UNBIASED DISTINGUISHING OF TRANSLATION FAMILIES AND INTEGRABILITY WITH RESPECT TO A CONVOLUTION OF MEASURES

H. S. KONIJN and R. SACKSTEDER

(Received 11 July 1966)

1. Introduction

Let M_1 and M_2 be two sets of probability measures defined on R^n. A measurable R^l valued function h ($l \geqq 1$) is said to *distinguish* M_1 *from* M_2 *unbiasedly* if there are numbers or vectors I_1 and I_2 ($I_1 \neq I_2$) such that $\int_{R^n} h(x)m(dx) = I_i$ if m is in M_i ($i = 1, 2$). Here we shall be concerned with the case where M_1 and M_2 are translation families, in that all of the elements of M_i are translates of a single measure m_i. This means that if, for any t in R^n, m_i^t is the measure defined by $m_i^t(E) = m_i(E-t)$, where $E-t = \{x-t : x \in E\}$, then $M_i = \{m_i^t : t \in T_i\}$, where T_i is a subset of R^n. If M_1 and M_2 are of this type, we will investigate the conditions under which there does not exist a function to distinguishing M_1 from M_2 unbiasedly. A case of special interest arises if $m_2(E) = m_1(BE) = m_1$ ($\{Bx : x \in E\}$), with B a non-degenerate $n \times n$ matrix, and particularly a nonzero multiple (*scale parameter*) of the identity matrix, cf. [1], [2]. For simplicity, take $l = 1$.

Much of our investigation requires an answer to the following question: To find condition on m_1 and m_2 such that m_1- and m_2-integrability of any function h implies integrability of h with respect to m, the *convolution* of m_1 and m_2.

2. Analytical results

Let $m = m_1^* m_2$ denote the convolution of m_1 with m_2. Standard arguments show that a measurable function $f : R^n \to R^1$ is m-integrable if and only if the function $g : R^{2n} \to R^1$ defined by $g(x, y) = f(x+y)$ is integrable with respect to the product measure $m_1 \times m_2$. The support of the measure m_i will be denoted by S_i, the support of m by S.

PROPOSITION 1. *Suppose that* E_1 *and* E_2 *satisfy*

$$m_1(E_2) = m_2(E_1) = 0,\ S_1 \subset T_2 \cup E_2,\ S_2 \subset T_1 \cup E_1,$$

and that $h : R^n \to R^1$ *is* m*-integrable. Then* h *does not distinguish* M_1 *from* M_2 *unbiasedly.*

https://doi.org/10.1017/S1446788700005243 Published online by Cambridge University Press

PROOF. $I_1 = \int_{R^n} h(y) m_1^x(dy) = \int_{T_2} h(x+y) m_1(dy)$

for every x in T_1, hence

$$I_1 = \int_{T_1} \left\{ \int_{T_2} h(x+y) m_1(dy) \right\} m_2(dx).$$

Similarly,

$$I_2 = \int_{T_2} \left\{ \int_{T_1} h(x+y) m_2(dx) \right\} m_1(dy).$$

Let $g(z) = h(z)$ if

$$z \in T_1 + T_2 = \{x+y : x \in T_1, y \in T_2\}$$

and $g(z) = 0$ otherwise. Then

$$I_1 = \int_{R^n} \left\{ \int_{R^n} g(x+y) m_1(dx) \right\} m_2(dy) \text{ and } I_2 = \int_{R^n} \left\{ \int_{R^n} g(x+y) m_2(dy) \right\} m_1(dx).$$

But since $|g| \leqq |h|$, g is $m_1 \times m_2$ integrable and Fubini's theorem implies that the iterated integrals for I_1 and I_2 are equal. This proves Proposition 1.

Propositions 2 and 3 show how the hypothesis of Proposition 1 that h is m-integrable can be satisfied. They are concerned with measures m_1 and m_2 which satisfy the following readily recognizable conditions: There is a sigma-finite measure μ on a sigma field $\mathscr{F}$ of subsets of R^n (with $\mathscr{F}$ including $\mathscr{B}$, the Borel sets of R^n) and such that:

(i) there is a constant a such that

(α) $\mu(E-s) \leqq a\mu(E)$ for $s \in S_2$, $E \in \mathscr{B}$, $E \subset S_1 + s$,

(β) $\mu(E+s) \leqq a\mu(E)$ for $s \in S_2$, $E \in \mathscr{B}$, $E \subset S_1$, and for $s \in S$, $E \in \mathscr{B}$, $-E \subset S_1 \cap S_2$,

(γ) $\mu(-E) \leqq a\mu(E)$ for $E \in \mathscr{B}$, $-E \subset S_1 \cap S_2$ and $E \subset S_1 \cap S_2$;

(ii) for $i = 1$ and 2 there are numerical constants a_i, a'_i, b_i and c, such that, if r_i is the density (= Radon-Nikodym derivative) of m_i with respect to μ, and x any element of S

(α) $r_i(y) \leqq a_i r_i(b_i x)$ for $y \in S_2$ (when $i = 2$) or $y \in (x - S_2) \cap S_1$ (when $i = 1$), and $c^2 \leqq |\frac{1}{2}x|^2 \leqq \frac{1}{2}x \cdot y$,

(β) $r_i(b_i x) \leqq a'_i r_i(x)$.

