Ann. Inst. Statist. Math.
32 (1980), Part A, 241-245

NONPARAMETRIC ESTIMATION OF MATUSITA'S MEASURE OF AFFINITY BETWEEN ABSOLUTELY CONTINUOUS DISTRIBUTIONS

IBRAHIM A. AHMAD*

(Received Dec. 14, 1977; revised July 26, 1979)

Abstract

Let F and G be two distribution functions defined on the same probability space which are absolutely continuous with respect to the Lebesgue measure with probability densities f and g, respectively. Matusita [3] defines a measure of the closeness, affinity, between F and G as: $\rho=\rho(F, G)=\int [f(x)g(x)]^{1/2}dx$. Based on two independent samples from F and G we propose to estimate ρ by $\hat{\rho}=\int [\hat{f}(x)\hat{g}(x)]^{1/2}dx$, where $\hat{f}(x)$ and $\hat{g}(x)$ are taken to be the kernel estimates of $f(x)$ and $g(x)$, respectively, as given by Parzen [5].

In this note sufficient conditions are given such that
(i) $\mathrm{E}(\hat{\rho}-\rho)^2\to 0$ as $x\to\infty$ and (ii) $\hat{\rho}\to\rho$ with probability one, as $n\to\infty$.

1. Introduction

Let F and G be two distribution functions (d.f.'s) defined on the same probability space. Assume that F and G admit densities f and g, respectively, with respect to a measure μ. Matusita [3] defines a of the closeness between F and G as:

$$\rho=\rho(F, G)=\int [f(x)g(x)]^{1/2}d\mu(x) . \tag{1.1}$$

Matusita [3] studies certain decision problems based on estimates of ρ when μ is the counting measure, while Matusita [4] gives an extensive account of the mathematical properties of ρ. In several other papers, Matusita applied ρ to various inferential problems such as classification,

* Research supported in part by the National Research Council of Canada and by McMaster University Science and Engineering Research Board.

The author is presently with the Department of Mathematical Sciences, Memphis State University, Memphis, Tennessee 38152.

independence, among others. He also furthered his mathematical studies of ρ in other publication, see Ahmad [2] for a list of references. Ahmad and Van Belle [1] introduced another measure of affinity when f and g are square integrable, viz.,

$$\lambda=\lambda(F, G)=2\int f(x)g(x)d\mu(x)\Big/\left\{\int f^2(x)d\mu(x)+\int g^2(x)d\mu(x)\right\} . \tag{1.2}$$

When μ is the counting measure, Ahmad and Van Belle [1] propose certain test statistics based on estimates of ρ and λ, while when μ is the Lebesgue measure, Ahmad [2] proposes a nonparametric estimate of λ using the kernel estimates of $f(x)$ and $g(x)$ and present its large sample properties and its applications in hypothesis testing. Note that λ is defined only when f and g are square integrable. This restriction motivates the development of inference about ρ. In this note a large class of nonparametric estimates of ρ is shown to be, under certain conditions, consistent in the second mean, and under a bit stronger conditions it is strongly consistent.

Let $X_1,\cdots,X_n$ and $Y_1,\cdots,Y_n$ be two independent random samples from F and G respectively. Assume that F and G admit probability density functions (p.d.f.'s) f and g, respectively, thus

$$\rho=\rho(F, G)=\int [f(x)g(x)]^{1/2}dx . \tag{1.3}$$

Furthermore, let k be a known p.d.f. satisfying the following conditions:

$$\sup_x k(u)<\infty \quad \text{and} \quad |u|k(u)\to 0 \qquad \text{as } |u|\to\infty , \tag{1.4}$$

and let $\{a_n\}$ be a sequence of nonnegative real numbers such that $a_n \to 0$ as $n\to\infty$. The kernel estimates of $f(x)$ and $g(x)$ are given by:

$$\hat{f}(x)=a_n^{-1}\int k[(x-u)/a_n]dF_n(u)=(na_n)^{-1}\sum_{i=1}^{n} k[(x-X_i)/a_n] , \tag{1.5}$$

and

$$\hat{g}(x)=a_n^{-1}\int k[(x-u)/a_n]dG_n(u)=(na_n)^{-1}\sum_{i=1}^{n} k[(x-Y_i)/a_n] . \tag{1.6}$$

The estimates (1.5) and (1.6) are due to Rosenblatt [6] and Parzen [5], and are called the kernel estimates. Thus a nonparametric estimate of ρ may be given by:

$$\hat{\rho}=\int [\hat{f}(x)\hat{g}(x)]^{1/2}dx . \tag{1.7}$$

It should be mentioned here that the results of this note are readily extendable to affinity of several distribution, cf, Matusita [4]..

