Convergence of Conditional Expectations in Banach Function Spaces

Masato Kikuchi

Department of Mathematics, Toyama University, 3190 Gofuku, Toyama 930-8555, Japan

Submitted by Ulrich Stadtmueller

Received October 5, 1998

We give a necessary and sufficient condition that assures the convergence of conditional expectations in a Banach function space. This extends the result of A. Alonso and F. Brambila-Paz [J. Math. Anal. Appl. 221 (1998), 161-176]. We also consider the convergence of conditional expectations in a weighted $L^p$-spaces.

Key Words: conditional expectation; Banach function space.

1. INTRODUCTION

Let $(\Omega, \mathcal{A}, \mu)$ be a complete probability space and $\mathcal{F} = \{\mathcal{F}_n\}_{n=1}^\infty$ be a sequence of sub-$\sigma$-fields of $\mathcal{A}$. Alonso and Brambila-Paz [1] recently gave a necessary and sufficient condition in order that the sequence $\{E[f | \mathcal{F}_n]\}_{n=1}^\infty$ converges in $L^p$ for every $f \in L^p$. Following the notation in [1], we write $\mathcal{F}_n \overset{L}{\longrightarrow} \mathcal{F}$ if, for every $A \in \mathcal{F}$, there exists a sequence $\{A_{n_k}\}_{k=1}^\infty$ of sets such that $A_{n_k} \in \mathcal{F}_n$ for each $n \geq 1$ and $1_{A_{n_k}} \to 1_A$ in $L^1$ (or equivalently in probability), and $\mathcal{F}_n \overset{\mu}{\rightarrow} \mathcal{F}$ if $1_{A_{n_k}} - E[1_{A_{n_k}} | \mathcal{F}_n] \to 0$ weakly in $L^2$ (or weakly in $L^1$ equivalently) for any sequence $\{A_{n_k}\}_{k=1}^\infty$ with $A_{n_k} \in \mathcal{F}_n$, $n \geq 1$. Let $\mathcal{F}_\mu$ denote the $\sigma$-field consisting of all $A \in \mathcal{A}$ such that $1_{A_{n_k}} \to 1_A$ in $L^1$ for some $\{A_{n_k}\}_{k=1}^\infty$ with $A_{n_k} \in \mathcal{F}_n$, $n \geq 1$, and $\mathcal{F}_\perp$ the $\sigma$-field generated by functions $g \in L^2$ such that $1_{A_{n_k}} \to g$ weakly in $L^2$ for some $\{A_{n_k}\}_{k=1}^\infty$ such that $A_{n_k} \in \mathcal{F}_n$. Then we have:

**Theorem A** ([1]). Let $\mathcal{F} = \{\mathcal{F}_n\}_{n=1}^\infty$ be a sequence of sub-$\sigma$-fields of $\mathcal{A}$ and $\mathcal{G}$ a sub-$\sigma$-field of $\mathcal{A}$. The following conditions are equivalent:

(i) If $1 \leq p < \infty$, then $E[f | \mathcal{F}_n] \to E[f | \mathcal{G}]$ in $L^p$ for every $f \in L^p$;

(ii) $\mathcal{F}_n \overset{L}{\longrightarrow} \mathcal{G}$ and $\mathcal{F}_n \overset{\mu}{\rightarrow} \mathcal{G}$;

(iii) $\mathcal{F}_\mu = \mathcal{F}_\perp$. 

193
We shall use only the equivalence between (i) and (ii) in what follows; an alternative proof of the equivalence will be given in Appendix.

In this note, we shall investigate the convergence of conditional expectations in more general Banach spaces $X$ of measurable functions, including Orlicz space, Lorentz spaces, weighted $L^p$-spaces, and so on. In particular, when $X$ is rearrangement invariant, we shall show that (ii) of Theorem A implies that $\|E[f \mid \mathcal{F}_n] - E[f \mid \mathcal{F}]\|_X \to 0$ for all $f \in X$ if and only if $X$ has an absolutely continuous norm. For example, this is the case if $X$ is the Orlicz space $L^\Phi$ with a Young function $\Phi$ satisfying the $\Delta_2$-condition, or Lorentz space $L^{p,q}$ with $1 < p < \infty$ and $1 \leq q < \infty$. We shall consider also the convergence of conditional expectations in weighted $L^p$-spaces in the last section. It will be proved that the condition $A_p$, introduced in Izumisawa and Kazamaki [6], is essential.

2. PRELIMINARIES

Let $(X, \| \cdot \|_X)$ be a Banach space (of equivalence classes) of measurable functions on $\Omega$. (Throughout this note, we shall consider Banach spaces over the real-number field.) $X$ is said to be a **Banach function space** if it has the following properties:

(i) $L^\infty \hookrightarrow X \hookrightarrow L^1$;

(ii) if $|f| \leq |g|$ a.s. and $g \in X$, then $f \in X$ and $\|f\|_X \leq \|g\|_X$;

(iii) if $0 \leq f_n \uparrow f$ a.s. and $\sup_n \|f_n\|_X < \infty$, then $f \in X$ and $\|f\|_X = \sup_n \|f_n\|_X$.

Property (iii) is called the **Fatou property**. In fact, (iii) gives the inequality

$$\left\| \liminf_n f_n \right\|_X \leq \liminf_n \|f_n\|_X$$

for any sequence $\{f_n\}$ of measurable functions. For the general theory of Banach function spaces, we refer to Chap. 1 of [2].

Let $f$ be a measurable function on $\Omega$. The **decreasing rearrangement** of $f$, denoted by $f^*$, is the decreasing function on the unit interval $I = [0, 1]$ given by

$$f^*(t) = \inf\{\lambda > 0 : \mu(|f| > \lambda) \leq t\}, \quad t \in I.$$ 

Then $|f|$ and $f^*$ have the same distribution. Let $f$ and $g$ be two measurable functions on $\Omega$. We write $f < g$ if

$$\int_0^t f^*(s) \, ds \leq \int_0^t g^*(s) \, ds, \quad t \in I.$$
A Banach function space $X$ is said to be **rearrangement invariant** (r.i.) if $X$ satisfies the condition:

(iv) if $g \in X$ and $f^* = g^*$, then $f \in X$ and $\|f\|_X = \|g\|_X$.

