The Annals of Applied Statistics
2009, Vol. 3, No. 4, 1266–1269
DOI: 10.1214/09-AOAS312A
Main article DOI: 10.1214/09-AOAS312

DISCUSSION OF: BROWNIAN DISTANCE COVARIANCE

BY PETER J. BICKEL[1] AND YING XU[1]

University of California, Berkeley

Szekely and Rizzo present a new interesting measure of correlation. The idea of using $\int |\phi_n(u,v) - \phi_n^{(1)}(u)\phi_n^{(2)}(v)|^2 \, d\mu(u,v)$, where ϕ_n, $\phi_n^{(1)}$, $\phi_n^{(2)}$ are the empirical characteristic functions of a sample (X_i, Y_i), $i = 1, \dots, n$, of independent copies of X and Y is not so novel. A. Feuerverger considered such measures in a series of papers [4]. Aiyou Chen and I have actually analyzed such a measure for estimation in [3] in connection with ICA.

However, the choice of $\mu(\cdot,\cdot)$ which makes the measure scale free, the extension to $X \in \mathbb{R}^p$, $Y \in \mathbb{R}^q$ and its identification with the Brownian distance covariance is new, surprising and interesting.

There are three other measures available, for general p, q:

1. The canonical correlation ρ between X and Y.
2. The rank correlation r (for $p = q = 1$) and its canonical correlation generalization.
3. The Renyi correlation R.

All vanish along with the Brownian distance (BD) correlation in the case of independence and all are scale free. The Brownian distance and Renyi covariance are the only ones which vanish iff X and Y are independent.

However, the three classical measures also give a characterization of total dependence. If $|\rho| = 1$, X and Y must be linearly related; if $|r| = 1$, Y must be a monotone function of X and if $R = 1$, then either there exist nontrivial functions f and g such that $\mathbb{P}(f(X) = g(Y)) = 1$ or at least there is a sequence of such nontrivial functions f_n, g_n of variance 1 such that $\mathbb{E}(f_n(X) - g_n(Y))^2 \to 0$.

In this respect, by Theorem 4 of Szekely and Rizzo, for the common $p = q = 1$ case, BD correlation does not differ from Pearson correlation.

Although we found the examples varied and interesting and the computation of p values for the BD covariance effective, we are not convinced that the comparison with the rank and Pearson correlations is quite fair, and think a comparison to R is illuminating.

Intuitively, the closer the form of observed dependence is to that exhibited for the extremal value of the statistic, the more power one should expect. Example 1 has Y as a distinctly nonmonotone function of X plus noise, a situation where we would expect the rank correlation to be weak and, similarly, the other examples correspond to nonlinear relationships between X and Y in which we would expect

[1]Supported in part by NSF Grant DMS-09-06808.