Throughout this note we shall assume that the set of discontinuity points of f and g are, respectively, null sets.

2. Main results

THEOREM 2.1. *If* $na_n \to \infty$, *then*

$$\mathrm{E}\,(\hat{\rho}-\rho)^2 \to 0 \qquad \text{as } n\to\infty . \tag{2.1}$$

PROOF. Note that

$$\begin{aligned}\mathrm{E}\,(\hat{\rho}-\rho)^2 \leqq 2\,\mathrm{E}\left\{\int (\hat{f}(x))^{1/2}[(\hat{g}(x))^{1/2}-(g(x))^{1/2}]dx\right\}^2 \\ +2\,\mathrm{E}\left\{\int (g(x))^{1/2}[(\hat{f}(x))^{1/2}-(f(x))^{1/2}]dx\right\}^2 .\end{aligned} \tag{2.2}$$

It suffices to show that one term in the above right-hand side converges to 0 as $n\to\infty$. The second term may be shown to converge similarly. But using Fubini's theorem,

$$\begin{aligned}&\mathrm{E}\left\{\int (\hat{f}(x))^{1/2}[(\hat{g}(x))^{1/2}-(g(x))^{1/2}]dx\right\}^2 \\ &=\int\int \mathrm{E}\,[\hat{f}(x)\hat{f}(y)]^{1/2}\,\mathrm{E}\,[(\hat{g}(x))^{1/2}-(g(x))^{1/2}][(\hat{g}(y))^{1/2}-(g(y))^{1/2}]dxdy \\ &\leqq\int\int [\mathrm{E}\,\hat{f}(x)]^{1/2}[\mathrm{E}\,\hat{f}(y)]^{1/2}\,\mathrm{E}^{1/2}\,[(\hat{g}(x))^{1/2}-(g(x))^{1/2}]^2 \\ &\qquad\cdot\,\mathrm{E}^{1/2}\,[(\hat{g}(y))^{1/2}-(g(y))^{1/2}]^2dxdy \\ &=\left\{\int [\mathrm{E}\,\hat{f}(x)]^{1/2}\,\mathrm{E}^{1/2}\,[(\hat{g}(x))]^{1/2}dx\right\}^2 \\ &\leqq\int [\mathrm{E}\,\hat{f}(x)]dx\cdot\int \mathrm{E}\,[(\hat{g}(x))^{1/2}-(g(x))^{1/2}]^2dx \\ &=\int \mathrm{E}\,[(\hat{g}(x))^{1/2}-(g(x))^{1/2}]^2dx \\ &\leqq\int \mathrm{E}\,|\hat{g}(x)-g(x)|dx ,\end{aligned} \tag{2.3}$$

where the first equality follows since $\hat{f}$ and $\hat{g}$ are independent, the first and the second inequalities follow from Schwarz's inequality while the last inequality follows since $\int \mathrm{E}\,\hat{f}(x)dx=1$ for all $n\geqq 1$ and since for all $a, b\geqq 0$, $(a-b)^2\leqq|a^2-b^2|$. Since $\mathrm{E}\,|\hat{g}(x)-g(x)|\leqq \mathrm{E}^{1/2}\,[\hat{g}(x)-g(x)]^2\leqq \{\mathrm{Var}\,\hat{g}(x)+[\mathrm{E}\,\hat{g}(x)-g(x)]^2\}^{1/2}$, it follows from Theorems 1A and 2A of Parzen [5] that for each continuity point of g, $\mathrm{E}\,|\hat{g}(x)-g(x)|\to 0$ and $n\to\infty$, but $\mathrm{E}\,|\hat{g}(x)-g(x)|\leqq \mathrm{E}\,\hat{g}(x)+g(x)$ which is integrable for all $n\geqq 1$ and converges to $2g(x)$ as $n\to\infty$, again an integrable function. Hence the extended Lebesgue dominated convergence theorem, Royden [7] p. 89, applies and we have $\int E|\hat{g}(x)-g(x)|dx \to 0$ as $n\to\infty$. Similarly we have

$$\text{(2.4)}\qquad E\left\{\int (g(x))^{1/2}[(\hat{f}(x))^{1/2}-(f(x))^{1/2}]dx\right\}^2 \leqq \int E|\hat{f}(x)-f(x)|dx \to 0 \quad \text{as } n\to\infty .$$