Furthermore $X$ is said to be **universally rearrangement invariant** (u.r.i.) if $X$ satisfies the condition:

(v) if $g \in X$ and $f < g$, then $f \in X$ and $\|f\|_X \leq \|g\|_X$.

If $X$ is u.r.i., then it is r.i., and the converse is true when $\Omega$ contains no $\mu$-atom. Luxemburg [8] proved that $X$ is u.r.i. if and only if there exists a r.i. function space $\tilde{X}$ with the underlying measure space $(I, dt)$ such that $\|f\|_X = \|f^*\|_{\tilde{X}}$ for every $f \in X$. (Note that the original definition of universal rearrangement invariance in [8] is different from ours, cf. [2, p. 90].) For instance, $L^p$-spaces, Orlicz spaces, and Lorentz spaces are r.i.

Now let $X$ be an arbitrary Banach function space and $f \in X$. We say that $f$ has an **absolutely continuous norm in** $X$, if $A_n \in \mathcal{A}$, $A_n \downarrow \emptyset$ a.s. implies $\|f 1_{A_n}\|_X \rightarrow 0$. If every $f \in X$ has an absolutely continuous norm, we say that $X$ itself has an absolutely continuous norm or that the norm of $X$ is absolutely continuous. The following lemma will be used frequently in what follows.

**Lemma 1.** Let $X$ be a Banach function space and suppose that $f \in X$ has an absolutely continuous norm. Then:

(i) If $A_n \in \mathcal{A}$ and $\mu(A_n) \rightarrow 0$, then $\|f 1_{A_n}\|_X \rightarrow 0$.

(ii) If $|f_n| \leq |f|$ a.s. and $f_n \rightarrow f$ a.s., then $\|f_n - f\|_X \rightarrow 0$.

For the proof, see [2, pp. 15-16].

Let $X$ be an arbitrary Banach function space again. The **associate space** of $X$, denoted by $X'$, is given by

$$
\|f\|_{X'} := \sup \left\{ \int_{\Omega} |fg| \, d\mu : g \in X, \|g\|_X \leq 1 \right\};
$$

$$
X' := \{f \in L^1 : \|f\|_{X'} < \infty \}.
$$

The associate space is again a Banach function space. Every Banach function space $X$ is isometrically isomorphic to its second associate space $X''$, cf. [2, p. 10].

3. CONVERGENCE IN BANACH FUNCTION SPACES

Let $X$ and $Y$ be two Banach function spaces and $\mathcal{F} = \{\mathcal{F}_n\}_{n=1}^\infty$ a sequence of sub-$\sigma$-fields of $\mathcal{A}$ such that $\mathcal{F}_n \uparrow \mathcal{F}$ and $\mathcal{F}_n \downarrow \mathcal{G}$ for some $\mathcal{G}$. 
In this section, we shall give a sufficient condition which assures the convergence of \( E[f | \mathcal{F}_n] \) in \( Y \) for every \( f \in X \). It is natural to assume that, for each \( n \geq 1 \), \( E[f | \mathcal{F}_n] \in Y \) whenever \( f \in X \) for our aim. If we assume this, \( E[\cdot | \mathcal{F}_n] \) must be a bounded linear operator on \( X \) into \( Y \) by the closed graph theorem, since \( X \rightarrow L^1 \). Thus we shall assume the boundedness of \( E[\cdot | \mathcal{F}_n] \). Needless to say, this is the case if \( X = Y = L^p \), and more generally, if \( X = Y \) and \( X \) is rearrangement invariant: see Corollary 1. In general, however, not every conditional expectation operator is bounded on \( X \) even in the case \( X = Y \). To see this, let \( \Omega \) be the probability space \( I \times I = [0,1] \times [0,1] \) with Lebesgue measure \( \mu \), \( \mathcal{F} \) the \( \sigma \)-field consisting of all sets of the form \( A \times I \), where \( A \) is Lebesgue measurable, and \( X \) a Banach function space with the norm given by

\[
\|f\|_X = \left( \int_{I \times [0,a]} |f| \, d\mu + \left( \int_{I \times [a,1]} |f|^2 \, d\mu \right)^{1/2} \right),
\]

where \( 0 < a < 1 \) is a constant. Choose \( \varphi \in L^1(I) \setminus L^2(I) \) and set \( f(s, t) = \varphi(s)1_{I \times [0,a]}(s, t) \). Then we have \( f \in X \), while \( E[f | \mathcal{F}] \notin X \).

Let \( B_1 \) and \( B_2 \) be Banach spaces. We denote by \( \mathcal{L}(B_1, B_2) \) the Banach space consisting of all bounded linear operators on \( B_1 \) into \( B_2 \). We also write \( \mathcal{L}(B_1) \) instead of \( \mathcal{L}(B_1, B_1) \).

**Theorem 1.** Let \( X \) and \( Y \) be Banach function spaces and \( \mathcal{F} = \{ \mathcal{F}_n \}_{n=1}^\infty \) be a sequence of sub-\( \sigma \)-fields of \( \mathcal{F} \). Assume that both \( X \) and \( Y \) have an absolutely continuous norm.

(a) If \( E[\cdot | \mathcal{F}_n] \in \mathcal{L}(X, Y) \) for all \( n \geq 1 \) and \( \sup_n \|E[\cdot | \mathcal{F}_n]\|_{\mathcal{L}(X, Y)} < \infty \), then the following conditions are equivalent:

(i) \( \mathcal{F}_n \downarrow \mathcal{F} \) and \( \mathcal{F}_n \uparrow \mathcal{F} \);

(ii) \( E[f | \mathcal{F}_n] \rightarrow E[f | \mathcal{F}] \) in probability for every \( f \in L^1 \);

(iii) \( E[\cdot | \mathcal{F}] \in \mathcal{L}(X, Y) \) and \( E[f | \mathcal{F}_n] \rightarrow E[f | \mathcal{F}] \) in \( Y \) for every \( f \in X \).

