

GENERALIZED NORMAL CORRELATION

O. V. Sarmanov

1966

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196601.65333>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

UDC 519.21

MATHEMATICS

O. V. SARMANOV

GENERALIZED NORMAL CORRELATION AND TWO-DIMENSIONAL FRÉCHET CLASSES

(Presented by Academician S. N. Bernstein on 17 VIII 1965)

1. Consider the class of functions $N\{f(x, y)\}$ representable by bilinear expansions in Hermite polynomials

$$f(x, y) = \frac{\exp\left[-\frac{x^2+y^2}{2}\right]}{2\pi} \left[1 + \sum_{k=1}^{\infty} c_k H_k(x) H_k(y) \right]. \quad (1)$$

The series (1) is assumed to converge in the mean, which is equivalent to the condition

$$\sum_{k=1}^{\infty} c_k^2 < \infty, \quad (2)$$

where $H_k(x)$ are orthogonal and normal with weight $e^{-x^2/2}/\sqrt{2\pi}$ on the whole real axis, i.e.

$$H_k(x) = \frac{(-1)^k}{\sqrt{k!}} e^{x^2/2} \frac{d^k}{dx^k} (e^{-x^2/2}) = \frac{1}{\sqrt{k!}} \left[x^k - \frac{k(k-1)}{2} x^{k-2} + \dots \right]. \quad (3)$$

It follows from (1) that

$$\int_{-\infty}^{\infty} f(x, y) dx = \frac{e^{-y^2/2}}{\sqrt{2\pi}}; \quad \int_{-\infty}^{\infty} f(x, y) dy = \frac{e^{-x^2/2}}{\sqrt{2\pi}}. \quad (4)$$

Definition. When the sum of the series (1) is nonnegative for all real x and y , we shall say that there is a generalized normal correlation between x and y .

2. Theorem. In order that the sum of the series (1) be the density of a generalized normal correlation, it is necessary and sufficient that the coefficients c_k

be the moments of some probability distribution concentrated inside the interval $[-1, 1]$.

Necessity. Let the sum of the series (1) be nonnegative for all x and y . Find the conditional mathematical expectation of the quantity $(y/x)^k$ for fixed $x \neq 0$. Express y^k through the polynomials (3):

$$y^k = \sqrt{k!} H_k(y) + a_{k-2} H_{k-2}(y) + a_{k-4} H_{k-4}(y) + \dots,$$

where $a_{k-2}, a_{k-4}, \dots$ are fully determined constants.

$$\begin{aligned} M_x y^k &= \sqrt{k!} c_k H_k(x) + a_{k-2} c_{k-2} H_{k-2}(x) + \dots \\ &= c_k x^k + b_{k-2} x^{k-2} + b_{k-4} x^{k-4} + \dots, \end{aligned}$$

where $b_{k-2}, b_{k-4}, \dots$ are also certain constants. Thus,

$$M_x (y/x)^k = c_k + b_{k-2}/x^2 + b_{k-4}/x^4 + \dots. \quad (5)$$

and for sufficiently large x , c_k differs arbitrarily little from the k -th moment of the variable y/x ; moreover, by virtue of (2), c_k cannot be a moment of a variable that, with positive probability, goes outside the interval $[-1, 1]$ or assumes the values ± 1 .

Sufficiency. Let $\{c_k\}$ form the sequence of moments of some variable ξ , concentrated inside the interval $[-1, 1]$ and given by the distribution function $G(\xi)$; then for each fixed $|\xi| < 1$ consider the well-known expansion for the density of normal correlation with correlation coefficient ξ

$$\frac{\exp \left[-\frac{x^2 + y^2 - 2\xi xy}{2(1-\xi^2)} \right]}{2\pi\sqrt{1-\xi^2}} = \frac{\exp \left[-\frac{x^2 + y^2}{2} \right]}{2\pi} \left[1 + \sum_{k=1}^{\infty} \xi^k H_k(x) H_k(y) \right], \quad (6)$$

where, by assumption,

$$\int_{-1}^1 \xi^k dG(\xi) = c_k, \quad k = 1, 2, \dots \quad (7)$$

From (6) and (7) it follows that

$$\frac{\exp \left[-\frac{x^2 + y^2}{2} \right]}{2\pi} \left[1 + \sum_{k=1}^{\infty} c_k H_k(x) H_k(y) \right] =$$

$$= \frac{1}{2\pi} \int_{-1}^1 \frac{\exp \left[-\frac{x^2+y^2-2\xi xy}{2(1-\xi^2)} \right]}{\sqrt{1-\xi^2}} dG(\xi) > 0, \quad (8)$$

which completes the proof.

