stichting
mathematisch
centrum

AFDELING MATHEMATISCHE STATISTIEK SW 14/72 JANUARY

F.H. RUYMGAART

NON-NORMAL BIVARIATE DENSITIES WITH NORMAL MARGINALS AND LINEAR REGRESSION FUNCTIONS

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat 49, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

NON-NORMAL BIVARIATE DENSITIES WITH NORMAL MARGINALS AND LINEAR REGRESSION FUNCTIONS

by

F.H. Ruymgaart

Summary. It is well known that for a bivariate density in order to be bivariate normal, the possession of univariate normal marginal densities alone is not sufficient. In this paper it will be shown by means of some counterexamples that the additional requirement of linearity of the regression functions does not supply a sufficient condition either.

1. Introduction. Throughout this paper attention will be restricted to bivariate densities with standard normal marginal densities. As usual the univariate standard normal density is denoted by

$$(1.1) \qquad \phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}, \qquad -\infty < t < \infty.$$

This assumption implies the finiteness of moments of any order. For such densities $h(x,y)$ the correlation coefficient ρ equals $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyh(x,y)dxdy$, the covariance. For ease of reference the above defined densities will be called densities of type I.

Suppose a density of type I is given. Let us write $r(x) = \int_{-\infty}^{\infty} y[h(x,y)/\phi(x)]dy$ for the regression of y on x and $s(y) = \int_{-\infty}^{\infty} x[h(x,y)/\phi(y)]dx$ for the regression of x on y. If these regression functions are given to be linear, say

$$(1.2) \qquad \begin{aligned} r(x) &= ax + b, \qquad -\infty < x < \infty, \\ s(y) &= cy + d, \qquad -\infty < y < \infty, \end{aligned}$$

where a, b, c, d are given constants, then a necessary condition for bivariate normality of the density h of type I is that the above constants satisfy the relations

$$(1.3) \qquad a = c = \rho, \quad b = d = 0.$$

We now define a density to be of type II if it is of type I and if, moreover, it has linear regression functions as given in (1.2), satisfying (1.3).

Some well known non-bivariate normal densities of type I are easily seen to be also of type II. This suggests that it would be generally known that densities of type II are not necessarily bivariate normal, but an explicit statement of this fact is unknown to the author.

2. Densities of type I, that are not of type II. In this section, by way of a further introduction, two densities of type I will be given that are not of type II, and that consequently are not bivariate normal. Let us first introduce a special notation for the bivariate normal density with standard normal marginals and correlation coefficient $-1 < \rho < 1$. This density will be written

$$(2.1) \qquad \phi_\rho(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\, e^{-\frac{x^2-2\rho xy+y^2}{2}}, \qquad -\infty < x,\ y < \infty.$$

Let us start with a simple example communicated by SHORACK [3]. Consider the density ϕ_0 (see (2.1)). Draw a circle in each of the four quadrants of the (x,y)-plane as indicated in Fig. 1. Transfer the mass contained in circle 2 to

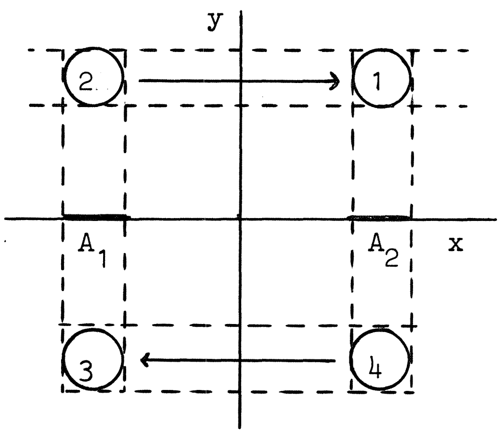

Fig. 1. Construction of non-normal bivariate density with normal marginals.

