Ann. Inst. Statist. Math.
32 (1980), Part A, 121-123

A NOTE ON MINIMUM DISTANCE ESTIMATES

CONSTANTINE A. DROSSOS AND ANDREAS N. PHILIPPOU*

(Received June 25, 1979)

Abstract

It is shown that minimum distance estimates enjoy the invariance property of maximum likelihood estimates.

The minimum distance method has been developed by Wolfowitz [4], [5], [6], Matusita [2], [3], and others, in order to provide strongly consistent estimates and optimal decision rules. This inference method includes as special cases all the commonly used ones, i.e. maximum likelihood, chi-square, Kolmogorov-Smirnov, least squares, and moments (see, e.g. Blyth [1]). In this note, it is our objective to show that the invariance property of maximum likelihood estimates, established by Zehna [7], is also enjoyed by minimum distance estimates. The proof is simple, and the result may be of interest to some mathematical statisticians.

Denote by $\mathcal{F}$ the class of distribution functions and let $d(\cdot,\cdot)$ be any non-negative "distance" function defined on $\mathcal{F}\times\mathcal{F}$, so that $d(F, G)$ measures "how far apart" the distribution functions F and G are. It is often desirable that $d(\cdot,\cdot)$ be a proper distance, as is for example the Kolmogorov-Smirnov distance $d(F, G)=\sup_x |F(x)-G(x)|$, but it is not always required. In the sequel, we shall use the word distance for reasons of consistency with relevant statistical practice, even though "distance" appears to be more appropriate.

DEFINITION. Let $X_1,\cdots,X_n$ be a random sample from a distribution function $F(x;\theta)$, where θ is an unknown parameter belonging to $\Theta\subseteq R^k$ $(k\geqq 1)$, and denote by $F_n(x)$ the empirical distribution function based on $X_1,\cdots,X_n$. If there exists a $\hat{\theta}=\hat{\theta}(X_1,\cdots,X_n)$ in Θ, such that

$$(1)\qquad d[F(x;\hat{\theta}), F_n(x)]=\inf\{d[F(x;\theta), F_n(x)];\ \theta\in\Theta\}\ ,$$

it is called a minimum distance estimate of θ.

* Supported under Grant No. 18-2773 of the American University of Beirut.
Key words: Estimates, minimum distance, maximum likelihood, invariance.

Let $\lambda=u(\theta)$, where $u(\cdot)$ is an arbitrary transformation from Θ onto $\Lambda\subseteq R^m$ $(1\leqq m\leqq k)$, and set

$$(2) \qquad \Theta_\lambda=\{\theta\in\Theta:\ u(\theta)=\lambda\} \qquad (\lambda\in\Lambda)\,.$$

In accordance with Zehna [7], we define the distance function induced by $u(\cdot)$ as follows,

$$(3) \qquad d_u[F(x;\lambda), F_n(x)]=\inf\{d[F(x;\theta), F_n(x)];\ \theta\in\Theta_\lambda\} \qquad (\lambda\in\Lambda)\,,$$

and we show that the following theorem holds.

THEOREM. *Let $\hat\theta=\hat\theta(X_1,\cdots,X_n)$ be a minimum distance estimate of θ based on a random sample $X_1,\cdots,X_n$ from the distribution function $F(x;\theta)$, and set $\hat\lambda=u(\hat\theta)$. Then $\hat\lambda$ is a minimum distance estimate of λ.*

PROOF. It suffices to show that $\hat\lambda\in\Lambda$, and

$$d_u[F(x;\hat\lambda), F_n(x)]\leqq d_u[F(x;\lambda), F_n(x)] \qquad (\lambda\in\Lambda)\,.$$

Since $\hat\theta\in\Theta$, it follows that $\hat\lambda=u(\hat\theta)\in\Lambda$, and this implies that $\hat\theta\in\Theta_{\hat\lambda}$, because of (2). Therefore,

$$\begin{aligned}
&d_u[F(x;\hat\lambda), F_n(x)]\\
&\quad=\inf\{d[F(x;\theta), F_n(x)];\ \theta\in\Theta_{\hat\lambda}\}\,, \qquad \text{by (3)}\,,\\
&\quad=d[F(x;\hat\theta), F_n(x)]\,, \qquad \text{by (1) and the fact that } \hat\theta\in\Theta_{\hat\lambda}\,,\\
&\quad=\inf\{d[F(x;\theta), F_n(x)];\ \theta\in\Theta\}\,, \qquad \text{by (1)}\,,\\
&\quad\leqq\inf\{d[F(x;\theta), F_n(x)];\ \theta\in\Theta_\lambda\} \qquad (\lambda\in\Lambda),\ \text{by (2)}\,,\\
&\quad=d_u[F(x;\lambda), F_n(x)]\,, \qquad \text{by (3)}\,. \qquad\qquad \text{Q.E.D.}
\end{aligned}$$

In ending, we note that estimates provided by the commonly used inference methods mentioned above are invariant under $u(\cdot)$. This is a corollary of the theorem, and it may be obtained by appropriately defining the non-negative distance function $d(\cdot,\cdot)$. For example, in the case of maximum likelihood estimates, based on a random sample $X_1,\cdots,X_n$ from a distribution function $F(x;\theta)$ (with probability density function $f(x;\theta)$), we define $d(\cdot,\cdot)$ as

$$d[F(x;\theta), F_n(x)]=\left[\prod_{j=1}^{n} f(x_j;\theta)\right]^{-1}.$$

UNIVERSITY OF PATRAS
AMERICAN UNIVERSITY OF BEIRUT

References

[1] Blyth, C. R. (1970). On the inference and decision models of statistics, *Ann. Math. Statist.*, **41**, 1034-1058.

[2] Matusita, K. (1953). On the estimation by the minimum distance method, *Ann. Inst. Statist. Math.*, **5**, 57-65.

[3] Matusita, K. (1964). Distance and decision rules, *Ann. Inst. Statist. Math.*, **16**, 305-315.

[4] Wolfowitz, J. (1952). Consistent estimators of the parameters of a linear structural relation, *Skand. Aktuarietidskr.*, **35**, 132-151.

[5] Wolfowitz, J. (1953). Estimation by the minimum distance method, *Ann. Inst. Statist. Math.*, **5**, 9-23.

[6] Wolfowitz, J. (1957). The minimum distance method, *Ann. Math. Statist.*, **28**, 75-87.

[7] Zehna, P. W. (1966). Invariance of maximum likelihood estimators, *Ann. Math. Statist.*, **37**, 744.