Remark. In normal correlation, $\xi = R$ with probability 1, and $c_k = R^k$. If ξ assumes the two values $\pm R$ with probabilities $1/2$, so that $c_{2k} = R^{2k}$, $c_{2k-1} = 0$, $k = 1, 2, \dots$, the corresponding density was considered in ⁽¹⁾ and was given the name pseudonormal correlation; finally, if all $c_k = 0$, then x and y are independent.

Corollary 1. If the sum (1) has constant sign and the variables x and y are dependent, then $c_{2k} > 0$ for all $k = 1, 2, \dots$.

Corollary 2. It is easy to verify that the characteristic function of the density (8) has the form

$$\varphi(t_1, t_2) = \exp \left[-\frac{t_1^2 + t_2^2}{2} \right] \int_{-1}^1 \exp[-Rt_1 t_2] dG(R). \quad (9)$$

In particular, if it is known that the correlation between x and y is normal, but the correlation coefficient is not known exactly and is uniformly distributed on the interval $[R - \varepsilon, R + \varepsilon]$, $R - \varepsilon > -1$, $R + \varepsilon < 1$, then the characteristic function of such a generalized normal correlation is expressed with the aid of (9) as follows:

$$\varphi(t_1, t_2, \varepsilon) = \exp \left[-\frac{t_1^2 + t_2^2 + 2Rt_1 t_2}{2} \right] \frac{\text{sh}(\varepsilon t_1 t_2)}{\varepsilon t_1 t_2}. \quad (10)$$

2. M. Fréchet in ^(2, 3) posed the following general problem: to find the class of two-dimensional distribution functions $\{P(x, y)\}$ having given partial (extreme or marginal) one-dimensional distributions $F(x)$ and $F_1(y)$. In the preceding section an important part of the normal Fréchet class has been found.

Morgenstern and Gumbel ⁽⁴⁾ considered the following example of functions belonging to the Fréchet class $K\{F(x), F_1(y)\}$ for arbitrary given $F(x)$ and $F_1(y)$:

$$P(x, y) = F(x)F_1(y) [1 + \lambda(1 - F(x))(1 - F_1(y))], \quad -1 < \lambda < 1. \quad (11)$$

If, however, densities exist (denoted by the corresponding lowercase letters), then

$$p(x, y) = f(x)f_1(y) [1 + \lambda(2F(x) - 1)(2F_1(y) - 1)]; \quad (12)$$

in particular, in the normal case,

$$p(x, y) = \frac{\exp\left[-\frac{x^2+y^2}{2}\right]}{2\pi} [1 + 4\lambda\Phi(x)\Phi(y)], \quad (13)$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt.$$

3. Let us consider the following more general method of constructing functions from the Fréchet class $k\{f(x), f_1(y)\}$, specified by the densities $f(x)$ and $f_1(y)$ on $L = [a \leq x \leq b]$ and $L_1 = [a_1 \leq y \leq b_1]$. Let $\varphi(x)$ be an arbitrary function (not equal to a constant) bounded on L , and let $\varphi_1(y)$ be an arbitrary function bounded on L_1 . Then the quantities

$$\begin{aligned} m &= \int_a^b \varphi(x) f(x) dx, & m_1 &= \int_{a_1}^{b_1} \varphi_1(y) f_1(y) dy, \\ \sigma^2 &= \int_a^b [\varphi(x) - m]^2 f(x) dx, & \sigma_1^2 &= \int_{a_1}^{b_1} [\varphi_1(y) - m_1]^2 f_1(y) dy \end{aligned} \quad (14)$$

are meaningful.

Denote

$$h = \sup_{x \in L} \frac{|\varphi(x) - m|}{\sigma}, \quad h_1 = \sup_{y \in L_1} \frac{|\varphi_1(y) - m_1|}{\sigma_1}. \quad (15)$$

The one-parameter family of densities

$$p_1(x, y, \lambda) = f(x) f_1(y) \left[1 + \frac{\lambda}{hh_1} \frac{[\varphi(x) - m][\varphi_1(y) - m_1]}{\sigma\sigma_1} \right], \quad (16)$$

where $-1 \leq \lambda \leq 1$, belongs to the Fréchet class $k\{f(x), f_1(y)\}$, since $\varphi(x) - m$ and $\varphi_1(y) - m_1$ are orthogonal to unity with weights $f(x)$ and $f_1(y)$, and the sum (16) is nonnegative.

We note that, by virtue of (14), $h \geq 1$ and $h_1 \geq 1$, and

$$\rho = \lambda/hh_1 \quad (17)$$

is the maximal correlation coefficient (5) between x and y .

