GLM Permutation - Nonparametric Inference for Arbitrary General Linear Models

Thomas Nichols[1,2,3], Gerard Ridgway[4], Matthew Webster[2], Stephen Smith[2]

[1]GlaxoSmithKline Clinical Imaging Centre, London, UK,   [2]FMRIB Centre, University of Oxford, Headington, UK,
[3]Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA,   [4]Centre for Medical Image Computing, University College London, London, UK

Introduction

Permutation methods are finding growing use in neuroimaging data analyses (e.g. randomise in FSL, SnPM in SPM, XBAMM/BAMM/CAMBA, etc). These methods provide exact control of false positives, make only weak assumptions, and allow nonstandard types of statistics (e.g. smoothed variance t-test). With fast and inexpensive computing, there would seem few reasons *not* to use nonparametric methods.

A significant limitation of these methods, however, is the lack of flexibility with respect to the experimental design and nuisance variables. Each specific design dictates the type of exchange-ability of null data, and hence how to permute. Nuisance effects (e.g. age) render data non-exchangeable even when the effect of interest is null. Hence, even something as simple as ANCOVA has no exact permutation test.

Recently there has been an active literature on approximate–but accurate–permutation tests for 2-variable regression, one effect of interest, one nuisance (see review by Anderson & Robinson [1]). Here we extend and evaluate these methods for use with an arbitrary General Linear Model (GLM).

Methods

Smith et al. [2] show that for any contrast, a GLM can be reformulated in a partitioned form

$$Y = X\beta + Z\gamma + \epsilon \tag{1}$$

where X contains solely the effects of interest corresponding to the contrast, and Z are all remaining (nuisance) effects. Hence testing $\mathcal{H}_0 : \beta = 0$ in (1) is equivalent to testing the contrast in the original GLM.

Name	Response			Predictor	
				Interest	Nuisance
Exact [1]	P	$(Y$	$- Z\gamma)$	X	Z
Freedman-Lane [1]	P	R_Z	Y	X	Z
ter Braak [1]	P	R_{XZ}	Y	X	Z
Kennedy [1]	P	R_Z	Y	$R_Z X$	
Smith [3]			Y	$P R_Z X$	Z
Shuffle-X [4]			Y	$P \quad X$	Z

Table 1: GLM permutation methods summarized by how response and predictors are adjusted and permuted. P is a $n \times n$ permutation matrix, R_Z is the residual-forming matrix using Z, and R_{XZ} is the residuals-forming matrix for the full model. Each method proceeds by fitting the response on a design matrix of predictors, repeated with different random permutation matrices.

We have extended six 2-variable regression permutation methods [1,3,4] to use an arbitrary GLM (see Table 1). Each method differs slightly in how they adjust the data for the nuisance effects, and exactly what quantities they permute. The Exact method cannot be used as it assumes the unknown true nuisance effect γ, but shows the spirit of most of the methods: Discount effect of nuisance, permute, then fit full model. While all of the methods can be re-formulated to avoid data permutation (an important computational consideration), only Smith & Shuffle-X allow the raw data to be modeled.

Since "truth" isn't available for real data we instead use Monte Carlo simulations. Simulations used 10,000 realizations with $n=6$ or 9 (all methods will be accurate with large n), 1 or 2 X regressors of interest either continuous or categorical, 1 or 2 Z nuisance regressors, and Gaussian and cubed mono-exponential errors (to induce extreme skew). Strong correlation is induced between X and Z to "stress" the methods (X & Z orthogonal should be accurate with any method.) Null hypothesis P-value distributions are plotted, and accuracy is measured with (1) K-S test statistic, (2) Bias (P-value minus nominal α) and (3) Mean Squared Error (MSE).