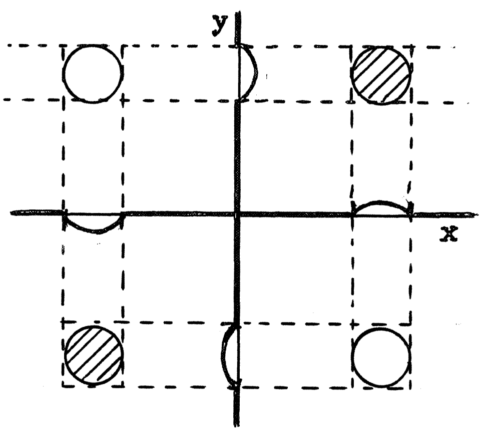

Fig. 2. The graphs of the regression functions corresponding to the density of Fig. 1.

circle 1, and transfer the mass contained in circle 4 to circle 3. (By this procedure a positive correlation is introduced.) The resulting density $h_1(x,y)$ clearly is a probability density and still has standard normal marginal densities. However, it evidently cannot be bivariate normal, since e.g. it assigns mass 0 to circles 2 and 4. This non-normality follows also from the character of the regression functions $r_1(x)$ and $s_1(y)$. It will intuitively be made clear that these functions are not linear (see Fig. 2). For instance consider $r_1(x)$. For x-values not contained in A_1 or A_2 (the projections of circles 2, 3 and 1, 4 on the x-axis respectively, (see Fig. 1) we have $h_1(x,y) = \phi_0(x,y)$ for all y, which is a function of y symmetric about y = 0. This implies $r_1(x) = 0$ for all x not contained in A_1 or A_2. For x-values in A_1 however, the function $h_1(x,y)$ is a function of y which is no longer symmetric about y = 0: it assigns more weight to the set of negative y-values. This results in a deviation from 0 of the function $r_1(x)$ in the negative direction (see Fig. 2). Similarly the deviation of $r_1(x)$ from 0 for x-values in A_2 is in the positive direction. An analogous reasoning holds for the regression function $s_1(y)$. As it were, the extra mass contained in circles 1 and 3 "attracts" the regression functions. The conclusion is that the density h_1 is of type I but not of type II, by non-linearity of the regression lines. Hence, the density is non-normal.

The reader may object that this example is rather artificial, because the density involved is discontinuous on the boundaries of the circles. To meet this objection from now on only completely continuous bivariate densities will be considered. If a density of type I has at least one continuous but not linear regression function, it is not of type II and it cannot be normal. Whether the density is of type I or type II, by the assumption of continuity it is bivariate normal only if $h(x,y) = \phi_\rho(x,y)$ for all $-\infty < x, y < \infty$. Here ρ is the correlation coefficient of h. A geometrical argument for rejecting the normality of a continuous bivariate density h is found by observing that it is necessary for normality that the intersection of the surface $z = h(x,y)$ in 3-dimensional space with any plane, perpendicular to the (x,y)-plane, is a symmetric and unimodal curve. In both arguments the assumption that h is continuous everywhere is essential.

The second example is a special case of an example given in

FELLER [1], page 99, which is due to E. NELSON. Let us introduce the function

$$u(t) = \sin t \qquad \text{for } -2\pi \leq t \leq 2\pi,$$

(2.2)

$$u(t) = 0 \qquad \text{for } t < -2\pi \text{ or } t > 2\pi.$$

The function $u(t)$ is continuous on the real line and so is the function $u(x) \cdot u(y)$ on the plane. Hence the function

$$h_2(x,y) = \phi_0(x,y) + \lambda\, u(x)\, u(y) \tag{2.3}$$

is continuous too. Here λ is a constant satisfying $0 < \lambda < \min_{-2\pi\leq x,y\leq 2\pi} \phi_0(x,y) = \phi_0(2\pi,2\pi)$ to prevent h_2 from assuming negative values. Since $\int_{-\infty}^{\infty} u(t)dt = 0$ it follows that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(x,y)dxdy = 1$, $\int_{-\infty}^{\infty} h_2(x,y)dy = \phi(x)$ and $\int_{-\infty}^{\infty} h_2(x,y)dx = \phi(y)$ (see (1.1)). Hence h_2 is a continuous density of type I. By observing that $\int_{-\infty}^{\infty} tu(t)dt = -4\pi$ we obtain for the regression function $r_2(x)$:

$$\begin{aligned} r_2(x) &= \int_{-\infty}^{\infty} y[h_2(x,y)/\phi(x)]dy \\ &= \int_{-\infty}^{\infty} y[\phi_0(x,y)/\phi(x)]dy + \lambda[u(x)/\phi(x)] \int_{-\infty}^{\infty} yu(y)dy \\ &= -\lambda 4\pi[u(x)/\phi(x)]. \end{aligned}$$

This function is continuous but not linear, hence the density is not of type II and consequently it is not normal.