Integrating (16), we pass to the distribution functions

$$P_1(x, y, \lambda) = F(x)F_1(y) + \frac{\lambda}{hh_1} \int_a^x \frac{\varphi(u) - m}{\sigma} dF(u) \int_{a_1}^y \frac{\varphi_1(v) - m_1}{\sigma_1} dF_1(v). \quad (18)$$

Since $F(x)$ and $F_1(y)$ are bounded functions, one may put $\varphi(x) \equiv F(x)$, and $\varphi_1(y) \equiv F_1(y)$; then, for continuous $F(x)$ and $F_1(y)$, we obtain that

$$\frac{1}{2} = \int_a^b F(x) dF(x) = m = m_1 = \sup_{x \in L} |F(x) - m| = \sup_{y \in L_1} |F_1(y) - m_1|,$$

$$\begin{aligned} P_1(x, y, \lambda) &= F(x)F_1(y) + \lambda(F^2(x) - F(x))(F_1^2(y) - F_1(y)) = \\ &= F(x)F_1(y)[1 + \lambda(F(x) - 1)(F_1(y) - 1)], \end{aligned} \quad (19)$$

as in example (11). The assumption of continuity is not a restriction, for, if it is proved that (19) is a distribution function for

for arbitrary discontinuous $F(x)$ and $F_1(y)$, then (19) gives a distribution function also for arbitrary distribution functions $F(x)$ and $F_1(y)$.

4. If the densities $p_k(x, y)$ belong to some Fréchet class, then a linear combination of them of the form $\sum_k \varepsilon_k p_k(x, y)$, where $\varepsilon_k > 0$ and $\sum_k \varepsilon_k = 1$, also belongs to the same class. With the help of this remark and (16), one can construct new examples of densities and distribution functions belonging to a given class (the possibility of similar bilinear constructions was indicated in ⁶).

Let us give several examples

$$p(x, y) = f(x)f_1(y) \left[1 + \sum_{k=1}^{\infty} \varepsilon_k \cos 2k\pi(F(x) - 1/2) \cos 2k\pi(F_1(y) - 1/2) \right], \quad (20)$$

where $\sum_{k=1}^{\infty} |\varepsilon_k| \leq 1$, and, integrating (20), we obtain the corresponding family of distribution functions

$$P(x, y) = F(x)F_1(y) + \frac{1}{4\pi^2} \sum_{k=1}^{\infty} \frac{\varepsilon_k}{k^2} \sin 2\pi k(F(x) - 1/2) \sin 2\pi k(F_1(y) - 1/2). \quad (21)$$

If in (18) we put $\varphi(x) = F^k(x)$, $\varphi_1(y) = F_1^k(y)$, and then consider the linear combination of the functions obtained, putting $\varepsilon_k = \varepsilon^k$, $k = 1, 2, \dots$, $0 \leq \varepsilon \leq 1/2$,

then we arrive at the following expression, also giving functions from the Fréchet class:

$$P_2(x, y, \varepsilon) = F(x)F_1(y) \left[1 + \frac{l(\varepsilon)}{\varepsilon} \right] - \frac{F(x)}{\varepsilon} l[\varepsilon F_1(y)] - \frac{F_1(y)}{\varepsilon} l[\varepsilon F(x)] + \frac{1}{\varepsilon} l[\varepsilon F(x)F_1(y)], \quad (22)$$

where

$$l(z) = - \int_0^z \frac{\ln(1-t)}{t} dt = \sum_{k=1}^{\infty} \frac{z^k}{k^2}, \quad \text{and} \quad P_2(x, y, 0) = F(x)F_1(y).$$

In conclusion, let us give one example of discontinuous densities from the normal Fréchet class

$$f(x, y, \lambda) = \begin{cases} \frac{\exp \left[-\frac{x^2 + y^2}{2} \right]}{2\pi} (1 + \lambda), & \text{for } xy > 0, \\ \frac{\exp \left[-\frac{x^2 + y^2}{2} \right]}{2\pi} (1 - \lambda), & \text{for } xy \leq 0, \\ -1 \leq \lambda \leq 1. \end{cases} \quad (23)$$

For the distribution (23), λ is the maximal correlation coefficient.

Steklov Mathematical Institute
Academy of Sciences of the USSR

Received
17 VIII 1965

REFERENCES

- ¹ O. V. Sarmanov, DAN, **132**, No. 2 (1960).
- ² M. Fréchet, Ann. Univ. Lyon, ser. III, **A14**, 53 (1951).
- ³ M. Fréchet, C. R., **242**, No. 20, 2426 (1956).
- ⁴ E. Gumbel, C. R., **246**, No. 19, 2717 (1958).
- ⁵ O. V. Sarmanov, DAN, **121**, No. 1 (1958).
- ⁶ S. Kantorovitz, Bull. Res. Council. Israel, **F.7**, No. 1, 43 (1957).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.