

Nguyen Van Ho

On the continuity of invariant statistics

Aplikace matematiky, Vol. 23 (1978), No. 5, 381--388

Persistent URL: <http://dml.cz/dmlcz/103763>

Terms of use:

© Institute of Mathematics AS CR, 1978

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

ON THE CONTINUITY OF INVARIANT STATISTICS

NGUYỄN VĂN HỒ

(Received February 21, 1977)

INTRODUCTION

The continuity of some estimates of the location and the location vector was proved in [3], Theorem 1, Hodges-Lehmann (1963), and in [4], Theorem 6.2.2, Puri-Sen (1971). All these estimates are translation invariant but the theorems do not characterize the interesting property of the estimates.

The aim of this paper is to establish theorems on the continuity of translation as well as scale invariant statistics in general, from which the above mentioned results in [3] and [4] follow.

Let us discuss the proof of the first assertion of Theorem 1 in [3] by Hodges-Lehmann (the first assertion of Theorem 6.2.2 in [4] by Puri-Sen is dealt with similarly). The authors have concluded $P(\Delta^{**} = c) = 0$ on the basis of the Fubini theorem and of the fact that each line L of the family of all lines parallel to the direction of $\mathbf{e} = (0, \dots, 0, 1, \dots, 1)$ with m zeros and n units intersects the set $S = \{\Delta^{**}(X_1, \dots, X_m, Y_1, \dots, Y_n) = c\}$ in a single point which has probability zero by the assumed continuity of the cdf H of $(X_1, \dots, X_m, Y_1, \dots, Y_n)$. Assume $X_1, \dots, X_m, Y_1, \dots, Y_n$ are mutually independent. The assumption makes it possible to apply the Fubini theorem, as the probability measure P is the product of its projections onto the axes $X_1, \dots, X_m, Y_1, \dots, Y_n$. Then the Fubini theorem could lead to the result $P(\Delta^{**} = c) = 0$ if L were parallel to some of the axes. However, L is not so. One may try to change axes by rotating them in order to make L parallel to some new axis. But after the rotation the probability measure P is not the product of its projections onto the new axes (cf. Example 1 below) and the result $P(\Delta^{**} = c) = 0$ cannot follow from the Fubini theorem (cf. Example 2 below).

These proofs would remain true if the following postulate were true: "the measurability of S and the fact that each L parallel to \mathbf{e} intersects S in a single point imply that S consists of a finite or denumerable family of measurable sets $\{S_i\}$ such that for each S_i there is an axis such that each line parallel to the axis intersects S_i in at most a single point".

On the basis of this postulate Theorem 1 in Section IV would remain true if the words “absolutely continuous” were replaced by “continuous” and $T(X, Y)$ were of the form $T = (T_1(X_1, Y_1), \dots, T_k(X_k, Y_k))$ with the assumption that X^1, \dots, X^N are mutually independent.

Similarly, on the basis of the following postulate: “the measurability of S and the fact that each line L lying in the hyperplanes $X_1 = x_1^0, \dots, X_m = x_m^0$, where x_1^0, \dots, x_m^0 are arbitrary real numbers, and containing the point $(x_1^0, \dots, x_m^0, 0, \dots, 0)$ intersects S in a single point imply the assertion on S as the above postulate”, Theorem 2 in Section IV would remain true if the words “absolutely continuous” were replaced by “continuous”, T were of the form $T = (T_1(X_1, Y_1), \dots, T_k(X_k, Y_k))$ with the assumption that X^1, \dots, X^N are mutually independent and the condition $T_i(X, Y) = 0$ iff $Y = 0^{(nk)}$, $1 \leq i \leq k$, were replaced by $T_i(X_i, Y_i) = 0$ iff $Y_i = 0^{(n)}$, $1 \leq i \leq k$.

Although the proof of the first assertion of Theorem 1 in [3] is dubious its second assertion on the absolute continuity proved in the same way remains still true, since the Lebesgue measure \mathcal{L} in R^N according to axes X_1, \dots, X_N is always the product of its projections onto the new axes ξ_1, \dots, ξ_N obtained by a rotation or by any transformation of axes X_1, \dots, X_N with the Jacobian $J = J_1(\xi_1) \dots J_N(\xi_N)$, and so the Fubini theorem leads to $\mathcal{L}(A^{**} \in A) = 0$ provided $\mathcal{L}(A) = 0$.

