The Annals of Statistics
1976, Vol. 4, No. 3, 668–672

WHEN ARE THE MEAN AND THE STUDENTIZED DIFFERENCES INDEPENDENT?

BY LENNART BONDESSON

University of Lund and Institute Mittag-Leffler

Let $X_1, \cdots, X_n$ be i.i.d. rv's. Let further $\bar{X} = \sum X_i/n$, $S^2 = \sum (X_i - \bar{X})^2$, and $U = ((X_1 - \bar{X})/S, \cdots, (X_n - \bar{X})/S)$. If the variables X_i are normally distributed or distributed as linearly transformed Gamma variables, $\bar{X}$ and U are independent. In this paper we show that also the converse must hold.

1. Introduction. Let $X_1, \cdots, X_n$ be n independent identically distributed random variables. It is well known that the independence between $\bar{X} = \sum X_i/n$ and $Y = (X_1 - \bar{X}, \cdots, X_n - \bar{X})$, $n \geqq 2$, is a characteristic property for the normal law. This is just a special case of the Darmois–Skitovich theorem (see e.g. Lukacs and Laha [6], page 75). If X_i are strictly positive nondegenerate variables, then $\bar{X}$ and $Z = (X_1/\bar{X}, \cdots, X_n/\bar{X})$, $n \geqq 2$, are independent if and only if the variables are gamma-distributed. A somewhat more general theorem was shown by Lukacs [4]. Let now $S^2 = \sum (X_i - \bar{X})^2$ and $U = ((X_1 - \bar{X})/S, \cdots, (X_n - \bar{X})/S)$. Of course, U is only well-defined if the distribution function of X_i is continuous. While $(Y\ (Z))$ is the maximal invariant under change of location (scale), U is the maximal invariant under change of both location and scale. In 1970 Kelker and Matthes [3] proved that, for $n \geqq 4$, $(\bar{X}, S^2)$ and U can be independent only if the variables follow the normal law. It is an interesting problem to find out for what distributions just $\bar{X}$ is independent of U. One type of distribution for which this is true is the normal one but also other distributions are possible. For, if X_i are gamma-distributed, then $\bar{X}$ is independent of Z and therefore also independent of U since U is a function of Z. As a linear transformation $X_i' = aX_i + b$ cannot change the independence relation, this must hold also for linearly transformed gamma-distributed variables. The purpose of this paper is to show that these mentioned cases are the only possible ones. In fact, this conclusion is proved to be true under the (partly) weaker assumption that $\bar{X}$ has constant regression on U. Our result answers a question asked on page 461 in the recent monograph by Kagan, Linnik and Rao [2] and is (partly) an extension of the Kagan–Linnik–Rao theorem and Theorem 6.2.1 given in this book on pages 155 and 197, respectively. The result is also closely related to some other results in estimation theory obtained by the author [1].

2. The theorem.

THEOREM. *Let $X_1, \cdots, X_n$ be independent identically distributed random variables*

Received June 1973; revised November 1973.

AMS 1970 *subject classification.* 62E10.

Key words and phrases. Constant regression, Cauchy's functional equation, characteristic function, analytic function.

with continuous distribution function. We assume that $n \geqq 6$. *If* $E[\bar{X} \mid U] = constant$ a.s., *then* X_i *are either normally distributed or distributed as linearly transformed gamma variables.*

PROOF. Let g be any bounded measurable function of U. From the regression assumption it follows that

$$E[(\sum X_i - n\mu)g(U)] = 0 . \tag{1}$$

Here $\mu = E[X_i]$. We first set

$$g(U) = (|X_1 - X_2|)^{it_1}(|X_3 - X_4|)^{it_2}(|X_5 - X_6|)^{it_3} ,$$

where $i = (-1)^{\frac{1}{2}}$ and t_1, t_2 and t_3 are real numbers having sum equal to zero. Since the right side is invariant under change of location and scale of the variables, it is really a function of U. Observe also that $(|X_1 - X_2|)^{it}$ is a well-defined (complex-valued) random variable for all t since $|X_1 - X_2| > 0$ a.s. Putting

