*The Annals of Applied Statistics*
2009, Vol. 3, No. 4, 1299–1302
DOI: 10.1214/09-AOAS312G
Main article DOI: 10.1214/09-AOAS312


# DISCUSSION OF: BROWNIAN DISTANCE COVARIANCE

BY CHRISTOPHER R. GENOVESE

*Carnegie Mellon University*

Congratulations to Professors Székely and Rizzo for such an exciting and enjoyable contribution. It is not often that one of our most basic techniques is given so fundamental, and so successful, a rethinking. Although using distance covariance requires giving up some useful properties associated with linearity—directionality/sign, exact expressions for the variance and covariance of sums, direct connection to the multivariate normal distribution—it offers useful properties in exchange. Distance covariance gives a true indicator of independence even for non-normal distributions, applies directly in multivariate settings (even when "$p \gg n$"), is the basis for general and powerful tests, can be adapted to use ranks, provides conditions for central limit theorems, and is straightforward to compute. That seems to be a favorable trade. In this discussion I will focus on the meaning of Brownian covariance, but first I want to raise a few questions to the authors (and the field).

The paper adapts the statistic in examples to derive resampling techniques and tests for nonlinearity and extends the covariance definition in several ways. Perhaps the authors can comment on how general these derived techniques are. For instance, what additional conditions, if any, are required for the test of nonlinearity in Example 6 (based on $\mathrm{dCov}(X, (I - X(X^T X)^{-1} X^T)Y))$ to be consistent? Also, the computations would appear to be $O(n^2)$, which can be burdensome for very large $n$. Are there speed-ups or approximations that yield comparable results more quickly? And are rates of convergence available for the empirical statistics, perhaps under stronger moment conditions?

But these are details. Even though the Pearson correlation is entrenched in the practice of several fields, including our own, what reason do we have not to aggressively introduce distance covariance and correlation into our practice *and* our teaching, even at the introductory level? It is rare in practice that we want a measure of linear association *per se*, more typically we use Pearson correlation as a proxy. Distance covariance provides most of what we do want in these cases with attendant theory and convenience that is hard to beat. And teaching about the difference between "uncorrelated" and "independent" is a thorn in the side of anyone who has had to do so. Distance covariance would require no more sophisticated ideas than what we already use in teaching correlation, without that complication. The statistic is expressed in terms of distances which are easy to understand, and it would free us from undue emphasis on Normal examples. It is interesting to ponder what it would take to change practice at this level.