Results

(a) 2 Discrete Predictors, n=9, Exp³ Errors   (b) Discrete Predictor, n=6, Gaussian Errors   (c) Discrete Predictor, n=6, Gaussian Errors
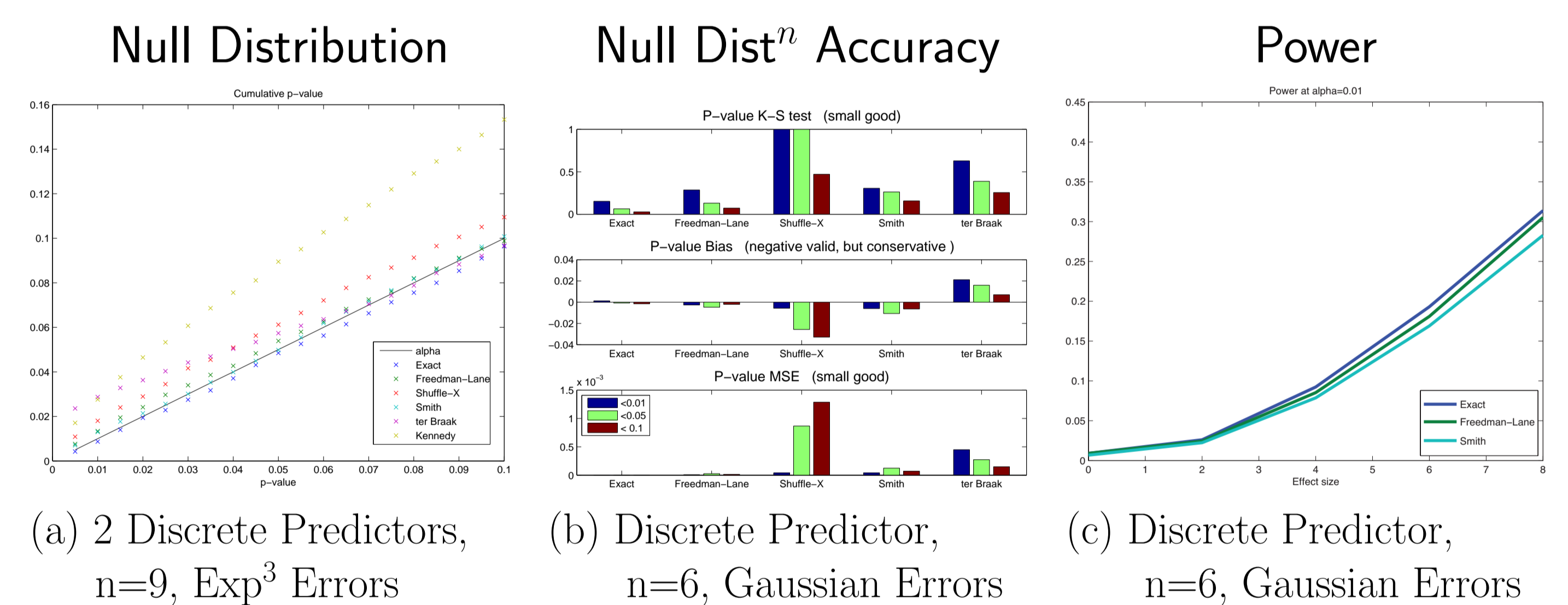
Figure 1: Example simulation results. (a) Cumulative distribution of P-values, showing the invalid behaviour of Kennedy, Shuffle-X and ter Braak. (b) Three summary measures for P-value distributions of small P-values (< 0.01, < 0.05, & < 0.1, as indicated), again showing poor performence of Shuffle-X and ter Braak, but similar quality of Freedman-Lane & Smith. (c) Power comparisons showed very similar performance between Freedman-Lane & Smith.

For continuous predictor of interest with Gaussian errors (not shown) all the methods performed quite similarly except for Kennedy, which had inflated Type I error rates. Under more challenging settings, i.e. discrete covariate with non-Gaussian errors, Kennedy remains anticonservative and Shuffle-X & ter Braak are found to perform erratically (Fig. 1a & 1b). This leaves Freedman-Lane and Smith, which perform similarly but Freeman-Lane has slightly better power with Gaussian errors (Fig. 1c).

Conclusion

We have developed and evaluated permutation methods for the GLM. Kennedy was found to be the worst method (see Software Notes below), and Freedman-Lane & Smith the best in the settings considered.

References: [1] Anderson & Robinson, *Australian and New Zealand Journal of Statistics*, 43:75-88, 2001. [2] Smith et al., *NeuroImage*, 34:127-136, 2007. [3] O'Gorman, *Communications in Statistics - Simulation and Computation*, 34:895-908, 2005. (Though developed independently, our "Smith" method is attributed to a referee on p. 897.) [4] Kennedy & Cade, *Communications in Statistics - Simulation and Computation*, 25:923-936, 1996. (Source of "Shuffle-X" method, renamed from "Shuffle-Z" for consistency with our notation.)
Software Notes: SnPM (http://www.sph.umich.edu/~nichols/SnPM) uses Shuffle-X, while FSL's randomise (http://www.fmrib.ox.ac.uk/fsl/randomise) used Kennedy up through FSL 4.0, and uses Freedman-Lane as of FSL 4.1.