PROPOSITION 2. *Assume* (i) *and* (ii) *with* $c = 0$. *Then if* h *is* m_i*-integrable for* $i = 1, 2$, h *is* m*-integrable* ($m = m_1^* m_2$).

PROOF. We first show that (iα) and (iβ) imply that m is absolutely continuous with respect to μ. If E is any Borel set, the part outside S has m-measure 0, and if E is any Borel set in S, then by (iα)

$$m(E) = \int_{R^n} m_1(E-y) m_2(dy) \leqq a \int_{R^n} \left\{ \int_E r_1(x-y) \mu(dx) \right\} m_2(dy),$$

https://doi.org/10.1017/S1446788700005243 Published online by Cambridge University Press

and so this holds for any Borel set of R^n. Since $r_1 \geqq 0$ and since, by taking $E = R^n$, it follows from (iβ) that the inner integral is at most a, the iterated integral is finite; therefore we can change the order of integration, obtaining

$$m(E) \leqq a \int_E \left\{ \int_{R^n} r_1(x-y) m_2(dy) \right\} \mu(dx) < \infty,$$

which shows the absolute continuity of m with respect to μ.

Therefore, there exists a density, r, of m with respect to μ, which satisfies (when $S_3 = S_2 \cap \{y \in S_2 : x-y \in S_1\}$)

$$r(x) \leqq a \int_{S_3} r_1(x-y) r_2(y) \mu(dy). \tag{2.1}$$

Let $x \in S$ be fixed and let $U = \{y \in S_3 : |\frac{1}{2}x|^2 \leqq \frac{1}{2}x \cdot y\}$; its complement is $V = \{y \in S_3 : |\frac{1}{2}x|^2 < \frac{1}{2}x \cdot (x-y)\}$.

If $y \in V$, $y' = x-y$ belongs to $(x-S_2) \cap S_1$, satisfies $|\frac{1}{2}x|^2 < \frac{1}{2}x \cdot y'$ and so by (iiα) $r_1(x-y) \leqq a_1 r_1(b_1 x)$, hence

$$\int_V r_1(x-y) r_2(y) \mu(dy) \leqq a_1 \int_V r_1(b_1 x) r_2(y) \mu(dy) \leqq a_1 r_1(b_1 x).$$

If $y \in U$, $r_2(y) \leqq a_2 r_2(b_2 x)$ and

$$\begin{aligned}\int_U r_1(x-y) r_2(y) \mu(dy) &\leqq a_2 \int_U r_1(x-y) r_2(b_2 x) \mu(dy) \\ &\leqq a_2 r_2(b_2 x) \int_{S_2} r_1(x-y) \mu(dy) \leqq a_2 a^2 r_2(b_2 x),\end{aligned}$$

where the last part follows by an application of (iβ) and (iγ).

Therefore, it has been shown that, for $x \in S$,

$$r(x) \leqq a a_1 r_1(b_1 x) + a^3 a_2 r_2(b_2 x).$$

Finally, by (iiβ), this shows that, for $x \in S$,

$$r(x) \leqq a_0 r_1(x) + a_0' r_2(x) \tag{2.2}$$

with $a_0 = a a_1 a_1'$, $a_0' = a^3 a_2 a_2'$. This implies the conclusion of Proposition 2.

PROPOSITION 3. *Assume* (i) *and* (ii) (*with* $c > 0$) *and suppose that* r_1 *or* r_2 *is bounded and that* r_1 *is lower semi-continuous. Let* T_2 *be an open set containing* S_2 *and suppose that* h *is* m_1^t*-integrable for every* t *in* T_2, *and is* m_1*- and* m_2*-integrable. Suppose also that* $\mu(E-s) \leqq a\mu(E)$ *holds also for all* s *in* T_2, $E \in \mathscr{B}$, $E \subset S_1 + s$. *Then* h *is* m*-integrable.*

PROOF. We showed in the proof of Proposition 2 that m has a density satisfying (2.1), and that, for $|x| \geqq 2c$, (2.2) holds. It remains to consider $r(x)$ on the set of $x \in S$ for which $|x| < 2c$; call its closure S_0.