(b) If \( E[\cdot | \mathcal{F}_n] \in \mathcal{L}(X, Y) \) for all \( n \geq 1 \) and \( \{E[f | \mathcal{F}_n]\} \) converges in \( Y \) for any \( f \in X \), then \( \sup_n \|E[\cdot | \mathcal{F}_n]\|_{\mathcal{L}(X, Y)} < \infty \) and one of (and hence all of) (i)–(iii) is true for some \( \mathcal{F} \).

Recall that, when \( \{f_n\} \) is a uniformly integrable sequence of measurable functions, \( \{f_n\} \) converges in probability if and only if it converges in \( L^1 \) (cf., e.g., [10, p. 116]). Hence the equivalence between (i) and (ii) follows from Theorem A.

**Remark 1.** We cannot remove the hypothesis \( \sup_n \|E[\cdot | \mathcal{F}_n]\|_{\mathcal{L}(X, Y)} < \infty \) in (a) of Theorem 1. In other words, the condition \( E[\cdot | \mathcal{F}_n] \in \mathcal{L}(X, Y) \),
$n \geq 1$ does not imply the boundedness of these conditional expectations in $\mathcal{L}(X, Y)$, even in the case where $X = Y$ and $\mathcal{F}$ is increasing. To see this, let $\Omega$ be the probability space $I = [0, 1]$ with Lebesgue measure $dt$. For each integer $n \geq 1$, let $\mathcal{F}_n$ denote the $\sigma$-field generated by the set $A_n = \{0, \frac{1}{n}\} \cup \{1 - \frac{1}{n}, 1\}$ and the Lebesgue measurable subsets of $\Omega \setminus A_n$.

Then $\mathcal{F}_n$ increases with $n$ and therefore $\mathcal{F}_n \overset{dt}{\to} \mathcal{F}$ and $\mathcal{F}_n \overset{\mathcal{A}}{\to} \mathcal{A}$, where $\mathcal{A}$ denotes the $\sigma$-field of Lebesgue measurable sets in $\Omega$. Let $1 \leq p < \infty$ and define $(X, \| \cdot \|_X)$ by

$$
\|f\|_X = \left(\int_0^1 |f(t)|^p \, dt\right)^{1/p};
$$

$$
X = \{f \in L'(\Omega, \mathcal{A}, dt) : \|f\|_X < \infty\}.
$$

Clearly $X$ is a Banach function space over $(\Omega, \mathcal{A}, dt)$. An elementary calculation shows that $E[f \mid \mathcal{F}_n] \in \mathcal{L}(X)$, $n \geq 1$, while $\sup_n \|E[\cdot \mid \mathcal{F}_n]\|_{\mathcal{L}(X)} = \infty$, since $\|E[1_{[1-(1/n), 1]} \mid \mathcal{F}_n]\|_X / \|1_{[1-(1/n), 1]}\|_X \to \infty$ as $n \to \infty$. Furthermore if we set

$$
f(t) = \begin{cases} (1 - t)^{-1/2p}, & \text{if } \frac{1}{2} \leq t \leq 1, \\ 0, & \text{otherwise,} \end{cases}
$$

then $\|E[f \mid \mathcal{F}_n] - f\|_X \geq \|E[f \mid \mathcal{F}_n]1_{[0, 1/n]}\|_X = p(2p - 1)^{-1}$ and hence $\|E[f \mid \mathcal{F}_n] - f\|_X \to 0$.

**Proof of Theorem 1.** (a) As mentioned above, Theorem A shows that (i) and (ii) are equivalent. Assume (iii). It is clear from the fact $L^\infty \hookrightarrow X, Y \hookrightarrow L^1$ that $E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{F}]$ in $L^1$ for every $f \in L^\infty$. This remains valid for every $f \in L^1$, since $L^\infty$ is dense in $L^1$ and

$$
\|E[f \mid \mathcal{F}_n] - E[f \mid \mathcal{F}]\|_1 \leq 2\|f - g\|_1 + \|E[g \mid \mathcal{F}_n] - E[g \mid \mathcal{F}]\|_1
$$

holds for any $g \in L^\infty$; thus (iii) $\Rightarrow$ (ii).

Now we prove (ii) $\Rightarrow$ (iii). Suppose first that $f \in L^\infty$. For the sake of simplicity, we write $f_n = E[f \mid \mathcal{F}_n]$ and $f_\infty = E[f \mid \mathcal{F}]$. For given $\varepsilon > 0$ and $g \in Y'$ with $\|g\|_{Y'} \leq 1$, we have

$$
\int_\Omega |f_n - f_\infty| \|g\| \, d\mu = \left(\int_{\{f_n - f_\infty\} > \varepsilon} + \int_{\{f_n - f_\infty\} \leq \varepsilon}\right)|f_n - f_\infty| \|g\| \, d\mu 
$$

$$
\leq 2\|f\|_\infty \int_{\{f_n - f_\infty\} > \varepsilon}\|g\| \, d\mu + \varepsilon \|g\|_1.
$$

As $Y \hookrightarrow L^1$, it follows that

$$
\|f_n - f_\infty\|_Y = \sup_{\|g\|_{Y'} \leq 1} \int_\Omega |f_n - f_\infty| \|g\| \, d\mu \leq 2\|f\|_\infty \|1_{\{f_n - f_\infty\} > \varepsilon}\|_Y + c\varepsilon,
$$

where $c$ is a constant depending only on $\varepsilon$.
where \( c > 0 \) is a constant. The norm of \( 1_{\{|f_n - f| > \varepsilon\}} \) on the right-hand side tends to zero by (ii) and Lemma 1 (i). Since \( \varepsilon > 0 \) is arbitrary, we have

\[
\|E[f | \mathcal{F}_n] - E[f | \mathcal{F}]\|_Y \to 0,
\]

provided \( f \in L^\infty \).

