PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 100, Number 4, August 1987

ASYMPTOTIC BEHAVIOR OF p-PREDICTIONS FOR VECTOR VALUED RANDOM VARIABLES

JUAN A. CUESTA AND CARLOS MATRÁN

ABSTRACT. Let (Ω, σ, μ) be a probability space and let X be a B-valued μ-essentially bounded random variable, where $(B, \| \, \|)$ is a uniformly convex Banach space. Given α, a sub-σ-algebra of σ, the p-prediction $(1 < p < \infty)$ of X is defined as the best L_p-approximation to X by α-measurable random variables.

The paper proves that the Pólya algorithm is successful, i.e. the p-prediction converges to an "∞-prediction" as $p \to \infty$. First the proof is given for p-means (p-predictions given the trivial σ-algebra), and the general case follows from the characterization of the p-prediction in terms of the p-mean of the identity in B with respect to a regular conditional probability. Notice that the problem was treated in [**7**], but the proof is not satisfactory (as pointed out in [**4**]).

1. Introduction. Throughout this paper (Ω, σ, μ) denotes a probability space, $(B, \| \, \|)$ is a uniformly convex Banach space, and $L_p(\sigma) = L_p(\Omega, \sigma, \mu, B)$, $1 \leq p \leq \infty$, represents the abstract Lebesgue-Bochner L_p-space. If α is a sub-σ-algebra of σ, $L_p(\alpha)$ denotes the (closed) subspace of $L_p(\sigma)$ consisting in all the equivalence classes in $L_p(\sigma)$ containing an α-measurable function. In this notation we will not make any distinction between a random variable and the equivalence class it represents. Recall that random variables in $L_p(\alpha)$ are strongly α-measurable, i.e. they are a.s. limits of finite valued α-measurable random variables.

In [**1**] Ando and Amemiya have introduced the p-predictions given a σ-algebra for real valued random variables. In an analogous way, taking into account that $L_p(\sigma)$, $1 < p < \infty$, is uniformly convex, we may consider the p-prediction of a variable $X \in L_p(\sigma)$ given the sub-σ-algebra α as the (unique) best L_p-approximation to X by elements of $L_p(\alpha)$. Therefore the p-prediction will be continuous in $L_p(\sigma)$. However there exist important differences between the real and the abstract cases. For instance, it is well known that if $B = R$ and $p = 2$, the 2-prediction given α coincides with the conditional mean given α, while this is not true for general uniformly convex spaces. In fact the conditional mean is always linear (see Diestel and Uhl [**8**, p. 122]) but the 2-prediction is not linear unless B is a Hilbert space or Ω is the union of two μ-atoms (Herrndorf [**10**]).

This paper deals with the study of the limit of p-predictions as $p \to \infty$. In the remainder of this work we consider a fixed μ-essentially bounded random variable X (i.e. $X \in L_\infty(\sigma)$) and we will prove that the p-prediction of X given the (fixed) sub-σ-algebra α converges to a best L_∞-approximation of X by elements of $L_\infty(\alpha)$ (or ∞-prediction of X given α).

For real valued random variables this result was proved in [**3**]. Notice that the study of the convergence as $p \to \infty$ of p-predictions on uniformly convex spaces

Received by the editors June 1, 1985.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 60G30, 28B05, 60B12.

was carried out in [**7**], but the proof of the existence of the limit is not satisfactory as pointed out in [**4**]. (Also in connection with the Pólya algorithm see [**5**].) Our proof consists of two stages. First we prove that the p-prediction, $1 < p < \infty$, given α can be obtained as the p-prediction (given the trivial σ-algebra) with respect to a regular conditional probability. For notational convenience, the p-prediction given the trivial σ-algebra or best L_p-approximation by constants will be called p-mean, in connection with the real case. At a second stage we will prove that the p-mean converges, as $p \to \infty$, to the Chebysev center of the probabilistic support of X (the Midrange in the real case).

2. p-predictions and regular conditional probabilities. This section is devoted to establish the relation between the p-prediction of X and the conditional probability distribution of X given α.

Given any metric space (E, d), β_E denotes the Borel σ-algebra on E. The weak convergence of measures will be denoted by $\xrightarrow{w}$.