$$\psi(t) = E[(X_1 + X_2 - 2\mu)(|X_1 - X_2|)^{it}]$$

and

$$\phi(t) = E[(|X_1 - X_2|)^{it}] ,$$

we can after some manipulations write equation (1) as

$$\psi(t_1)\phi(t_2)\phi(t_3) + \psi(t_2)\phi(t_1)\phi(t_3) + \psi(t_3)\phi(t_1)\phi(t_2) = 0 , \tag{2}$$

whenever $t_1 + t_2 + t_3 = 0$. Notice that ψ and ϕ are continuous functions satisfying $\psi(0) = 0$, $\phi(0) = 1$, $\psi(t) = \overline{\psi(-t)}$, and $\phi(t) = \overline{\phi(-t)}$. Since $E[|X_i|] < \infty$, we have $E[|X_1 - X_2|] < \infty$. Therefore, and as $|X_1 - X_2| > 0$ a.s., $\phi(z) = E[(|X_1 - X_2|)^{iz}]$ is a function that is analytic when $-1 < \operatorname{Im} z < 0$ and continuous when $-1 \leqq \operatorname{Im} z \leqq 0$. Hence $\phi(t)$ cannot vanish identically on any real interval (see e.g. Titchmarsh [7], page 157). Let $S \subset R$ be the set of all nonzeroes of ϕ. Thus, S is a dense subset of R. Using this, we are going to show that the solution of equation (2) is given by

$$\psi(t) = i\tau t\phi(t) , \qquad t \in R ,$$

where τ is a real constant. For $t \in S$, we set $h(t) = \psi(t)/\phi(t)$. We can then for t_1, t_2, and $-(t_1 + t_2)$ all belonging to S write equation (2) as

$$h(t_1) + h(t_2) = -h(-t_1 - t_2) .$$

Putting $t_2 = 0$, we see that $h(t_1) = -h(-t_1)$ and hence

$$h(t_1) + h(t_2) = h(t_1 + t_2) , \qquad t_1, t_2, (t_1 + t_2) \in S . \tag{3}$$

Let $(-\ell, \ell)$ be the largest open interval containing the origin on which $\phi(t)$ vanishes nowhere. Equation (3) is Cauchy's functional equation and therefore (observe that $h(0) = 0$)

$$h(t) = i\tau t , \qquad -\ell < t < \ell .$$

The constant τ is real since $h(t) = \overline{h(-t)}$. Now we set for $t \in S$

$$g(t) = h(t) - i\tau t .$$

The function g is zero on $(-\ell, \ell)$. We get

$$g(t_1) + g(t_2) = g(t_1 + t_2)\,, \qquad t_1, t_2, (t_1 + t_2) \in S\,.$$

Letting t_1 be fixed and $t_2 \in (-\ell, \ell)$, we conclude that $g(t)$ is constant for all $t \in (t_1 - \ell, t_1 + \ell) \cap S$. As $g(0) = 0$ and S is a dense subset of R, it is easily realized that $g(t) = 0$ for all $t \in S$. Hence

$$\psi(t) = i\tau t\phi(t)\,, \qquad t \in S\,.$$

Since both sides are continuous functions and S is a dense subset of R, equality must hold for all t. This is the result wanted. Equivalently,

$$E[(X_1 + X_2 - 2\mu)(|X_1 - X_2|)^{it}] = i\tau t \cdot E[(|X_1 - X_2|)^{it}]\,. \tag{4}$$

However, this relation cannot be directly used to produce the final result. Instead, we have to utilize the regression assumption once more. Let s be any fixed real number. Then, for a dense set of points t in R, $-(t + s)$ is not a zero of ϕ. For such points t we now set

$$g(U) = L^{it}M^{is}(|X_4 - X_5|)^{-i(t+s)}\,.$$

Here

$$L = |a_1X_1 + a_2X_2 + a_3X_3| + \tfrac{1}{2}(a_1X_1 + a_2X_2 + a_3X_3)$$

and

$$M = |a_1X_1 + a_2X_2 + a_3X_3|\,,$$

where a_1, a_2 and a_3 are arbitrary real numbers not all zero but satisfying $a_1 + a_2 + a_3 = 0$. From equation (1) we obtain

