A CHARACTERIZATION OF DEVIATION FROM NORMALITY UNDER CERTAIN MOMENT ASSUMPTIONS

W.R. McGillivray and C. L. Kaller

(received January 16, 1966)

1. If F_n is the distribution function of a distribution with moments up to order n equal to those of the standard normal distribution, then from Kendall and Stuart [1, p. 87],

$$\lim_{n \to \infty} F_n(x) = \Phi(x),$$

where Φ is the distribution function of the standard normal distribution. It is of practical interest to provide some indication of the possible deviation from normality of a distribution which has a limited number of its moments equal to the corresponding normal moments. This is of particular significance with respect to the estimation of population parameters when a population is assumed to be normally distributed.

2. Define a symmetric probability density function f_{2n} on the real numbers, R_1, by $f_{2n}(x) = p_{2n}(x) \emptyset(x)$ for $n = 2, 3, \ldots,$ where $p_{2n}(x) = a_o + a_2 x^2 + \ldots + a_{2n} x^{2n}$, $a_{2i} \in R_1$ for $i = 0, 1, 2, \ldots, n$, and $\emptyset$ is the standard normal probability density function. Then

(1) $$p_{2n}(x) \geq 0 \quad \text{for all } x \in R_1,$$

and

(2) $$\int_{-\infty}^{\infty} f_{2n}(x) = 1.$$

https://doi.org/10.4153/CMB-1966-061-5 Published online by Cambridge University Press

We now impose the moment constraints:

$$(3) \qquad \int_{-\infty}^{\infty} x^{2r} f_{2n}(x)\,dx = \int_{-\infty}^{\infty} x^{2r} \phi(x)\,dx = \frac{(2r)!}{2^r r!},$$

where $r = 1, 2, \ldots, k \le n$, and $n = 2, 3, \ldots$.

Then, since $\frac{(2r)!}{2^r r!}$ is the $2r^{th}$ central moment (u_{2r}^N) of the standard normal distribution, the moments of the distribution with density function f_{2n} are identical to the standard normal moments up to order $2k$. Equation (2) and the k simultaneous equations given by (3) yield $k + 1$ simultaneous equations in the coefficients $\{a_{2i}\}_{i=0}^{n}$.

3. If $k = n$, we have the unique solution $a_o = 1$ and $a_{2i} = 0$ for $i = 1, 2, \ldots, n$.

If $k = n-1$, we have the following matrix form, where $\bar{r}$ is defined by

$$\bar{r} = \begin{cases} r(r-2)\ldots 2, & \text{if } r \text{ is an even integer,} \\ r(r-2)\ldots 1, & \text{if } r \text{ is an odd integer.} \end{cases}$$

$$(4) \qquad \begin{bmatrix} 1 & \bar{1} & \ldots & \overline{2n-3} \\ \bar{1} & \bar{3} & \ldots & \overline{2n-1} \\ \vdots & \vdots & & \vdots \\ \overline{2n-3} & \overline{2n-1} & \ldots & \overline{4n-5} \end{bmatrix} \begin{bmatrix} a_o \\ a_2 \\ \vdots \\ a_{2n-2} \end{bmatrix} = \begin{bmatrix} 1 - \overline{2n-1}\,a_{2n} \\ \bar{1} - \overline{2n+1}\,a_{2n} \\ \vdots \\ \overline{2n-3} - \overline{4n-3}\,a_{2n} \end{bmatrix}$$

Let this array be designated by $M_n A_n = B_n$, where $A_n^T = (a_o, a_2, \ldots, a_{2n-2})$. Either by showing that $|M_{n+1}| = (2n)!\,|M_n|$ (where $|A|$ is the determinant of A) or by considering the quadratic form

$$A_n^T M_n A_n = E(a_o + a_2 X + \ldots + a_{2n-2} X^{2n-1})^2,$$

(where X is a standard normal random variable) it can be shown

https://doi.org/10.4153/CMB-1966-061-5 Published online by Cambridge University Press

that M_n is non-singular.

Denoting by H_k the Hermite Polynomial of order k defined by $\phi(x)H_k(x) = (-1)^k \phi^{(k)}(x)$, we have the following

THEOREM 1. The probability density function $f_{2n}(x) = p_{2n}(x)\,\phi(x)$ is given by

$$f_{2n}(x) = \phi(x)\{1 + a_{2n} H_{2n}(x)\}, \tag{5}$$

where $n = 2, 3, \ldots,$ and a_{2n} is chosen so that $f_{2n}(x) \geq 0$ for all $x \in R_1$.

Proof. Since M_n is non-singular, there exists a unique solution for (4) given by $A_n = (M_n)^{-1} B_n$. This solution is in terms of the parameter a_{2n}. It is evident that (5) has the properties required by (1), (2) and (3) and hence (5) is the required expression for f_{2n}. q.e.d.

The condition in Theorem 1 that $f_{2n}(x) \geq 0$ for all $x \in R_1$ is fulfilled by choosing a_{2n} such that $0 \leq a_{2n} \leq A_{2n}$, where

$$A_{2n} = \frac{1}{\left|\inf_x H_{2n}(x)\right|} \quad \text{for } n = 2, 3, \ldots . \tag{6}$$

A more general equation than (5) may be obtained by considering the polynomial

$$q_n(x) = \sum_{i=0}^{n} a_i x^i \quad \text{for } n = 1, 2, \ldots .$$

Requiring $f_n(x) = q_n(x)\,\phi(x)$ to satisfy $\int_{-\infty}^{\infty} f_n(x)dx = 1$, and

$$\int_{-\infty}^{\infty} x^i f_n(x)dx = u_n^N, \quad i = 1, 2, \ldots, n-1,$$

produces $f_n(x) = \phi(x)\{1 + a_n H_n(x)\}$. The only equations of this form which permit a non-zero a_n to be chosen so that $f_n(x) \geq 0$ for all $x \in R_1$ are those where n is even, as is the case in (5)

https://doi.org/10.4153/CMB-1966-061-5 Published online by Cambridge University Press

above.

