Math. Scientist **36**, 105–106 (2011)
Printed in England

UNDERSTANDING AND USING FISHER'S p: A FOUR-PART ARTICLE

K. R. W. BREWER * AND
GENEVIEVE HAYES,** *Australian National University*
A. N. GILLISON,*** *Center for Biodiversity Management*

Summary of Parts 1–4

Whilst most investigators using Fisher's p may still be confident that they are using a tried and trusted statistic, others are becoming less certain. Their doubts are well justified. Generally speaking, the p-statistic works well wherever there is an actual or implied diffuse null hypothesis, but in Part 1 'Countering the p-statistic fallacy' (see pp. 107–116) we show that where that null hypothesis is precise, the two-sided p-statistic is commonly misinterpreted.

It is also shown in Part 1 that the single-parameter case of the Schwarz-inspired Bayesian information criterion (BIC), which is well known to be unduly parsimonious, can be remedied by introducing certain additional but clearly needed penalty terms, which in turn can be seen to coalesce into a simple function of T (the large sample limiting case of Student's t).

The relevance of this new information criterion (IC) is confirmed by several features. First, that it is a simple function of the T-statistic only. A simple transformation of this new IC also provides a realistic reinterpretation for Fisher's p. Finally, even when it is generalised to D free parameters, it is still for almost all practical purposes intermediate in size between the BIC itself and the corresponding, but *insufficiently* parsimonious, Akaike's AIC.

In Part 2 'A reference Bayesian hypothesis test' (see pp. 117–125), we develop a reference Bayesian hypothesis test which is fully compatible with the single-parameter case of this new IC. An important role is played here by a hitherto rather neglected (and initially purely empirical) law of numbers (see Benford (1938), which itself had been based on Newcomb (1881)). When the T-statistic is observed to be equal to one, that reference Bayesian test is indifferent between the null and alternative hypotheses. When the new IC is extended to small samples, as a function of Student's t-statistic, another important role is played by Fisher's p; this time in setting an upper bound to the false discovery rate, regardless of the number of degrees of freedom involved.

In Part 3 'Examining an empirical data set' (to appear in *Math. Scientist* **37**), a set of 1294 regression slopes from a biodiversity survey of tropical landscape mosaics in two hemispheres (see Gillison *et al.* (2010)) is used to provide empirical support for the new IC, our earlier theoretical findings are confirmed, and certain additional conclusions are drawn, including two in particular that were initially unexpected.

In Part 4 'Do we even need to specify a prior measure at H_0?' (to appear in *Math. Scientist* **37**), the approximate results derived in the three earlier parts are supplemented by exact results that can be obtained using a somewhat similar approach, but one that dispenses entirely with

Received 21 September 2010; revision received 21 December 2010.
* Postal address: School of Finance, Actuarial Studies and Applied Statistics, College of Business and Economics, Australian National University, Canberra, ACT 0200, Australia. Email address: ken.brewer@anu.edu.au
** Email address: ghayes17@gmail.com
*** Postal address: Center for Biodiversity Management, PO Box 120, Yungaburra, 4884, Queensland, Australia. Email address: andyg@cbmglobe.org

the need for any explicit null hypothesis. Finally, we suggest some likely consequences of the recognition that, when the implied null hypothesis is precise, much smaller values of p than currently envisaged are needed to supply any useful false discovery rate.

Acknowledgements for Parts 1–4

We have received invaluable IT assistance from Tamie Anakotta (née Henderson) of the Australian Bureau of Statistics, from Pirra Elford of the ANU's College of Business and Economics' IT unit, and from Tim Higgins of the School of Finance, Actuarial Studies and Applied Statistics (FAS) in the College of Business and Economics of the ANU. Tim Higgins, Steven Stern, and Steven Roberts (all from FAS) have further provided useful help with the paper in general. The very pertinent findings in Berger and Delampady (1987) and in Sellke *et al.* (2001) were pointed out to us by Johann Song, who was at the time an Honours Year student in FAS. Teresa Neeman of the ANU's Statistical Consulting Unit provided useful advice on current practice in clinical trials.