



---

Strong Law of Large Numbers with Respect to a Set-Valued Probability Measure

Authors(s): Madan L. Puri and Dan A. Ralescu

Source: *The Annals of Probability*, Vol. 11, No. 4 (Nov., 1983), pp. 1051-1054

Published by: Institute of Mathematical Statistics

Stable URL: <http://www.jstor.org/stable/2243518>

Accessed: 29-03-2016 15:59 UTC

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://about.jstor.org/terms>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



*Institute of Mathematical Statistics* is collaborating with JSTOR to digitize, preserve and extend access to *The Annals of Probability*

<http://www.jstor.org>

# STRONG LAW OF LARGE NUMBERS WITH RESPECT TO A SET-VALUED PROBABILITY MEASURE<sup>1</sup>

BY MADAN L. PURI AND DAN A. RALESCU

*Indiana University and University of Cincinnati*

In this paper we define the expected value of a random vector with respect to a set-valued probability measure. The concepts of independent and identically distributed random vectors are appropriately defined, and a strong law of large numbers is derived in this setting. Finally, an example of a set-valued probability useful in Bayesian inference is provided.

**1. Introduction.** This research is motivated by the following consideration: there are instances in Bayesian estimation when the prior probability is not known precisely. In such situations DeRobertis and Hartigan (1981) suggest using an interval of measures rather than a single prior, and extend the Bayes theorem in this setting. This idea is reminiscent of upper and lower probabilities (see Koopman, 1940, and Dempster, 1967). The risk  $R(\theta, \delta)$  associated with a decision function  $\delta$  is a random variable in the Bayes setting (since the unknown parameter  $\theta$  is assumed to be a random variable). The main question then is: how one can evaluate the average risk when the prior measure is not known precisely. The concept which seems to be useful in such situations is that of a set-valued measure (see Debreu and Schmeidler, 1970, and Artstein, 1972) with respect to which the expectation of a random variable is evaluated.

In Section 2 we give some preliminaries on set-valued measures, and we define the expected value. In Section 3 we prove a strong law of large numbers with respect to a set-valued probability. In Section 4 we give an example of a set-valued probability measure.

**2. Expectation with respect to a set-valued probability measure.** The concept of a set-valued measure was defined in connection with the integral of a set-valued function (see Debreu and Schmeidler, 1970).

Let  $\Omega$  be a set,  $\mathcal{A}$  a  $\sigma$ -algebra of subsets of  $\Omega$ , and  $\mathcal{P}(\mathbb{R}^n)$  the collection of all subsets of  $\mathbb{R}^n$ . A set-valued measure is a function  $\Pi: \mathcal{A} \rightarrow \mathcal{P}(\mathbb{R}^n)$ , such that (i)  $\Pi(A) \neq \emptyset$  for every  $A \in \mathcal{A}$ , (ii)  $\Pi(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \Pi(A_j)$  for every disjoint family  $\{A_j\}_j$ ,  $A_j \in \mathcal{A}$ .

Here the sum  $\sum_{j=1}^{\infty} B_j$  of subsets of  $\mathbb{R}^n$  is defined as the collection of all vectors  $b = \sum_{j=1}^{\infty} b_j$  where  $b_j \in B_j$  and  $\sum_{j=1}^{\infty} \|b_j\| < \infty$ .

In what follows we consider only bounded set-valued measures (such that  $\Pi(\Omega)$  is bounded). It follows that for such measures,  $\Pi(\emptyset) = \{0\}$ .

A selection  $\mu$  of  $\Pi$  is a vector-valued measure  $\mu: \mathcal{A} \rightarrow \mathbb{R}^n$ , such that  $\mu(A) \in \Pi(A)$  for every  $A \in \mathcal{A}$ .

An atom of the set-valued measure  $\Pi$  is an event  $A \in \mathcal{A}$  with  $\Pi(A) \neq \{0\}$  and such that  $A_1 \subset A$  implies  $\Pi(A_1) = \{0\}$  or  $\Pi(A \setminus A_1) = \{0\}$ . A set-valued measure with no atoms is called nonatomic.

The following theorem due to Artstein (1972) will be used in the sequel:

**THEOREM 2.1.** (a) *If  $\Pi$  is bounded, nonatomic set-valued measure, then  $\Pi(A)$  is convex for every  $A \in \mathcal{A}$ .*

(b) *If  $\Pi$  is bounded set-valued measure, then, for every  $A \in \mathcal{A}$  and  $x \in \Pi(A)$ , there exists a selection  $\mu$  of  $\Pi$  such that  $\mu(A) = x$ .*

---

Received October 1982.