$$\begin{aligned} E[(X_1 + X_2 + X_3 - 3\mu)L^{it}M^{is}] \cdot E[(|X_4 - X_5|)^{-i(t+s)}] \\ + E[(X_4 + X_5 - 2\mu)(|X_4 - X_5|)^{-i(t+s)}] \cdot E[L^{it}M^{is}] = 0\,, \end{aligned}$$

so using (4) (with t replaced by $-(t + s)$ and X_1, X_2 replaced by X_4, X_5), we get after division by $E[(|X_4 - X_5|)^{-i(t+s)}]$

$$E[(X_1 + X_2 + X_3 - 3\mu)L^{it}M^{is}] = i\tau(t + s)E[L^{it}M^{is}]\,. \tag{5}$$

Observing that both sides are continuous functions of t, we easily conclude that (5) holds for all t and s. Let $V = \log L$ and $W = \log M$. (Notice that L and M are strictly positive a.s.) Hence

$$E[(X_1 + X_2 + X_3 - 3\mu)\exp\{itV + isW\}] = i\tau(t + s)E[\exp\{itV + isW\}]\,. \tag{6}$$

Equivalently,

$$\begin{aligned} E[(X_1 + X_2 + X_3 - 3\mu)\exp\{itV + isW\}] \\ = \tau \cdot E\left[\left(\frac{\partial}{\partial V} + \frac{\partial}{\partial W}\right)\exp\{itV + isW\}\right]. \end{aligned}$$

Since t and s are arbitrary, it should be possible to replace $\exp\{itV + isW\}$ by any sufficiently smooth function of (V, W). Let $f(V, W)$ be a function belonging to C_0^∞. We define the Fourier transform $\hat{f}$ by

$$f(V, W) = \iint \hat{f}(t, s) \exp\{itV + isW\}\, dt\, ds .$$

Multiplying both sides of (6) by $\hat{f}(t, s)$, integrating over (t, s), and then letting the operations integration and taking expected value change order, we get

$$E[(X_1 + X_2 + X_3 - 3\mu)f(V, W)] = \tau \cdot E\Big[\Big(\frac{\partial}{\partial V} + \frac{\partial}{\partial W}\Big)f(V, W)\Big] . \tag{7}$$

Let now

$$h(V, W) = \exp\{2it(e^V - e^W)\} = \exp\{it(a_1X_1 + a_2X_2 + a_3X_3)\} ,$$

where the real number t is arbitrary and of course not the same one as before. Then $h(V, W)$ has finite expected value and so has

$$\Big(\frac{\partial}{\partial V} + \frac{\partial}{\partial W}\Big) h(V, W) = it(a_1X_1 + a_2X_2 + a_3X_3) \exp\{it(a_1X_1 + a_2X_2 + a_3X_3)\} .$$

We set

$$f_N(V, W) = h(V, W) \cdot \chi(V/N, W/N) .$$

Here $\chi \in C_0^\infty$, has support contained in $|V| + |W| \leqq 2$, and is 1 on $|V| + |W| \leqq 1$. Since (7) holds for $f = f_N$, we easily find by letting N tend to infinity that it also holds for $f = h$. Setting $t_j = ta_j$, $j = 1, 2, 3$, we therefore have

$$\begin{aligned} &E[(X_1 + X_2 + X_3 - 3\mu) \exp\{it_1X_1 + it_2X_2 + it_3X_3\}] \\ &\quad = \tau \cdot E[(it_1X_1 + it_2X_2 + it_3X_3) \exp\{it_1X_1 + it_2X_2 + it_3X_3\}] , \qquad t_1 + t_2 + t_3 = 0 . \end{aligned}$$

Equivalently, if t_1, t_2 and t_3 are small enough,

$$\begin{aligned} &\frac{E[(X_1 - \mu - i\tau t_1X_1) \exp\{it_1X_1\}]}{E[\exp\{it_1X_1\}]} + \frac{E[(X_2 - \mu - i\tau t_2X_2) \exp\{it_2X_2\}]}{E[\exp\{it_2X_2\}]} \\ &\quad + \frac{E[(X_3 - \mu - i\tau t_3X_3) \exp\{it_3X_3\}]}{E[\exp\{it_3X_3\}]} = 0 , \qquad \text{when} \quad t_1 + t_2 + t_3 = 0 . \end{aligned}$$