Geometrically it can be seen that h_2 is not normal by plotting the graph of $h_2(x,\pi/2) = \phi_0(x,\pi/2) + \lambda u(x)$ as a function of x. From Fig. 3 it may be seen that this function is not symmetric, and hence a necessary condition for normality is not fulfilled.

In the next section it will be shown that a slight modification of this example gives a non-normal density of type II.

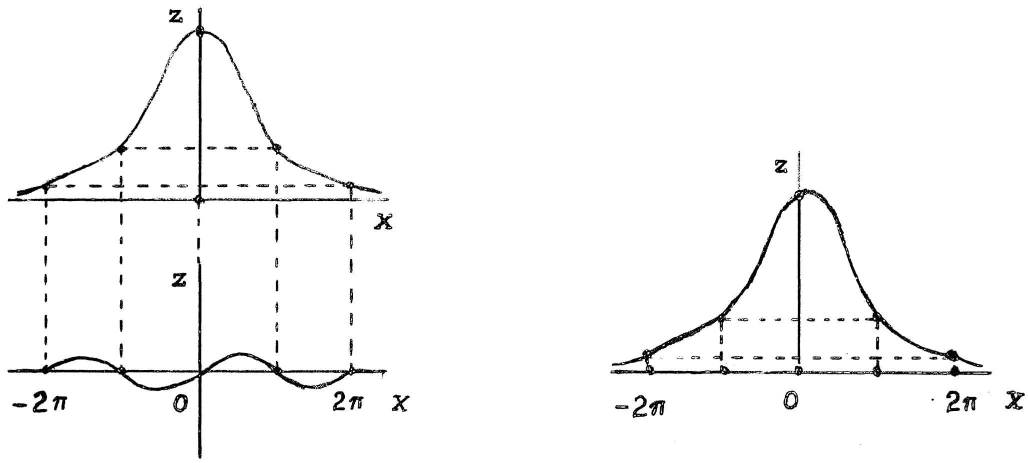

Fig. 3. The graph of the function $h_2(x,\pi/2)$, defined in (2.3).

3. Two non-normal densities of type II. Instead of $u(t)$ let us consider $u(|t|)$ (see (2.2)) and define

$$h_3(x,y) = \phi_0(x,y) + \lambda u(|x|)\, u(|y|). \tag{3.1}$$

Here λ is the same constant as in (2.3). This function is continuous and from $\int_{-\infty}^{\infty} u(|t|)dt = 0$ it follows that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(x,y)dxdy = 1$, $\int_{-\infty}^{\infty} h_3(x,y)dy = \phi(x)$ and $\int_{-\infty}^{\infty} h_3(x,y)dx = \phi(y)$. Hence h_3 is a continuous density of type I. For the regression function we now have

$$\begin{aligned} r_3(x) &= \int_{-\infty}^{\infty} y[h_3(x,y)/\phi(x)]dy \\ &= \int_{-\infty}^{\infty} y[\phi_0(x,y)/\phi(x)]dy + \lambda[u(|x|)/\phi(x)] \int_{-\infty}^{\infty} yu(|y|)dy \\ &= 0, \end{aligned}$$

and analogously

$$s_3(y) = 0,$$

because $\int_{-\infty}^{\infty} tu(|t|)dt = 0$. For the same reason we find $\rho_3 = 0$. Thus the regressions are linear (they satisfy (1.2)) and (1.3) holds. Consequently h_3 is a continuous density of type II.

