The Annals of Probability
1982, Vol. 10, No. 2, 496–499

POSITIVELY CORRELATED NORMAL VARIABLES ARE ASSOCIATED[1]

By Loren D. Pitt

University of Virginia

It is shown that normal variables are associated if and only if their correlations are nonnegative.

1. Introduction. A real function $f(\bar{x}) = f(x_1, \cdots, x_k)$ of k real variables will be called increasing if it is a nondecreasing function of each of the separate variables $x_1, \cdots, x_k$. If $\bar{X} = (X_1, \cdots, X_k)$ is a random vector and if the inequality

$$\mathrm{Cov}[f(\bar{X}), g(\bar{X})] \geq 0 \tag{1}$$

holds for each pair of bounded Borel measurable increasing functions f and g, then the variables $X_1, \cdots, X_k$ are called associated (Esary, Proschan, and Walkup, 1967).

In this note we prove the following.

THEOREM. *Let $\bar{X}$ be multivariate normal with mean vector* 0 *and covariance matrix*

$$\Sigma = (\sigma_{ij} = \mathrm{Cov}[X_i, X_j]).$$

The condition

$$\sigma_{ij} \geq 0 \quad \text{for} \quad 1 \leq i, j \leq k \tag{2}$$

is necessary and sufficient for the variables to be associated.

Condition (2) is obviously necessary. Several people have conjectured (2) is also sufficient and this has been proved for $k \leq 4$ (T. Savits, private communication). The best previous sufficient condition which is independent of k is that Σ be non-singular and that

$$\Sigma^{-1} \equiv C = (c_{ij})$$

satisfies

$$c_{ij} \leq 0 \quad \text{for} \quad i \neq j; \tag{3}$$

see Barlow and Proschan (1975) and Kemperman (1977). Condition (3) implies (2) but for $k \geq 3$ the converse does not hold.

The method of proof is an adaptation from Pitt (1977). In joint unpublished work with I. Herbst, these ideas have proved to be an effective tool in establishing general correlation inequalities of a type related to the FKG inequalities.

The ideas of the proof are all given in Section 2 where the proof is exposited in a special case. Section 3 gives the technical approximation arguments necessary in the general case.

2. The heart of the proof. In this section it will be assumed that Σ is non-singular and that the functions f and g are continuously differentiable with bounded partial derivatives $\partial f/\partial x_i$ and $\partial g/\partial x_i$, $1 \leq i \leq k$.

Bring in an independent copy $\bar{Z}$ of $\bar{X}$ and for $0 \leq \lambda \leq 1$ introduce the random vector

$$\bar{Y}(\lambda) \doteq \lambda \bar{X} + (1 - \lambda^2)^{1/2} \bar{Z}.$$

Received April 1981.

[1] Research supported in part by National Science Foundation Grant NSF-MCS-76-06427A04.

AMS 1970 *subject classifications.* Primary 62H99.

Key words and phrases. Normal variables, positive correlation, association of random variables.

Note that for each fixed λ, $\bar{Y}(\lambda)$ is normal with the same covariance matrix Σ and that

$$\operatorname{Cov}[X_i, Y_j(\lambda)] = \lambda\sigma_{ij}.$$

Next we set

$$F(\lambda) = Ef(\bar{X})g(\bar{Y}(\lambda)).$$

We observe that $F(\lambda)$ is continuous in λ and that $F(0) = Ef(\bar{X})Eg(\bar{X})$ and $F(1) = Ef(\bar{X})g(\bar{X})$. It thus suffices to show that $F'(\lambda)$ exists and $F'(\lambda) \geq 0$ for $0 \leq \lambda < 1$. To this end observe that the density of $\bar{X}$ is

$$\phi(\bar{x}) = (2\pi)^{-k/2}(\det C)^{-1/2}\exp\left\{-\frac{1}{2}\sum_{i,j=1}^{k} c_{ij}x_ix_j\right\}$$

and that the conditional density of $\bar{Y}(\lambda)$ given $\bar{X} = \bar{x}$ is

$$p(\lambda; \bar{x}, \bar{y}) = (1-\lambda^2)^{-k/2}\phi((1-\lambda^2)^{-1/2}(\lambda\bar{x} - \bar{y})).$$

Thus

$$F(\lambda) = \int_{R^k} \phi(\bar{x})f(\bar{x})g(\lambda, \bar{x})\, d\bar{x},$$

where

(4) $$g(\lambda, \bar{x}) = \int_{R^k} p(\lambda; \bar{x}, \bar{y})g(\bar{y})\, d\bar{y}.$$