By hypothesis either there is a constant k_1 such that $r_1(x) \leqq k_1$ for all x, whence, by (2.1), $r(x) \leqq ak_1$ for all x; or there is a constant k_2 such that $r_2(y) \leqq k_2$ for all y, whence by (2.1), $r(x) \leqq ak_2 \int_{S_3} r_1(x-y)\mu(dy) \leqq a^3 k_2$,

https://doi.org/10.1017/S1446788700005243 Published online by Cambridge University Press

where the last part follows from (iβ) and (iγ) as already noted in the proof of Proposition 2. So in either case there is a constant k_0 such $r(x) \leqq k_0$ for all x.

Now S_0 is closed; and if x is in S_0, $x = x_1+x_2$, where x_i is in S_i. There is an $\varepsilon = \varepsilon(x_2) > 0$ such that if $|y-x_2| < \varepsilon$, y is in T_2, and there is a y_1 such that $|y_1-x_1| < \varepsilon$ and $r_1(y_1) > 0$. Let $y_2 = x_1+x_2-y_1$ so that $x = y_1+y_2$, $|y_2-x_2| < \varepsilon$ and so y_2 is in T_2. Let δ and η be such that $r_1(y) \geqq \eta > 0$ if $|y-y_1| < \delta$. Then $r_1(y-y_2) \geqq \eta$ if $|y-x| = |y-y_1-y_2| < \delta$. Since x was an arbitrary point of S_0, S_0 can be covered by a finite number, k, of disks, the j^{th} disk having radius δ^j and center x^j, such that, if $|y-x^j| < \delta^j$,

$$r_1(y-y^j) \geqq \eta^j > 0,$$

where y^j is in T_2. Then, writing $\eta_0 = \eta^1+\cdots+\eta^k$, we have that for y in S_0

$$r(y) \leqq k_0 \leqq (k_0/\eta_0) \sum_{j=1}^{k} r_1(y-y^j).$$

But the last hypothesis of the Proposition implies that for any $E \in \mathscr{B}$, $E \subset S_1+y^j$,

$$\int_E r_1(x-y^j)\mu(dx) \leqq am_1^t(E) \text{ with } t = y^j \ (j = 1, \cdots, k).$$

So since h was to be m_1^t-integrable for $t \in T_2$, $\int_{R^n} h(y)r_1(y-y^j)\mu(dy)$ exists and is finite for $j = 1, \cdots, k$. Hence $\int_{R^n} h(y)r(y)\mu(dy)$ exists and is finite, which proves the Proposition.

3. An example

This section is devoted to an example (in R^1) of two sets of measures $M_1 = \{m_1^j : j \in J\}$ and $M_2 = \{m_2^j : j \in J\}$ which can be distinguished unbiasedly, even though the union of the S_{1j} coincides with the union of the S_{2j} (where S_{1j} is the support of m_1^j, S_{2j} the support of m_2^j).

Let J denote the integers. The measure m_i is defined as follows: Let $\sum_{k=1}^{\infty} a_k$ be a series of positive terms whose sum is 1, for example take $a_k = (\frac{1}{2})^k$. Then m_1^j assigns the mass a_k at the integer 2^k+j and m_2^j assigns the mass a_k at $3 \cdot 2^k+j$. A function h will be constructed which distinguishes M_1 and M_2 unbiasedly, by requiring that for every j in J it satisfy:

$$(3.1)_j \qquad \int_{-\infty}^{+\infty} |h(x)|m_1^j(dx) = \sum_{k=1}^{\infty} |h(2^k+j)|a_k < \infty,$$

$$(3.2)_j \qquad \int_{-\infty}^{+\infty} h(x)m_1^j(dx) = \sum_{k=1}^{\infty} h(2^k+j)a_k = 1,$$

$$(3.3)_j \qquad \int_{-\infty}^{+\infty} |h(x)| m_2^j(dx) = \sum_{k=1}^{\infty} |h(3\cdot 2^k + j)| a_k < \infty,$$

$$(3.4)_j \qquad \int_{-\infty}^{+\infty} h(x) m_2^j(dx) = \sum_{k=1}^{\infty} h(3\cdot 2^k + j) a_k = 0.$$

It can be shown that if $i \neq j$, the sets $S_{1i} \cap S_{1j}$ and $S_{2i} \cap S_{2j}$ contain at most one element, and that, for all i and j, $S_{1j} \cap S_{2j}$ contains at most two elements. Now h can easily be defined by induction on $|j|$ to satisfy $(3.1)_j$, $(3.2)_j$, $(3.3)_j$, and $(3.4)_j$. The idea of the proof is that h can be defined first on S_{10}, then on S_{20}, S_{11}, S_{21}, $S_{1,-1}$, $S_{2,-1}$, S_{12}, and so on. At each step, where h is to be defined on some S_{1j} (or S_{2j}), h will have been defined on only finitely many points of S_{1j} (or S_{2j}) in previous steps, hence it will be possible to satisfy $(3.1)_j$ and $(3.2)_j$ (or $(3.3)_j$ and $(3.4)_j$), since $a_k > 0$ for every k.

