The Annals of Statistics
1974, Vol. 2, No. 6, 1226-1237

BAYESIAN RECONSTRUCTIONS OF m, n-PATTERNS[1]

BY MARC MOORE

Ecole Polytechnique, Montreal

The notion of m, n-pattern is introduced—namely, a division of the unit interval into at most n cells (intervals or points), each having one of m colors. Given an unknown m, n pattern, it is desired to produce a reconstruction of the pattern using $r \geqq 1$ sample points (fixed or chosen at random) where the color is determined. The problem is studied from a decision-theoretic point of view. A way to obtain all the probability measures on the set of m, n-patterns is given. The notion of a Bayesian reconstruction rule (B.R.R.) is introduced. It is proved that when B.R.R.'s are considered, it is sufficient to use certain fixed sample points. A complete class of reconstruction rules is obtained. Finally an example of a B.R.R. is given for 2, 2-patterns.

1. Introduction and summary. Switzer (1967) studied the reconstruction of patterns from sample data. We recall Switzer's definition of a pattern: A Euclidean k-dimensional region A having unit volume is partitioned into m Lebesgue-measurable subregions $A_1, A_2, \cdots, A_m$ each having positive volume and an identifying color. It is supposed that this partition of A is a realization of a random process with certain stationarity and isotropy properties. This pattern is unobservable except at a fixed set of n sample points $s_1, \cdots, s_n$. The color of these points is used to produce an estimated reconstruction of the pattern.

In this paper, a different notion of pattern on the interval [0, 1] is introduced; these patterns have at most m colors and at most n cells, a cell being a connected subset, possibly a singleton, of [0, 1] the points of which are all of the same color. These patterns are called m, n-patterns. The differences between the two definitions are the following: the m, n-patterns are defined on [0, 1] only, the notion of cells is introduced, there may be one (or more) subregion A_i with length zero and it is not supposed, in the definition, that the division of [0, 1] is a realization of a random process; however, when Bayesian methods are considered a priori probability measures are placed on the set of m, n-patterns.

The problem considered is to reconstruct a fixed unknown m, n-pattern after observing (without error) the colors at a selected set of sample points with fixed or random locations. This reconstruction problem is introduced from a decision-theoretic point of view.

Received April 1972; revised November 1973.

[1] This paper was written with the partial support of the National Research Council of Canada, Grant A8211.

AMS 1970 *subject classifications.* Primary, 62C10, 62C99; Secondary, 62C07, 62F99.

Key words and phrases. Pattern, cell, color, sample points, reconstruction rule, measure, Bayes, Wald's decision theory.

In Section 2, a way to obtain all the probability measures on the set of m, n-patterns is given. The notion of a Bayesian reconstruction rule is defined in Section 3; it is proved that when Bayesian reconstruction rules are considered, it is sufficient to use certain fixed sample points. The selection of sample points corresponds to the choice of experimental design in ordinary statistical problems, and this result may therefore be interpreted as implying no advantage for a randomized choice of design. This is done by showing that every random choice of the sample points may be replaced, without loss, by the selection of a fixed set of sample points. Only the existence of this set is proved. Section 4 is devoted to the application of Wald's decision theory to the reconstruction problem from fixed sample points; a complete class of reconstruction rules is obtained. Finally, in Section 5, an example is given for the case of 2, 2-patterns.

2. m, n-Patterns. Let m, n be the two integers ($m > 1, n \geqq 1$) and $K = \{k_1, \cdots, k_m\}$ be a set of m different colors. An m, n-pattern is a pair (α, C), where α is a function from $[0, 1]$ into K and C is a partition of $[0, 1]$ into ordered connected subsets $C_1, \cdots, C_N$, satisfying:

(1) $x, y \in C_i$ implies $\alpha(x) = \alpha(y) = k(C_i)$ for each $i = 1, \cdots, N$, i.e. all points in C_i have the same color $k(C_i)$;

(2) $1 \leqq N \leqq n$, and C_1 has zero at left end point;

(3) $k(C_i) \neq k(C_{i+1}), i = 1, \cdots, N - 1$, i.e. no two adjacent sets of the partition have the same color.

The subsets $C_1, \cdots, C_N$ are called "cells;" a cell reduced to a point is called "point cell." Denote by A the set of all m, n-patterns (m, n, K fixed).

Bayes procedures for the reconstruction of an m, n-pattern are introduced in the next sections. A way to get all the probability measures on A is now given.

This is done in several steps: the set A is partitioned into five categories of m, n-patterns, a measurable subset B of R^{4n} and a measurable partition B^* of B are defined, a one-to-one correspondence is established between B^* and A, and finally the objective is reached by the definition of probability measures on B^*.

Let α be a fixed element of A. The pattern α is said to be of:

Category 1: if there are no point cells in α.

Category 2: if $N = n$ and there are s point cells ($s \geqq 1$) in α.

Category 3: if $1 < N < n$ and there is exactly one point cell in α, this cell being at point one.

Category 4: if $1 < N < n$ and there are s point cells ($s \geqq 1$) in α, no point cell being at point one.

Category 5: if $1 < N < n$ and there are s point cells ($s > 1$) in α, one being at point one.