Since $\lambda > 0$, it follows that $h_3(\pi/2,\pi/2) = \phi_0(\pi/2,\pi/2) + \lambda$ is unequal to $\phi_0(\pi/2,\pi/2)$. Hence h_3 cannot be normal (see Section 2).

A geometrical argument disproving the normality is found by plotting the graph of the function $h_3(x,\pi/2) = \phi_0(x,\pi/2) + \lambda u(|x|)$.

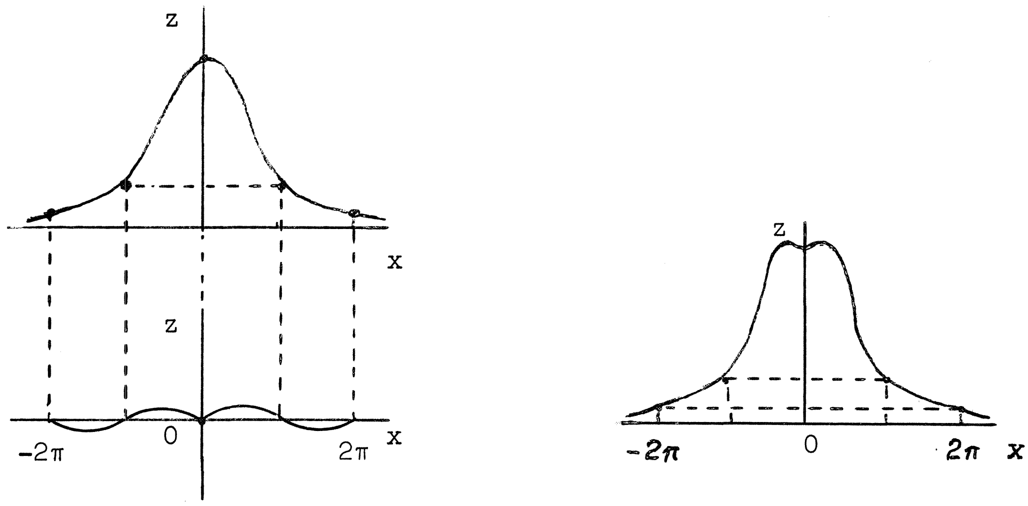

Fig. 4. The graph of the function $h_3(x,\pi/2)$, defined in (3.1).

From Fig. 4 it follows that this function is not unimodal as is required for normality.

One may still object that these examples are rather artificial. A more natural example is a special case of FELLER [1], page 99. Let us consider the mixture

$$h_4(x,y) = [\phi_{-\frac{1}{2}}(x,y) + \phi_{\frac{1}{2}}(x,y)]/2. \tag{3.2}$$

In [1] this density is presented only as an example of a non-normal density of type I. It is easy to see that h_4 is a continuous density of type I. Since ϕ_ρ has regression lines ρx and ρy we have

$$r_4(x) = \int_{-\infty}^{\infty} y[h_4(x,y)/\phi(x)]dy$$

$$= [\int_{-\infty}^{\infty} y[\phi_{-\frac{1}{2}}(x,y)/\phi(x)]dy + \int_{-\infty}^{\infty} y[\phi_{\frac{1}{2}}(x,y)/\phi(x)]dy]/2$$

$$= [-x/2 + x/2]/2$$

$$= 0,$$

and analogously

$$s_4(y) = 0.$$

Moreover, $\rho_4 = (-\frac{1}{2} + \frac{1}{2})/2 = 0$ and hence the regressions satisfy both (1.2) and (1.3). This proves that h_4 is a continuous density of type II.

This density is not normal because $h_4(0,0) = 1/\pi\sqrt{3}$ which is unequal to $\phi_0(0,0) = 1/2\pi$ (see Section 2).

Geometrically it can also be seen that this density is not normal. Let us take an arbitrary $\eta > 0$ and let us consider the graph of the function $h_4(x,\eta)$ as a function of x. This function may be written (see (2.1))

$$h_4(x,\eta) = \frac{2e^{-\frac{3\eta^2}{8}}}{\sqrt{6\pi}} \left[\frac{1}{2} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{(x+\eta/2)^2}{2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\eta/2)^2}{2}} \right) \right]. \tag{3.3}$$

The expression within the brackets is a mixture of two univariate normal densities having expectations $-\eta/2$ and $\eta/2$ respectively and unit variances. From WESSELS [4], Section 4, it follows that for η sufficiently large this mixture, and hence $h_4(x,\eta)$, is bimodal (see Fig. 5). According to Section 2 unimodality is necessary for normality.