We set $\phi_\lambda(\bar{x}) = (1-\lambda^2)^{-k/2}\phi((1-\lambda^2)^{-1/2}\bar{x})$ and note that

$$g(\lambda, \bar{x}) = (\phi_\lambda * g)(\lambda\bar{x}) = \int_{R^k} g(\lambda\bar{x} - \bar{y})\phi_\lambda(\bar{y})\, d\bar{y}.$$

This shows that partial derivatives $\partial g(\lambda, \bar{x})/\partial x_i$ exist and are bounded. Moreover, since g is increasing and $\lambda > 0$ we have

(5) $$\frac{\partial g}{\partial x_i}(\lambda, \bar{x}) \geq 0.$$

Next, an explicit computation using the heat equations

$$\frac{\partial \phi}{\partial \sigma_{ii}} = \frac{1}{2}\frac{\partial^2 \phi}{\partial x_i^2}, \qquad \frac{\partial \phi}{\partial \sigma_{ij}} = \frac{\partial^2 \phi}{\partial x_i \partial x_j}, \quad i \neq j,$$

of Plackett (1954) shows that

$$\frac{\partial p}{\partial \lambda} = \frac{p}{(1-\lambda^2)}\left\{k\lambda - \sum_{i,j} x_i c_{ij}(\lambda x_j - y_j) - \frac{\lambda}{(1-\lambda^2)}\sum_{i,j}(\lambda x_i - y_i)c_{ij}(\lambda x_j - y_j)\right\}$$
$$= -\frac{1}{\lambda}\left\{\sum_{i,j}\sigma_{ij}\frac{\partial^2 p}{\partial x_i \partial x_j} - \sum_i x_i \frac{\partial p}{\partial x_i}\right\}.$$

Thus

$$F'(\lambda) = -\frac{1}{\lambda}\int_{R^k} \phi(\bar{x})f(\bar{x})\left\{\sum_{i,j}\sigma_{ij}\frac{\partial^2 g(\lambda, \bar{x})}{\partial x_i \partial x_j} - \sum_i x_i \frac{\partial g(\lambda, \bar{x})}{\partial x_i}\right\} d\bar{x},$$

and an integration by parts gives

$$F'(\lambda) = \frac{1}{\lambda}\int_{R^k} \phi(\bar{x})\left\{\sum_{i,j}\sigma_{ij}\frac{\partial f(\bar{x})}{\partial x_i}\frac{\partial g(\lambda, \bar{x})}{\partial x_j}\right\} d\bar{x}.$$

Because $f(\bar{x})$ is increasing, we have $\partial f/\partial x_i \geq 0$. Combining this with (2) and (5) shows that $F'(\lambda) \geq 0$ and completes the proof in this special case.

3. Technicalities. If we assume that $f(\bar{x})$ and $g(\bar{x})$ are bounded and continuous, we can remove our assumption that they are continuously differentiable. In fact, if $\psi_\varepsilon \geq 0$ is a C^∞ approximate identity and if f is a bounded increasing continuous function, then $\psi_\varepsilon * f$ is C^∞, increasing, has bounded derivatives and $\psi_\varepsilon * f \to f$ in the sense of uniformly bounded pointwise convergence.

The assumption that Σ is non-singular is not serious: For bounded continuous functions f and g the covariance $\mathrm{Cov}[f(\bar{X}), g(\bar{X})]$ is a continuous function of Σ and for any covariance Σ and any $\varepsilon > 0$, $\Sigma_\varepsilon = \Sigma + \varepsilon I$ is non-singular.

The problem of removing the condition that f and g are continuous is less routine when Σ is singular. One option is simply to quote the result of Esary, Proschan and Walkup (1967) to the effect that (1) holds for (Borel) measurable increasing f and g if and only if it holds for increasing continuous f and g. A second related option, which we choose to follow here and which leads to the same conclusion, is based on the following lemma. This shows that continuous increasing functions are dense in the space of measurable increasing functions.