Let B be the set of points $(X_1, Y_1, K_1, S_1, \cdots, X_n, Y_n, K_n, S_n)$ in R^{4n}, such that:

(1) $0 = X_1 \leqq Y_1 = X_2 \leqq Y_2 = X_3 \leqq \cdots = X_n \leqq Y_n = 1$,

(2) $K_i \in \{1, 2, \cdots, m\}$ $(i = 1, \cdots, n)$,
(3) $S_1 = 2$, $S_i \in \{1, 2\}$ $(i = 2, \cdots, n)$,
(4) if $X_i < Y_i = X_{i+1} < Y_{i+1}$ for some $1 \leqq i \leqq n - 1$ then $K_i \neq K_{i+1}$,
(5) if $X_i = Y_i = X_{i+1} = Y_{i+1} = \cdots = X_j = Y_j = X_{j+1} < Y_{j+1}$ for some $1 \leqq i \leqq j < n$ and $X_{i-1} < Y_{i-1} = X_i$ (if $i > 1$) then $K_{i-1} \neq K_i$ (if $i > 1$), $K_i = K_{i+1} = \cdots = K_j \neq K_{j+1}$, $S_i = S_{i+1} = \cdots = S_j = 2$, $S_{j+1} = 1$,
(6) if $X_{i-1} < Y_{i-1} = X_i = Y_i = \cdots = Y_n = 1$ for some $1 < i \leqq n$ then $K_i = K_{i+1} = \cdots = K_n$; furthermore, if $K_{i-1} \neq K_i$ then $S_i = 2$ and $S_{i+1} = \cdots = S_n = 1$, if $K_{i-1} = K_i$ then $S_i = S_{i+1} = \cdots = S_n = 1$.

Let α be a fixed element of A. Consider the sequence of N triplets $(Z_1, T_1, H_1) \cdots (Z_N, T_N, H_N)$ where Z_i and T_i are respectively the left and the right end points of C_i, and H_i is the index of $k(C_i)$ $(i = 1, \cdots, N)$. To associate a point in B with α we use the following idea: the X_i's, Y_i's and K_i's correspond respectively to the Z_i's, T_i's and H_i's, S_i is put equal to 1(2) if the ith cell is open (closed) at his left end. These correspondences have to be described more precisely because the number N of cells might be less than n. Without going into the details (they may be obtained from Moore (1971)) we give some indications.

If α is of category 1 and $N < n$ we put $X_{N+1} = Y_{N+1} = \cdots = Y_n = 1$ (i.e. $X_i = Z_i$, $Y_i = T_i$, $i = 1, \cdots, N$ and the remaining components are all put equal to 1) and $K_{N+1} = H_N$; this specifies exactly one point in B. If a is of category 2 or 3, by the same method exactly one point in B is associated with a.

Suppose now that α is of category 4. In this case we must add n-N components "X" and n-N components "Y", to the one naturally defined by the Z_i's and T_i's, to get a point in B. This is done by adding equal components, these components being equal to 1 or to the Z_i's corresponding to point cells. The number of equal components added at each place is arbitrary, the only restriction being that the total number of equal components added must be n-N. The components "K" corresponding to a sequence of added components "X" and "Y" are all put equal to the component "K" corresponding to the last components "X" and "Y" naturally defined by the Z_i's and T_i's. The components "S" corresponding to a sequence of added components "X" and "Y" are all put equal to 2 if the added components are not 1 and, all put equal to 1 if the added components are 1.

The following simple example illustrates the procedure. Let α be the 3, 5-pattern where $C_1 = [0, \frac{1}{2})$, $C_2 = \{\frac{1}{2}\}$, $C_3 = (\frac{1}{2}, 1]$, $k(C_1) = k_1$, $k(C_2) = k_2$, $k(C_3) = k_3$. To this pattern we could associate the point in B where

$X_1 = 0$	$X_2 = \frac{1}{2}$	$X_3 = \frac{1}{2}$	$X_4 = \frac{1}{2}$	$X_5 = \frac{1}{2}$
$Y_1 = \frac{1}{2}$	$Y_2 = \frac{1}{2}$	$Y_3 = \frac{1}{2}$	$Y_4 = \frac{1}{2}$	$Y_5 = 1$
		added		
$K_1 = 1$	$K_2 = 2$	$K_3 = 2$	$K_4 = 2$	$K_5 = 3$
$S_1 = 2$	$S_2 = 2$	$S_3 = 2$	$S_4 = 2$	$S_5 = 1$

or the point in B where

$X_1 = 0$	$X_2 = \frac{1}{2}$	$\boxed{X_3 = \frac{1}{2}}$	$X_4 = \frac{1}{2}$	$\boxed{X_5 = 1}$
$Y_1 = \frac{1}{2}$	$Y_2 = \frac{1}{2}$	$\boxed{Y_3 = \frac{1}{2}}$	$Y_4 = 1$	$\boxed{Y_5 = 1}$
		added		added
$K_1 = 1$	$K_2 = 2$	$K_3 = 2$	$K_4 = 3$	$K_5 = 3$
$S_1 = 2$	$S_2 = 2$	$S_3 = 2$	$S_4 = 1$	$S_5 = 1$.

In this example we must add 2 components "X" and 2 components "Y"; in the first case both are put equal to $\frac{1}{2}$ (there is a point cell at $\frac{1}{2}$), in the second case one is put equal to $\frac{1}{2}$ and one is put equal to 1. We could also put both components "X" and both components "Y" equal to 1.

Using the same procedure we associate a point in B with a pattern of category 5.

To a pattern of category 4 or 5, it is possible to associate many points in B. Grouping these points in a class we can form a measurable partition B^* of B and establish a one-to-one correspondence T between B^* and A.

Let β be the class of Borel subsets of B, τ the σ-algebra of subsets C of B^* being such that the set of points in B forming C is an element of β. On A, consider the σ-algebra γ given by the images (T) of the elements of τ. Considering all the $4n$-dimensional cumulative distribution functions Q such that $\mu_Q(B) = 1$ (μ_Q being the probability measure associated with Q), all the probability measures on (B, β) are obtained, then also all the probability measures on (B^*, τ), and finally, T being one to one, all the probability measures on (A, γ).