4. Equation (5) may be used to characterize possible deviation from normality when $2n-2$ moments of a distribution are identical to the normal moments $(n = 2, 3, \ldots)$. It is noteworthy that (5) is in fact the first two terms of the Hermite Polynomial form of the Gram-Charlier expansion of a density function with standard normal moments up to order $2n-2$ and the parameter a_{2n} added.

THEOREM 2. For f_{2n} defined by (5) and A_{2n} defined by (6),

$$\sup_x |f_{2n}(x) - \phi(x)| \le \frac{A_{2n} u_{2n}^N}{\sqrt{2\pi}}, \tag{7}$$

where u_{2n}^N is the $2n^{th}$ central moment of the standardized normal distribution, $n = 2, 3, \ldots$.

Proof. Transforming an inequality given by Uspensky [4, p. 594], we obtain $|H_{2n}(x)| \le (\overline{2n-1})e^{x^2/2}$. Then, since $u_{2n}^N = (\overline{2n-1})$ and $|f_{2n}(x) - \phi(x)| \le A_{2n}\phi(x)|H_{2n}(x)|$, we obtain

$$\sup_x | f_{2n}(x) - \phi(x)| \le \frac{A_{2n} u_{2n}^N}{\sqrt{2\pi}}. \qquad \text{q.e.d.}$$

THEOREM 3. Let F_{2n} be the distribution function of f_{2n} as given by (5), and let A_{2n} be defined by (6). Then if Φ is the distribution function of the standard normal distribution,

$$\sup_x |F_{2n}(x) - \Phi(x)| \le A_{2n} \sup_x \{\phi(x)|H_{2n-1}(x)|\}, \tag{8}$$

for $n = 2, 3, \ldots$.

Proof. Integrating (5), we obtain

$$F_{2n}(x) = \Phi(x) + a_{2n}\int_{-\infty}^{x} \phi(t)H_{2n}(t)\,dt \le \Phi(x) + A_{2n}\phi(x)H_{2n-1}(x).$$

https://doi.org/10.4153/CMB-1966-061-5 Published online by Cambridge University Press

Hence

$$\sup_x |F_{2n}(x) - \Phi(x)| \leq A_{2n} \sup_x \{\phi(x) |H_{2n-1}(x)|\} . \qquad \text{q.e.d.}$$

5. The use of the result $H_k'(x) = kH_{k-1}(x)$ and the roots of Hermite Polynomials given by Smith [3, p. 357] enable us to determine A_{2n} defined by (6). Similar analysis permits us to evaluate $\sup_x\{\phi(x)|H_{2n-1}(x)|\}$ in (8). Table 1 presents values for A_{2n} and for the inequalities (7) and (8) for $n = 2, 3, 4$.

TABLE 1

n	A_{2n}	$\sup_x \lvert f_{2n}(x) - \phi(x)\rvert$	$\sup_x \lvert F_{2n}(x) - \Phi(x)\rvert$
2	.1667	.20000	.10
3	.009686	.05812	.03
4	.0003298	.01385	.005

6. If, in practice, we have a population whose standardized density function is unimodal and bell-shaped (as are the density functions f_{2n}), it is commonly assumed that the population is approximately normally distributed. Estimates of the population mean and variance are then obtained and the resulting normal distribution with this mean and variance is assumed to be a satisfactory approximation to the true population distribution.

Utilizing the density function f_{2n} which has been defined, the data in Table 1 illustrate the discrepancy which may exist between the true population distribution function and the normal distribution function when only the first two moments are considered, that is for $n = 2$. To safely assume the population is normally distributed, it appears necessary that additional information about the population must be known. Table 1 indicates that there is a significant decrease in possible deviation from normality when the population central moment of order four is identical to the fourth central normal moment. The decrease in possible deviation when the sixth central moments are identical is even more significant.

https://doi.org/10.4153/CMB-1966-061-5 Published online by Cambridge University Press

It is well known that knowledge of higher moments strengthens an assumption made as to the form of the distribution of a population. However, it is significant that deviations of the magnitude of those in Table 1 have been produced using a polynomial with only one free parameter. Permitting more than one free parameter in this polynomial would only serve to increase the maximum possible deviation shown in Table 1.

REFERENCES

1. M.G. Kendall and A. Stuart, The Advanced Theory of Statistics: Volume 1, Distribution Theory. London, England, Charles Griffin and Company, Ltd. (1958).

2. W.R. McGillivray, Deviations from Normality under Certain Moment Assumptions. Unpublished Master's Thesis, Department of Mathematics, University of Saskatchewan, Saskatoon, Saskatchewan, (1965).

3. E.R. Smith, Zeros of the Hermitian Polynomials. The American Mathematical Monthly, 43 (1936), pages 354-358.

4. J.V. Uspensky, On the Development of Arbitrary Functions of Hermite's and Laguerre's Polynomials. Annals of Mathematics, (2)28, (1927), pages 593-619.

Regina, Saskatchewan

https://doi.org/10.4153/CMB-1966-061-5 Published online by Cambridge University Press