LEMMA. *Let $d\mu(\bar{x})$ be a finite Borel measure on R^k and let $f(\bar{x})$ be a bounded increasing Borel function on R^k with*

$$\sup\{|f(\bar{x})| : \bar{x} \in R^k\} \equiv M. \tag{6}$$

Then there exists a sequence $\{f_n(\bar{x})\}$ of continuous increasing functions each of which is bounded by M and which satisfies

$$\lim_{n\to\infty} f_n(\bar{x}) = f(\bar{x}) \text{ a.e. } [\mu].$$

PROOF. We may assume $f(\bar{x}) \geq 0$. We let $A(\alpha) = \{\bar{x} : f(\bar{x}) \geq \alpha\}$ and denote the indicator function of $A(\alpha)$ with $l_{A(\alpha)}$. If $f(\bar{x})$ is increasing so is $l_{A(\alpha)}(\bar{x})$ increasing and from $f(\bar{x}) = \lim n^{-1} \sum_{j=1}^{\infty} l_{A(j/n)}(\bar{x})$ we see that it is sufficient to treat the special case when

$$f(\bar{x}) = l_A(\bar{x})$$

is an increasing indicator function.

Because μ is regular, there will exist compact sets K_n and C_n and open sets O_n with $K_n \subseteq A \cap C_n$ and $A \subseteq O_n$ such that

$$\mu(O_n - K_n) \leq n^{-2} \quad \text{and} \quad \mu(R^k - C_n) \leq n^{-2}.$$

For $\varepsilon > 0$ and $\bar{x} \in R^k$ we introduce the octant

$$R(\bar{x}, \varepsilon) = \{\bar{y} \in R^k : y_i \geq x_i - \varepsilon \quad \text{for} \quad 1 \leq i \leq k\}.$$

Since $l_A(\bar{x})$ is increasing and since $O_n \supset A$ is open, for each $\bar{x} \in K_n \subseteq A$ we can find an $\varepsilon > 0$ satisfying

$$C_n \cap R(\bar{x}, 2\varepsilon) \subset O_n. \tag{7}$$

By the compactness of K_n we can choose an $\varepsilon_n > 0$ so that (7) holds for all $\varepsilon \leq \varepsilon_n$ and all $\bar{x} \in K_n$. We let $\{R(\bar{x}_i, \varepsilon_n) : 1 \leq i \leq N_n\}$ be a finite cover of K_n and set

$$L_n = \cup\{R(\bar{x}_i, \varepsilon_n) : 1 \leq i \leq N_n\}.$$

To complete the argument we set

$$Q(\bar{x}) = \{\bar{y} \in R^k : 0 \leq y_i - x_i \leq \varepsilon_n \quad \text{for} \quad 1 \leq i \leq k\}$$

and we denote the Lebesgue measure of a set A with $|A|$. The desired sequence of functions is

$$f_n(\bar{x}) = \varepsilon_n^{-k} |Q(\bar{x}) \cap L_n|.$$

It is elementary to check that $f_n(\bar{x})$ is continuous and increasing with $0 \le f_n(\bar{x}) \le 1$, with $f_n(\bar{x}) = 1$ on $L_n \supseteq K_n$ and $f_n(\bar{x}) = 0$ if $\bar{x} \in C_n - O_n$. Thus

$$\int_{R^k} |f_n(\bar{x}) - l_A(\bar{x})| \, d\mu(\bar{x}) \le \int_{C_n} |f_n(\bar{x}) - l_A(\bar{x})| \, d\mu(\bar{x}) + \int_{R^k - C_n} |f_n(\bar{x}) - l_A(\bar{x})| \, d\mu(\bar{x})$$

$$\le \mu(O_n - K_n) + \mu(R^k - C_n) \le 2n^{-2}.$$

This shows that $f_n(\bar{x}) \to l_A(\bar{x})$ a.e. $[\mu]$ and the proof is complete.

REFERENCES

BARLOW, R. E., PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing.* Holt, Rinehart and Winston, New York.

ESARY, J. D., PROSCHAN, F., WALKUP, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38** 1466–1474.

KEMPERMAN, J. H. B. (1977). On the FKG-inequalities on a partially ordered space. *Proc. Koninklijke Nederlandse Akad. van Wetenschappen* **80** 313–331.

PITT, L. D. (1977). A Gaussian correlation inequality for symmetric convex sets. *Ann. Probability* **5** 470–474.

PLACKETT, R. L. (1954). A reduction formula for normal multivariate integrals. *Biometrika* **41** 351–360.

UNIVERSITY OF VIRGINIA
DEPARTMENT OF MATHEMATICS
MATH ASTRONOMY BLDG.
CABELL DRIVE
CHARLOTTESVILLE, VA 22903