This is again Cauchy's functional equation and hence we get for all t small enough

$$E[(X_1 - \mu - i\tau tX_1) \exp\{itX_1\}] = i\rho t \cdot E[\exp\{itX_1\}] , \tag{8}$$

where ρ is a real constant. Let $\phi(t) = E[\exp\{itX_1\}]$. Then after some manipulations (8) transforms into (ϕ' is the derivative of ϕ)

$$\phi'(t)/\phi(t) = (i\mu - \rho t)/(1 - i\tau t) .$$

Integrating up, we obtain

$$\begin{aligned} \phi(t) &= \exp\{i\mu t - \rho t^2/2\} , && \text{if} \quad \tau = 0 , \\ &= \Big(\frac{1}{1 - i\tau t}\Big)^{(\rho/\tau^2 + \mu/\tau)} \cdot \exp\{-i\rho t/\tau\} , && \text{if} \quad \tau \neq 0 . \end{aligned}$$

Since the right-hand sides are analytic functions and ϕ is a characteristic function, equality must hold for all real t (see e.g. Lukacs [5], page 132). Observe that when $\tau \neq 0$, necessarily $\rho/\tau^2 + \mu/\tau \geqq 0$, for otherwise ϕ cannot be a characteristic function since $\phi(t)$ tends to infinity when t does so. The same argument shows that $\rho \geqq 0$ when $\tau = 0$. We then recognize the functions on the right side as the characteristic function for a normal distribution and the characteristic function for a linearly transformed gamma variable, respectively. The proof is complete.

REMARK 1. A thorough analysis of the proof of the theorem shows that the condition $E[\bar{X} \mid U] =$ constant can be replaced by the somewhat weaker condition $E[\bar{X} \mid U'] =$ constant, where

$$U' = ((X_1 - X_2)/|X_5 - X_6|, (X_1 - X_3)/|X_5 - X_6|, (X_3 - X_4)/|X_5 - X_6|) .$$

Observe that U' has dimension 3 while U is $(n-2)$-dimensional. This result leads to the suspicion that it should be sufficient to assume $n \geqq 5$ in the theorem. Other heuristic arguments indicate that $n \geqq 4$ should be the necessary and sufficient condition. It is easy to construct a counterexample for $n = 2$. For $n = 3$ this seems to be much harder.

REMARK 2. The assumption that X_i are identically distributed can be removed. However, the conclusion will then be slightly changed.

REMARK 3. If $\bar{X}$ and U are independent, then it is not automatically true that $\bar{X}$ has constant regression on U since $E[\bar{X}]$ could fail to exist. Therefore the theorem does not completely answer the question given in the title.

REMARK 4. It is also possible to characterize the normal distribution by the property that S^2 has constant regression on U. (Compare Theorem 4.1 in [1].) The precise formulation and the proof of this result will be given in another paper.

REFERENCES

[1] BONDESSON, L. (1973). Characterizations of the normal and the gamma distributions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **26** 335–344.

[2] KAGAN, A. M., LINNIK, YU. V. and RAO, C. R. (1973). *Characterization Problems in Mathematical Statistics.* Wiley, New York.

[3] KELKER, D. and MATTHES, T. K. (1970). A sufficient statistics characterization of the normal distribution. *Ann. Math. Statist.* **41** 1086–1090.

[4] LUKACS, E. (1955). A characterization of the gamma distribution. *Ann. Math. Statist.* **26** 319–324.

[5] LUKACS, E. (1960). *Characteristic Functions.* Griffin, London.

[6] LUKACS, E. and LAHA, R. G. (1964). *Applications of Characteristic Functions.* Griffin, London.

[7] TITCHMARSH, E. C. (1939). *The Theory of Functions*, 2nd ed. Oxford Univ. Press.

STATISTISKA CENTRALBYRÅN
A/SK
FACK
S-10250 STOCKHOLM 27
SWEDEN