<sup>1</sup> Research supported by the National Science Foundation Grant IST-7918468.

AMS 1980 subject classification. Primary, 60B12; secondary, 60F15.

Key words and phrases. Set-valued measure, strong law of large numbers, interval of measures.

A *set-valued probability* on  $\Omega$  is a set-valued measure  $\Pi: \mathcal{A} \rightarrow \mathcal{P}([0, 1])$  such that  $1 \in \Pi(\Omega)$ .

A *set-valued probability space* is a triple  $(\Omega, \mathcal{A}, \Pi)$  where  $\Pi$  is a set-valued probability.

Without loss of generality, one can assume that  $\Pi$  is absolutely continuous with respect to a probability measure  $P$  on  $\Omega$ ;  $\Pi \ll P$ , that is, for a set  $A \in \mathcal{A}$  for which  $P(A) = 0$ , we have  $\Pi(A) = \{0\}$ . (see Artstein, 1972).

Let  $X: \Omega \rightarrow \mathbb{R}^n$  be a random vector such that  $E_P(\|X\|) = \int_{\Omega} \|X\| dP < \infty$ . The *expected value* of  $X$  with respect to  $\Pi$  is defined as  $\int_{\Omega} X d\Pi = \{\int_{\Omega} X d\mu: \mu \text{ is a selection of } \Pi\}$ . According to Theorem 2.1 (b) it is clear that  $\int_{\Omega} X d\Pi \neq \phi$  if  $E_P(\|X\|) < \infty$ .

**3. Strong law of large numbers.** Let  $(\Omega, \mathcal{A}, \Pi)$  be a set-valued probability space, and let  $X: \Omega \rightarrow \mathbb{R}^n$  be a random vector. Then  $X$  induces a set-valued probability on the Borel sets in  $\mathbb{R}^n$  (denoted by  $\mathcal{B}_n$ ) in the following way:

$$B \in \mathcal{B}_n, \quad \Pi_X(B) = \Pi(X \in B).$$

The random vectors  $X_i, i \geq 1$  defined on  $(\Omega, \mathcal{A}, \Pi)$  are *independent* if  $\Pi(X_1 \in B_1, X_2 \in B_2, \dots, X_i \in B_i) = \Pi(X_1 \in B_1) \cdots \Pi(X_i \in B_i)$  where the product of subsets  $M$  and  $N$  of  $[0, 1]$  is defined by  $MN = \{mn: m \in M, n \in N\}$ . They are *identically distributed* if  $\Pi_{X_1} = \dots = \Pi_{X_i} = \dots$ . Clearly these concepts generalize the classical concepts of independent and identically distributed random vectors (with respect to an ordinary probability measure).

Finally, we need another notation: if  $x \in \mathbb{R}^n$ , and  $A \subset \mathbb{R}^n$ , then

$$d(x, A) = \inf_{a \in A} \|x - a\|.$$

We now prove our main theorem.

**THEOREM 3.1.** *Let  $X_i, i \geq 1$  be independent and identically distributed random vectors defined on a set-valued probability space  $(\Omega, \mathcal{A}, \Pi)$  such that  $\Pi \ll P$  where  $P$  is a probability measure. If  $E_P(\|X_1\|) < \infty$ , then  $d((1/n) \sum_{j=1}^n X_j, \int_{\Omega} X_1 d\Pi) \rightarrow 0$  almost everywhere with respect to  $\Pi$ .*

**PROOF.** Clearly, if  $\mu$  is a selection of  $\Pi$  (which exists according to Theorem 2.1 (b)), it is not true in general that  $X_i, i \geq 1$  are independent and identically distributed with respect to  $\mu$ .

To prove the theorem, we will show the following:

There exists a probability measure  $Q$  on  $\Omega$  which is a selection of  $\Pi$  and such that  $X_i, i \geq 1$  are independent and identically distributed with respect to  $Q$ , and  $E_Q(\|X_1\|) < \infty$ .

Let  $Q(A) = \sup \Pi(A)$  for every  $A \in \mathcal{A}$ . The fact that  $Q$  is a probability measure follows from Proposition 3.1 of Artstein (1972). Also it is clear that  $Q(X_i \in B) = Q(X_1 \in B)$  for every  $B \in \mathcal{B}$ , and  $i \geq 1$  i.e.,  $X_i, i \geq 1$  are identically distributed.

To prove that  $X_i, i \geq 1$  are independent, it suffices to show that  $Q(X_1 \in B_1, X_2 \in B_2) = Q(X_1 \in B_1)Q(X_2 \in B_2)$  for every  $B_1, B_2 \in \mathcal{B}_n$ . By the definition of independence, it suffices to show that  $\sup(MN) = \sup M \sup N$  where  $M, N \subset [0, 1]$ . This being easy to establish, the desired independence follows.

Since  $\Pi \ll P$ , it follows from classical results that  $E_P(\|X_1\|) < \infty$  implies  $\int_{\Omega} \|X_1\| dQ < \infty$ .

We now prove that  $Q$  is a selection of  $\Pi$ . From Theorem 2.1(b) there exists a probability measure  $Q_1$  which is a selection of  $\Pi$ . Clearly  $Q_1(A) \leq Q(A)$  for every  $A \in \mathcal{A}$ , but this implies that  $Q_1 = Q$ .

Now from the classical law of large numbers, it follows that  $(1/n) \sum_{j=1}^n X_j \rightarrow E_Q(X_1)$  almost everywhere with respect to  $Q$ . Thus

$$d\left(\frac{1}{n} \sum_{j=1}^n X_j, \int_{\Omega} X_1 d\Pi\right) \leq \left\| \frac{1}{n} \sum_{j=1}^n X_j - E_Q(X_1) \right\| \rightarrow 0,$$

and so  $d((1/n) \sum_{j=1}^n X_j, \int_{\Omega} X_1 d\Pi) \rightarrow 0$  almost everywhere ( $Q$ ).

The definition of  $Q$  implies that the above convergence actually holds almost everywhere with respect to  $\Pi$ , that is,  $\Pi(\sum_{j=1}^n X_j/n \not\rightarrow \int_{\Omega} X_1 d\Pi) = \{0\}$ . This completes the proof.

**4. An example.** The strong law of large numbers proved in Section 3 is a generalization of the classical law of large numbers.

The simplest example of a set-valued probability measure is provided by an interval of measures (as studied by DeRobertis and Hartigan, 1981). More precisely let  $P_1$  and  $P_2$  be two finite measures on  $(\Omega, \mathcal{A})$  such that  $P_1(A) \leq P_2(A)$  for every  $A \in \mathcal{A}$ , and let  $P_2$  be a probability measure. Let  $\Pi: \mathcal{A} \rightarrow \mathcal{P}([0, 1])$  be defined as

$$(4.1) \quad \Pi(A) = [P_1(A), P_2(A)], \quad A \in \mathcal{A}.$$

Clearly  $\Pi(\phi) = \{0\}$  and  $1 \in \Pi(\Omega)$ .

Let  $\{A_j\}_j$  be a disjoint family,  $A_j \in \mathcal{A}$ . We must show that  $\Pi(\cup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \Pi(A_j)$ . This is equivalent to

$$\sum_{j=1}^{\infty} [P_1(A_j), P_2(A_j)] = [\sum_{j=1}^{\infty} P_1(A_j), \sum_{j=1}^{\infty} P_2(A_j)].$$

The above equality follows from the formulas

$$\inf \sum_{j=1}^{\infty} [P_1(A_j), P_2(A_j)] = \sum_{j=1}^{\infty} P_1(A_j),$$

$$\sup \sum_{j=1}^{\infty} [P_1(A_j), P_2(A_j)] = \sum_{j=1}^{\infty} P_2(A_j),$$

and from the convexity of  $\sum_{j=1}^{\infty} [P_1(A_j), P_2(A_j)]$ .

Thus  $\Pi$  defined by (4.1) is a set-valued probability. Also  $\Pi$  is absolutely continuous with respect to  $P_2$ .

If  $X: \Omega \rightarrow \mathbb{R}^n$  is a random vector such that  $E_{P_2}(\|X\|) < \infty$ , then the expected value of  $X$  (as defined in Section 2) is given by  $\int_{\Omega} X d\Pi = \{E_P(X) : P_1 \leq P \leq P_2\}$  where  $P$  is a finite measure.

If  $X_i, i \geq 1$  are independent and identically distributed with respect to  $\Pi$  (given by (4.1)) and note that the latter condition is equivalent to the fact that  $X_i, i \geq 1$  are identically distributed with respect to  $P_1$  and  $P_2$ , then the law of large numbers given by Theorem 3.1 implies

$$\inf_{P_1 \leq P \leq P_2} \left\| \frac{1}{n} \sum_{j=1}^n X_j - E_P(X_1) \right\| \rightarrow 0 \quad \text{almost everywhere with respect to } P_2.$$

It is interesting to note that, under certain hypotheses, every set-valued probability is of the form (4.1).

**THEOREM 4.1.** *Let  $\Pi: \mathcal{A} \rightarrow \mathcal{P}([0, 1])$  be a nonatomic set-valued probability measure such that  $\Pi(\Omega)$  is closed. Then  $\Pi(A) = [P_1(A), P_2(A)]$  for every  $A \in \mathcal{A}$ , where  $P_1$  is a measure and  $P_2$  is a probability measure such that  $P_1(A) \leq P_2(A), A \in \mathcal{A}$ .*

**PROOF.** Denote  $P_1(A) = \inf \Pi(A)$  and  $P_2(A) = \sup \Pi(A), A \in \mathcal{A}$ . We show that  $P_1$  and  $P_2$  are measures. Let  $\{A_j\}_j$  be a disjoint family of sets in  $\mathcal{A}$ . Then, clearly

$$(4.2) \quad P_1(\cup_{j=1}^{\infty} A_j) = \inf(\sum_{j=1}^{\infty} \Pi(A_j)) \geq \sum_{j=1}^{\infty} \inf \Pi(A_j) = \sum_{j=1}^{\infty} P_1(A_j).$$

Let  $\varepsilon > 0$ . Then there exists  $x_j \in \Pi(A_j)$  such that  $x_j < \inf \Pi(A_j) + \varepsilon/2^j, j \geq 1$ . Thus  $\sum_{j=1}^{\infty} x_j \leq \sum_{j=1}^{\infty} \inf \Pi(A_j) + \varepsilon$ . From (4.2), the series  $\sum_{j=1}^{\infty} \inf \Pi(A_j)$  is convergent. So  $\sum_{j=1}^{\infty} x_j \in \sum_{j=1}^{\infty} \Pi(A_j)$ . Consequently

$$(4.3) \quad \inf(\sum_{j=1}^{\infty} \Pi(A_j)) \leq \sum_{j=1}^{\infty} \inf \Pi(A_j) + \varepsilon.$$

Since  $\varepsilon > 0$  is arbitrary, it follows from (4.2) and (4.3) that

$$P_1(\cup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} P_1(A_j).$$

Now  $P_1(\Omega) = \inf \Pi(\Omega) \in \Pi(\Omega)$  since  $\Pi(\Omega)$  is closed. From Theorem 2.1(b), there exists a selection  $Q_1$  of  $\Pi$  such that  $Q_1(\Omega) = P_1(\Omega)$ . Since  $P_1 \leq Q_1$ , we have  $P_1 = Q_1$ , so  $P_1$  is a

selection of  $\Pi$ . Similarly  $P_2$  is a selection of  $\Pi$ , and obviously  $P_2(\Omega) = 1$ . Finally, since (from Theorem 2.1 (a))  $\Pi(A)$  is convex for every  $A \in \mathcal{A}$ , it follows that  $\Pi(A) = [P_1(A), P_2(A)]$ ,  $A \in \mathcal{A}$ , which was to be proved.

**REMARK.** It may be noted that the intervals of probability measures (DeRobertis and Hartigan, op. cit.) are restandardized when computing expectations and posterior probabilities, so that the ranges of expectations for set-valued probabilities and for intervals of probabilities do not coincide.

## REFERENCES

- [1] ARTSTEIN, Z. (1972). Set-valued measures. *Trans. Amer. Math. Soc.* **165** 103–125.
- [2] DEBREU, G. and SCHMEIDLER, D. (1970). The Radon-Nikodym derivative of a correspondence. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.*, 41–56. Univ. of California Press.
- [3] DEMPSTER, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Statist.* **36** 325–339.
- [4] DEROBERTIS, L. and HARTIGAN, J. A. (1981). Bayesian inference using intervals of measures. *Ann. Statist.* **9** 235–244.
- [5] KOOPMAN, B. O. (1940). The axioms and algebra of intuitive probability. *Ann. Math.* **41** 269–278.

DEPARTMENT OF MATHEMATICS  
INDIANA UNIVERSITY  
BLOOMINGTON, INDIANA 47405

DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF CINCINNATI  
CINCINNATI, OHIO 45221