THEOREM 2.2. *If k is a continuous functions of bounded variation, if for any $\gamma>0$, $\sum_{n=1}^{\infty}\exp(-\gamma n a_n^2)<\infty$, if $\inf_x f(x)>0$ and $\inf_x g(x)>0$, and if $\int f(x)g(x)dx<\infty$, then*

$$\text{(2.5)}\qquad \hat{\rho}\to\rho \quad \textit{with probability one as } n\to\infty .$$

PROOF. Note that,

$$\begin{aligned}\text{(2.6)}\qquad |\hat{\rho}-\rho| &= \left|\int [\hat{f}(x)\hat{g}(x)]^{1/2}dx-\int [f(x)g(x)]^{1/2}dx\right| \\ &= \left|\int [|\hat{f}(x)\hat{g}(x)-f(x)g(x)|]\cdot\{[\hat{f}(x)\hat{g}(x)]^{1/2}+[f(x)g(x)]^{1/2}\}^{-1}dx\right| \\ &\leqq [\inf f(x)\inf g(x)]^{-1/2}\int |\hat{f}(x)\hat{g}(x)-f(x)g(x)|dx \\ &\leqq C\left\{\int \hat{f}(x)|\hat{g}(x)-g(x)|dx+\int g(x)|\hat{f}(x)-f(x)|dx\right\} \\ &\leqq C\left\{\sup_x|\hat{g}(x)-E\,\hat{g}(x)|+\int \hat{f}(x)|E\,\hat{g}(x)-g(x)|dx\right. \\ &\qquad \left.+\sup_x|\hat{f}(x)-E\,\hat{f}(x)|+\int g(x)|E\,\hat{f}(x)-f(x)|dx\right\} .\end{aligned}$$

Under the condition of the theorem it follows as in Lemma 2.2 (iii) of Ahmad [2] that the first and third terms above converge to 0 with probability one as $n\to\infty$. The second term is majorized by

$$\sup_x|\hat{f}(x)-E\,\hat{f}(x)|\int |E\,\hat{g}(x)-g(x)|dx+\int E\,\hat{f}(x)|E\,\hat{g}(x)-g(x)|dx ,$$

where the first term converge to 0 with probability one as $n\to\infty$, since as seen in the proof of Theorem 2.1, $\int |E\,\hat{f}(x)-f(x)|dx\to 0$, as $n\to\infty$. Also since $E\,\hat{f}(x)|E\,\hat{g}(x)-g(x)|\to 0$ as $n\to\infty$ at all continuity points x of f and g and since $E\,\hat{f}(x)|E\,\hat{g}(x)-g(x)|\leqq E\,\hat{f}(x)|E\,\hat{g}(x)+g(x)|$ which is integrable and converges to $2f(x)g(x)$ again and integrable function, thus the extended Lebesgue dominated convergence theorem applies and $\int E\,\hat{f}(x)|E\,\hat{g}(x)-g(x)|dx\to 0$ as $n\to\infty$. The fourth term in the right-hand side of (2.6) also converge to 0 as $n \to \infty$.

Remark 2.1. An interesting and open question would be to discuss the limiting distribution of $\hat{\rho}$.

McMaster University

References

[1] Ahmad, I. A. and Van Belle, G. (1974). Measuring affinity of distributions. *Reliability and Biometry, Statistical Analysis of Life Testing,* (eds. F. Proschan and R. J. Serfling), SIAM, Philadelphia, 651-668.

[2] Ahmad, I. A. (1980). Nonparametric estimation of an affinity measure between two absolutely continuous distributions with hypothesis testing applications, *Ann. Inst. Statist. Math.*, **32**, 223-240.

[3] Matusita, K. (1955). Decision rules based on the distance for the problem of fit, two samples and estimation, *Ann. Math. Statist.*, **26**, 631-640.

[4] Matusita, K. (1967). On the notion of affinity of several distributions and some of its applications, *Ann. Inst. Statist. Math.*, **19**, 181-192.

[5] Parzen, E. (1962). On estimation of a probability density function and mode, *Ann. Math. Statist.*, **33**, 1065-1076.

[6] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function, *Ann. Math. Statist.*, **27**, 832-837.

[7] Royden, H. L. (1968). *Real Analysis,* (Second Edition), Macmillan, New York.