In view of Lemma 2(iii) in Section II, the second assertion of Theorem 6.2.2 in [4] must be modified as follows: “the condition (6.2.8) and the absolute continuity of $F_\alpha(x)$ for each $\alpha = 1, \dots, n$ imply the absolute continuity of the cdf’s of each component of $\hat{\theta}_n$ ”.

Example 1. Let P be a Gaussian measure with the density $f(x, y) = (2\pi ab)^{-1} \cdot \exp\{-\frac{1}{2}(x^2/a^2 + y^2/b^2)\}$, $a \neq b > 0$. Thus $P = P_x \times P_y$, where P_x, P_y are projections of P onto the axes x and y . Let (ξ, η) be the new axes obtained by rotating (x, y) by an angle $-\pi/4$. Then $P \neq P_\xi \times P_\eta$.

Example 2. Suppose X and Y are dependent, $X = Y$ a.s., and X is uniformly distributed on $(0, 1)$. The probability measure P generated by (X, Y) has a continuous cdf

$$\begin{aligned}
 F(x, y) &= 0 && \text{if } x \leq 0 \text{ or } y \leq 0, \\
 &x && \text{if } 0 < x < y \text{ and } 0 < x < 1, \\
 &y && \text{if } 0 < y < x \text{ and } 0 < y < 1, \\
 &1 && \text{if } x \geq 1 \text{ and } y \geq 1.
 \end{aligned}$$

Let $S = \{x - y = 0\}$. Each line parallel to y intersects S in a single point of probability zero by the continuity of F but $P(S) = P(X = Y) = 1$.

Lemmas in Section II emphasize and summarize the continuity relations between the joint cdf of a random vector and its marginal cdf’s.

II. LEMMAS

Let (Ω, \mathcal{A}, P) be a probability space and let ξ_1, \dots, ξ_k be random variables (a.s. finite) defined on it. Let $F_1(x), \dots, F_k(x)$ be cdf's of ξ_1, \dots, ξ_k and let $F(x_1, \dots, x_k)$ be the cdf of (ξ_1, \dots, ξ_k) .

Lemma 1. *The following assertions are equivalent:*

- (i) $F_1(x), \dots, F_k(x)$ are continuous,
- (ii) $F_1(x), \dots, F_k(x)$ are uniformly continuous,
- (iii) $F(x_1, \dots, x_k)$ is continuous,
- (iv) $F(x_1, \dots, x_k)$ is uniformly continuous.

Proof.

(i) \Leftrightarrow (ii) and (iv) \Rightarrow (iii) are evident.

(ii) \Rightarrow (iv): For $x' = (x'_1, \dots, x'_k)$ and $x'' = (x''_1, \dots, x''_k)$ let $y' = (y'_1, \dots, y'_k)$, where $y'_i = \min(x'_i, x''_i)$, $1 \leq i \leq k$, and $y'' = (y''_1, \dots, y''_k)$, where $y''_i = \max(x'_i, x''_i)$, $1 \leq i \leq k$. The conclusion follows from

$$\begin{aligned}
 |F(x'') - F(x')| &\leq F(y'') - F(y') = \\
 &= P\left\{ \bigcap_{i=1}^k [\xi_i \leq y'_i] \setminus \bigcap_{j=1}^k [\xi_j \leq y'_j] \right\} = \\
 &= P\left(\bigcap_i [\xi_i \leq y'_i] \cap \left(\bigcup_j [\xi_j > y'_j] \right) \right) \leq \\
 &\leq P\left\{ \bigcup_{j=1}^k \left([\xi_j > y'_j] \cap [\xi_j \leq y'_j] \right) \right\} \leq \\
 &\leq \sum_{j=1}^k P\{y'_j < \xi_j \leq y''_j\} = \sum_{j=1}^k |F_j(x''_j) - F_j(x'_j)|.
 \end{aligned}$$

