The Annals of Mathematical Statistics
1970, Vol. 41, No. 4, 1344–1346

A CONSISTENT NONPARAMETRIC MULTIVARIATE DENSITY ESTIMATOR BASED ON STATISTICALLY EQUIVALENT BLOCKS[1]

BY M. P. GESSAMAN

Ithaca College

1. Introduction and summary. Let $x_1, x_2, \cdots, x_m$ be a random sample from a p-dimensional random variable $X = (X_1, X_2, \cdots, X_p)$ with probability distribution P. It is assumed that P is absolutely continuous with respect to Lebesgue measure, and that the corresponding probability density function is denoted by f. If $z = (z_1, z_2, \cdots, z_p)$ is a point at which f is both continuous and positive, an estimator for $f(z)$ based on statistically equivalent blocks is suggested and its consistency is shown.

This estimator grew out of work on the nonparametric discrimination problem. Fix and Hodges [2] showed how density estimation could be used in this problem and demonstrated a consistent estimator at points such as z. Loftsgaarden and Quesenberry [4] proposed another estimator which is consistent at points such as z; their estimator was based on statistically equivalent blocks. Although this estimator is easier to use in practice than that suggested by Fix and Hodges, it does require separate calculations if the sample is to be used to estimate the density at two or more points, and gives complex regions on which the estimate is constant if it is desired to estimate f on some subset of the entire space. The estimator suggested in this paper is consistent at all points at which the two above estimators are consistent and allows the investigator to estimate the density at every point of p-dimensional Euclidean space from one construction, as well as providing rectangular regions on which the estimate is constant.

2. A consistent density estimator based on statistically equivalent blocks. For the purpose of this discussion, it is assumed that $p = 2$; the extension or restriction to any p-dimensional Euclidean space is immediate. The details of the theory of statistically equivalent blocks and coverages can be found in Wilks [5], Fraser [3], and Anderson [1]. Let $h_1(x), h_2(x), \cdots, h_m(x)$ be m real-valued functions of X such that the distribution of $h_j(x), j = 1, 2, \cdots, m$, is continuous when X has probability distribution P; and let $(j_1, j_2, \cdots, j_m)$ be a permutation of $(1, 2, \cdots, m)$. Use the function $h_{j_1}(x)$ to order the x_i, $i = 1, 2, \cdots, m$, and define $x^{(j_1)}$ as the j_1th value in this ordering. Then the cut $h_{j_1}(x) = h_{j_1}(x^{(j_1)})$ defines two blocks

$$B_{1 \cdots j_1} = \{X; h_{j_1}(x) \leqq h_{j_1}(x^{(j_1)})\}$$
$$B_{j_1+1 \cdots m+1} = \{X; h_{j_1}(x) > h_{j_1}(x^{(j_1)})\}.$$

Proceeding in the obvious fashion, the functions $h_{j_2}(x)$, $i = 2, \cdots, m$, and the sample values can be used to partition the plane into $m+1$ unit blocks. The

Received June 30, 1969.

[1] Research for this paper was carried out in part under a research grant-in-aid from the Research Council of the College Center of the Finger Lakes.

density $f(x)$ will be estimated uniformly on rectangles which are unions of these unit blocks.

Let k_m be any sequence of positive integers such that

$$\lim_{m\to\infty} k_m = \infty \quad \text{and} \quad \lim_{m\to\infty} k_m/m = 0. \tag{2.1}$$

Partition the space into "horizontal" blocks by $[(m/k_m)^{\frac{1}{2}}]-1$ cuts made as evenly spaced as possible on the first coordinate X_1. There are $[(m/k_m)^{\frac{1}{2}}]$ such blocks, each the union of at least $(mk_m)^{\frac{1}{2}}$ unit blocks. In turn, partition each of these into $[(m/k_m)^{\frac{1}{2}}]$ "vertical" blocks by cuts on the second coordinate X_2; these blocks are each unions of k_m or k_m+1 unit blocks. The estimate of $f(x)$ will be made uniformly on each of the resulting m/k_m rectangular blocks. A numerical example will illustrate the method of construction.

EXAMPLE. Let $m = 75{,}000$. If k_m is the greatest integer less than or equal to $m^{\frac{1}{2}}$, then $k_m = 42$. Make the cuts on the first coordinate so that 12 blocks contain 1785 observations and 30 blocks contain 1786 observations, e.g. let $j_1 = 1786$, $j_2 = 3572$, $j_3 = 5357, \cdots, j_6 = 10{,}715$, $j_7 = 12{,}500, \cdots, j_{41} = 73{,}215$, $j_{42} = 1$, $j_{43} = 2, \cdots, j_{75{,}000} = 75{,}000$ and $h_{j_2}(x) = X_1$ for $i = 1, \cdots, 41$, and $h_{j_2}(x) = X_2$ otherwise. By this method, the plane is partitioned first into 42 blocks, each of which is a union of 1785 or 1786 unit blocks. Using unions of consecutively ranked unit blocks, each of the first k_m blocks can be partitioned into k_m rectangular blocks such that each is a union of k_m or k_m+1 unit blocks. The density is then estimated uniformly on each of the resulting $(k_m)^2 = 1764$ blocks.

