A CHARACTERIZATION OF THE NORMAL DISTRIBUTION

By RYOICHI SHIMIZU

(June 25, 1961)

1. It has been shown by several authors [1], [2], [3] that the independence of sample mean and some quadratic statistics characterizes the normal distribution.

In this note we shall give a more general result.

2. Let $x_1, x_2, \cdots, x_n$ be independent and identically distributed random variables with density $p(x)$ and variance σ^2. Set

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i, \qquad s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 .$$

By a kernel of a symmetric statistic $U(x_1, x_2, \cdots, x_n)$, we mean any statistic $u(x_1, x_2, \cdots, x_m)$ $(n\geqq m)$ such that,

$$U(x_1, x_2, \cdots, x_n)=\frac{(n-m)!}{n!}\sum u(x_{i_1}, x_{i_2}, \cdots, x_{i_m})$$

where the sum is taken over all permutation $(i_1, i_2, \cdots, i_m)$ of m integers such that $1\leqq i_k\leqq n$, $i_k\neq i_l$ if $k\neq l$ $(k, l=1, 2, \cdots, m)$. The symmetric statistic U is refered to as a U-statistic. Note that if $x_1, x_2, \cdots, x_n$ are independent and identically distributed, we have

$$(1) \qquad E(f(x_1, \cdots, x_n)\cdot u(x_1, \cdots, x_m))=E(f(x_1, \cdots, x_n)\cdot U(x_1, \cdots, x_n))$$

for any symmetric statistic $f(x_1, x_2, \cdots, x_n)$.
Now we have,

THEOREM: *Let $h(x_1, x_2, \cdots, x_m)$ be any distribution-free unbiased estimate (d.f.u.e.) of variance σ^2.*

(i) *If $\bar{x}$ any h are independent, $p(x)$ is the normal density.*

(ii) *If, conversely, $p(x)$ is the normal density, $\bar{x}$ and h are uncorrelated.*

(iii) *If, in addition, h is invariant under the group of all translations $x_i\rightarrow x_i+c$, which will be denoted by G, then it is independent of $\bar{x}$ under the assumption of normality.*

The proof of this theorem is based on the following lemmas.

LEMMA 1: *A statistic* $h(x_1, \cdots, x_m)$ *is a* d.f.u.e. *of variance if and only if it is a kernel of* s^2.

PROOF: If part is obvious. If, conversely, h is a d.f.u.e. of variance then the U-statistic corresponding to h is a symmetric d.f.u.e. of variance.

But by [4], the order statistic is complete sufficient for the class of distributions over R^n corresponding to each coordinate having the same distribution function which is any absolutely continuous distribution. Hence by [5] (Theorem 5.1), s^2 is the only symmetric d.f.u.e. of variance and we have $U=s^2$ as was to be proved.

LEMMA 2: *Let* $x_1, x_2, \cdots, x_n$ *be independent and normally distributed, and* $g(x_1, \cdots, x_m)$ *be a statistic.*

(i) *If* g *is invariant under* G, *then it is independent of* $\bar{x}$.

(ii) *If* g *is a kernel of a U-statistic which is invariant under* G, $\bar{x}$ *and* g *are uncorrelated.*

PROOF: (i) g is independent of the location parameter of which $\bar{x}$ is the complete sufficient statistic. (see e.g. [6]).

(ii) In view of (i), $\bar{x}$ and U are independent, hence they are uncorrelated. The desired result follows from (1) with $f(x_1, x_2, \cdots, x_n)=\bar{x}$.

PROOF OF THE THEOREM: (i) We have, from lemma 1 and relation (1),

$$\begin{aligned}(2)\qquad &\int\cdots\int s^2\cdot\exp(it_1\bar{x})\cdot p(x_1)\cdots p(x_n)dx_1\cdots dx_n\\ &=\int\cdots\int U\cdot\exp(it_1\bar{x})\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n\\ &=\frac{(n-m)!}{n!}\sum\int\cdots\int h(x_{i_1},\cdots, x_{i_m})\cdot\exp(it_1\bar{x})\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n\\ &=\int\cdots\int h(x_1, \cdots, x_m)\cdot\exp(it_1\bar{x})\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n\end{aligned}$$

The joint characteristic function of $\bar{x}$ and h is

$$\varphi(t_1, t_2)=\int\cdots\int\exp(it_1\bar{x})\cdot\exp(it_2h)\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n$$

Therefore, if $\bar{x}$ and h are independent, we have

$$\begin{aligned}&\int\cdots\int h\cdot\exp(it_1\bar{x})\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n\\ &=\int\cdots\int h\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n\cdot\int\cdots\int\exp(it_1\bar{x})\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n\end{aligned}$$

But by (2) this is equivalent to

$$\int\cdots\int s^2\cdot\exp(it_1\bar{x})\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n$$
$$=\int\cdots\int s^2\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n\cdot\int\cdots\int\exp(it_1\bar{x})\cdot p(x_1)\cdots p(x_n)\,dx_1\cdots dx_n,$$

which leads, as was shown in [1], to the differential equation

$$-\psi(t)\frac{d^2\psi(t)}{dt^2}+\left(\frac{d\psi(t)}{dt}\right)^2=\sigma^2(\psi(t))^2$$

with

$$\psi(t)=\int e^{itx}p(x)dx, \qquad t=t_1/n$$

the solution of which is the characteristic function of the normal distribution.