(iii) \Rightarrow (i): Suppose there is a discontinuous cdf, say $F_1(x)$. Let x_0 be a discontinuity point of $F_1(x)$. Denote $A = \xi_1^{-1}(x_0) = \{\omega : \xi_1(\omega) = x_0\} \in \mathcal{A}$. Then

$$P(\xi_1 = x_0) = F_1(x_0) - F_1(x_0 - 0) = P(A) = p > 0.$$

Since F is continuous,

$$\begin{aligned}
 0 &= F(x_0, x_2, \dots, x_k) - F(x_0 - 0, x_2, \dots, x_k) = \\
 &= P(\xi_1 = x_0, \xi_2 \leq x_2, \dots, \xi_k \leq x_k)
 \end{aligned}$$

for any real x_2, \dots, x_k . Therefore either

$$P\{\omega \in A : \xi_2(\omega) \leq x_2, \dots, \xi_k(\omega) \leq x_k\} = 0$$

for any real x_2, \dots, x_k ,

$$\text{or } \sum_{i=2}^k P(\xi_i = \infty) \geq P(A) = p > 0,$$

which contradicts the a.s. finiteness of the random variables ξ_2, \dots, ξ_k . Q.E.D.

Corollary 1. *Lemma 1 remains true when ξ_1, \dots, ξ_k are random vectors of arbitrary dimensions n_1, \dots, n_k , respectively.*

Remark 1. The equivalence introduced in Lemma 1 is a special property of cdf's. For bounded and nondecreasing multivariate functions the equivalence does not hold in general. Let us consider

$$\begin{aligned} G(x, y) &= 0 \quad \text{if } x \leq 0 \quad \text{or } y \leq 0, \\ &(1 - e^{-y})e^{yx} \quad \text{if } 0 < x \leq e^{-y} \quad \text{and } y > 0, \\ &1 - e^{-y} \quad \text{if } x > e^{-y} \quad \text{and } y > 0. \end{aligned}$$

Clearly, $G(-\infty, -\infty) = 0$, $G(+\infty, +\infty) = 1$ and $G(x, y)$ is nondecreasing and continuous in x and y , but $G(x, y)$ is not a cdf, as

$$\Delta G = G(x_2, y_2) - G(x_1, y_2) - G(x_2, y_1) + G(x_1, y_1) < 0$$

for

$$0 < x_1 < e^{-y_1} < x_2 \quad \text{and} \quad 0 < y_1 < -\ln(x_1) < y_2.$$

One has

$$\begin{aligned} G_1(x) = G(x, \infty) &= 0 \quad \text{if } x \leq 0, \\ &1 \quad \text{if } x > 0, \end{aligned}$$

i.e., $G_1(x)$ is not continuous and moreover, $G(x, y)$ is not uniformly continuous, as $G(x, 1 - \ln(x)) - G(0, 1 - \ln(x)) > 1 - e^{-1} > \frac{1}{2}$ for any x , $0 < x < 1$.

Lemma 2.

- (i) *If $F(x_1, \dots, x_k)$ is absolutely continuous, then $F_1(x), \dots, F_k(x)$ are so as well.*
- (ii) *If $F_1(x), \dots, F_k(x)$ are absolutely continuous and ξ_1, \dots, ξ_k are mutually independent, then $F(x_1, \dots, x_k)$ is so as well.*
- (iii) *If $F_1(x), \dots, F_k(x)$ are absolutely continuous, then $F(x_1, \dots, x_k)$ is not generally so but it is continuous even if ξ_1, \dots, ξ_k are mutually dependent.*

Proof.

(i) and (ii) are well-known.

(iii) The continuity of F follows from Lemma 1. In order to prove that F is generally not absolutely continuous it is sufficient to form a counterexample. Let $k = 2$. Let $\xi_1 = \xi_2$ a.s. and let each of them be uniformly distributed on $(0, 1)$. From

Example 2 one has clearly $\partial^2 F / (\partial x_1 \partial x_2) = 0$ a.e. with respect to the Lebesgue measure \mathcal{L} in R^2 , and $\iint_s dF(x_1, x_2) = 1$ where $s = \{(x, x), 0 < x < 1\}$ with $\mathcal{L}(s) = 0$. It means that $F(x_1, x_2)$ is not absolutely continuous (but is continuous), while $F_1(x)$ and $F_2(x)$ are so. Q.E.D.