3. Bayesian reconstruction rules. Let α be a fixed unknown element of A, i.e. α has at most n cells and at most m colors, these colors being in a fixed known set K of m colors; nothing more is known about α.

Let F be an r-dimensional ($r \geqq 1$) cumulative distribution function such that $\mu_F(U) = 1$ where $U = [0, 1] \times \cdots \times [0, 1]$ is the r-dimensional unit cube. Following F a point $(y_1, \cdots, y_r)$ in U is chosen and the points $y_1, \cdots, y_r$ on $[0, 1]$ are considered; at each of these points, called sample points, the position and the color of the point is observed. The sample space is then Ω, the r-fold Cartesian product of $[0, 1] \times K$. Using this information, it is desired to construct an element of A which will be considered as a reconstruction (estimation) of the unknown m, n-pattern α. This problem will be studied from a decision theoretic point of view where the set of "states of nature" is A and the "decision space" is also A.

Let $\mathscr{W}$ be the set of finite measures W on $[0, 1]$. The loss functions considered are the functions ℓ of the form, $\ell(\alpha, \delta) = W(\alpha \triangle \delta)$ where $\alpha \triangle \delta = \{x: \alpha(x) \neq \delta(x)\}$, $W \in \mathscr{W}$. The number $\ell(\alpha, \delta)$ is the loss which occurs when the pattern α is reconstructed by the pattern δ. It is easy to see that for each $W \in \mathscr{W}$ the function ℓ defines a pseudometric on A. Denote by T_W the topology generated by the pseudometric ℓ and by ϕ the smallest σ-algebra of subsets of A, containing all the open spheres (T_W).

To define precisely a reconstruction rule, a σ-algebra of subsets of Ω must be defined. Denote by $\mathscr{A}$ the class of Borel subsets of [0, 1], by $\mathscr{K}$ the class of subsets of K, by $\mathscr{Y}$ the product σ-algebra $\mathscr{A} \times \mathscr{K}$. Consider the σ-algebra $\mathscr{Y}^r$ of subsets of Ω, which is the product σ-algebra of $\mathscr{Y}$, r times by itself.

DEFINITION. *A* reconstruction rule δ (abbreviated as R.R.) is a function which assigns to each $Y \in \Omega$ a probability measure G_Y on (A, ϕ) in such a way that for every $C \in \phi$ the application $\Omega \to_\delta G.$ (C) is $\mathscr{Y}^r$-measurable.

Denote by D the set of reconstruction rules. The following proposition is proved in the Appendix.

PROPOSITION 3.1. *For each* $W \in \mathscr{W}$, *each set in* ϕ *is in* γ.

Then, the way to get all the probability measures on (A, γ) introduced in Section 2, gives also all the probability measures on (A, ϕ).

Let $\alpha \in A$, $Y \in \Omega$, $\delta \in D$. The expected loss which occurs when the observation is Y, and δ is used to reconstruct α, is given by

$$L(\alpha, \delta(Y)) = \int_A \ell(\alpha, \ell)\, dG(\ell)$$

where G is the probability measure on (A, ϕ) assigned to Y by δ.

It can be proved (using Proposition 4.2) that for each $W \in \mathscr{W}$ and each $\delta \in D$, the real function defined on $A \times \Omega$ by $L(\cdot, \delta(\cdot))$ is $\gamma \times \mathscr{Y}^r$-measurable.

The risk associated with δ, when δ is used to reconstruct α, is defined by

$$R(\alpha, \delta) = \int_\Omega L(\alpha, \delta(Y))\, dP_\alpha(Y)$$

where P_α is the probability measure given by

$$P_\alpha[y_1 \in B_1, k(y_1) = k_{i_1}, \cdots, y_r \in B_r, k(y_r) = k_{i_r}] = \int_{B_1 \cap \alpha_{i_1}} \cdots \int_{B_r \cap \alpha_{i_r}} dF(y_1, \cdots y_r)\,,$$

$B_1, \cdots, B_r$ being Borel sets in [0, 1] and $\alpha_{i_j} = \{x\colon \alpha(x) = k_{i_j}\}$.

Given an a priori distribution (probability measure) λ on (A, γ), the risk corresponding to δ relative to λ is

$$r(\lambda, \delta) = \int_A R(\alpha, \delta)\, d\lambda(\alpha)\,.$$

DEFINITION. Given F and λ, a R.R. δ^* is called a Bayesian reconstruction rule (abbreviated as B.R.R.) relative to λ, if

$$r(\lambda, \delta^*) = \inf_{\delta \in D} r(\lambda, \delta)\,.$$

In the above definition F is fixed; if F is modified δ^* may change. So the F for which $r(\delta^*, \lambda)$ is as small as possible should be used in practice.

THEOREM 3.1. *For each* F *and* λ *(for which a* B.R.R. *relative to* λ *exists) there is a fixed point* $\bar{y} \in U$ *(which depends on* F *and* λ*) such that*

$$r_{\bar{y}}(\lambda, \delta_{\bar{y}}^*) \leqq r_F(\lambda, \delta_F^*)$$

where $r_{\bar{y}}(\lambda, \delta_{\bar{y}}^*)$ *is the risk corresponding to the* B.R.R. *relative to* λ, *when the information is taken from* $\bar{y}$, *and* $r_F(\lambda, \delta_F^*)$ *is the risk corresponding to the* B.R.R.

relative to λ, when the information is taken from a random point chosen on U following F.

