*The Annals of Applied Statistics*
2009, Vol. 3, No. 4, 1282–1284
DOI: 10.1214/09-AOAS312D
Main article DOI: 10.1214/09-AOAS312


# DISCUSSION OF: BROWNIAN DISTANCE COVARIANCE

BY ANDREY FEUERVERGER

*University of Toronto*

Concepts of dependence are central in the theory of statistics and to most of its applications. It is therefore a pleasure to commend the authors—hereafter SR—for their theoretical contributions to our understanding of some of its subtler aspects, and for their provocatively interesting data analysis examples.

Standard measures, such as the product moment correlation, Spearman's rank correlation, Kendall's tau or Fisher–Yates' normal scores statistic, are all deficient. These only measure dependence of a "monotone character" and will not be effective even in such simple situations as when $Y$ has a nonmonotone regression on $X$ and $X$ is sampled randomly. Another simple example where such measures fail is when $X_i = V_i Z_i$ and $Y_i = V_i Z_i'$, where the "innovations" $Z_i$, $Z_i'$ are, say, independent standard normal variates, but the $X_i$, $Y_i$ share a common random scaling $V_i$; such structures arise in the stochastic volatility models of finance.

Owing to their importance, consistent measures of dependence—and, in particular, measures which in principle admit sample analogues on the basis of which tests consistent against all dependence alternatives can be constructed—have appeared previously, and at least as far back as Renyi (1953). Renyi's measure has, of course, ideal theoretical properties, but implementing its sample analogues is not straightforward, and for that reason it has not become a mainstay in applications. [See, e.g., Buja (1990).] In that respect the dependence measure (which predate's Renyi's) introduced by Hoeffding (1948), and later rediscovered in a more transparent form by Blum, Keifer and Rosenblatt (1961), has been more successful. See also Csörgő (1985).

There is also some precedent for the measures proposed at (2.4) and (2.6) in SR (at least for the case when $\alpha = 1$), although these appear here in a substantially extended form, and based on a novel approach with fresh interpretations. For example, Feuerverger (1993)—hereafter F93—proposed measures based on

$$\int\int \frac{|f_{X,Y}^n(s,t) - f_X^n(s) f_Y^n(t)|^2}{(1-e^{-s^2})(1-e^{-t^2})} W(s,t)\,ds\,dt, \tag{1}$$

with $W(s,t)$ a suitable weight function. In F93, the denominator term in (1) was suggested on the basis of its being (proportional to) the limiting variance of the term within the modulus in the numerator, under the null hypothesis of independence in the case of standard normality. The ratio within the integral in (1) is defined by continuity at the limiting values of $s = 0$ and/or $t = 0$. Using the bell-shaped weight function

$$W(s,t) = \left(\frac{1-e^{-s^2}}{s^2}\right)\left(\frac{1-e^{-t^2}}{t^2}\right) \tag{2}$$

leads to the form (2.6) of SR. In fact, simple modifications to the weight function (2) can lead to interesting and potentially useful variants of the $T_1, T_2, T_3$ statistics defined by SR at (2.11), wherein the absolute value functions are replaced by more general functions; see, for example, the computations leading up to (4.11) in F93. It should be noted, however, that in F93 the variates $X$ and $Y$ are univariate, while SR deal with the case where these are random vectors of dimensions $p$ and $q$.

The case of univariate $X$ and $Y$ affords another advantage, and, in particular, with respect to desiderata one might wish to place on a dependence measure. Thus, it is desirable that a dependence measure should not require moment conditions on the variables—not even a finite first moment. And second, it is desirable to go beyond the stated scale invariance $(X, Y) \to (\varepsilon X, \varepsilon Y)$, not only so as to allow the values of $\varepsilon$ applied to $X$ and $Y$ to differ, but also to have the invariance $(X, Y) \to (\phi(X), \psi(Y))$ with respect to strictly monotone transformations $\phi$ and $\psi$. In the univariate case, this may be achieved by replacing the $X$'s and $Y$'s by, say, their normal scores. [In fact, this is the reason behind the choice of denominator in (1).] The resulting rank-type test will then also have the advantage of being $H_0$-distribution free as SR note in Section 4.3. Furthermore, since the empirical marginal distributions will then no longer be random, the term $T_3$ in (2.11) of SR can then be dispensed with, while the term $T_2$ can be reduced, resulting in substantially simplified computations. Of related note, the representation (2.8) of SR is particularly interesting.

While it would certainly be of interest to examine what can be done along such lines when $X$ and $Y$ are not restricted to be univariate, there is a further problem of a multivariate character that arises. This refers to the case where we seek to assess mutual independence among more than two variables, for example $X$, $Y$ and $Z$, with each of these being either univariate or multivariate. This problem was alluded to briefly in F93, and perhaps some progress may be possible on the basis of decompositions along lines indicated, for example, in Deheuvels (1981). However, this further multivariate problem still remains largely unresolved.

Progress sometimes consists of seeing a familiar object in a totally new way, and the notion of Brownian covariance—as indeed of covariances relative to other stochastic processes, particularly the fractional Brownian motions—introduced in this article by SR is particularly novel. Certainly the fact that $\mathcal{W}$ should equal $\mathcal{V}$ seems at least as surprising as SR purport it to be. But no small part of that surprise stems from the fact that Brownian covariance should lead to a statistic that happens to be consistent against all alternatives. Why should this have happened? What role does normality of the process play in it? Could the essential condition for consistency be that the process be of full rank in the sense of requiring a complete set of basis functions to represent it? Or is it enough that some separating class of functions [e.g., Breiman (1968), page 165ff] should underly the process in some sense? Here SR leave us with a nice mystery which seems surely worthwhile to try to resolve.

## REFERENCES

BLUM, J. R., KEIFER, J. and ROSENBLATT, M. (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.* **32** 485–498. MR0125690

BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading, MA. MR0229267

BUJA, A. (1990). Remarks on functional canonical variates, alternating least squares methods, and ACE. *Ann. Statist.* **18** 1032–1069. MR1062698

CSÖRGŐ, S. (1985). Testing independence by the empirical characteristic function. *J. Multivariate Anal.* **16** 290–299. MR0793494

DEHEUVELS, P. (1981). An asymptotic decomposition for multivariate distribution-free tests of independence. *J. Multivariate Anal.* **11** 102–113. MR0612295

FEUERVERGER, A. (1993). A consistent test for bivariate dependence. *Int. Statist. Rev.* **61** 419–433.

HOEFFDING, W. (1948). A non-parametric test of independence. *Ann. Math. Statist.* **19** 546–557. MR0029139

Renyi, A. (1953). On measures of dependence. *Acta Math. Hungar.* **10** 441–451.

DEPARTMENT OF STATISTICS
UNIVERSITY OF TORONTO
180 ST. GEORGE STREET
TORONTO, ONTARIO
CANADA M55 363
E-MAIL: andrey@utstat.toronto.edu