Computing Weighted Chi-Squared Distributions and Related Quantities

Raffaello Seri[a], and Christine Choirat[b]

[a]Dipartimento di Economia, Università degli Studi dell'Insubria, Via Monte Generoso 71, 21100 Varese, Italy

[b]IQSS Harvard University, 1737 Cambridge Street, CGIS Knafel Building, Cambridge, MA 02138, USA

Under general conditions, the asymptotic distribution of degenerate second order $U-$ and $V-$statistics is an (infinite) weighted sum of chi-squared random variables whose weights are the eigenvalues of an integral operator associated with the statistic. Also their behaviour in terms of statistical power can be characterised as a function of the eigenvalues and eigenfunctions of the same integral operator. Unfortunately, not all operators corresponding to statistics of interest have a known eigendecomposition. A general algorithm computing these quantities starting from the operator associated with the statistic would therefore be useful. Some algorithms for specific cases have been proposed in Schilling (1983, p. 23) and Csörgő (1986, p. 719), but a general algorithm is still lacking. This is complicated by the fact that the kernels of the operators in use in statistics are generally semi-smooth, i.e. they are continuous but have discontinuous derivatives along the main diagonal.

In this context we discuss the use of the Wielandt-Nyström method of approximation of an integral operator based on quadrature. The algorithm can be used to approximate (as precisely as needed) the eigenvalues and eigenfunctions of the operator, as well as the asymptotic distribution and the power of the test statistic, and to build several measures of performance for tests based on $U-$ and $V-$statistics. The algorithm can be used with several methods of numerical integration. An extensive numerical study shows that, among classical quadrature methods, Gauss-Legendre and Clenshaw-Curtis rules perform quite well and have similar behaviours, as shown by Trefethen (2008). Both of them are outperformed by the Clenshaw-Curtis quadrature proposed by Kang et al. (2003) that is especially designed for semi-smooth kernels.

Csörgő, S. (1986) Testing for normality in arbitrary dimension, *Ann. Statist.*, **14**(2), 708–723.

Kang, S.Y., Koltracht, I., Rawitscher, G. (2003) Nyström-Clenshaw-Curtis quadrature for integral equations with discontinuous kernels, *Math. Comp.*, **72**(242), 729–756.

Schilling, M.F. (1983) Goodness of fit testing in $\mathbf{R}^m$ based on the weighted empirical distribution of certain nearest neighbor statistics, *Ann. Statist.*, **11**(1), 1–12

Trefethen, L.N. (2008) Is Gauss quadrature better than Clenshaw–Curtis? *SIAM Review*, **50**(1), 67–87.