PROOF. We can write

$$r_F(\lambda, \delta_F^*) = \int_U \sum_{K^r} \int_A \int_A \mathcal{L}(\alpha, \mathcal{A})\, dG_{Y,F}(\mathcal{A})\, d\lambda(\alpha \mid (y, k(y)))\, dQ(k(y) \mid y)\, dF(y)$$

where: K^r is the r-fold Cartesian product of K,

$$y = (y_1, \cdots, y_r) \in U, \qquad k(y) = (k(y_1), \cdots, k(y_r)),$$

$G_{Y,F}$ is the probability measure associated to $Y = (y, k(y))$ by

$$\delta_F^*, Q(k(y) \in D \mid y) = \lambda(\{\alpha : (\alpha(y_1), \cdots, \alpha(y_r)) \in D\}), \qquad D \subset K^r.$$

So $r_F(\lambda, \delta_F^*)$ is the expectation, with respect to F, of a nonnegative function. Then there is at least one fixed point $\bar{y} \in U$ such that

$$\begin{aligned} r_F(\lambda, \delta_F^*) &\geqq \sum_{K^r} \int_A \int_A \mathcal{L}(\alpha, \mathcal{A})\, dG_{\bar{Y},F}(\mathcal{A})\, d\lambda(\alpha \mid (\bar{y}, k(\bar{y})))\, dQ(k(\bar{y}) \mid \bar{y}) \\ &= r_{\bar{y}}(\lambda, \delta_F^*) \end{aligned}$$

where $r_{\bar{y}}(\lambda, \delta_F^*)$ is the risk corresponding to the rule δ_F^* used with the information obtained from $\bar{y}$. In the next section, it is proved that, when the sample points are fixed, for any λ there is a B.R.R. relative to λ. Then

$$r_{\bar{y}}(\lambda, \delta_F^*) \geqq r_{\bar{y}}(\lambda, \delta_{\bar{y}}^*). \quad \square$$

Using Theorem 3.1 as motivation, it may be supposed that the sample points are fixed (F degenerate). In that case, the sample space is $\Omega' = K^r$ and a R.R. is a function which assigns to each $Y \in \Omega'$ a probability measure G on (A, ϕ); denote by D' this set of R.R. In the next section, it is supposed that the observations are taken at fixed points; it would be necessary to find the fixed point $\bar{y} \in U$ for which the risk $r(\lambda, \delta^*)$ is minimum. An example will be given in Section 5, but this problem will not be considered in general here.

4. Application of Wald's decision theory. In this section we first verify the first six assumptions of Wald's decision theory (Wald (1950) Chapter 3), and then, applying Wald's results, we obtain a complete class of reconstruction rules.

Among the first six assumptions of Wald's decision theory, the first three and the fifth assumptions are trivially satisfied here (when the sample points are fixed).

The intrinsic distance between two elements α, $\mathcal{A}$ of A is defined by

$$\Delta(\alpha, \mathcal{A}) = \sup_c |\mathcal{L}(\alpha, c) - \mathcal{L}(\mathcal{A}, c)|.$$

Wald's fourth assumption is: "A is compact relative to Δ." This assumption is satisfied here as follows from Propositions 4.1 and 4.2.

PROPOSITION 4.1. *For each $\alpha, \mathcal{A} \in A$, $\mathcal{L}(\alpha, \mathcal{A}) = \Delta(\alpha, \mathcal{A})$.*

PROOF. The function $\mathscr{A}$ is a pseudometric on A; then for each $\mathscr{C} \in A$

$$|\mathscr{A}(\mathscr{A}, \mathscr{C}) - \mathscr{A}(\mathscr{B}, \mathscr{C})| \leqq \mathscr{A}(\mathscr{A}, \mathscr{B})$$

and likewise for the supremum over $\mathscr{C}$. It is clear that

$$\sup_{\mathscr{C}} |\mathscr{A}(\mathscr{A}, \mathscr{C}) - \mathscr{A}(\mathscr{B}, \mathscr{C})| \geqq |\mathscr{A}(\mathscr{A}, \mathscr{B}) - \mathscr{A}(\mathscr{B}, \mathscr{B})| = \mathscr{A}(\mathscr{A}, \mathscr{B})$$

and hence the result. □

PROPOSITION 4.2. *For each $W \in \mathscr{W}$, A is compact relative to the topology T_W.*

PROOF. It suffices to show every infinite sequence $\langle \mathscr{A}_\nu \rangle$ of elements in A has a convergent subsequence (s.s.) (Royden (1963), page 142, Corollary 14).

Let N_ν be the number of cells in $\mathscr{A}_\nu$, $(Z_{\nu 1}, T_{\nu 1}, H_{\nu 1}), \cdots, (Z_{\nu N_\nu}, T_{\nu N_\nu}, H_{\nu N_\nu})$ the sequence of N_ν triplets, as defined in Section 2, corresponding to $\mathscr{A}_\nu$.

Consider the sequence of varying length vectors $\langle (N_\nu, H_{\nu 1}, \cdots, H_{\nu N_\nu}) \rangle$. Since $1 \leqq N_\nu \leqq n$ and $H_{\nu p} \in \{1, \cdots, m\}$ $p = 1, \cdots, N_\nu$ for all ν, there are only a finite number of values which $(N_\nu, H_{\nu 1}, \cdots, H_{\nu N_\nu})$ can assume. Hence there must be at least one value, say $(N_0, H_{01}, \cdots, H_{0N_0})$ which occurs infinitely often. Let $\langle {}_1\mathscr{A}_i \rangle$ be an infinite s.s. of $\langle \mathscr{A}_\nu \rangle$ such that $({}_1N_i, {}_1H_{i1}, \cdots, {}_1H_{i\,{}_1N_i}) = (N_0, H_{01}, \cdots, H_{0N_0})$ for all i.