Recall that an α-measurable function is strongly α-measurable iff its image is a.s. contained in a separable subset of B.

Let P_X denote the probability measure induced by X on (B, β_B). There exists a unique smaller closed set in B of P_X-probability one. Moreover this set, which we denote by $S(X)$, is separable (it is called the support of the probability P_X). Therefore the set $\Gamma_{S(X)}$ of all probability measures defined on $\beta_{S(X)}$ is separable and metrizable with the topology associated to the weak convergence of measures (see Parthasarathy [**11**, p. 43]).

Now, considering the sub-σ-algebra α of σ, since $S(X)$ is separable, there exists (see Ash [**2**, p. 265]) a function $Q_\alpha \colon \Omega \times \beta_{S(X)} \to R$ such that:

(a) For every $w \in \Omega$, $Q_\alpha(w, \cdot)$ is a probability measure on $\beta_{S(X)}$ (hence we can also consider it as a probability on β_B).

(b) For every $A \in \beta_{S(X)}$, $Q_\alpha(\cdot, A)$ is a version of $\mu[X \in A/\alpha]$, the conditional distribution of X given α.

Also, since X is μ-essentially bounded, we can choose Q_α verifying

(c) For every $w \in \Omega$ the identity map on B is $Q_\alpha(w, \cdot)$-essentially bounded.

Q_α will be called a regular conditional probability (R.C.P.) of X given α, and we have

PROPOSITION 2.1. *The mapping $T \colon \Omega \to \Gamma_{S(X)}$ defined by $T(w) = Q_\alpha(w, \cdot)$ is strongly α-measurable (we consider on $\Gamma_{S(X)}$ the Borel σ-algebra associated with the weak convergence of measures).*

PROOF. Let F be the set of all continuous bounded real valued functions defined on B. Then the family of subsets of $\Gamma_{S(X)}$,

$$\left\{Q \in \Gamma_{S(X)}, \left| \int f\,dQ - \int f\,dP \right| < \delta \right\}, \qquad \delta > 0,\ f \in F,\ P \in \Gamma_{S(X)},$$

is a subbase of the topology on $\Gamma_{S(X)}$, and, since $\Gamma_{S(X)}$ is separable, it suffices to show that for $\delta > 0$, $f \in F$, and $P \in \Gamma_{S(X)}$:

$$\left\{ w/ \left| \int f(t) Q_\alpha(w, dt) - \int f(t) P(dt) \right| < \delta \right\} \in \alpha.$$

But this is obvious because the map $w \to \int f(t) Q_\alpha(w, dt)$ is α-measurable. $\square$

In the next proposition we use the Skorohod representation theorem for weak convergence of probability measures (see Skorohod [**12**]), stated for convenience in the following way.

"Let P_n, $n = 0, 1, 2, \ldots$, belong to $\Gamma_{S(X)}$ and $P_n \xrightarrow{w} P_0$. Then there exist $S(X)$-valued random variables $Y_0, Y_1, Y_2, \ldots$ defined on an appropriate probability space (W, Φ, L) such that:

(i) Y_n, $n = 0, 1, 2, \ldots$, induces on $(S(X), \beta_{S(X)})$ the probability P_n.

(ii) $Y_n \to Y_0$ L-a.s."

Now, let $Q \in \Gamma_{S(X)}$. Define $H_p(Q)$, $1 < p < \infty$, to be the p-mean, computed in $L_p(B, \beta_B, Q, B)$, of the identity on B. Then

PROPOSITION 2.2. *$H_p \colon \Gamma_{S(X)} \to B$ is a continuous map.*

PROOF. Let $Q_n \in \Gamma_{S(X)}$, $n = 0, 1, 2, \ldots$, such that $Q_n \xrightarrow{w} Q_0$ (recall that $\Gamma_{S(X)}$ is metrizable) and let Y_n, $n = 0, 1, 2, \ldots$, be the $S(X)$-valued random variables obtained from Skorohod's theorem. Also observe that the p-mean of a variable depends only on the probability induced by the variable (the law of the random variable).