In the above construction $\{[(m/k_m)^{\frac{1}{2}}]-2\}^2$ of the blocks are bounded. On the unbounded blocks estimate the value of $f(z)$ at zero. If z is contained in a bounded block, let $B_z{}^m$ denote this block and let $A_z{}^m$ denote the area of $B_z{}^m$. Then an estimator of $f(z)$ is given by

$$f_m{}^*(z) = \frac{k_m}{(m+1)A_z{}^m}. \tag{2.2}$$

It should be noted that the upper and right boundaries of the blocks, when they exist, belong to the block; the estimate of the density on these boundaries is then that pertaining to the appropriate block. Hence the density is estimated at every point in the plane.

THEOREM. *If z is a point at which f is continuous and positive, then the estimator $f_m{}^*(z)$ given by* (2.2) *is consistent for $f(z)$.*

PROOF. With arbitrarily high probability the point z is interior to a bounded rectangle if m is large enough; in the following discussion it is assumed that m is large enough that this condition is satisfied.

Let $S = \{R;\ R$ is a bounded rectangle and z is an interior point of $R\}$; for each $R \in S$, let $A(R)$ and $||e(R)||$ denote the area of R and the longer side of R, respectively. Since f is continuous at z, for each $\varepsilon > 0$ there corresponds a $\sigma > 0$ such that

$$\text{if} \quad ||e(R)|| < \sigma \quad \text{for} \quad R \in S, \quad \text{then} \quad |[P(R)/A(R)]-f(z)| < \varepsilon. \tag{2.3}$$

The initial cuts on X_1 produced a block $B_z^{m,1}$ which is a union of at least $(mk_m)^{\frac{1}{2}}$ unit blocks. The probability of $B_z^{m,1}$ under P is a beta random variable with parameters (approximately) $(mk_m)^{\frac{1}{2}}$ and $m-(mk_m)^{\frac{1}{2}}+1$. Since the sequence $\{(mk_m)^{\frac{1}{2}}\}$ satisfies the conditions of (2.1), it is obvious that $P(B_z^{m,1}) \to_p 0$. But, under the assumptions on z, $P(B_z^{m,1}) \to_p 0$ only if the length of $B_z^{m,1}$ on X_1 converges in probability to zero. With arbitrarily high probability the length of $B_z^{m,1}$ on X_1 (the length of the X_1 side on B_z^m) is less than σ for m sufficiently large.

Given the original cuts on X_1 which produced $B_z^{m,1}$, the cuts made later on this block are based on X_2 only. The probability of B_z^m, a union of k_m or k_m+1 unit blocks, converges in probability to zero. As in the case of $B_z^{m,1}$, above, this can occur only if the length on X_2 of B_z^m converges to zero. Therefore, for m sufficiently large, the length of B_z^m on both X_1 and X_2 is less than σ with arbitrarily high probability, i.e. if $||e(B_z^m)||$ denotes the longer side of B_z^m, then

$$||e(B_z^m)|| \to_p 0. \tag{2.4}$$

From (2.3) and (2.4) $P(B_z^m)/A_z^m \to_p f(z)$ or

$$[\{(m+1)P(B_z^m)\}/k_m]/[\{(m+1)A_z^m\}/k_m] \to_p f(z). \tag{2.5}$$

If it is shown that

$$\{(m+1)P(B_z^m)\}/k_m \to_p 1, \tag{2.6}$$

then the denominator

$$\{(m+1)A_z^m\}/k_m \to_p 1/f(z). \tag{2.7}$$

But, (2.7) is equivalent to showing that $f_m^*(z)$ is consistent for $f(z)$. Recall that $P(B_z^m)$ is a beta random variable with parameters k_m and $m-k_m+1$ (or k_m+1 and $m-k_m$). By a simple application of Chebyshev's Inequality, (2.6) follows and the proof is complete.

Acknowledgment. The author wishes to express appreciation to the referee and the associate editor for their helpful suggestions with regard to some needed comments and clarifications.

REFERENCES

[1] ANDERSON, T. W. (1966). Some nonparametric multivariate procedures based on statistically equivalent blocks. *Multivariate Analysis: Proceedings of an International Symposium*, Ed. Paruchuri R. Krishnaiah. Academic Press, New York.

[2] FIX, E. and HODGES, J. L., JR. (1951). Discriminatory analysis, nonparametric discrimination: consistency properties. Report No. 4, Project No. 21-49-004, USAF School of Aviation Medicine.

[3] FRASER, D. A. S. (1957). *Nonparametric Methods in Statistics*. Wiley, New York.

[4] LOFTSGAARDEN, D. O. and QUESENBERRY, C. P. (1965). A nonparametric estimate of a multivariate density function. *Ann. Math. Statist.* **36** 1049–1051.

[5] WILKS, S. S. (1962). *Mathematical Statistics*. Wiley, New York.