ELSEVIER

Statistics & Probability Letters 20 (1994) 235–238

STATISTICS & PROBABILITY LETTERS

Multivariate normality via conditional normality

M. Ahsanullah[a,*], Jacek Wesolowski[b,1]

[a] Department of Management Sciences, Rider University, Lawrenceville, NJ, 08648-3099, USA
[b] Mathematical Institute, Warsaw University of Technology, Pl. Politechniki 1, 00-661 Warsaw, Poland

Received April 1993; revised October 1993

Abstract

A new characterization of the multivariate normal distribution using conditional normality is given.

1. Introduction

Problems of characterizing multivariate Gaussian measures by their conditional structures have attracted more attention in recent years. One method of dealing with this question is by assuming some normal conditional(s), see Hamedani (1992) for a recent review of the subject. The other approach is by investigation of measures with Gaussian conditional structure of the second-order (linear conditional expectations and non-random conditional variances), see Wesolowski (1991) for a brief survey and some new results. There is also a very recent interest in combining these two ideas (Arnold, 1993). In this paper we are concerned with the first approach.

For some $n \geqslant 2$, consider the following two conditions.

(1) The conditional distribution of X_n given $X_1, \ldots, X_{n-1}$ is normal $N(\alpha_0 + \sum_{j=1}^{n-1} \alpha_j X_j, \sigma^2)$, where $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}, \sigma^2$ are some real constants and $\sigma^2 > 0$.

(2) The r.v.'s $X_1, \ldots, X_n$ are identically distributed.

In the case $n = 2$, Ahsanullah (1985) proved that (1) and (2) imply joint normality.

For $n > 2$, (1) and (2) do not characterize the joint normality of the X's. Counterexamples were given for $n = 3$ in Ahsanullah and Sinha (1986) and independently in Arnold and Pourahmadi (1988) (in the sequel we refer to this paper by AP).

The main problem, we address here, is to change (2) to some other equidistribution relation to obtain a characterization of the multivariate normal distribution by (1). This new relation is discussed in Section 2, while the characterization is given in Section 3. An example of such a result was presented in Ahsanullah and Sinha (1986), where instead of (1), a very strong condition of exchangeability was assumed. The proof was

* Corresponding author. E-mail: Ahsan @ rider.bitnet.

[1] Research partially supported by the Stefan Batory Foundation, Warsaw, Poland

SSDI 0167-7152(93)E0165-P

based on considering the form of the joint density. A much stronger result of this kind was obtained in AP with (2) changed into $(X_1, \ldots, X_{n-1}) \stackrel{d}{=} (X_2, \ldots, X_n)$. The general approach they applied refers to existence and uniqueness of stationary distributions for indecomposable Markov processes. Our result, being parallel to that from AP, since we assume some other version of condition (2), is also another straightforward extension of both the results from Ahsanullah (1985) ($n = 2$) and Ahsanullah and Sinha (1986) (exchangeability). The method of the proof lies in applying an idea from Ahsanullah (1985) to conditional characteristic functions.

It should be emphasized that the Markov approach is not valid for the equidistribution condition, we consider, since the resulting process is not indecomposable.

All the problems discussed in this paper are invariant under univariate changes of scale or location.

2. Discussion of equidistribution conditions

Instead of (2) we are interested here in the following condition.

(3) $$(X_0, X_1, \ldots, X_k) \stackrel{d}{=} (X_0, X_1, \ldots, X_{k-1}, X_{k+1}), \quad k = 1, 2, \ldots, n-1,$$

where $X_0 = 0$ a.s. At first glance it looks like (3) is stronger than the AP assumption.

(4) $$(X_1, \ldots, X_{n-1}) \stackrel{d}{=} (X_2, \ldots, X_n).$$

However they are actually of similar nature. To see this take $n = 4$. Then both of the conditions yield identical distribution of all univariate marginals (it is fulfilled for any $n > 2$) and additionally from (3), we have

$$(X_1, X_2) \stackrel{d}{=} (X_1, X_3) \stackrel{d}{=} (X_1, X_4) \qquad (X_2, X_3) \stackrel{d}{=} (X_2, X_4), \qquad (X_1, X_2, X_3) \stackrel{d}{=} (X_1, X_2, X_4),$$

while (4) implies

$$(X_1, X_2) \stackrel{d}{=} (X_2, X_3) \stackrel{d}{=} (X_3, X_4) \qquad (X_1, X_3) \stackrel{d}{=} (X_2, X_4), \qquad (X_1, X_2, X_3) \stackrel{d}{=} (X_2, X_3, X_4),$$

Hence the conditions (3) and (4) have parallel forms.