But, as $S(X)$ is bounded, $Y_n \to Y_0$ also in L_p (a distinct L_p!), so the result is obtained from the continuity of the p-mean (recall that $L_p(W, \Phi, L, B)$ is also uniformly convex). □

COROLLARY 2.3. *The map $M_p \colon \Omega \to B$ defined by $M_p(w) = H_p(Q_\alpha(w, \cdot))$ is strongly α-measurable.*

PROOF. By Propositions 2.1 and 2.2 M_p is α-measurable, but moreover $M_p(\Omega) \subset H_p(\Gamma_{S(X)})$ which is separable as the image of a separable set by a continuous function. □

The fundamental theorem in this section is based in the following result.

LEMMA 2.4. *Let X be a B-valued random variable on (Ω, σ, μ) and let H: $B \times \Omega \to R$ be $\beta_B \times \alpha$-measurable such that $H(X, \mathrm{Id}) \in L_1(\Omega, \sigma, \mu, R)$. Then $\int H(t, \cdot) Q_\alpha(\cdot, dt)$ is a version of the conditional mean of $H(X, \mathrm{Id})$ given the sub-σ-algebra α of σ.*

PROOF. A standard reasoning proves the statement (begin with $H = I_{D \times A}$, where $D \in \beta_B$, $A \in \alpha$). □

THEOREM 2.5. *The random variable M_p defined in Corollary 2.3 is a version of the p-prediction of X given α (i.e. the p-prediction of X given α coincides with the p-mean of the identity with respect to Q_α).*

PROOF. Let $g \in L_p(\alpha)$. By Lemma 2.4,

$$\begin{aligned} \|X - g\|_p^p &= \int \left[\int \|t - g(w)\|^p Q_\alpha(w, dt) \right] \mu(dw) \\ &\geq \int \left[\int \|t - M_p(w)\|^p Q_\alpha(w, dt) \right] \mu(dw) = \|X - M_p\|_p^p; \end{aligned}$$

hence Corollary 2.3 and the uniqueness of p-predictions prove the result. □

3. The Pólya algorithm for p-predictions (vector valued case). The existence and uniqueness of the best L_∞-approximation by constants in uniformly convex spaces is well known (see Garkavi **[9]** or Singer **[13]**). In fact, the Chebysev center of $S(X)$ (which we denote by π_∞) is the best L_∞-approximation to X. In this section we will prove that the p_n-predictions given α converge on an "∞-prediction given α" as $p_n \to \infty$. Hence the Pólya algorithm holds for these L_p-approximations.

Some additional notation will be employed: $\pi_p(X/\alpha)$ is the p-prediction of X given α, π_p is the p-mean of X $(1 < p < \infty)$, and $V_p = \inf\{\|X - h\|_p,\ h \in B\}$ $(1 < p \leq \infty)$. Note that, obviously, $V_\infty \geq V_p$ for all p.

We need a previous theorem.

THEOREM 3.1. *If $0 < m < V_\infty$ then there exists a $\delta > 0$ such that for every $h \in B$, $\mu[w/\|X(w) - h\| \leq V_\infty - m] \leq 1 - \delta$.*

PROOF. Suppose not. Then there exist $m < V_\infty$ and a sequence $(h_k)_k$ in B such that for each k: $\mu\{\|X - h_k\| \leq V_\infty - m\} > 1 - 1/k$.

Moreover, since $S(X)$ is bounded, we can choose the sequence verifying $h_k \to h_0$ weakly in B.

Let $t \in S(X)$ and let $\tau < m/2$; the definition of $S(X)$ entails that

$$\mu\{w/\|X(w) - t\| < \tau\} > 0.$$

Therefore there exists n_0 such that for $k \geq n_0$,

$$\{\|X - h_k\| \leq V_\infty - m\} \cap \{\|X - t\| < \tau\} \neq \varnothing,$$

so $\|h_k - t\| \leq V_\infty - m/2$. But this implies $\|h_0 - t\| \leq \lim_k \|h_k - t\| \leq V_\infty - m/2$ for all $t \in S(X)$, contradicting the definition of V_∞. □

THEOREM 3.2. *There exists $\pi_\infty(X/\alpha) \in L_\infty(\alpha)$ such that $\pi_{p_n}(X/\alpha) \to \pi_\infty(X/\alpha)$ a.s. as $p_n \to \infty$, and $\|X - \pi_\infty(X/\alpha)\|_\infty \leq \|X - h\|_\infty$ for every $h \in L_\infty(\alpha)$.*

PROOF. From Theorem 2.5 it is obvious that for proving the convergence it suffices to consider the case in which α is the trivial σ-algebra. We will prove that $\pi_{p_n} \to \pi_\infty$ (the Chebysev center of $S(X)$).