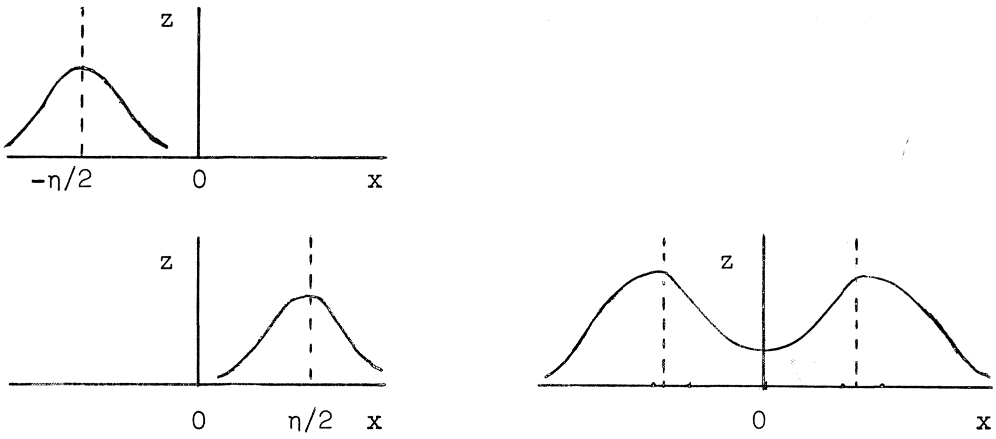

Fig. 5. The graph of the function $h_4(x,\eta)$, defined in (3.3).

So far all bivariate densities considered in this section have correlation coefficient zero. This restriction is inessential. Especially the last example, which is the most appealing one, can straightforwardly be extended. It can be shown that

$$(3.4) \qquad h^*(x,y) = \lambda\phi_\rho(x,y) + (1-\lambda)\,\phi_{\tilde{\rho}}(x,y),$$

where $0 < \lambda < 1$ and $-1 < \rho$, $\tilde{\rho} < 1$ with $\rho \neq \tilde{\rho}$, is a continuous non-normal density of type I (see [1]). It is easily seen that the regression functions are

$$(3.5) \qquad \begin{aligned} r^*(x) &= \lambda\rho x + (1-\lambda)\tilde{\rho}x, \\ s^*(y) &= \lambda\rho y + (1-\lambda)\tilde{\rho}y. \end{aligned}$$

Moreover $\rho^* = \lambda\rho + (1-\lambda)\tilde{\rho}$, so that the regression functions again satisfy (1.2) and (1.3). Consequently the class $\{h^* = h^*_{\lambda,\rho,\tilde{\rho}} : 0<\lambda<1,\ -1<\rho,\ \tilde{\rho}<1,\ \rho\neq\tilde{\rho}\}$ is a class of continuous non-normal densities of type II. For most values of λ, ρ, $\tilde{\rho}$ the corresponding ρ^* is unequal to zero.

4. Some literature. Apart from the references given above it may be useful to refer to the recent book of MARDIA [2] on bivariate distributions. Possibly with help of some general techniques reviewed in this book, a less heuristic approach to the problem can be found.

Acknowledgement. The author would like to thank Professor HEMELRIJK for useful comments.

REFERENCES

[1] FELLER, W. (1966). An introduction to Probability Theory and its Applications 2. John Wiley and Sons, Inc., New York.

[2] MARDIA, K.V. (1970). Families of Bivariate Distributions. Griffin, London.

[3] SHORACK, G.R. Personal communication.

[4] WESSELS, J. (1964). Multimodality in a family of probability densities, with application to a linear mixture of two normal densities. Statistica Neerlandica 18, 267-282.