Next let \( f \in X \) be arbitrary. From (ii) we see that there exists a subsequence \( \{E[f | \mathcal{F}_{n_k}]\}^\infty_{k=1} \) which converges to \( E[f | \mathcal{F}] \) a.s. It follows from the Fatou property that

\[
\|E[f | \mathcal{F}]\|_Y \leq \limsup_{k \to \infty} \|E[f | \mathcal{F}_{n_k}]\|_Y \leq C\|f\|_X
\]

where \( C = \sup_n \|E[\cdot | \mathcal{F}_n]\|_{\mathcal{F}(X,Y)} \), and hence that \( E[\cdot | \mathcal{F}] \in \mathcal{L}(X,Y) \) and \( \|E[\cdot | \mathcal{F}]\|_{\mathcal{F}(X,Y)} \leq C \). Note that \( L^\infty \) is dense in \( X \) by Lemma 1. Let \( \varepsilon > 0 \) and choose \( g \in L^\infty \) so that \( \|f - g\|_X < \varepsilon \); then we have

\[
\|E[f | \mathcal{F}] - E[f | \mathcal{F}]\|_Y \\
\leq \|E[(f - g) | \mathcal{F}_n]\|_Y + \|E[g | \mathcal{F}_n] - E[g | \mathcal{F}]\|_Y \\
+ \|E[(f - g) | \mathcal{F}]\|_Y \\
\leq 2C\varepsilon + \|E[g | \mathcal{F}_n] - E[g | \mathcal{F}]\|_Y.
\]

Since \( E[g | \mathcal{F}_n] \) tends to \( E[g | \mathcal{F}] \) in \( Y \), as proved above, and \( \varepsilon > 0 \) is arbitrary, we obtain (iii).

(b) Suppose that \( E[\cdot | \mathcal{F}_n] \in \mathcal{L}(X,Y) \) for all \( n \geq 1 \) and \( \{E[f | \mathcal{F}_n]\}^\infty_{n=1} \) converges in \( Y \) for any \( f \in X \). Then the Banach–Steinhaus theorem shows that \( \sup_n \|E[\cdot | \mathcal{F}_n]\|_{\mathcal{F}(X,Y)} < \infty \). Let \( T \) denote the operator given by

\[
Tf = \lim_{n \to \infty} E[f | \mathcal{F}_n] \text{ in } Y, \quad f \in X.
\]

We must prove that \( T \) is a conditional expectation operator. To this end, it suffices to show that \( T \) extends to an idempotent linear contraction on \( L^1 \) (see [9, p. 14]). A standard argument shows that \( T \) extends to a linear contraction on \( L^1 \) and \( Tf = \lim_n E[f | \mathcal{F}_n] \) in \( L^1 \), since \( Y \hookrightarrow L^1 \) and \( X \) is dense in \( L^1 \). Furthermore we have \( T(Tf) = Tf \), because

\[
\|E[f | \mathcal{F}_n] - E[Tf | \mathcal{F}_n]\|_1 = \|E[E[f | \mathcal{F}_n] - Tf | \mathcal{F}_n]\|_1 \\
\leq \|E[f | \mathcal{F}_n] - Tf\|_1 \to 0 \quad (n \to \infty).
\]

This completes the proof.  

**COROLLARY 1.** Suppose that \( X \) is universally rearrangement invariant (or rearrangement invariant if \( \Omega \) contains no \( \mu \)-atom) and \( X \) has an absolutely
CONDITIONAL EXPECTATIONS

continuous norm. Then \( E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}] \) in \( X \) for every \( f \in X \) if and only if \( \mathcal{F}_n \uparrow \mathcal{G} \) and \( \mathcal{F}_n \downarrow \mathcal{G} \).

**Proof.** In view of Theorem 1, it suffices to show that every conditional expectation operator is a linear contraction on \( X \) into itself. But this is an immediate consequence of Calderón’s work [4, Theorem 3], provided that \( \Omega \) contains no \( \mu \)-atom.

A more lucid way is to use the following formula

\[
\int_0^t f^*(s) \, ds = \inf \{ \| f_1 \|_1 + t \| f_2 \|_\infty : f = f_1 + f_2, f_1 \in L^1, f_2 \in L^\infty \}, \tag{1}
\]

which is valid for every \( f \in L^1 \). For the proof, see, e.g., [2, p. 74]. From (1) we see that if \( f \in X, \ f = f_1 + f_2, f_1 \in L^1 \) and \( f \in L^\infty \), then

\[
\int_0^t E[f \mid \mathcal{F}]^*(s) \, ds \leq \| E[f_1 \mid \mathcal{F}] \|_1 + t \| E[f_2 \mid \mathcal{F}] \|_\infty \\
\leq \| f_1 \|_1 + t \| f_2 \|_\infty
\]

holds for any sub-\( \sigma \)-field \( \mathcal{F} \) of \( \mathcal{G} \). Taking the infimum of the right-hand side, we get \( E[f \mid \mathcal{F}] \prec f \), and therefore \( \| E[f \mid \mathcal{F}] \|_X \leq \| f \|_X \), since \( X \) is u.r.i. This completes the proof. \( \blacksquare \)

In Corollary 1 the hypothesis of absolute continuity of the norm of \( X \) is essential. In fact, Fetter [5] gave an example of a bounded martingale which does not converge in \( L^\infty \). More generally we have:

**THEOREM 2.** Suppose that \( \Omega \) contains no \( \mu \)-atom. If \( X \) is a rearrangement invariant function space, the following conditions are equivalent:

(i) if \( \mathcal{F}_n \uparrow \mathcal{G} \) and \( \mathcal{F}_n \downarrow \mathcal{G} \), then \( E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}] \) in \( X \) for every \( f \in X \);

(ii) \( X \) has absolutely continuous norm.

**Remark 2.** As we shall see below, we can replace (i) by

(i') if \( \mathcal{F}_n \uparrow \mathcal{G} \), then \( E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}] \) in \( X \) for every \( f \in X \).