Another example which is not so special is the following

Example 3. Let ξ_1 be uniformly distributed on $(0, 1)$ and $P(\xi_2 \in (0, 1)) = 1$. Let $\xi_2 = \xi_1$ for $0 < \xi_1 \leq \frac{1}{2}$, and let ξ_2 be uniformly distributed on $(\frac{1}{2}, 1)$ for $\frac{1}{2} < \xi_1 \leq 1$, i.e. the conditional cdf of ξ_2 for a given ξ_1 is of the form

$$F_2(x_2 | x_1) = 0 \text{ if } 0 < x_1 \leq \frac{1}{2} \text{ and } x_2 < x_1, \\ 2x_2 - 1 \text{ if } \frac{1}{2} < x_1 < 1 \text{ and } \frac{1}{2} < x_2 < 1, \\ 1 \text{ if } 0 < x_1 \leq \frac{1}{2} \text{ and } x_2 \geq x_1, \text{ or } \frac{1}{2} < x_1 < 1 \text{ and } x_2 \geq 1.$$

Therefore

$$F(x_1, x_2) = 0 \text{ if } x_1 \leq 0 \text{ or } x_2 \leq 0, \\ x_1 \text{ if } 0 < x_1 \leq \frac{1}{2} \text{ and } x_2 \geq x_1, \text{ or } \frac{1}{2} < x_1 < 1 \text{ and } x_2 \geq 1, \\ x_2 \text{ if } 0 < x_2 \leq \frac{1}{2} \text{ and } x_1 \geq x_2, \text{ or } \frac{1}{2} < x_2 < 1 \text{ and } x_1 \geq 1, \\ \frac{1}{2} + 2(x_1 - \frac{1}{2})(x_2 - \frac{1}{2}) \text{ if } \frac{1}{2} < x_1 < 1 \text{ and } \frac{1}{2} < x_2 < 1, \\ 1 \text{ if } x_1 \geq 1 \text{ and } x_2 \geq 1.$$

Thus we obtain

$$F_2(x_2) = 0 \text{ if } x_2 \leq 0, \\ x_2 \text{ if } 0 < x_2 < 1, \\ 1 \text{ if } x_2 \geq 1,$$

i.e. ξ_2 is also uniformly distributed on $(0, 1)$ as ξ_1 , while

$$\frac{\partial^2 F}{\partial x_1 \partial x_2} = \begin{cases} 2 & \text{if } \frac{1}{2} < x_1 < 1 \text{ and } \frac{1}{2} < x_1 < 1, \\ 0 & \text{otherwise,} \end{cases} \\ \iint_{R^2} \frac{\partial^2 F}{\partial x_1 \partial x_2} dx_1 dx_2 = \frac{1}{2} \text{ and } \iint_s dF(x_1, x_2) = \frac{1}{2},$$

where $s = \{(x, x), 0 < x < \frac{1}{2}\}$ with $\mathcal{L}(s) = 0$, i.e. F is not absolutely continuous (but is continuous).

Corollary 2. Lemma 2 can be generalized for the case of ξ_1, \dots, ξ_k being random vectors.

III. NOTATION AND DEFINITIONS

Let $X^j = (X_1^j, \dots, X_k^j)$, $1 \leq j \leq N$ be k -dimensional random vectors. Let $X = (X^1, \dots, X^m)$, $Y = (X^{m+1}, \dots, X^{m+n})$, $Z = (X, Y)$, where $m \geq 0$, $n \geq 1$, $m + n = N$. For the case $m = 0$, $Y = Z$, let x, y, z, \dots be representations of X, Y, Z, \dots , respectively. Denote $Z_i = (X_i, Y_i)$, $1 \leq i \leq k$, where

$$X_i = (X_i^1, \dots, X_i^m), \quad Y_i = (X_i^{m+1}, \dots, X_i^{m+n}).$$

Let $T = T(Z) = T(X, Y) = (T_1(X, Y), \dots, T_k(X, Y))$ be a k -dimensional statistic. For $a = (a_1, \dots, a_s) \in R^s$, $b = (b_1, \dots, b_s) \in R^s$, let $a^{(p)}$ stand for $(a, \dots, a) \in R^{ps}$ and $a * b = (a_1 b_1, \dots, a_s b_s)$.