It is not difficult to see that they are essentially different. To this end take

$$X_1 = U, \qquad X_2 = (U + V)/2, \qquad X_3 = (U + W)/2,$$

where U, V, W are i.i.d. standard Cauchy r.v.'s. Then $(X_1, X_2) \stackrel{d}{=} (X_1, X_3)$, $X_1 \stackrel{d}{=} X_2 \stackrel{d}{=} X_3$ and no renumeration is possible to obtain the AP combination. To see it observe that characteristic functions ϕ_{ij} of $(X_i, X_j), i, j = 1, 2, 3, i \neq j$, have the following forms:

$$\phi_{12}(s, t) = \phi_{13}(s, t) = \exp[-\tfrac{1}{2}(|2s + t| + |t|)],$$

$$\phi_{21}(s, t) = \phi_{31}(s, t) = \exp[-\tfrac{1}{2}(|s + 2t| + |t|)],$$

$$\phi_{23}(s, t) = \phi_{32}(s, t) = \exp[-\tfrac{1}{2}(|s + t| + |s| + |t|)].$$

Similarly, one can easily construct measures which have property (4) but not (3).

To emphasize the difference between both the combinations again take $n = 4$. In this case, the only Gaussian measures which are characterized in AP are those which have the following properties: $\rho_{12} = \rho_{23} = \rho_{34}, \rho_{13} = \rho_{24}$ while we are interested in those fulfilling: $\rho_{12} = \rho_{13} = \rho_{14}, \rho_{23} = \rho_{24}$, where ρ_{ij} is the correlation coefficient of X_i and X_j.

3. Characterizations

Our main result allows one to replace the equidistribution condition (2) by (3).

Theorem 3.1. *If the conditions* (1) *and* (3) *are fulfilled for some* $n \geqslant 2$, *then* $(X_1, \ldots, X_n)$ *has a multivariate normal distribution.*

Proof. Without any loss of generality we assume that $\alpha_0 = 0$. Consider the conditional characteristic functions $\phi_k(s) = E(e^{isX_k} | X_0, \ldots, X_{k-1})$, $k = 1, 2, \ldots, n$, $s \in \mathbb{R}$. Observe that it suffices to show that ϕ_k is normal with mean $\alpha_{0,k}X_0 + \alpha_{1,k}X_1 + \cdots + \alpha_{k-1,k}X_{k-1}$ and non-random variance σ_k^2, where $\alpha_{0,k}, \ldots, \alpha_{k-1,k}, \sigma_k$ are some real numbers and $\sigma_k > 0$, $k = 1, \ldots, n$. It is an obvious consequence of the fact that one can rebuild the joint distribution knowing $X_1, X_2 | X_1, \ldots, X_n | X_1, \ldots, X_{n-1}$.

To prove that the conditional distributions have the required form, we apply backward induction with respect to k. For $k = n$, the result holds by assumption (1). Assume now that $X_k | X_0, X_1, \ldots, X_{k-1}$ is $N(\alpha_{0,k}X_0 + \cdots + \alpha_{k-1,k}X_{k-1}, \sigma_k^2)$ for some $k = 2, \ldots, n$. Then by (3), we have

$$E(e^{isX_{k-1}} | X_0, \ldots, X_{k-2}) = E(e^{isX_k} | X_0, \ldots, X_{k-2})$$

and consequently

$$\phi_{k-1}(s) = E(\phi_k(s) | X_0, \ldots, X_{k-2}).$$

By induction, we obtain

$$\phi_{k-1}(s) = \phi_{k-1}(\alpha_{k-1,k}s)\exp[is(\alpha_{0,k}X_0 + \cdots + \alpha_{k-2,k}X_{k-2})]\exp(-\tfrac{1}{2}\sigma_k^2 s^2).$$

After m iterations of the above formula, we have

$$\begin{aligned}(5)\qquad \phi_{k-1}(s) = {}& \phi_{k-1}(\alpha_{k-1,k}^m s)\\ &\cdot\exp[is(\alpha_{0,k}X_0 + \cdots + \alpha_{k-2,k}X_{k-2})(1 + \alpha_{k-1,k} + \cdots + \alpha_{k-1,k}^{m-1})]\\ &\cdot\exp[-\tfrac{1}{2}\sigma_k^2 s^2(1 + \alpha_{k-1,k}^2 + \cdots + \alpha_{k-1,k}^{2(m-1)})].\end{aligned}$$

Assume now that $|\alpha_{k-1,k}| < 1$ and take in (5) limit for $m \to \infty$. Then

$$\phi_{k-1}(s) = \exp[is(\alpha_{0,k-1}X_0 + \cdots + \alpha_{k-2,k-1}X_{k-2})]\exp(-\tfrac{1}{2}\sigma_{k-1}^2 s^2),$$

where

$$\alpha_{j,k-1} = \alpha_{j,k}\frac{1}{1 - \alpha_{k-1,k}}, \quad j = 0, 1, \ldots, k-2,$$

$$\sigma_{k-1}^2 = \sigma_k^2 \frac{1}{1 - \alpha_{k-1,k}^2} > 0.$$

Hence $X_{k-1} | X_0, \ldots, X_{k-2}$ is normal.