That is to say, when \( \Omega \) contains no atom, \( X \) has an absolutely continuous norm if and only if every uniformly integrable martingale \( (f_n)_{1 \leq n \leq \infty} \) such that \( f_\infty \in X \) converges in \( X \). For example, if \( \varphi \) is an increasing concave function such that \( \varphi(0 + ) > 0 \), then the Lorentz space \( M(\varphi) \) (see [2, p. 69]) is a r.i. space which does not have an absolutely continuous norm. Hence there exists a uniformly integrable martingale \( (f_n) \) which is not convergent in \( M(\varphi) \) and such that \( f_\infty \in M(\varphi) \).
To prove Theorem 2, we shall use the following elementary lemma.

**Lemma 2.** Suppose that Ω contains no μ-atom. If \( A, B \in \mathcal{A} \) and \( \mu(A) = \mu(B) > 0 \), then for every measurable function \( f \), there is another measurable function \( g \) such that

\[
\mu\{\omega \in A : |f(\omega)| > \lambda\} = \mu\{\omega \in B : |g(\omega)| > \lambda\}, \quad \lambda > 0. \tag{2}
\]

**Proof.** The assertion is easily proved by the fact that, if Ω contains no atom, then every right continuous decreasing function on \( I = [0,1] \) is the decreasing rearrangement of a function on Ω. For the proof, see [3, p. 44].

Let \( A, B \) and \( f \) be as in the statement. By the fact mentioned above, there exits a measurable function \( g \) such that \( f_A^* = g_B^* \), where \( f_A^* \) and \( g_B^* \) denote the decreasing rearrangement of \( f \) and \( g \), with respect to the conditional probability measures \( \mu(\cdot|A) \) and \( \mu(\cdot|B) \), respectively. In other words, \( |f| \) and \( |g| \) have the same distribution with respect to \( \mu(\cdot|A) \) and \( \mu(\cdot|B) \), respectively. This implies (2), since \( \mu(A) = \mu(B) \).

**Proof of Theorem 2.** Corollary 1 shows that (ii) implies (i). To prove the converse, suppose that (ii) is false; that is, there is an \( f \in X \) which is not of an absolutely continuous norm in \( X \). Then we can find a decreasing sequence \( \{A_n\}_{n=1}^\infty \) of sets in \( \mathcal{A} \) such that \( A_n \downarrow \emptyset \) a.s. and

\[
\|f\|_{A_n} \geq \varepsilon
\]

for every \( n \geq 1 \) with some \( \varepsilon > 0 \). We may assume that \( f \geq 0 \) a.s., \( \{f > 0\} \subset A_1 \) and \( \mu(A_1) \leq 1/2 \). As Ω contains no μ-atom, there is a decreasing sequence \( \{B_n\}_{n=1}^\infty \) of sets in \( \mathcal{A} \) such that \( A_1 \cap B_1 = \emptyset \) and \( \mu(A_n) = \mu(B_n) \) for all \( n \geq 1 \) (cf. [3, p. 44]). Using Corollary 2, choose non-negative measurable functions \( g_n \) so that

\[
\mu\{\omega \in A_n \setminus A_{n+1} : f(\omega) > \lambda\} = \mu\{\omega \in B_n \setminus B_{n+1} : g_n(\omega) > \lambda\}, \quad \lambda > 0.
\]

Define the function \( g \) by \( g = \sum_{n=1}^\infty g_n \mathbf{1}_{B_n \setminus B_{n+1}} \) on Ω. Then \( f \) and \( g \) have the same distribution: for each \( \lambda > 0 \)

\[
\mu\{\omega \in \Omega : f(\omega) > \lambda\} = \sum_{n=1}^\infty \mu\{\omega \in A_n \setminus A_{n+1} : f(\omega) > \lambda\} = \sum_{n=1}^\infty \mu\{\omega \in B_n \setminus B_{n+1} : g_n(\omega) > \lambda\} = \mu\{\omega \in \Omega : g(\omega) > \lambda\}.
\]
Therefore we have \( g \in X \), since \( X \) is r.i. In the same way we see that \( f1_{A_n} \) and \( g1_{B_n} \) have the same distribution, and hence that
\[
\int_{A_n} f \, d\mu = \int_{B_n} g \, d\mu, \quad n \geq 1. \tag{4}
\]

Let \( \mathcal{T}_n \) denote the \( \sigma \)-field generated by \( A_n \cup B_n \) and the measurable subsets of \( \Omega \setminus (A_n \cup B_n) \); then \( A_n \cup B_n \) is a single \( \mu \)-atom for \( \mathcal{T}_n \). Since \( A_n \cup B_n \) \( \downarrow \emptyset \) a.s., \( \mathcal{T}_n \) increases with \( n \) and \( \mathcal{A} = \bigvee_{n=1}^{\infty} \mathcal{T}_n \). It follows from the martingale convergence theorem and Theorem A that \( \mathcal{T}_n \uparrow \mathcal{A} \) and \( \mathcal{T}_n \downarrow \mathcal{A} \). Let \( h_n = f-g \) and \( h = (h_n)_{1 \leq n \leq \infty} \) denote the martingale induced by \( h_n \): \( h_n = E[h_n \mid \mathcal{T}_n] \), \( n \geq 1 \). Since \( \{f > 0\} \subset A_1 \), \( \{g > 0\} \subset B_1 \) and \( A_1 \cap B_1 = \emptyset \), we have by (4)
\[
\begin{align*}
h_n &= \frac{1_{A_n \cup B_n}}{\mu(A_n \cup B_n)} \int_{A_n \cup B_n} h_\infty \, d\mu + h_\infty 1_{\Omega \setminus (A_n \cup B_n)} \\
&= \frac{1_{A_n \cup B_n}}{\mu(A_n \cup B_n)} \left( \int_{A_n} f \, d\mu - \int_{B_n} g \, d\mu \right) + h_\infty 1_{\Omega \setminus (A_n \cup B_n)} \\
&= h_\infty 1_{\Omega \setminus (A_n \cup B_n)} .
\end{align*}
\]

It follows that
\[
|h_n - h_\infty| = |h_\infty| 1_{A_n \cup B_n} = f1_{A_n} + g1_{B_n} \geq f1_{A_n} \text{ a.s.},
\]
and therefore, from (3) we obtain \( \|h_n - h_\infty\|_X \geq \|f1_{A_n}\|_X \geq \varepsilon \) for every \( n \geq 1 \). This shows that (i) is false; thus (i) \( \Rightarrow \) (ii). □

