a restructuring, resulting in the assets being returned to the public sector, to Transport for London (TfL). In November 2010, ESTEEM won the prestigious Institution of Engineering and Technology Innovation Award, in the Asset Management category. In its entry submission, London Underground Ltd described the aim of ESTEEM as "an approach to optimising investment that aids training, promotes culture change and improves decision making" [2].

The ESTEEM protocol was followed in particular for all of the maintainable components in every one of the stations and bridges in that portion of the London Underground network originally controlled by Metronet. This comprised about 2/3 of the network and many thousands of components, all of whose uncertain characteristics were assessed by experts, leading to a full uncertainty specification and analysis within the ABLE structure. This analysis was used as the basis for developing and comparing whole-life maintenance strategies for all of these assets, as part of the decision support structure outlined in the Esteem documents. The ESTEEM Civil Assets least WLC predictions substantiated a basis for long term investment in the asset base and justified inclusion of preventative maintenance in a new performance contract for maintaining the assets. A particular example of the benefits reaped from this project is in the waterproofing of structures. Prior to ESTEEM, this was thought to be too expensive to justify. However, ESTEEM predictions anticipated a 20% savings in maintenance costs over a 30-year period, a saving of £5m p.a. The water-proofing was thus implemented at the end of 2009 for all concrete and masonry structures and continues to this day. A further example is the information systems used in London Underground stations. ESTEEM has become a critical operational system used to maintain and update these systems, and was fully implemented by summer 2011. It includes a hand held asset survey system used across London Underground for stations, maintains the asset register, reports condition and produces deterministic predictions of WLC for budgetary purposes on an ongoing basis. Currently, London Underground is maintaining condition state reports electronically for an intended Bayes linear update of the Civil Assets degradation predictions. In summary, the ESTEEM project, and the Bayesian work underpinning it, has been of great benefit to TfL, with current cost savings in the order of £5m-£10m for Civil assets and development of investment policy options for stations that have enabled prioritisation of future investment at levels that are sustainable[3].

Our second example of the impact of the methodology in **[1]** derives from **[3]**. Goldstein was a PI in the Basic Technology funded "Managing Uncertainty for Complex Models" consortium of universities. A postdoc within this consortium learned to apply Bayes linear methodology; he then left the consortium to join FERA, the arm of DEFRA dealing with regulation, policy and risk, as a statistician. He applied these methods there, for example within the project Food, Additives, Food Contact Material and Exposure Task (FACET), an 8.9 million euro project, involving 20 research organisations, funded by the European Commission, under the Seventh Framework Programme, which ran for four years from September 2008[4]. Project objectives were to deliver to the European Community a sustainable system to monitor intake of chemicals from food among European populations. Databases on food intake, chemical occurrence and chemical concentration were linked in algorithms for the estimation of probabilistic exposure to target food chemical intake. The experts struggled to specify full probability distributions across this complex space, but they had some experience of average consumption rates with standard deviations and there had been some studies into correlations between food types and across countries. As a result, the Bayes linear approach was judged a good fit for modelling food consumption databases for building up this model[4,5].

Unilever and FERA collaborated on a hazard assessment model based on the Bayes linear kinematic methodology, as part of Unilever's overall research effort to find novel approaches for assuring customer safety. The model considers the potency of chemicals that cause human sensitisation when applied to the skin, resulting in an undesired immune response known as allergic contact dermatitis. This presents clinically as a rash, skin lesion, papules or blistering at the site of exposure. Risk assessors in this area must weigh up several lines of evidence from in vivo and in vitro experiments when characterising the potency for a new chemical product in

order to determine a safe dose for exposed individuals. Beginning in 2010, Unilever applied the Bayes model in a series of assessments, based around products such as cinnamic aldehyde (used to give products a cinammon aroma, and a known skin sensitiser). This provided for Unilever estimates such as ingredient dose on skin that would induce an allergic response in certain percentages of consumers. The Bayes linear kinematic provided the framework for modelling the assessors' expectations and uncertainties and updating those beliefs in the light of the competing data sources. This approach to synthesising multiple lines of evidence and estimating hazard was judged to provide a transparent mechanism to construct, defend and communicate risk management decisions. Its value to Unilever is reflected in the fact that the company is working on extending the model to incorporate population variance, the uncertainty in the amount of product/ingredient that the consumer applies and the probability that the amount applied exceeds the adverse effect threshold for a given consumer. A published account of the details from the study that are publically available is provided in a 2013 paper[6] and also the document [5]. The Unilever internal documents on the outcomes are confidential and not for public dissemination, but we have been allowed to quote the following, from two internal reports, as illustrations of the role of Bayes linear methodology at Unilever[7]:

“Due to this feature, Bayes linear theory is applied to solve the skin sensitization risk assessment problem, as in such a problem, the available information is usually not enough for the specification of a full probability distribution. On the basis of the Bayes linear theory, the Bayes Linear method is applied when the new information is deterministic while the Bayes Linear Kinematic method is applied when there are uncertainties existing in the new information."
[Progress towards modelling a population-level risk metric for skin allergy risk assessment: page 5]

"What we've done to date: In 2011, a model was developed with FERA to predict median human threshold, i.e., the threshold to sensitize 50% of population under specific clinical exposure conditions, according to the data from both in vivo and in vitro tests. [The 50% threshold is not chosen as the protection goal for sensitization incidence but as the easiest percentile for experts to consider when judging correlation to other assay results.] The objective of this model was to make transparent, coherent and robust, an expert 'weight of evidence' analysis using the Bayes Linear method. This model was to allow comparison to be made between tests on how informative they were on the median threshold in humans."
[Skin allergy risk assessment document: page 1]

5. Sources to corroborate the impact

1. A Report for LUL Nominee BCV Limited, Trading as Metronet Rail BCV. Asset Management Consulting Ltd, 12 May 2009. (PDF file).

2. Entry form submitted by London Underground Ltd. for the IET Innovation Awards 2010. (Hard copy).

3. Letter signed by the ESTEEM project technical lead and project manager, 2007-2011, confirming details of impact.

4. Details of the FACET project objectives, funding and partners, on the CORDIS website http://cordis.europa.eu/projects/rcn/87815_en.html, and FACET project final report [pdf].

5. Document on role of Bayes linear methods at FERA written and provided to us by the relevant statistician at FERA.

6. Gosling,J.P., Hart, A. et al (2013) *A Bayes linear approach to weight of evidence risk assessment for skin allergy,* Bayesian Analysis, 8, 169-186 [Detailed public document on the Unilever, Fera collaboration]

7. Text from Toxicology Risk Modeller, Unilever Safety & Environmental Assurance Centre (SEAC), confirming the above description of the impact on Unilever and the quotes from two Unilever internal reports related to Bayes linear risk assessment.