Next consider the sequence of $2N_0$-tuples $\langle {}_1\zeta_i \rangle$ where ${}_1\zeta_i = ({}_1Z_i, {}_1T_{i1}, \cdots, {}_1Z_{iN_0}, {}_1T_{iN_0})$. Since ${}_1\zeta_i \in C^{2N_0}$ (the $2N_0$ dimensional unit cube) for all i and C^{2N_0} is compact, there exists a $2N_0$-tuple, say $\zeta = (Z_{01}, T_{01}, \cdots, Z_{0N_0}, T_{0N_0})$, and an infinite s.s. $\langle {}_2\mathscr{A}_j \rangle$ of $\langle {}_1\mathscr{A}_i \rangle$ such that ${}_2\zeta_j$ converges to ζ, where ${}_2\zeta_j = ({}_2Z_{j1}, {}_2T_{j1}, \cdots, {}_2Z_{jN_0}, {}_2T_{jN_0})$ (Royden (1963), page 142, Corollary 14).

By construction it is clear that

$$({}_2Z_{j1}, {}_2T_{j1}, {}_2H_{j1}, \cdots, {}_2Z_{j\,{}_2N_j}, {}_2T_{j\,{}_2N_j}, {}_2H_{j\,{}_2N_j})$$

converges to

$$(Z_{01}, T_{01}, H_{01}, \cdots, Z_{0N_0}, T_{0N_0}, H_{0N_0}) = \tau\,;$$

however, τ need not correspond to an element in A. (Suppose $m = n = 4$. Consider the sequence $\langle \mathscr{A}_\nu \rangle$, where $N_\nu = 4$, $H_{\nu 1} = 1$, $H_{\nu 2} = 2$, $H_{\nu 3} = 3$, $H_{\nu 4} = 4$ for all ν and $Z_{\nu 1} = 0$, $T_{\nu 1} = Z_{\nu 2} = \frac{1}{2} - 1/4\nu$, $T_{\nu 2} = Z_{\nu 3} = \frac{1}{2}$, $T_{\nu 3} = Z_{\nu 4} = \frac{1}{2} + 1/4\nu$, $T_{\nu 4} = 1$. In this case $(Z_{\nu 1}, T_{\nu 1}, H_{\nu 1}, \cdots, Z_{\nu 4}, T_{\nu 4}, H_{\nu 4})$ converges to $(0, \frac{1}{2}, 1, \frac{1}{2}, \frac{1}{2}, 2, \frac{1}{2}, \frac{1}{2}, 3, \frac{1}{2}, 1, 4)$ which cannot correspond to any element of A.)

Now let $0 = \alpha_1 < \cdots < \alpha_d = 1$ denote the distinct elements among $Z_{01}, T_{01}, \cdots, Z_{0N_0}, T_{0N_0}$. Instead of writing $\mathscr{A}(x) = k_i$, we now write $\mathscr{A}(x) = i$, the index of the color of x in $\mathscr{A}$. Consider the sequence $\langle ({}_2\mathscr{A}_j(\alpha_1), \cdots, {}_2\mathscr{A}_j(\alpha_d)) \rangle$. Again, since there are only finitely many values which $({}_2\mathscr{A}_j(\alpha_1), \cdots, {}_2\mathscr{A}_j(\alpha_d))$ can assume, there must be one, say $(\mathscr{A}_1, \cdots, \mathscr{A}_d)$, which occurs infinitely often. Let $\langle {}_3\mathscr{A}_k \rangle$ be an infinite s.s. of $\langle {}_2\mathscr{A}_j \rangle$ such that $({}_3\mathscr{A}_k(\alpha_1), \cdots, {}_3\mathscr{A}_k(\alpha_d)) = (\mathscr{A}_1, \cdots, \mathscr{A}_d)$ for all k.

Define the integers $\eta_1, \cdots, \eta_d$ by $\eta_i =$ the largest integer j such that $Z_{0j} = \alpha_i$, and define

$$I_{\eta_i} = (Z_{0\eta_i}, T_{0\eta_i})$$

(note that $Z_{0\eta_i} = \alpha_i$ and $T_{0\eta_i} = \alpha_{i+1}$).

Let $\alpha_0 \in A$ be defined by

$$\begin{aligned} \alpha_0(x) &= H_{0\eta_i} && \text{if} \quad x \in I_{\eta_i}, \\ &= \alpha_i && \text{if} \quad x = \alpha_i . \end{aligned}$$

Finally, define functions $f; f_1, f_2, \cdots$ from [0, 1] to [0, 1] by

$$\begin{aligned} f(x) &= 0 && \text{for all} \quad x, \\ f_k(x) &= 0 && \text{if} \quad {}_3\alpha_k(x) = \alpha_0(x), \\ &= 1 && \text{if} \quad {}_3\alpha_k(x) \neq \alpha_0(x). \end{aligned}$$

By construction, it is clear that

$$\lim_{k\to\infty} f_k(x) = f(x) \qquad \text{for all} \quad x .$$

Hence,

$$f_k \to f \qquad \text{a.e.} \quad (W)$$

for each $W \in \mathscr{W}$; then by the Lebesgue convergence theorem

$$\lim_{k\to\infty} W({}_3\alpha_k \triangle \alpha_0) = \lim_{k\to\infty} \int_0^1 f_k(x)\, dW(x) = 0 . \qquad \square$$

DEFINITION. A sequence $\langle \delta_i \rangle$ of R.R. converges, in the regular sense, to the R.R. δ_0 if

$$\lim_{i\to\infty} G_{iY}(C) = G_{0Y}(C)$$

for each $Y \in \Omega'$ and for each open (T_W) subset C of A whose boundary has probability zero according to G_{0Y}; let $G_{iY}(G_{0Y})$ denote the probability measure associated with Y by $\delta_i(\delta_0)$.