4. CONVERGENCE IN WEIGHTED \( L^p \)-SPACES

In this section, we shall consider the convergence in weighted \( L^p \)-spaces. Let \( u \) and \( v \) be positive integrable functions on \( \Omega \). For \( 1 \leq p < \infty \), we denote by \( L^p(u) \) (resp., \( L^p(v) \)) the \( L^p \)-space relative to the measure \( u \, d\mu \) (resp., \( v \, d\mu \)); that is, the norm of \( L^p(u) \) is given by
\[
\|f\|_{L^p(u)} = \left( \int_{\Omega} |f|^p u \, d\mu \right)^{1/p} .
\]

It follows from Hölder's inequality that if \( 1 < p < \infty \) and \( u^{-1/(p-1)} \in L^1 \), then \( L^\infty \rightarrow L^p(u) \rightarrow L^1 \), where \( L^1 \) and \( L^\infty \) are the spaces of integrable and essentially bounded functions with respect to \( \mu \). Thus \( L^p(u) \) is a Banach function space over \( (\Omega, \mathcal{A}, \mu) \). As for \( p = 1 \), \( L^1(u) \) is a Banach function space, provided \( u^{-1} \in L^\infty \).
Let $\mathcal{F} = \{\mathcal{F}_n\}_{n=1}^\infty$ be a sequence of sub-$\sigma$-fields of $\mathcal{A}$. For $1 < p < \infty$ we say that the pair $(u, v)$ satisfies $A_p$ (with respect to $\mathcal{F}$) and write $(u, v) \in A_p$ (or $(u, v) \in A_p(\mathcal{F})$) if

$$\sup_{n \geq 1} E\left[ v^{-1/(p-1)} \mid \mathcal{F}_n \right] E\left[ u \mid \mathcal{F}_n \right]^{1/(p-1)} \leq K \text{ a.s.},$$

(A$_p$)

where $K > 0$ is a constant. When $p = 1$, the pair is said to satisfy $A_1$ and written as $(u, v) \in A_1(= A_1(\mathcal{F}))$ if

$$\sup_{n \geq 1} v^{-1} E\left[ u \mid \mathcal{F}_n \right] \leq K \text{ a.s.}$$

(A$_1$)

Condition $A_p$ was introduced by Izumisawa and Kazamaki [6], in the case where $u = v$ and $\mathcal{F} = \{\mathcal{F}_n\}$ is increasing in $n$, for their study of the weighted Doob inequality. Their work was the trigger that brought the successive works on the weight theory for martingales, such as Uchiyama [12], Sekiguchi [11], Kazamaki and Kikuchi [7], and so on. We shall show that $A_p$ is essential for the convergence of conditional expectations in the weighted $L^p$-space.

**Theorem 3.** Let $u$ and $v$ be positive integrable functions satisfying $u^{-1/(p-1)} \in L^1$, $v^{-1/(p-1)} \in L^1$ or $u^{-1} \in L^\infty$, $v^{-1} \in L^\infty$, according as $1 < p < \infty$ or $p = 1$. Suppose that $\mathcal{F} = \{\mathcal{F}_n\}_{n=1}^\infty$ is a sequence of sub-$\sigma$-fields of $\mathcal{A}$ such that $\mathcal{F}_n \uparrow \mathcal{G}$ and $\mathcal{F}_n \downarrow \mathcal{G}$ for some $\mathcal{G}$. Then the following conditions are equivalent:

(i) if $f \in L^p(v)$, then $E[f \mid \mathcal{F}_n]$, $E[f \mid \mathcal{G}] \in L^p(u)$ and $E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}]$ in $L^p(u)$;

(ii) $(u, v) \in A_p(\mathcal{F})$.

In view of Theorem 1, (i) is true if and only if

$$\sup_n \|E[\cdot \mid \mathcal{F}_n]\|_{L^p(v), L^p(u)} < \infty.$$ 

Therefore the above theorem immediately follows from the next lemma.

**Lemma 3.** If $u$ and $v$ are as in Theorem 3 and $\mathcal{F}$ is a sub-$\sigma$-field of $\mathcal{A}$, then the following conditions are equivalent:

(i) $E[\cdot \mid \mathcal{F}] \in L(L^p(v), L^p(u))$ and the norm of $E[\cdot \mid \mathcal{F}]$ does not exceed $K$;

(ii) $E[v^{-1/(p-1)} \mid \mathcal{F}]^{p-1} E[u \mid \mathcal{F}] \leq K^p$ or $v^{-1} E[u \mid \mathcal{F}] \leq K$, according as $1 < p < \infty$ or $p = 1$. 

Suppose that \(1 < p < \infty\) and (i) is true; then
\[
E\left[ E[f \mid \mathcal{F}]^p u \right] \leq K^p E[f^p u]
\] (5)
holds for every non-negative \(f\). Setting \(f = v^{-1/(p-1)}1_{(v > e) \cap A}\) in (5), where \(A \in \mathcal{F}\) and \(e > 0\), we get
\[
E\left[ E\left[ v^{-1/(p-1)}1_{(v > e)} \mid \mathcal{F} \right]^p E[u \mid \mathcal{F}] 1_A \right]
= E\left[ E\left[ v^{-1/(p-1)}1_{(v > e) \cap A} \mid \mathcal{F} \right]^p u \right]
\leq K^p E\left[ v^{-1/(p-1)}1_{(v > e) \cap A} \right].
\]
As \(A \in \mathcal{F}\) is arbitrary, it follows that
\[
E\left[ v^{-1/(p-1)}1_{(v > e)} \mid \mathcal{F} \right]^p E[u \mid \mathcal{F}] \leq K^p E\left[ v^{-1/(p-1)}1_{(v > e)} \mid \mathcal{F} \right],
\]
and hence that
\[
E\left[ v^{-1/(p-1)}1_{(v > e)} \mid \mathcal{F} \right]^{p-1} E[u \mid \mathcal{F}] \leq K^p.
\]
Letting \(e \downarrow 0\), we obtain \(E[u^{-1/(p-1)}] \mid \mathcal{F}]^{p-1} E[u \mid \mathcal{F}] \leq K^p\), as desired.