Definition 1. The statistic $T(X, Y)$ is said to be translation invariant iff

$$(1) \quad T(X, Y + b^{(n)}) = T(X, Y) + b \quad \text{for all } b = (b_1, \dots, b_k) \in R^k.$$

Definition 2. The statistic $T(X, Y)$ is said to be scale invariant of the first type or of the second type iff

$$(2) \quad T(X, tY) = t T(X, Y) \quad \text{for all } t \in R^1, \text{ or}$$

$$(3) \quad T(X, a^{(n)} * Y) = a * T(X, Y) \quad \text{for all } a = (a_1, \dots, a_k) \in R^k,$$

respectively.

Definition 3. The statistic T is said to be linear invariant of the first type or of the second type iff it is translation invariant as well as scale invariant of the first type or of the second type, i.e. iff

$$(4) \quad T(X, tY + b^{(n)}) = t T(X, Y) + b \quad \text{for all } t \in R^1 \text{ and all } b \in R^k, \text{ or}$$

$$(5) \quad T(X, [a^{(n)} * Y] + b^{(n)}) = [a * T(X, Y)] + b \quad \text{for all } a, b \in R^k,$$

respectively.

Remark 2. In some cases (if necessary) Definitions 2 and 3 may be modified in the following way: (2), (3), (4), and (5) are replaced by

$$(2') \quad T(tX, tY) = t T(X, Y) \quad \text{for all } t \in R^1,$$

$$(3') \quad T(a^{(m)} * X, a^{(n)} * Y) = a * T(X, Y) \quad \text{for all } a \in R^k,$$

$$(4') \quad T(tX, tY + b^{(n)}) = t T(X, Y) + b \quad \text{for all } t \in R^1 \text{ and all } a \in R^k,$$

$$(5') \quad T(a^{(m)} * X, [a^{(n)} * Y] + b^{(n)}) = [a * T(X, Y)] + b \quad \text{for all } a, b \in R^k.$$

Note that Definitions 2 and 3 of the second type are stronger than those of the first type. Statistics satisfying one or all the Definitions would be formulated in estimating location parameters, see e.g. Hodges-Lehmann (1963) [3], Puri-Sen (1971) [4], Bickel-Lehmann (1975) [1], [2], ...

Remark 3. If $T(X, Y) = (T_1(X_1, Y_1), \dots, T_k(X_k, Y_k))$, T is scale or linear invariant of the second type iff it is scale or linear invariant of the first type, respectively, as obtained easily from the Definitions.

IV. THEOREMS

Let us keep the notation of Section III.

Theorem 1. *Let $T(X, Y) = (T_1(X, Y), \dots, T_k(X, Y))$ be translation invariant. Then the cdf's of T_1, \dots, T_k are absolutely continuous provided the cdf $F(x, y)$ of (X, Y) is so.*

Proof. For $i, 1 \leq i \leq k$ and $A \subset R^1$ with $\mathcal{L}(A) = 0$, put $R_i^A = \{(x, y) \in R^{Nk} : T_i(x, y) \in A\}$. R_i^A is measurable. Consider in R^{Nk} the family of all lines parallel to the direction of the vector $I = (0^{(mk)}, 1^{(nk)})$:

$$\mathcal{L}_I = \{L(x, y) = \{(x, y + t^{(nk)}), t \in R^1\}, (x, y) \in R^{Nk}\}.$$

For each $L = L(x^0, y^0) \in \mathcal{L}_I$, one has

$$[(x, y) \in L(x^0, y^0) \cap R_i^A] \Leftrightarrow [(x, y) = (x^0, y^0 + [c - T_i(x^0, y^0)]^{(nk)}), c \in A]$$