If not, then there exist $\tau > 0$ and a subsequence, which we denote as the initial, such that $\|\pi_{p_n} - \pi_\infty\| > \tau$ for all n.

Let $m > 0$ and suppose that t is in B and verifies $\|t - \pi_\infty\| \leq V_\infty$ and $\|t - \pi_{p_n}\| \leq V_\infty + m$. Then, if we call θ to the modulus of convexity of B,

$$\left\|t - \tfrac{1}{2}(\pi_{p_n} + \pi_\infty)\right\| \leq [1 - \theta(\tau/(V_\infty + m))](V_\infty + m).$$

Take $m_0 > 0$. As θ is a nondecreasing function, choosing $0 < m < m_0$ small enough, there exists $r > 0$ such that

$$(*) \qquad \left\|t - \tfrac{1}{2}(\pi_\infty + \pi_{p_n})\right\| \leq V_\infty - r.$$

Since $\mu\{w/\|X(w) - \pi_\infty\| \leq V_\infty\} = 1$ by definition of V_∞, $(*)$ and Theorem 3.1, with $\frac{1}{2}(\pi_\infty + \pi_{p_n})$ as h, imply $\mu\{w/\|X(w) - \pi_{p_n}\| \leq V_\infty + m\} \leq 1 - \delta(r)$. Therefore $V_{p_n} \geq (V_\infty + m)(\delta(r))^{1/p_n}$ which contradicts that $V_\infty \geq V_p$ for all p.

Finally, the fact that $\pi_\infty(X/\alpha)$ is an ∞-prediction is obtained from Egoroff's theorem with similar techniques to those employed in **[3]**. □

REFERENCES

1. T. Ando and I. Amemiya, *Almost everywhere convergence of prediction sequences in* L_p $(1 < p < \infty)$, Z. Wahrsch. Verw. Gebiete **4** (1965), 113–120.
2. R. Ash, *Real analysis and probability*, Academic Press, New York, 1972.
3. J. A. Cuesta and C. Matrán, *Semirecorrido condicionado (expresion asintotica de la r-esperanza condicionada)*, Trabajos Estadíst Investigación Oper. **34** (1983), 54–60.
4. ——, *Conditional bounds and best* L_∞*-approximations in probability spaces*, J. Approx. Theory (to appear).
5. ——, *On the Pólya algorithm in isotonic approximation*, Preprint.
6. R. B. Darst, *Convergence of* L_p*-approximations as* $p \to \infty$, Proc. Amer. Math. Soc. **81** (1981), 433–436.
7. R. B. Darst, D. A. Legg, and D. W. Townsend, *The Pólya algorithm in* L_∞ *approximation*, J. Approx. Theory **38** (1983), 209–220.
8. J. Diestel and J. J. Uhl, *Vector measures*, Amer. Math. Soc., Providence, R.I., 1977.
9. A. L. Garkavi, *The best possible net and the best possible cross-selection of a set in a normed space*, Amer. Math. Soc. Transl. (2) **39** (1964), 111–132.
10. N. Herrndorf, *Approximation of vector-valued random variables by constants*, J. Approx. Theory **37** (1983), 175–181.
11. K. R. Parthasarathy, *Probability measures on metric spaces*, Academic Press, New York, 1967.
12. A. V. Skorohod, *Limit theorems for stochastic processes*, Theory Probab. Appl. **1** (1956), 261–290.
13. I. Singer, *Best approximation in normed linear spaces by elements of linear subspaces*, Springer-Verlag, Berlin and New York, 1970.

DEPARTAMENTO DE MATEMATICAS, UNIVERSIDAD DE SANTANDER, SANTANDER, SPAIN
(Current address of J. A. Cuesta)

DEPARTAMENTO DE ESTADISTICA, UNIVERSIDAD DE VALLADOLID, VALLADOLID, SPAIN

Current address (Carlos Matrán): Departamento de Matematicas, Universidad Autonoma de Madrid, Madrid, Spain