Next, suppose that \(p = 1\) and (i) is true, that is, that
\[
E\left[ E[f \mid \mathcal{F}] u \right] \leq E[f v]
\]
holds for every non-negative \(f\). Setting \(f = v^{-1}1_A(\in L^\infty)\) for \(A \in \mathcal{F}\), we have
\[
E\left[ v^{-1} E[u \mid \mathcal{F}] 1_A \right] = E\left[ E\left[ v^{-1}1_A \mid \mathcal{F} \right] u \right] \leq KE[1_A],
\]
which implies that \(v^{-1} E[u \mid \mathcal{F}] \leq K\). Thus we have proved that (i) implies (ii). Now assume that \(1 < p < \infty\) and (ii) is true. Then Hölder’s inequality gives that, for every \(f \in L^p(v)\),
\[
|E[f \mid \mathcal{F}]|^p E[u \mid \mathcal{F}] \leq E[|f| v^{1/p} v^{-1/p} \mid \mathcal{F}]^p E[u \mid \mathcal{F}]
\]
\[
\leq E[|f| v^{1/p} \mid \mathcal{F}] E[v^{-1/(p-1)} \mid \mathcal{F}]^{p-1} E[u \mid \mathcal{F}]
\]
\[
\leq K^p E[|f| v^{1/p} \mid \mathcal{F}].
\]
It follows that
\[
E[|E[f \mid \mathcal{F}]|^p u] = E[|E[f \mid \mathcal{F}]|^p E[u \mid \mathcal{F}]] \leq K^p E[|f| v^{1/p}],
\]
and therefore \(\|E[f \mid \mathcal{F}]\|_{L^p(v)} \leq K \|f\|_{L^p(v)}\), which gives (i).
If $p = 1$ and $\nu^{-1}E[u | \mathcal{F}] \leq K$, then

$$|E[f | \mathcal{F}]|E[u | \mathcal{F}] \leq E[|f|\nu^{-1}E[u | \mathcal{F}] | \mathcal{F}] \leq KE[|f|\nu | \mathcal{F}].$$

This gives that $\|E[f | \mathcal{F}]\|_{L^1(u)} \leq K\|f\|_{L^1(\nu)}$. In any case, (ii) implies (i). The lemma is established.

APPENDIX

We shall prove the equivalence between (i) and (ii) of Theorem A.

Step 1. If $\mathcal{F}_n \uparrow \mathcal{F}$, each $f_n \in L^\infty$ is $\mathcal{F}_n$ measurable and $\sup_n \|f_n\|_\infty < \infty$, then $f_n - E[f_n | \mathcal{F}] \to 0$ weakly in $L^2$.

We may assume that $\sup_n \|f_n\|_\infty \leq 1$. For each pair of integers $n \geq 1$ and $k \geq 1$, define the function $f_{n,k}$ by

$$f_{n,k} = \sum_{j=-k}^k \frac{1}{j/k} 1_{j/k \leq f_n < (j+1)/k}.$$

Clearly each $f_{n,k}$ is a simple $\mathcal{F}_n$-measurable function and $\|f_{n,k} - f_n\|_\infty \leq 1/k$ for all $n \geq 1$. Since $\mathcal{F}_n \uparrow \mathcal{F}$, we see that $f_{n,k} - E[f_{n,k} | \mathcal{F}] \to 0$ weakly in $L^2$ as $n \to \infty$. For every $g \in L^2$ we have

$$\left| \int_{\Omega} (f_n - E[f_n | \mathcal{F}])g d\mu \right|$$

$$\leq \left| \int_{\Omega} (f_n - f_{n,k})g d\mu \right| + \left| \int_{\Omega} (f_{n,k} - E[f_{n,k} | \mathcal{F}])g d\mu \right|$$

$$+ \left| \int_{\Omega} E[f_{n,k} - f_n | \mathcal{F}]g d\mu \right|$$

$$\leq 2\|f_{n,k} - f_n\|_2\|g\|_2 + \left| \int_{\Omega} (f_{n,k} - E[f_{n,k} | \mathcal{F}])g d\mu \right|$$

Letting $n \to \infty$, we get

$$\lim_{n \to \infty} \left| \int_{\Omega} (f_n - E[f_n | \mathcal{F}])g d\mu \right| \leq 2\|g\|_2/k,$$

for any $k \geq 1$, and hence the assertion.

Step 2. If $\mathcal{F}_n \uparrow \mathcal{F}$ and $\mathcal{F}_n \not\uparrow \mathcal{F}$, then $E[f | \mathcal{F}_n] \to E[f | \mathcal{F}]$ weakly in $L^2$ for every $f \in L^\infty$. 

If \( f \in L^\infty \), then \( \{E[f \mid \mathcal{F}_n]\}_{n=1}^{\infty} \) is bounded in \( L^2 \) and hence there exists a subsequence \( \{E[f \mid \mathcal{F}_{n_k}]\} \) which converges weakly in \( L^2 \). Let \( h \) be the weak limit of \( E[f \mid \mathcal{F}_{n_k}] \). Then Step 1 shows that \( E[f \mid \mathcal{F}_{n_k}] - E[E[f \mid \mathcal{F}_{n_k}] \mid \mathcal{G}] \to 0 \) weakly in \( L^2 \). It then follows that for every \( A \in \mathcal{G} \),

\[
\int_A h \, d\mu = \lim_{k \to \infty} \int_A E[f \mid \mathcal{F}_{n_k}] \, d\mu = \lim_{k \to \infty} \int_A E[f \mid \mathcal{F}_{n_k}] \mid \mathcal{G} \, d\mu
\]

\[
= \lim_{k \to \infty} \int_{\mathcal{G}} E[1_A \mid \mathcal{G}] E[f \mid \mathcal{F}_{n_k}] \, d\mu = \int_{\mathcal{G}} E[h \mid \mathcal{G}] \, d\mu
\]

