MEMO-RANDOM NUMBER 1
ON COMONOTONICITY

ROGER KOENKER

1. Introduction

Dependence among random variables is a nasty business and we usually ignore the nastiness (at our peril) by employing measures of linear association like the Pearson correlation coefficient. A more attractive category of dependence concepts are those based on the copula function.

Definition 1. *A copula is a multivariate distribution function whose univariate marginal distributions are all $U[0,1]$, i.e., uniformly distributed on $[0,1]$.*

For continuous* m-variate distributions we may associate a copula $C(\cdot)$ with a distribution function $F(\cdot)$:

$$F(x) = C(F_1(x_1), \ldots, F_m(x_m))$$

where $F_i(\cdot)$ denotes the ith marginal distribution of F. And consequently, we have

$$C(u) = F(F_1^{-1}(u_1), \ldots, F_m^{-1}(u_m))$$

The copula represents the dependence among the components of the random vector $X = (X_1, \ldots, X_m)'$ while abstracting from its idiosyncratic marginal behavior. Consequently, it is well suited to the task of constructing measures of dependence. The classical measures of this type are Spearman (1904) and Kendall (1938) rank correlations. Dependence measures based on the copula function are invariant under strictly increasing transformations of the coordinates and are consequently sometimes called "measures of concordance."

Our main interest is to explore the extreme case of perfectly concordant random variables. The following classical result provides bounds on the distribution function in terms of its marginals.

Theorem 1. *(Fréchet) Let $F(x)$ be an m-variate distribution function with univariate marginals $F_1, \ldots, F_m$. Then for all $x \in \Re^m$,*

$$\max\{0, F_1(x_1) + \ldots + F_m(x_m) - (m-1)\} \leq F(x_1, \ldots, x_m) \leq \min\{F_1(x_1), \ldots, F_m(x_m)\}$$

Version September 19, 2002.

*For discrete m-variate distributions we encounter problems of nonuniqueness in the representations given above. It may be useful to consider *continuous* approximations to such discrete cases, say by convolution with a smooth density with scale tending to zero.

Reformulating this result in terms of the copula function for bivariate cases we have

$$\max\{0, u_1 + u_2 - 1\} \leq C(u_1, u_2) \leq \min\{u_1, u_2\}$$

In the bivariate case both bounds are valid copulas and correspond to bivariate df's for the random vectors (U, U) and $(U, 1 - U)$, with $U \sim U[0, 1]$. The upper bound characterizes the situation in which a pair of random variables has a maximal degree of monotone dependence. Following Schmeidler (1986) and an extensive recent literature in finance we say they are *comonotonic.*

Denneberg (1994) provides several equivalent definitions of comonotonicity, one of which can be easily interpreted for random variables.

Definition 2. *The two functions $X, Y : \Omega \to \Re$ are comonotonic if there exists a third function $Z : \Omega \to \Re$ and increasing functions f and g such that $X = f(Z)$ and $Y = g(Z)$.*

From our point of view the crucial property of comonotonic random variables is the behavior of quantile functions of their sums. For comonotonic random variables X,Y, we have

$$F^{-1}_{X+Y}(u) = F^{-1}_X(u) + F^{-1}_Y(u)$$

By comonotonicity we have a $U \sim U[0, 1]$ such that $Z = g(U) = F^{-1}_X(U) + F^{-1}_Y(U)$ where g is left continuous and increasing, so by monotone invariance, $F^{-1}_{g(U)} = g \circ F^{-1}_U$.

Thus, in a rather strange way, comonotonicity plays the same role for quantile functions that independence does for variances. For comonotonic random variables quantile functions of sums are sums of quantile functions, just as variances of sums are sums of variances for independent random variables. And yet, comonotonicity and independence are diametrically opposite notions, the former describing a state of maximal dependence. The extremal nature of comonotone random variables is clarified somewhat by following result.

Theorem 2. *(Major (1978)) Let X, Y be random variables with marginal distribution functions F and G, respectively, and finite first absolute moments. Let $\rho(x)$ be a convex function on the real line, then*

$$\inf E\rho(X - Y) = \int_0^1 \rho(F^{-1}(t) - G^{-1}(t))dt.$$

where the inf *is over all joint distributions, H, for (X, Y) having marginals F and G.*

Mallows (1972) formulates this result for $\rho(u) = u^2$, and notes that it implies among other things that the maximal Pearson correlation of X and Y occurs at the Fréchet bound $H(x, y) = \min\{F(x), G(y)\}$,

$$\max \int xy dF(x, y) = \int_0^1 F^{-1}(t)G^{-1}(t)dt.$$

Bickel and Freedman (1981) elaborate on the case $\rho(x) = x^2$ in considerable further detail.