This example shows that one cannot take for granted the validity of the interchange in order of integration in Proposition 1. m_1^0 and m_2^0 have their supports in J and are absolutely continuous with respect to the measure which assigns the mass one to each integer and so satisfies (i). However, for any c, b_1 and b_2 there are always points $x \geqq 2c$ in the support of $m_1^0 * m_2^0$ for which, for $i = 1$ or 2, $r_i(b_i x) = 0$ but $r_i(y) \neq 0$ for some y satisfying the condition of (iiα), or $r_i(b_i x) \neq 0$ but $r_i(x) = 0$. Note also that m_1^0 and m_2^0 differ by a scale parameter.

4. Application to statistical estimation problems

Let $M = \{m_k : k \in K\}$ be a set of probability measures on R^n. If $\tau : K \to R^1$ is any function, a statistic $h : R^n \to R^1$ is said to be an *unbiased estimator* of τ in the family of probability distributions

$$\Omega = \{m_k^y : k \in K, y \in T_k\},$$

where $T_k \subset R^n$, if for every k in K,

$$\tau(k) = \int_{R^n} h(x) m_k^y(dx)$$

whenever y is in T_k. We are interested in showing that unbiased estimators do not exist in certain cases.

Case 1: *h is semi-bounded.* In some applications it is natural to require that any statistic used to estimate τ be non-negative or, more generally, semi-bounded. For example, if the range of τ is the set of possible values of a scale parameter and all of these values are non-negative, it would be reasonable to require that h be non-negative. Proposition 1 then shows that there can be no semi-bounded, unbiased estimator of τ, unless τ is constant.

Case 2. *Ω dominates $\Omega * \Omega$.* Ω is said to dominate $\Omega * \Omega$ if, whenever m_1 and m_2 are elements of Ω, their convolution $m_1 * m_2$ satisfies

https://doi.org/10.1017/S1446788700005243 Published online by Cambridge University Press

$$(m_1 * m_2)(E) \leqq \sum_{i=1}^{N} n_i(E),$$

where E is any Borel set and $n_1, \cdots, n_N$ are (not necessarily distinct) elements of Ω, which do not depend on E. For example, Ω might be a stable class. If h is an unbiased estimator of τ, h must be integrable with respect to $n_1, \cdots, n_N$, hence h is $m_1 * m_2$ integrable. Proposition 1 then shows that there is no unbiased estimator of τ unless τ is constant.

CASE 3. *Application of Propositions* 2 *and* 3. Condition (i) is automatically satisfied if μ is Lebesgue measure or if μ is the measure which assigns mass one to the set of integral lattice points of R^n. Condition (ii) will often be satisfied for some b_i's less than 1. This is clearly so for the discrete distributions generally encountered in the statistical literature.

Consider now distributions which have densities with respect to Lebesgue measure; in particular, for $i = 1, 2$, let r_i be continuous and positive on a disk of R^n of m_i-measure 1. Then (iiα) is satisfied if r_i becomes nonincreasing for large $|x|$, the case occurring in all examples considered in practice. Moreover, the approach to zero of the density as $|x|$ increases is in many cases such that (iiβ) is satisfied as well.

Thus Propositions 1, 2, 3, show that unbiased estimates of τ cannot exist in many of the frequently encountered cases. For the estimation of a scale parameter it is natural to require that for $i = 1$ and 2 the existence in the finite sense of $\int_{R^n} h(x)r_i(x)\mu(dx)$ implies that of $\int_{R^n} h(x)r_i(b_i x)\mu(dx)$ for some open range of b_i; in that case condition (iiβ) may be dispensed with.

5. Application to a testing problem

Let $M_1 = \{m_1^t : t \in T_1\}$ and $M_2 = \{m_2^t : t \in T_2\}$ be two sets of probability measures on R^n, as in the introduction. A measurable function $h : R^n \to R^1$ is called a (strictly) unbiased test of M_1 against M_2 of level α if $\int_{R^n} h(x)m_1^t(dx) \leqq \alpha$ for t in T_1 and $\int_{R^n} h(x)m_2^t(dx) > \alpha$ for t in T_2. There is a Proposition analogous to Proposition 1 which gives conditions in which tests of level α do not exist. The proof is essentially the same as that of Proposition 1.

References

[1] H. S. Konijn, 'On a theorem of Halmos concerning unbiased estimation of moments', *Australian Math. Soc.* **4** (1964), 229—232.

[2] H. S. Konijn, 'Statistical reproduction of orderings and translation subfamilies', *Ann. Math. Statist.* **37** (1966), 196—202.

Tel-Aviv University
and
The City University of New York