and therefore \( h = E[h \mid \mathcal{G}] \) is \( \mathcal{G} \)-measurable. Now let \( B \in \mathcal{G} \) and choose a sequence \( \{B_n\} \) so that \( B_n \in \mathcal{F}_n \) for all \( n \geq 1 \) and \( \mu(B \Delta B_n) \to 0 \). This is possible, since \( \mathcal{F}_n \xrightarrow{k} \mathcal{G} \). Then we have

\[
\left| \int_{B_{n_k}} f \, d\mu - \int_B h \, d\mu \right|
\]

\[
\leq \left| \int_{\mathcal{G}} \left( 1_{B_{n_k}} - 1_B \right) E[f \mid \mathcal{F}_{n_k}] \, d\mu \right| + \left| \int_B \left( E[f \mid \mathcal{F}_{n_k}] - h \right) \, d\mu \right|
\]

\[
\leq \|f\|_\infty \mu(B \Delta B_{n_k}) + \int_B \left| E[f \mid \mathcal{F}_{n_k}] - h \right| \, d\mu \to 0, \quad k \to \infty.
\]

Thus \( \int_B f \, d\mu = \lim_k \int_{B_{n_k}} f \, d\mu = \int_B h \, d\mu \), and hence \( E[f \mid \mathcal{G}] = h \); in other words,

\[
E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}] \text{ weakly in } L^2.
\]

Now we prove that \( E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}] \text{ weakly in } L^2 \); suppose conversely that this is false. Then there is a \( g \in L^2 \) and an \( \varepsilon > 0 \) such that

\[
\lim_{n \to \infty} \left| \int_{\mathcal{G}} (E[f \mid \mathcal{F}_n] - E[f \mid \mathcal{G}]) g \, d\mu \right| > \varepsilon.
\]

Select a subsequence \( \{n_j\} \) so that

\[
\left| \int_{\mathcal{G}} (E[f \mid \mathcal{F}_{n_j}] - E[f \mid \mathcal{G}]) g \, d\mu \right| > \varepsilon
\]

for all \( j \). Since \( \mathcal{F}_{n_j} \xrightarrow{k} \mathcal{G} \) and \( \mathcal{F}_{n_j} \xrightarrow{\mathcal{G}} \mathcal{G} \), there exists a subsequence \( \{m_k\} \) of \( \{n_j\} \) such that \( E[f \mid \mathcal{F}_{m_k}] \to E[f \mid \mathcal{G}] \text{ weakly in } L^2 \) by what we have proved above. This contradicts the choice of \( \{n_j\} \), and the assertion of Step 2 is proved.
Step 3. If $E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}]$ weakly in $L^2$ for every $f \in L^\infty$, then the convergence holds in $L^2$.

Let $f \in L^\infty$. Since $E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}]$ and $E[E[f \mid \mathcal{G}] \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}]$ weakly in $L^2$, we have $E[E[f \mid \mathcal{G}] \mid \mathcal{F}_n] - E[f \mid \mathcal{F}_n] \to 0$ weakly in $L^2$. Hence we have

$$
\|E[f \mid \mathcal{F}_n] - E[f \mid \mathcal{G}]\|_2^2 = \langle f, E[f \mid \mathcal{F}_n] - E[E[f \mid \mathcal{G}] \mid \mathcal{F}_n] \rangle \\
+ \langle E[f \mid \mathcal{G}], E[f \mid \mathcal{F}_n] - E[f \mid \mathcal{F}_n] \rangle \to 0,
$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of $L^2$.

Step 4. If $E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}]$ in $L^2$ for every $f \in L^\infty$, then $E[g \mid \mathcal{F}_n] \to E[g \mid \mathcal{G}]$ in $L^p$ for every $g \in L^p$, where $1 \leq p < \infty$.

Suppose that $1 \leq p \leq 2$ and $g \in L^p$. Choosing $g_k \in L^\infty$ so that $0 \leq |g_k| \leq |g|$ and $g_k \to g$ in $L^p$, we have

$$
\|E[g \mid \mathcal{F}_n] - E[g \mid \mathcal{G}]\|_p \leq 2\|g - g_k\|_p + \|E[g_k \mid \mathcal{F}_n] - E[g_k \mid \mathcal{G}]\|_2.
$$

Letting $n \to \infty$ and then $k \to \infty$, we obtain the assertion in the case where $1 \leq p \leq 2$. The extension to the case of $2 < p < \infty$ is standard; see the proof of Theorem 1.

Starting from Step 1, we have proved that (ii) of Theorem A implies (i). Now we prove the converse.

Step 5. If $E[f \mid \mathcal{F}_n] \to E[f \mid \mathcal{G}]$ in $L^1$ for every $f \in L^1$, then $\mathcal{F}_n \uparrow \mathcal{G}$ and $\mathcal{F}_n \downarrow \mathcal{G}$.

Let $A \in \mathcal{G}$ be arbitrary and set $A_n = \{E[1_A \mid \mathcal{F}_n] > 1/2\}$; then $A_n \in \mathcal{F}_n$ for each $n \geq 1$. Since $E[1_A \mid \mathcal{F}_n] \to 1_A$ in probability, we have

$$
\mu(A_n \Delta A) \leq \mu(|E[1_A \mid \mathcal{F}_n] - 1_A| \geq 1/2) \to 0
$$

as $n \to \infty$, and therefore $\mathcal{F}_n \uparrow \mathcal{G}$.

Now suppose that $A_n \in \mathcal{F}_n$ for all $n \geq 1$. If $f \in L^2$, then

$$
\left| \int_{\Omega} (1_{A_n} - E[1_{A_n} \mid \mathcal{G}]) f d\mu \right| = \left| \int_{A_n} (E[f \mid \mathcal{F}_n] - E[f \mid \mathcal{G}]) d\mu \right| \\
\leq \|E[f \mid \mathcal{F}_n] - E[f \mid \mathcal{G}]\|_1 \to 0,
$$

which means that $\mathcal{F}_n \downarrow \mathcal{G}$.

Thus the equivalence between (i) and (ii) of Theorem A is established.
REFERENCES