(ii) and (iii) are immediate consequences of lemma 2.

Examples.

(i) $$h=\frac{n}{n-1}\sum_{i=1}^{n}p_i(x_i-\bar{x})^2, \qquad \sum_{i=1}^{n}p_i=1$$

is a d.f.u.e. of variance

(ii) (J. N. K. Rao.[3])

$$h=d^2=\Big(\sum_{i=1}^{m}\sum_{j=1}^{n}l_{ij}^2\Big)^{-1}\cdot\sum_{i=1}^{m}(l_{i1}x_1+\cdots+l_{in}x_n)^2$$

where

$$\sum_{i=1}^{m}l_{ij}=0, \qquad \text{for } i=1, 2,\cdots, n,$$

is a d.f.u.e. of variance which is invariant under G.

(iii) $h(x_1, \cdots, x_n)$
$$=s^2+\sum_i a_i^1\varphi_1(x_i)+\sum a_{ij}^2\varphi_2(x_i, x_j)+\cdots+\sum a_{i_1\cdots i_{n-1}}^{n-1}\varphi_{n-1}(x_{i_1}, \cdots, x_{i_{n-1}})$$

where $\sum a_i^1=\sum a_{ij}^2=\cdots=\sum a_{i_1\cdots i_{n-1}}^{n-1}=0$, and $\varphi_1, \varphi_2,\cdots, \varphi_{n-1}$ are symmetric integrable functions,
is a d.f.u.e. of variance which is not necessarily invariant under G. If x_i's are replaced by $x_i-\bar{x}$, then h is invariant under G.

3. The same reasoning applies to the multivariate case, as was shown in [1] and [3].

Let $p(x_1, \cdots, x_r)$ be the density of r-variate $x_1, x_2, \cdots, x_r$, x_{ki} ($k=1, 2, \cdots n$, $i=1, 2, \cdots r$) the kth observation on the ith variate,

$$\bar{x}_i=\frac{1}{n}\sum_{k=1}^{n} x_{ki} \qquad i=1, 2, \cdots, r$$

and

$$s_{ij}=\frac{1}{n-1}\sum_{k=1}^{n}(x_{ki}-\bar{x}_i)(x_{kj}-\bar{x}_j) \qquad i, j=1, 2, \cdots, r.$$

Then, a statistic $h_{ij}(x_{1i}, x_{2i}, \cdots, x_{ni}; x_{1j}, x_{2j}, \cdots, x_{nj})$ is a d.f.u.e. of covariance σ_{ij} of x_i and x_j if and only if it is a kernel of s_{ij} and assuming that the distribution of h_{ij} is independent of the joint distribution of r sample means $(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_r)$, we obtain the equation

$$\frac{\psi_{ij}}{\psi}-\frac{\psi_i\psi_j}{\psi^2}=-\sigma_{ij} \tag{3}$$

where

$$\psi=\psi(t_1\cdots t_r)=\int\cdots\int e^{i(t_1x_1+\cdots+t_rx_r)}\cdot p(x_1\cdots x_r)\,dx_1\cdots dx_r$$

$$\psi_i=\frac{\partial\psi}{\partial t_i}, \qquad \psi_{ij}=\frac{\partial^2\psi}{\partial t_i\partial t_j}$$

If (3) is true for $i, j=1, 2, \cdots, r$, we have a set of partial differential equations which leads to the characteristic function of the multivariate normal distribution.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] E. Lukacs, "A charactrization of the normal distribution," *Ann. Math. Stat.*, Vol. 13 (1942) p p. 91~93.

[2] S. Geisser, "A note on the normal distribution," *Ann. Math. Stat.*, Vol 27 (1956), p p. 858~859.

[3] J. N. K. Rao, "A charactrization of the normal distribution," *Ann. Math. Stat.*, Val. 29 (1958) p p. 914~919.

[4] D. A. S. Fraser, "Completeness of order statistics," *Canadian Jour. Math.*, Vol. 6 (1954) p p. 42~45.

[5] E. L. Lehmann and Scheffé, "Completeness, similar region and unbiased estimation," *Sankhyā*, Vol. 10 (1950), p p. 305~340.

[6] D. Basu, "On statistics independent of a complete sufficient statistics," *Sankhyā*, Vol. 15 (1955), pp. 377~380.