In the present context Wald's sixth assumption may be formulated in the following way: "The set D' of R.R. is such that:

(1) for each $\delta_1, \delta_2 \in D'$ and each $0 < \alpha < 1$ there is an element $\delta_3 \in D'$ such that $G_{3Y}(C) = \alpha G_{1Y}(C) + (1 - \alpha) G_{2Y}(C)$ for each $C \in \phi$ and each $Y \in \Omega'$,

(2) D' is closed in the sense of the regular convergence given above."

It is easy to see that the first part is satisfied.

PROPOSITION 4.3. *D' is closed in the sense of regular convergence.*

PROOF. Let $\langle \delta_i \rangle$ be a sequence of R.R. which converges in the regular sense to δ_0. It will be proved that for each $Y \in \Omega'$, G_{0Y} is a probability measure on (A, ϕ), i.e. $\delta_0 \in D'$.

Let Y be a fixed element of Ω' and E be a finite intersection of open spheres (T_W), such that the boundary of E has probability zero according to G_{0Y}. Since each such set E is open (T_W), $\lim_{i\to\infty} G_{iY}(E) = G_{0Y}(E)$. Since A is separable relative to T_W, the sequence $\langle G_{iY} \rangle$ converges weakly to G_{0Y} (Billingsley (1968), page 15, Corollary 2). Since A is compact relative to T_W, from the Prohorov theorem (Billingsley (1968), page 37, Theorem 6.1), the sequence $\langle G_{iY} \rangle$ forms a relatively compact family of probability measures on (A, ϕ), i.e. there is a subsequence $\langle G_{i(j)Y} \rangle$ and a probability measure Q_{0Y} on (A, ϕ) such that $\langle G_{i(j)Y} \rangle$

converges weakly to Q_{0Y}. Then $G_{0Y} \equiv Q_{0Y}$ and G_{0Y} is a probability measure on (A, ϕ). This property being true for each $Y \in \Omega'$, we have $\delta_0 \in D'$. □

From the first six assumptions of Wald's theory, the following results can be formulated.

THEOREM 4.1. *For any a priori distribution* λ, *there is a* B.R.R. *relative to* λ (Wald (1950), Theorem 3.5).

THEOREM 4.2. *There is a minimax* R.R., *i.e. a* R.R. δ_0 *such that* $\sup_\alpha R(\alpha, \delta_0) = \inf_\delta \sup_\alpha R(\alpha, \delta)$ (Wald (1950), Theorem 3.7).

DEFINITION. δ_0 is a B.R.R. in the wide sense if there is a sequence of a priori distributions $\langle \lambda_i \rangle$ such that

$$\lim_{i\to\infty} [\inf_\delta r(\lambda_i, \delta) - r(\lambda_i, \delta_0)] = 0 .$$

THEOREM 4.3. *Any minimax* R.R. *is a* B.R.R. *in the wide sense* (Wald (1950), Theorem 3.8).

THEOREM 4.4. *The class of all* B.R.R. *in the wide sense is a complete class* (Wald (1950), Theorem 3.17).

DEFINITION. *A* sequence of patterns $\langle \alpha_i \rangle$ converges in the regular sense to a pattern α_0 if

$$\lim_{i\to\infty} P_{\alpha_i}(V) = P_{\alpha_0}(V)$$

uniformly in $V \subset \Omega'$ where, since the sample points $y_1, \cdots, y_r$ are fixed, $P_\alpha((k_{i_1}, \cdots, k_{i_r}))$ equals one if $\alpha(y_j) = k_{i_j}$ for $j = 1, \cdots, r$, and equals zero otherwise.

Wald's last assumption is the following: "*A* is compact in the sense of regular convergence. If $\langle \alpha_i \rangle$ converges to α_0 in the regular sense, then

$$\lim_{i\to\infty} \angle(\alpha_i, \angle) = \angle(\alpha_0, \angle)$$

uniformly in *b*."

This assumption is not always satisfied here, as shown by the following example. Consider the sequence of 2, 3-patterns $\langle \alpha_i \rangle$ where in α_i,

$$C_1 = [0, \tfrac{3}{8}) , \qquad C_2 = [\tfrac{3}{8}, \textstyle\sum_{j=1}^{i} (\tfrac{1}{2})^j) , \qquad C_3 = [\textstyle\sum_{j=1}^{i} (\tfrac{1}{2})^j, 1] ,$$

$K(C_1) = k_1$, $K(C_2) = k_2$, $K(C_3) = k_1$. Suppose that the sample points are fixed and all in $[0, \frac{3}{8})$. Let α_0 be the 2, 3-pattern where

$$C_1 = [0, \tfrac{1}{2}) , \qquad C_2 = [\tfrac{1}{2}, \tfrac{3}{4}) , \qquad C_3 = [\tfrac{3}{4}, 1] ,$$

$K(C_1) = k_1$, $K(C_2) = k_2$, $K(C_3) = k_1$. For each $V \subset \Omega'$, $P_{\alpha_i}(V) = P_{\alpha_0}(V)$ $i = 1, 2, \cdots$; consider the loss function given by Lebesgue measure on [0, 1], then $\lim_{i\to\infty} \angle(\alpha_i, \alpha_0) = \frac{3}{8} \neq \angle(\alpha_0, \alpha_0) = 0$.