This means that each $L \in \mathcal{L}_I$ intersects R_i^A in a set equivalent to A , therefore the set has the Lebesgue measure zero. Then $\mathcal{L}(R_i^A) = 0$, by the Fubini theorem applied to the Lebesgue measure \mathcal{L} in R^{Nk} . It follows from the absolute continuity of $F(x, y)$ that $P(T_i \in A) = \int_{R_i^A} \dots \int dF(x, y) = 0$ for all $A \subset R^1$ with $\mathcal{L}(A) = 0$, i.e. the cdf of T_i is absolutely continuous. Q.E.D.

Remark 4. In view of Lemma 2 (iii) one cannot obtain the absolute continuity of the joint cdf of $T = (T_1, \dots, T_k)$ under the assumptions of Theorem 1. The same argument explains why the second assertion of Theorem 6.2.2 in [4] mentioned in Section I is dubious.

Corollary 3. *The result of Theorem 1 holds for a statistic T which is linear invariant of the first type or of the second type.*

Theorem 2. *Let $T = (T_1(X, Y), \dots, T_k(X, Y))$ be scale invariant of the first type or of the second type and such that $T_i(x, y) = 0$ iff $y = 0^{(nk)}$, $1 \leq i \leq k$. Let $F(x, y)$ be absolutely continuous. Then the cdf's of T_1, \dots, T_k are so as well.*

Proof. Theorem 2 is proved similarly as Theorem 1, with \mathcal{L}_I replaced by

$$\mathcal{L}'_I = \{L(x, y) = \{(x, ty) : t \in R^1\}, (x, y) \in R^{Nk}, y \neq 0^{(nk)}\}.$$

The intersection of $L(x^0, y^0)$ with $R^A_i, A \subset R^1, \mathcal{L}(A) = 0$ is $\{(x^0, [c/(T(x^0, y^0))] y^0), c \in A\}$, which has the Lebesgue measure zero. The rest of the proof follows as in Theorem 1. Q.E.D.

Remark 5. Theorem 2 remains true for T which is scale invariant in the sense of the definition modified as in Remark 2 and such that $T_i(x, y) = 0$ iff $(x, y) = 0^{(Nk)}$. In order to prove it let us put $\mathcal{L}'_I = \{L(x, y) = \{t(x, y), t \in R^1\}, (x, y) \in R^{Nk} \setminus \{0^{(Nk)}\}\}$. Then the intersection of $L(x^0, y^0) \in \mathcal{L}'_I$ with $R^A_i, A \subset R^1, \mathcal{L}(A) = 0$ is $\{[c/(T_i(x^0, y^0))] (x^0, y^0), c \in A\}$ of the Lebesgue measure zero.

References

- [1] P. J. Bickel, E. L. Lehmann: Descriptive statistics for nonparametric models. I. Introduction. Ann. Statist. (1975), V. 3, N. 5, 1038—1044.
- [2] P. J. Bickel, E. L. Lehmann: Descriptive statistics for nonparametric models. II. Location. Ann. Statist. (1975), V. 3, N. 5, 1045—1069.
- [3] J. L. Hodges Jr., E. L. Lehmann: Estimates of location based on rank tests. Ann. Math. Stat. (1963), V. 34, 598—611.
- [4] M. L. Puri, P. K. Sen: Nonparametric methods in multivariate analysis. New York, 1971.

Souhrn

SPOJITOST INVARIANTNÍCH STATISTIK

NGUYỄN VĂN HỒ

Cílem článku je dokázat obecné věty o absolutní spojitosti statistik invariantních vzhledem ke změně polohy a měřítka, z nichž plynou příbuzné výsledky Hodgese-Lehmana a Puri-Sena. Vyšetřuje se také vztah mezi spojitostí sdružené distribuční funkce náhodného vektoru a jeho marginálních distribučních funkcí.

Author's address: Dr. Nguyễn Văn Hồ, Polytechnic Institute of Hanoi, Khoa Toán lý, Đại Học Bách Khoa, Hanoi, VSR.