Finally observe that $|\alpha_{k-1,k}| \geqslant 1$ is impossible since then upon taking limit in (5), we obtain $\phi_{k-1}(s) = 0$ for any $s \neq 0$ which contradicts the non-degeneracy assumption imposed on $X_k | X_0, \ldots, X_{k-1}$. $\square$

Remarks. (1) As a by product in the course of the proof we have obtained the following bounds $|\alpha_i| < 1$, $i = 1, 2, \ldots, n-1$.

(2) Observe that if all α's in (1) are equal to zero, then X_n has $N(0, \sigma^2)$ distribution and is independent of $(X_1, \ldots, X_{n-1})$. Hence by (3) it follows easily that $X_1, \ldots, X_n$ are i.i.d. The same observation holds also for the pair (1, 4).

(3) Note that you cannot weaken the assumption (3) by considering only $k = n - 1$. To see that take $n = 3$. Let a non-normal r.v. X_1 be independent of a bivariate normal vector (X_2, X_3), such that $X_2 \stackrel{d}{=} X_3$. Then $(X_1, X_2) \stackrel{d}{=} (X_1, X_3)$. Since $X_3 | X_1, X_2 \stackrel{d}{=} X_3 | X_2$ then the condition (1) is also fulfilled. However (X_1, X_2, X_3) is not normal jointly.

(4) All three pairs of conditions (1, 2), (1, 3) and (1, 4) have for $n = 2$ the same form and are not only sufficient but also necessary. For $n > 2$ (1, 2) is not sufficient (while, obviously, remains necessary) and the pairs (1, 3) and (1, 4) are no longer necessary (being sufficient). Hence we come to an interesting open problem of complementing (1) by some equidistribution assumption to obtain a pair which is both necessary and sufficient for any $n \geqslant 2$.

Combining our method and the AP Markovian approach the following general result may be obtained easily.

Instead of (2) consider a condition being a mixture of (3) and (4).

$$\text{For some } k,\ 1 \leqslant k \leqslant n,$$

$$(X_0, X_1, \ldots, X_{l-1}, X_l) \stackrel{d}{=} (X_0, X_1, \ldots, X_{l-1}, X_{l+1}), \quad l = k, \ldots, n-1, \quad \text{if } k < n, \tag{6}$$

$$(X_1, \ldots, X_{k-1}) \stackrel{d}{=} (X_2, \ldots, X_k) \quad \text{if } k > 1.$$

Theorem 3.2. *If the condition* (1) *and* (6) *are fulfilled for some* $n \geqslant 2$ *then* $(X_1, \ldots, X_n)$ *is multivariate normal.*

Observe that for $k = n$, Theorem 3.2 yields the AP characterization and for $k = 1$ our Theorem 3.1 follows. Also rephrasings of this result involving permutations of r.v.'s are possible.

Acknowledgement

The second author would like to express his sincere thanks to Rider University, Lawrenceville, NJ, for providing facilities for this research. The authors thank the referee for his helpful comments.

References

Ahsanullah, M. (1985), Some characterizations of the bivariate normal distribution, *Metrika* **32**, 215–218.

Ahsanullah, M. and B.K. Sinha (1986), On normality via conditional normality, *Calcutta Statist. Assoc. Bull.* **35**, 193–202.

Arnold, B.C. (1993), Characterization involving conditional specification, *The Third Eugene Lukacs Symp.*, Bowling Green, OH, Invited Lecture.

Arnold, B.C. and M. Pourahamadi (1988), Conditional characterizations of multivariate distributions, *Metrika* **35**, 99–108.

Hamedani, G.G. (1992), Bivariate and multivariate normal characterizations, *Comm. Statist. Theory Methods* **21**(9), 2665–2689.

Stoyanov, J.N. (1987), *Counter Examples in Probability* (Wiley, Chichester, UK).

Wesolowski, J. (1991), Gaussian conditional structure of the second order and the Kagan classification of multivariate distributions, *J. Multivariate Anal.* **39**, 79–86.