We need not restrict attention to differences in random variables. In the recent actuarial literature there has been considerable attention devoted to role of comonotonicity. Suppose, that we have the sum $Z = X_1 + X_2 + \ldots + X_n$ where the X_i denote possibly dependent random variables representing losses from various insured risks. And suppose we know the marginal distributions of the X_i, but not their joint distribution. Is there a way to bound the behavior of Z using the comonotonic version of their joint distribution? The following theorem of Kaas, Dhaene, Vyncke, Goovaerts, and Denuit (2001), shows that the comonotone version of the joint distribution provides a worst case scenario for any convex loss function.

Theorem 3. *Given $Z = X_1 + X_2 + \ldots + X_n$ with $X_i \sim F_i$, for $i = 1, \ldots, n$, and any convex function ρ,*

$$E\rho(Z) \leq E\rho(\tilde{Z}) = \int_0^1 \rho(\sum F_i^{-1}(t))dt$$

where $\tilde{Z} = \sum \tilde{X}_i = \sum F_i^{-1}(U)$ and $U \sim U[0,1]$.

The foregoing results raise some intriguing questions about the intrepretation of quantile regression models that we address in the next section.

2. Random Coefficients, Comonotonicity, and Quantile Regression

The quantile regression model

$$F_{Y|X}^{-1}(\tau|x) = x'\beta(\tau) \tag{2.1}$$

may be interpreted as a random coefficient model in which for any given design vector, x, we have conditional on $X = x$,

$$Y = x'\beta(U) \tag{2.2}$$

with U is a scalar uniform $U[0,1]$ random variable. This interpretation follows immediately from the fact that $Y = F_Y^{-1}(U)$ has the distribution, F_Y. The random coefficient interpretation of (2.1) is quite distinct from classical random coefficient specifications. Rather than assuming that the coordinates of β are independent random variables, we adopt the opposite viewpoint, that the coordinates of β are functionally perfectly dependent.

It is dangerous to jump immediately to the conclusion that the random vector $\beta(U)$ is comonotonic. In fact, in most parameterizations there is no reason to believe that the functions $\beta_i(\tau)$ are will be monotone. Nevertheless, what is crucial is that there exists a reparameterization that does exhibit comonotonicity. Recall that we can always reparameterize (2.1) as

$$F_{Y|X}^{-1}(\tau|x) = x'A^{-1}A\beta(\tau) = z'\gamma(\tau) \tag{2.3}$$

Suppose that we choose $p = \dim(\beta)$ design points $\{x^k : k = 1, ..., p\}$ where the model (2.1) holds. Now choose the matrix A so that $Ax^k = e^k$, the kth unit basis vector. Then for any x^k we have that conditional on $X = x^k$,

$$Y = (e^k)'\gamma(U) = \gamma_k(U). \tag{2.4}$$

And inside the convex hull of the x^k points, i.e. conditioning on on a point $x = \sum w_k x^k$ for $0 \leq w_k \leq 1$ with $\sum w_k = 1$, we have

$$Y = \sum w_k \gamma_k(U) \tag{2.5}$$

and we have a comonotonic random coefficient representation of the model. In effect, we have done nothing more than reparameterized the model so that the coordinates,

$$\gamma_k = F^{-1}_{Y|X}(\tau|x^k) \quad k = 1, ..., p,$$

are the conditional quantile functions of Y at the points x^k. The fact that quantile functions of sums of nonnegative comonotonic random variables are sums of the marginal quantile functions allows us to interpolate linearly between the chosen x^k. Of course, linear extrapolation is also possible, but we need to be cautious in this case about possible violations of the monotonicity requirement on sums of quantile functions.

References

BICKEL, P. J., AND D. A. FREEDMAN (1981): "Some Asymptotic Theory for the Bootstrap," *The Annals of Statistics*, 9, 1196–1217.

DENNEBERG, D. (1994): *Non-Additive Measure and Integral*. Kluwer Academic Publishers.

KAAS, R., J. DHAENE, D. VYNCKE, M. GOOVAERTS, AND M. DENUIT (2001): "A Simple Geometric Proof that comonotonic risks have the largest convex sum," preprint.

KENDALL, M. (1938): "A New Measure of Rank Correlation," *Biometrika*, 30, 81–93.

MAJOR, P. (1978): "On the Invariance Principle for Sums of Independent Identically Distributed Random Variables," *Journal of Multivariate Analysis*, 8, 487–517.

MALLOWS, C. L. (1972): "A Note on Asymptotic Joint Normality," *The Annals of Mathematical Statistics*, 43, 508–515.

SCHMEIDLER, D. (1986): "Integral representation without additivity," *Proceedings of Am Math Society*, 97, 255–261.

SPEARMAN, C. (1904): "The proof and measurement of association between two things," *Amer. J. Psych.*, 15, 72–101.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN