Z. Wahrscheinlichkeitstheorie verw. Geb. 21, 233-240 (1972)

Independent Conditional Expectations and $\mathscr{L}$ log $\mathscr{L}$

J. G. BASTERFIELD

1. Introduction

Let $(\Omega, \mathscr{F}, P)$ denote a probability triple, and $\{\mathscr{F}_n\}$ an independent sequence of sub-σ-algebras of $\mathscr{F}$ i.e., for $k=1, 2, \dots$ and $F_j \in \mathscr{F}_j$ $(1 \leqq j \leqq k)$,

$$P(F_1 \cap F_2 \cap \cdots \cap F_k) = P(F_1) P(F_2) \dots P(F_k).$$

We shall use X to denote a random variable on $(\Omega, \mathscr{F}, P)$ and X_n to denote $E(X|\mathscr{F}_n)$, its conditional expectation w.r.t. $\mathscr{F}_n$. All the X we shall consider will belong to $\mathscr{L}^1$, and we denote $E(X)$ by μ. Our principal result is the following

Theorem 1. *If* $E(|X| \log^+ |X|) < \infty$ *then*

$$P[X_n \to \mu \text{ as } n \to \infty] = 1.$$

We also give an example (due to David Williams) which shows that the theorem becomes false if $E(|X| \log^+ |X|) < \infty$ is replaced by $E(|X|) < \infty$. We conjecture but are unable to prove that the condition $E(|X| \log^+ |X|) < \infty$ is best possible.

These results appear to contrast with the characterisations of $\mathscr{L} \log \mathscr{L}$ discussed in Gundy [2] where the sequence of σ-algebras considered is essentially monotone, and the method of proof an application of martingale theory. We use the theory of Orlicz spaces to establish Theorem 1.

What does follow from the Martingale Convergence Theorem is that *if* $E(|X|) < \infty$, *then* $X_n \to \mu$ *in* $\mathscr{L}^1$ *norm and therefore in probability.* For let $Y_n = E(X|\mathscr{G}_n)$ where

$$\mathscr{G}_n = \bigvee_{m \geqq n} \mathscr{F}_m.$$

Then, by the martingale theorem, $Y_n \to Y_\infty$ with probability one. By Kolmogorov's Zero-One Law, Y_∞ is constant with probability one. Since also $Y_n \to Y_\infty$ in $\mathscr{L}^1$ norm, we see that $Y_\infty = \mu$ with probability one. However

$$X_n - \mu = E([Y_n - \mu] | \mathscr{F}_n)$$

so that

$$E(|X_n - \mu|) \leqq E(|Y_n - \mu|) \to 0.$$

2. Some Orlicz Space Results

Proof of Theorem 1. We shall need some results from the theory of Orlicz spaces. The function

$$\phi(u) = \begin{cases} 0 & \text{for } 0 < u \leqq 1 \\ \log u + 1 & \text{for } 1 < u < \infty \end{cases}$$

has inverse (if we observe the appropriate convention where the inverse in the strict sense is not defined)

$$\psi(v)=\begin{cases}0 & \text{for } v=0\\ 1 & \text{for } 0<v\leqq 1\\ \exp(v-1) & \text{for } 1<v<\infty.\end{cases}$$

Hence the functions

$$\Phi(u)=\int_0^u \phi(t)\,dt=\begin{cases}0 & \text{for } 0\leqq u\leqq 1\\ u\log u & \text{for } 1\leqq u<\infty\end{cases}$$

and

$$\Psi(v)=\int_0^v \psi(t)\,dt=\begin{cases}v & \text{for } 0\leqq v\leqq 1\\ \exp(v-1) & \text{for } 1<v<\infty\end{cases}$$

are complementary functions in the sense of Young, so we have (Theorem 5.4.1 of Zaanen [3]):

Young's Inequality

$$uv\leqq\Phi(u)+\Psi(v) \quad \text{for all } u, v\geqq 0.$$

We shall sometimes denote $\Phi(u)$ by $u\log^+ u$, as is usual.

Let $\mathscr{L}_\Phi^*$ denote the set of all measurable functions U on Ω such that $\int_\Omega \Phi(|U|)\,dP<\infty$, and similarly for $\mathscr{L}_\Psi^*$. Let $\mathscr{L}_\Phi=\mathscr{L}_\Phi(\Omega,\mathscr{F},P)$ be the set of all measurable functions U on Ω such that

$$\|U\|_\Phi=\sup\Big\{\int_\Omega |UV|\,dP\colon \int_\Omega \Psi(|V|)\,dP\leqq 1\Big\}<\infty$$

and (interchanging Φ and Ψ) similarly for $\mathscr{L}_\Psi$. Then (Theorems 5.5.1 and 5.6.1 of Zaanen [3]) $\mathscr{L}_\Phi$ and $\mathscr{L}_\Psi$ are Banach spaces, and if $U\in\mathscr{L}_\Phi^*$, $U\in\mathscr{L}_\Phi$ and

$$\|U\|_\Phi\leqq\int_\Omega \Phi(|U|)\,dP+1$$

with a corresponding result for $\mathscr{L}_\Psi$.

Also (Corollary to Theorem 5.5.2 of Zaanen [3]) the functions which belong to $\mathscr{L}_\Phi$ are precisely those which belong to $\mathscr{L}_\Phi^*$. (Note: the corresponding result for $\mathscr{L}_\Psi$ and $\mathscr{L}_\Psi^*$ is false.)

Lemma 1. *The bounded functions are dense in* $\mathscr{L}_\Phi$.

Proof. If $U\in\mathscr{L}_\Phi$ then $U\in\mathscr{L}_\Phi^*$ i.e. $\int_\Omega |U|\log^+|U|\,dP<\infty$. Choose $K>2$ such that

$$\int_{|U|>K} |U|\log^+|U|\,dP<4^{-r}.$$

Let $U=U_1+U_2$, where $U_1=0$ if $|U|<K$ and $U_2=0$ if $|U|\geqq K$. Then it is sufficient to show that we can make $\|U_1\|_\Phi$ small by a suitable choice of r.

Since

$$\Phi(0)=0$$

and

$$\Phi(2u)=2u(\log u+\log 2)<4\Phi(u) \quad \text{if } u>2,$$

then

$$\Phi(2^r|U_1|)\leqq 4^r\Phi(|U_1|) \quad \text{for all } \omega.$$

Now, by Young's inequality, for any V

$$\begin{aligned} 2^r|U_1 V| &\leqq \Phi(2^r|U_1|)+\Psi(|V|) \\ &\leqq 4^r \Phi(|U_1|)+\Psi(|V|). \end{aligned}$$

Thus

$$\begin{aligned} \int_\Omega |2^r U_1 V|\,dP &\leqq 4^r \int_\Omega \Phi(|U_1|)\,dP+\int_\Omega \Psi(|V|)\,dP \\ &\leqq 1+1 \quad \text{if } \int_\Omega \Psi(|V|)\,dP \leqq 1. \end{aligned}$$

Therefore $\|2^r U_1\|_\Phi \leqq 2$, so $\|U_1\|_\Phi \leqq 2^{1-r}$ which completes the proof of the lemma.

3. Application of a Theorem of Banach

We shall require the following theorem of Banach (Theorem IV.11.2 of Dunford and Schwartz [1]):

Let $\{T_n\}$ be a sequence of continuous linear maps from a Banach space $\mathscr{X}$ into the space $M(\Omega)$ of all measurable functions on Ω with the topology of convergence in probability. Suppose that for each $x \in \mathscr{X}$ we have $\sup_n |T_n(x,\omega)| < \infty$ a.s. Suppose also that for each x in a dense set in $\mathscr{X}$ the limit $\lim_{n\to\infty} T_n(x,\omega)$ exists a.s. Then for every $x\in\mathscr{X}$, $\lim_{n\to\infty} T_n(x,\omega)$ exists a.s.

We now establish that $E(.|\mathscr{F}_n)$: $\mathscr{L}_\Phi(\Omega)\to M(\Omega)$ is a continuous linear map. The linearity is obvious, so it suffices to prove continuity at 0. Since the constant function 1 has $\int_\Omega \Psi(1)\,dP=1$, it follows that

$$\|U\|_\Phi \geqq \int_\Omega |U\cdot 1|\,dP = \|U\|_1.$$

But $E(.|\mathscr{F}_n)$: $\mathscr{L}^1\to\mathscr{L}^1$ is of norm 1, so $\|U\|_\Phi \geqq \|E(U|\mathscr{F}_n)\|_1$. Hence $\|U\|_\Phi\to 0$ implies that $\|E(U|\mathscr{F}_n)\|_1\to 0$ and a fortiori that $E(U|\mathscr{F}_n)\to 0$ in probability.

Lemma 2 (J. Galambos)[1]. *If $E(X^2)<\infty$ then*

$$P[X_n\to\mu]=1.$$

Proof of Lemma 2. We may as well suppose $\mu=0$, or else replace X by $X-\mu$. Thus X belongs to the Hilbert space $\mathscr{L}_0^2=\mathscr{L}_0^2(\Omega,\mathscr{F},P)$ consisting of those elements of $\mathscr{L}^2$ which have mean zero. Now $X_n=E[X|\mathscr{F}_n]$ is the orthogonal projection of X onto the space $\mathscr{L}_0^2(\Omega,\mathscr{F}_n,P)$ of $\mathscr{F}_n$-measurable functions in $\mathscr{L}_0^2$. Because $\{\mathscr{F}_n\}$ is independent, the projections $E[.|\mathscr{F}_n]$ are orthogonal one to another. By Bessel's inequality,

$$E(X^2)=\|X\|_2^2\geqq \sum \|X_n\|_2^2=\sum E(X_n^2)=E\sum X_n^2.$$

Thus $\sum X_n^2$ converges with probability 1. Hence

$$P[X_n^2\to 0]=1.$$

which completes the proof of the lemma.

---

[1] Sieve methods in the theory of probability and in the theory of numbers. Doctoral thesis, L. Eötvös University, Budapest 1963.

$\mathscr{L}^2$ contains all the bounded functions and so (Lemma 1) is dense in $\mathscr{L}_\Phi$. Hence, applying the theorem of Banach quoted above, it follows that either $\lim X_n$ exists a.s. for all $X \in \mathscr{L}_\Phi$ or there exists an X in $\mathscr{L}_\Phi(\Omega)$ with

$$P[\sup_n |X_n(\omega)| = \infty] > 0. \tag{1}$$

If the first alternative obtains then $\lim X_n = \mu$ with probability one because, as we saw in the Introduction, $X_n \to \mu$ in probability. Thus the first alternative implies Theorem 1.

We shall show that the second alternative leads to a contradiction.

If (1) holds for X it must hold for X^+ or X^- (or both) so we may assume it to hold for a positive X (when the X_n may be taken positive also).

Since $X_n(\omega) < \infty$ a.s. we must have

$$P[\limsup_{n\to\infty} X_n(\omega) = \infty] > 0.$$

However

$$\{\omega \colon \limsup_{n\to\infty} X_n(\omega) > c\} \in \mathscr{F}_\infty$$

for any constant c, where

$$\mathscr{F}_\infty = \bigcap_{m \geqq 1} \bigvee_{n \geqq m} \mathscr{F}_n$$

is the tail σ-algebra of the sequence $\{\mathscr{F}_n\}$. By Kolmogorov's Zero-One Law all sets in $\mathscr{F}_\infty$ have probability either zero or one so

$$P[\limsup_{n\to\infty} X_n(\omega) = \infty] = 1.$$

Let us consider first the special case in which each $\mathscr{F}_n$ has the form $\mathscr{F}_n = \{\emptyset, A_n, \bar{A}_n, \Omega\}$ (where $A_n \in \mathscr{F}$ and $\bar{A}_n$ denotes $\Omega \setminus A_n$), and for notational convenience (interchanging A_n and $\bar{A}_n$ if necessary) take $E(X|A_n) \geqq E(X|\bar{A}_n)$. Then, since $E(X_n) = E(X)$, $E(X|\bar{A}_n) \leqq E(X)$. Denote $P(A_n)$ by α_n.

Lemma 3. *A subsequence $\{n_r\colon r = 1, 2, 3, \ldots\}$ may be selected such that*

(i) $$\sum_{r=1}^{\infty} \alpha_{n_r} = \infty$$

and

(ii) $$E(X|A_{n_r}) \to \infty \quad \textit{as} \quad r \to \infty.$$

These imply

$$P[\limsup_{r\to\infty} X_{n_r} = \infty] = 1.$$

Proof. For $c > E(X)$ let $N(c) = \{n\colon E(X|A_n) > c\}$. Note that $E(X|\bar{A}_n) \leqq E(X) < c$, so $X_n(\omega) > c$ implies (with the possible exception of a set of probability zero) that $\omega \in A_n$ and $E(X|A_n) > c$. Since

$$P[\limsup_{n\to\infty} X_n(\omega) = \infty] = 1$$

it follows from one of the Borel-Cantelli Lemmas that

$$\sum_{n \in N(c)} \alpha_n = \infty.$$

Put $c=2\mu$ and take $n_1<n_2<\cdots<n_k$, the first k members of $N(2\mu)$, where k is the least integer such that

$$\alpha_{n_1}+\alpha_{n_2}+\cdots+\alpha_{n_k}>1.$$

Now put $c=3\mu$. Since

$$\sum_{n\in N(3\mu),\, n>n_k} \alpha_n=\infty$$

it is possible to take $n_{k+1}<\cdots<n_l$, the first $l-k$ members of $N(3\mu)$ which exceed n_k, where l is the least integer such that

$$\alpha_{n_{k+1}}+\cdots+\alpha_{n_l}>1.$$

Now take $c=4\mu$ etcetera. The sequence

$$n_1,\ldots,n_k,n_{k+1},\ldots,n_l,n_{l+1},\ldots$$

has the properties

(i)
$$\sum_{r=1}^{\infty}\alpha_{n_r}=\infty$$

and

(ii)
$$E(X|A_{n_r})\to\infty \quad \text{as} \quad r\to\infty.$$

By virtue of (i) and the independence of the A_{n_r}, almost all ω belong to an infinite number of A_{n_r} (Borel-Cantelli), so (ii) ensures

$$\limsup_{r\to\infty} X_{n_r}=\infty \quad \text{a.s.}$$

as required to complete the proof of the lemma.

Thus it is sufficient to consider the case

$$E(X|A_n)\to\infty \quad \text{as} \quad n\to\infty, \quad \sum\alpha_n=\infty.$$

Trivially, we may demand that $\alpha_n>0$ for all n. Now,

$$E(X|A_n)=\frac{1}{\alpha_n}\int_{A_n} X\,dP.$$

So (since $\sum\alpha_n=\infty$)

$$\frac{\sum_{n=1}^{m}\int_{A_n} X\,dP}{\sum_{n=1}^{m}\alpha_n}\to\infty \quad \text{as} \quad m\to\infty,$$

i.e.,

$$\int_{\Omega}\frac{\sum_{n=1}^{m} Y_n}{\sum_{n=1}^{m}\alpha_n} X\,dP\to\infty \quad \text{as} \quad m\to\infty$$

(where Y_n is the characteristic function of A_n).

Thus, as $m \to \infty$, $\|Z_m X\|_1 \to \infty$, where

$$Z_m = \sum_{n=1}^{m} Y_n \Big/ \sum_{n=1}^{m} \alpha_n.$$

We now need the following result.

Lemma 4. $\|Z_m\|_\Psi < L < \infty$ *for all m, where the value of the constant L need only depend on* α_1.

We now complete the proof of Theorem 1 assuming Lemma 4, deferring the proof of Lemma 4 until later.

By Theorem 5.5.3 of Zaanen [3]

$$\|XZ_m\|_1 \leqq \|X\|_\Phi \|Z_m\|_\Psi.$$

So $\|XZ_m\|_1 \leqq \|X\|_\Phi L$ which contradicts the previous statement that $\|XZ_m\|_1 \to \infty$ as $m \to \infty$.

This contradiction completes the proof of Theorem 1 in the case when $\mathscr{F}_n$ has the form $\{\emptyset, A_n, \bar{A}_n, \Omega\}$.

Now suppose only that the $\mathscr{F}_n$ are independent and that $X \in \mathscr{L}_\Phi$. Put

$$A_n(\varepsilon) = \{\omega\colon X_n(\omega) \geqq \mu + \varepsilon\}.$$

Then

$$E[X|A_n(\varepsilon)] = E[X_n|A_n(\varepsilon)]$$

because $A_n(\varepsilon) \in \mathscr{F}_n$. Hence

$$E[X|A_n(\varepsilon)] \geqq \mu + \varepsilon.$$

But now, by the special case of the Theorem already proved,

$$P[\omega \in A_n(\varepsilon) \text{ i.o.}] = 0.$$

Hence

$$P[\limsup_{n\to\infty} X_n \leqq \mu + \varepsilon] = 1$$

and similarly

$$P[\liminf_{n\to\infty} X_n \geqq \mu - \varepsilon] = 1.$$

This proves the Theorem in the case of general independent $\mathscr{F}_n$, except that it remains to prove Lemma 4.

4. Proof of Lemma 4

Proof of Lemma 4. Since $Z_m \geqq 0$, $|Z_m| = Z_m$. We have

$$\|Z_m\|_\Psi \leqq \int_\Omega \Psi(Z_m)\, dP + 1$$

if the R.H.S. is finite, so it suffices to show that $E[\Psi(Z_m)]$ is bounded.

Since $\Psi(v) \leqq \exp v$ for all v, it further suffices to show that $E[\exp(Z_m)]$ is bounded.

Denote $\sum_{n=1}^{m} \alpha_n$ by s_m.

Then

$$E[\exp(Z_m)] = E\left[\exp\left(\sum_{n=1}^{m} Y_n/s_m\right)\right]$$
$$= E\left[\prod_{n=1}^{m} \exp(Y_n/s_m)\right]$$
$$= \prod_{n=1}^{m} E[\exp(Y_n/s_m)]$$
$$= \prod_{n=1}^{m} [1+\alpha_n\{\exp(1/s_m)-1\}].$$

So

$$\log E[\exp(Z_m)] = \sum_{n=1}^{m} \log[1+\alpha_n\{\exp(1/s_m)-1\}].$$

Now

$$\log(1+x) \leqq x \quad \text{for } x \geqq 0,$$

so

$$\log E[\exp(Z_m)] \leqq \sum_{n=1}^{m} \alpha_n\{\exp(1/s_m)-1\}.$$

But we may choose a constant K depending only on α_1 so that

$$e^x - 1 < Kx \quad \text{for } 0 \leqq x \leqq 1/\alpha_1.$$

Hence

$$\log E[\exp(Z_m)] \leqq \sum_{n=1}^{m} \alpha_n K/s_m$$
$$= K.$$

Therefore

$$E[\exp(Z_m)] \leqq L = e^K$$

as required.

5. A Counterexample

We now give an example to show that if we replace the condition

$$E(|X| \log^+ |X|) < \infty \quad \text{by} \quad E(|X|) < \infty$$

the conclusion of Theorem 1 does not hold.

Throughout the remainder of the paper let $(\Omega, \mathscr{F}, P)$ be a triple supporting a sequence $\{A_n\colon n=1,2,3,\ldots\}$ of independent events A_n with $P(A_n)=1/n$.

Let $\mathscr{F}_n = \{\emptyset, A_n, \bar{A}_n, \Omega\}$ and put $X = \sum_{m=1}^{\infty} m!\, W_{m+2}$ where W_m is the characteristic function of $\bigcap_{k=1}^{m} A_k$. Then

$$E(W_m) = P\left(\bigcap_{k=1}^{m} A_k\right) = \prod_{k=1}^{m} P(A_k) = 1/m!.$$

So

$$E(X) = \sum_{m=1}^{\infty} m!/(m+2)! = \sum_{m=1}^{\infty} \left(\frac{1}{m+1} - \frac{1}{m+2}\right) = \tfrac{1}{2}.$$

Now

$$E(X|A_n)=\frac{1}{P(A_n)}\int_{A_n} X\,dP.$$
$$=n\int_{A_n} X\,dP.$$

So

$$\frac{1}{n}E(X|A_n)=\int_{A_n} X\,dP$$
$$=\int_{A_n}\sum_{m=1}^{n-3} m!\,W_{m+2}\,dP+\int_{A_n}\sum_{m=n-2}^{\infty} m!\,W_{m+2}\,dP$$
$$=\frac{1}{n}\sum_{m=1}^{n-3} m!\,E(W_{m+2})+\sum_{m=n-2}^{\infty}\frac{m!}{(m+2)!}$$
$$=\frac{1}{n}\sum_{m=1}^{n-3}\frac{1}{(m+1)(m+2)}+\sum_{m=n-2}^{\infty}\frac{1}{(m+1)(m+2)}$$
$$=\frac{1}{n}\left(\frac{1}{2}-\frac{1}{n-1}\right)+\frac{1}{n-1}.$$

Hence

$$E(X|A_n)=\frac{1}{2}-\frac{1}{n-1}+\frac{n}{n-1}$$
$$=\tfrac{3}{2}.$$

Since $\sum_{n=1}^{\infty} P(A_n)=\infty$, almost all ω belong to an infinite number of A_n (Borel-Cantelli). Hence

$$\limsup_{n\to\infty} X_n=\tfrac{3}{2}+\mu,$$

which is the required result.

I should like to thank Dr. David Williams for much helpful discussion as well as for the counter-example, and Dr. F. Papangelou for pointing out the relevance of the theorem of Banach used in the proof. I should also like to thank Prof. R. Pyke for the proof (considerably shorter than my original one) of Lemma 4 given above.

References

1. Dunford, N., Schwartz, J.T.: Linear operators (part I). New York: Interscience 1958.
2. Gundy, R.F.: On the class $L\log L$, martingales, and singular integrals. Studia math. **33**, 109–118 (1969).
3. Zaanen, A.C.: Linear analysis. Amsterdam: North Holland 1953.

J.G. Basterfield
Department of Statistics
University College of Wales
Penglais, Aberystwyth, U.K.

(Received May 18, 1971)