5. Example. In this section, an example of a B.R.R. for 2, 2-patterns is given. To specify an a priori probability distribution on that set of patterns, denoted by *A*, it is sufficient to give a probability distribution on the set *B* of points

$(X_1, Y_1, K_1, S_1, X_2, Y_2, K_2, S_2)$ in R^8, as defined in Section 2. We consider the probability distribution on B given by the following process:

(1) following the uniform distribution on $[0, 1]$ a point s is chosen; let $Y_1 = X_2 = s$ and $S_2 = 1$ with probability one;

(2) with probability $\frac{1}{2}$, $K_1 = 1$ (i.e. the color assigned to $[0, s]$ is k_1) also with probability $\frac{1}{2}$, $K_1 = 2$.

The process just described corresponds to a distribution λ which is concentrated on the patterns with exactly two cells, each of positive length; denote by A^* that class of patterns. These petterns are well described by pairs $(s, \bar{k})$ where s is the division point of the cells, $[0, s]$, $(s, 1]$ and $\bar{k}$ is the color of $[0, s]$.

Let $0 \leqq y_1 < y_2 < \cdots < y_r \leqq 1$ be r fixed sample points. Consider only the reconstruction rules which are functions from Ω' to A^* and let W be Lebesgue measure on $[0, 1]$. In that case

$$r(\lambda, \delta) = \sum_{\Omega'} \int_0^1 \sum_{\bar{k}} l[(s, \bar{k}); (u, \bar{\bar{k}})] h[(s, \bar{k}) \mid k(y)]\, ds\, v(k(y))$$

where: $k(y) = (k(y_1), \cdots, k(y_r))$, $(u, \bar{\bar{k}})$ is the reconstruction of $(s, \bar{k})$ obtained by the R.R. δ,

$$l[(s, \bar{k}); (u, \bar{\bar{k}})] = |u - s| \quad \text{if} \quad \bar{k} = \bar{\bar{k}},$$
$$= 1 - |u - s| \quad \text{if} \quad \bar{k} \neq \bar{\bar{k}};$$

$$h[(s, k_1) \mid k(y_1) = k_1, \cdots, k(y_r) = k_1] = \frac{1}{1 + y_1 - y_r} \quad \text{if} \quad s \geqq y_r, \tag{5.1}$$
$$= 0 \quad \text{otherwise};$$

$$h[(s, k_2) \mid k(y_1) = k_1, \cdots, k(y_r) = k_1] = \frac{1}{1 + y_1 - y_r} \quad \text{if} \quad s < y_1, \tag{5.2}$$
$$= 0 \quad \text{otherwise};$$

$$h[(s, k_1) \mid k(y_1) = k_2, \cdots, k(y_r) = k_2] \equiv (5.2),$$
$$h[(s, k_2) \mid k(y_1) = k_2, \cdots, k(y_r) = k_2] \equiv (5.1);$$

$$h[(s, k_1) \mid k(y_1) = k_1, \cdots, k(y_l) = k_1, k(y_{l+1}) = k_2, \cdots, k(y_r) = k_2]$$
$$\equiv h[(s, k_2) \mid k(y_1) = k_2, \cdots, k(y_l) = k_2, k(y_{l+1}) = k_1, \cdots, k(y_r) = k_1]$$
$$= \frac{1}{y_{l+1} - y_l} \quad \text{if} \quad y_l \leqq s < y_{l+1},$$
$$= 0 \quad \text{otherwise};$$

$$h[(s, k_1) \mid k(y_1) = k_2, \cdots, k(y_l) = k_2, k(y_{l+1}) = k_1, \cdots, k(y_r) = k_1]$$
$$\equiv h[(s, k_2) \mid k(y_1) = k_1, \cdots, k(y_l) = k_1, k(y_{l+1}) = k_2, \cdots, k(y_r) = k_2]$$
$$\equiv 0,\ 1 \leqq l < r$$

and $v(\cdot)$ is the distribution on Ω'.

To find the Bayesian reconstruction of $(s, \bar{k})$ it is sufficient to minimize

$$\int_0^1 \sum_{\bar{k}} l[(s, \bar{k}); (u, \bar{\bar{k}})] h[(s, \bar{k}) \mid k(y)]\, ds$$

with respect to $(u, \bar{\bar{k}})$, for each $k(y) \in \Omega'$. The Bayesian reconstruction is as follows:

(1) if $k(y_i) = k_1(k_2)$ $i = 1, \cdots, r$ then

(a) if $0 \leqq y_1 + y_r < 1$, $\quad u = \dfrac{y_1 + y_r + 1}{2}$, $\quad \bar{\bar{k}} = k_1(k_2)$,

if $1 \leqq y_1 + y_r$, $\quad u = \dfrac{y_1 + y_r - 1}{2}$, $\quad \bar{\bar{k}} = k_2(k_1)$;

(2) if $k(y_1) = k_1(k_2), \cdots, k(y_l) = k_1(k_2)$

$$k(y_{l+1}) = k_2(k_1), \cdots, k(y_r) = k_2(k_1), \qquad 1 \leqq l < r$$

$$u = \frac{y_l + y_{l+1}}{2}, \qquad \bar{\bar{k}} = k_1(k_2).$$

It can be proved that, using this rule, the risk will be minimum if the fixed sample points $y_1, \cdots, y_r$ are such that:

(1)
$$2y_1 - y_2 - y_r = -1,$$
$$2y_r - y_1 - y_{r-1} = 1 \qquad \text{if } r > 2,$$
$$y_l = \frac{y_{l-1} + y_{l+1}}{2} \qquad (l = 2, \cdots, r-1)$$

(2) $\quad y_2 = y_1 + \frac{1}{2} \qquad$ if $r = 2$,

(3) $\quad y_1 = \frac{1}{2} \qquad$ if $r = 1$.

As a consequence of Theorem 3.1, these fixed sample points give a risk inferior or equal to the risk corresponding to any random choice of the sample points.

6. Appendix. We give here the proof of Proposition 3.1.

It is sufficient to prove that for each $\alpha_0 \in A$ and each $W \in \mathscr{W}$, the function $\mathscr{L}(\cdot, \alpha_0)$ is γ-measurable. To prove this, it is sufficient to consider two cases: first, W purely atomic, second, W has no atoms.

First case. Suppose that W is purely atomic. Since W is finite its set of atoms is countable; denote by $x_1, x_2, \cdots$ the atoms of W. Let $k_i^{(0)} = \alpha_0(x_i)$, $A_i = \{\alpha : \alpha(x_i) = k_i^{(0)}\}$ then $\mathscr{L}(\alpha, \alpha_0) = \sum_{i=1}^{\infty} W(x_i)\chi_{\bar{A}_i}(\alpha)$ where $\chi_{\bar{A}_i}$ is the characteristic function of the complement of A_i. It is sufficient to prove that $A_i \in \gamma$ for each $i = 1, \cdots$; we prove that for each $x \in [0, 1]$ and each $\bar{k} \in K$, $\{\alpha : \alpha(x) = \bar{k}\} \in \gamma$. Suppose that $0 < x < 1$; $\{\alpha : \alpha(x) = \bar{k}\} = \bigcup_{N=1}^{n} \bigcup_{i=1}^{N} A_{Ni}$ where $A_{Ni} = \{\alpha : \alpha$ has N cells, $x \in C_i$, $K(C_i) = \bar{k}\}$. For each $1 \leqq N \leqq n$ and each $1 \leqq i \leqq N$, $A_{Ni} \in \gamma$, because to that set of patterns correspond the elements of B^* formed by points in B for which there is a j, $i \leqq j \leqq n - (N - i)$, such that one of the following possibilities is satisfied:

(1) $X_j < x < Y_j$, $K_j = \bar{k}$
(2) $X_j < Y_j = x$, $K_j = \bar{k}$, $S_{j+1} = 1$
(3) $x = X_j < Y_j$, $K_j = \bar{k}$, $S_j = 2$
(4) $X_j = Y_j = x$, $K_j = \bar{k}$, $S_j = 2$, $X_{j-1} < Y_{j-1}$.

The set of these elements in B^* belongs to τ. By similar arguments, the same is proved for $x = 0$ and $x = 1$.

Second case. Suppose that W has no atoms. Let

$$\ell(\alpha, \alpha_0) = W[0, 1] - \int_0^1 g_0(\alpha, x)\, dW(x)$$

where

$$g_0(\alpha, x) = 1 \quad \text{if} \quad \alpha(x) = \alpha_0(x)\,,$$
$$= 0 \quad \text{if} \quad \alpha(x) \neq \alpha_0(x)\,.$$

For each integer $\nu > 0$, consider the points $0, 1/\nu, \cdots, (\nu - 1)/\nu$ and the functions from A to $\{0, 1\}$ given by

$$c_{\nu i}(\alpha) = 1 \quad \text{iff} \quad \alpha\left(\frac{i}{\nu}\right) = \alpha_0\left(\frac{i}{\nu}\right) \qquad i = 0, \cdots, \nu - 1\,.$$

By the considerations made in the first case, for each $\nu > 0$ the functions $c_{\nu i}$, $i = 0, \cdots, \nu - 1$, are γ-measurable.

For each fixed α consider the sequence of measurable functions

$$\sum_{i=0}^{\nu-1} c_{\nu i}(\alpha)\chi_{E_{\nu i}}(\cdot) \qquad \nu = 1, 2, \cdots$$

from $[0, 1]$ to $[0, 1]$ where $E_{\nu i} = [i/\nu, (i + 1)/\nu)$ if $i < \nu - 1$ and $E_{\nu,\nu-1} = [(\nu - 1)/\nu, 1]$. It can be easily proved that for almost all $x[W]$ and each α,

$$g_0(\alpha, x) = \lim_{\nu\to\infty} \sum_{i=0}^{\nu-1} c_{\nu i}(\alpha)\chi_{E_{\nu i}}(x)\,.$$

Using the Lebesgue convergence theorem,

$$\int_0^1 g_0(\cdot, x)\, dW(x) = \lim_{\nu\to\infty} \int_0^1 \left[\sum_{i=0}^{\nu-1} c_{\nu i}(\cdot)\chi_{E_{\nu i}}(x)\right] dW(x)\,,$$
$$= \lim_{\nu\to\infty} \sum_{i=0}^{\nu-1} c_{\nu i}(\cdot) W(E_{\nu i})\,.$$

For each ν and each $i = 0, \cdots, \nu - 1$, $c_{\nu i}$ is γ-measurable, hence $\int_0^1 g_0(\cdot, x)\, dW(x)$ is γ-measurable. □

Acknowledgment. I wish to express my gratitude to my thesis advisor, Professor Constance van Eeden, of the Université de Montréal, for her guidance and encouragement during the preparation of my thesis and also this paper which is based upon it. Thanks are also due to the associate editor and the referee for their helpful criticisms and assistance. The referee's suggestions have led to clarification of the proof of Proposition 4.2.

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

MOORE, M. (1971). Sur la reconstruction des schémas. Ph. D. dissertation, Université de Montréal.

ROYDEN, H. L. (1963). *Real Analysis.* MacMillan, New York.

SWITZER, P. (1967). Reconstructing patterns from sample data. *Ann. Math. Statist.* **38** 138–154.

WALD, A. (1950). *Statistical Decision Functions.* Wiley, New York.

DEPARTMENT OF MATHEMATICS
ECOLE POLYTECHNIQUE
2500 AVENUE MARIE-GUYARD
MONTREAL 250, QUEBEC, CANADA